Proceedings of

the 4th Conference of the European Social Simulation Association

(ESSA'07)

September 10-14, 2007 Toulouse, France

Frédéric Amblard (ed.)

ISBN : 978-2-9520326-7-4

IRIT Editions

Preface

For nearly 15 years, social simulation and computational approaches for understanding social phenomena has reached a rapidly growing audience from very different disciplines. Some researchers are attracted by new tools, methods or even paradigms to understand their object of research as is the case in sociology, economics or psychology. Others have found in simulation an interesting tool with which to communicate, negotiate or to support stakeholders' decision-making. This is the case in some political sciences, in management, and with the participative modelling approach for resource-management. Other researchers, mainly modellers, are interested by the richness of social phenomena and the challenges it presents in trying to capture even a small subset of this complexity with simple models; such is the case for computer scientists or physicists. ESSA'07, the fourth edition of the annual conference of the European Social Simulation Association (http://www.essa.eu.org) following the other editions is a meeting place at the crossroad of all these disciplines and motivations. The articles that will be presented during the conference are at the same at a very valuable level in their own disciplinary domain but are also accessible to a wider audience and bring then inspiration to other scientists.

These proceedings contain the papers presented during this conference held in Toulouse, France, September 10-14, 2007. The increasing popularity of the ESSA conferences enables this year to receive more papers than usual and then to increase the scientific quality of the conference. As an indicator, the selection rate for presentations is nearly 50% this year.

Finally, special thanks are due to the numerous researchers who submitted their best works to ESSA'07, presented a tutorial, reviewed submissions or volunteered their time in any other way.

Frédéric Amblard
Editor

Table of contents

Keynote Speakers

Changing the rules of the game: experiments with humans and virtual agents

Marco A. Janssen

School of Human Evolution and Social Change
Arizona State University
Tempe, USA
Marco.janssen@asu.edu

Many resource problems can be classified as commons dilemmas, a dilemma between the interest of the individual and the interest of the group as a whole. During the last decades substantial progress has been made in understanding how people can avoid the tragedy of the commons. However, we lack good understanding how people change institutional arrangements over time in an effective way in an environment with dynamic resources.

I will discuss the initial results of a project where we look at innovation of institutional arrangements in common pool resource management where we combine laboratory and field experiments with agent-based modeling. In laboratory experiments groups share resources in a dynamic spatially explicit virtual environment, while the pencil and paper field experiments in Colombia and Thailand include various types of resources (fishery, foresty and irrigation). Using the individual level data derived from the experiments we develop and test agent-based models to derive better understanding of the experimental data. We also use the agent-based models to explore the evolution of institutional rules in various contexts that we could not (yet) experiment with.

Going back and forth between experiments with humans and virtual agents is a fruitful way to develop empirically-based agent-based models. I will discuss methodological challenges experienced in this project as well as initial results of the various models.

Social Simulation for Self-Star Systems: An idea whose time has come?

David Hales

Department of Computer Sciences
University of Bologna
Bologna, Italy
dave@davidhales.com

I will talk about, what I believe, are the increasing similarities between open problems in distributed systems engineering and recent research topics in agent-based social simulation. Specifically, engineers of software components in massive (millions of entities) open systems are asking how those components can come to self-organise, self-repair and self-manage in a bottom-up manor: so called self-* systems. Inspiration has been drawn from biological systems and also, but more rarely, social systems. Yet many of the questions agent-based social simulators address are central to key engineering issues in self-* systems:

- Emergence and Self-organisation: Understanding the Micro to Macro AND Macro to Micro link
- Cooperation and Trust: Getting disparate components to "hang-together" even with bad guys around
- Evolving network structures: Constructing and maintaining functional topologies robustly
- Constructing adaptive / evolutionary heuristics rather than rational action models

I will argue that application of the wealth of ideas coming from agent-based social simulation to self-* systems engineering is now, more than ever, a viable research objective. An idea whose time has finally come. I will give an overview of the emerging Self-* research area and indicate areas where social simulation has already had, and may in the near future have, major impact. I will discuss briefly how the engineering and social simulation modeling methodologies may be combined productively. I will give a detailed example of some of my recent work applying novel group selection models for cooperation in peer-to-peer systems. I will also highlight some possible obstacles and worrying developments that could hold back progress in this area and how we, as a scientific community, might tackle them.

Plenary Sessions

Simulation of solidarity in a resource sharing situation

Pieter Bots[1,2], Olivier Barreteau[1], and Géraldine Abrami[1]

[1] Cemagref, Montpellier Regional Centre, BP 5095, F-34033 Montpellier, France
[2] Delft University of Technology, faculty of Technology, Policy and Management, PO Box 5015, NL-2600 GA Delft, The Netherlands
{pieter.bots, olivier.barreteau, geraldine.abrami}@cemagref.fr

Abstract. In this paper we present a first attempt to represent social behavior of actors in a resource sharing context in such a way that different forms of solidarity can be detected and measured. We expect that constructing agent-based models of water-related interactions at the interface of urban and rural areas, and running social simulations to study the occurrence and consequences of solidary behavior will produce insights that may eventually contribute to water and land resource management practice. We propose a typology for solidary behavior, present the agent-based architecture that we are using, show some illustrative results, and formulate some questions that will guide our future work.

Keywords: agent-based model; social simulation; solidarity; typology; values.

1 Introduction

Is it conceivable that a farmer voluntarily gives up irrigation to improve the surface water quality in the city downstream? Or agrees to let his fields be inundated to reduce flood damage in urban areas? And will upstream city dwellers of their own accord invest in a separate system of rain water retention to avoid sewer overflows? Intrigued by these and other questions concerning water-dependent individuals who live at the interfaces between urban and rural areas, we have started to explore the issue of solidarity around water sharing and maintenance of water resources during an interdisciplinary workshop in January 2007 in Montpellier[1]. The participating researchers brought a variety of case studies in very different geographical and political contexts to the table, and a variety of definitions of solidarity as well.

To sort out this diversity of situations of water-related interactions in which some form of solidarity becomes apparent or is called for, we decided to develop a decontextualised simulation model of solidarity with agents seeking to uphold various

[1] This workshop benefited from the participation of Géraldine Abrami, Olivier Barreteau, Bettina Blümling Pieter Bots, Catherine Carré, Flavie Cernesson, Raphaële Ducrot, Katrin Erdlenbruch, Patrice Garin, Patrick Le Goulven, and Cathy Werey. We wish to thank them for numerous contributions to the discussion, some of which provided inspiration for this paper. However, only the authors are responsible for the views and ideas put forward in this paper.

values, living in either an urban area or a rural area, and sharing a resource. The purpose of these models is to allow analysis how different types of solidarity (can) play a role in the dynamics of a system that comprises an urban area and a rural area linked by a shared resource. This analysis is expected to provide useful feedback to the initial interdisciplinary group, allowing the researchers to focus on common questions regarding solidarity about water, questions that need to be answered not only to improve water resource management, but also to strengthen the social cohesion of territories.

This paper presents the first stage of this work: a sketch of an urban perspective and a rural perspective on the role and use of land and water resources, a categorization of different types of solidarity, the So-Si-So modeling platform that we are developing to represent and analyze behavioral patterns of solidarity, and an embryonic version of Solid'Eau, a simulation model that eventually should allow us to study the social behavior of agents located in a system around a shared water resource. The first results obtained with this model show that the So-Si-So model architecture indeed allows identifying and measuring different types of solidarity, and lead to several questions that will guide our future work in this area.

2 Solidarity at the Interface between Urban and Rural Areas

The social phenomena that we are interested in spring from the tension between two opposing perspectives on water issues: an urban perspective and a rural perspective. At a first glance, this tension leads to attempts from one group to impose constraints on, or get compensations from, the other group. Looking more closely, a variety of more specific and refined stakeholder perceptions can be distinguished within each perspective. Moreover, at the level of individuals, the boundary between an urban community and a rural community is not so clear as it is sometimes claimed: there is always a peri-urban area where individuals are difficult to categorize as urban or rural, individuals migrate at various speeds from urban areas to rural areas and vice versa, and they are linked to both types of areas in intricate ways. To emphasize interdependency and the potential for conflict, however, the two perspectives are characterized as follows:

Urban perspective:

- Cities are the centers of modern human society. Their complex social structure is essential for technological, economic and cultural development.
- Water is a necessity of life. Citizens are entitled to adequate drinking water supply and sanitation.
- Rural areas are indispensable for food production, but large scale agriculture reduces the ecological and recreational value of the landscape.
- Rivers are a drinking water supply as well as a discharge channel for rainwater and wastewater, but also a potential threat that needs to be contained by adequate flood prevention measures.

- The natural water system has a high capacity for regeneration, but balances are threatened by the extraction of irrigation water and heavy use of fertilizers and pesticides in agricultural production.

Rural perspective:
- Not only historically, but also today, the modern, human society has its roots in small agricultural communities. Without agriculture no food, and without food no cities.
- Water supports natural vegetation and wildlife as well as agriculture and human life. The traditional rural life style is in harmony with nature and uses water resources in a sustainable way.
- The growth in population, which is strongest in urban areas, puts greater demands on agricultural production as well as on facilities for recreation.
- Urbanization creates a new hydrological environment (buildings and roads replace vegetation and soil, sewers replace stream channels), reducing infiltration and groundwater recharge while increasing runoff and the probability of flooding.
- Urban areas produce high volumes of faeces and other organic waste, heavy metals, and mineral oil products. Although sewage systems and wastewater treatment plants mitigate emission to surface water, high peak discharges generate sewer overflows. Also, industrial and household consumption of electricity causes thermal pollution that threatens aquatic wildlife.

These outlines of the urban and rural perspectives clearly show a number of interdependencies and tensions that underlie many water resource management issues. However, these interdependencies and tensions may not be obvious at the stakeholder level. Any attempt to cope with undesirable impacts from one perspective is likely to generate unexpected feedbacks on other aspects of the other perspective. For example, if an urban group succeeds in imposing constraints on rural water use for agriculture, this will impact on landscapes which are valuable for residents of urban areas.

Given these more or less straightforward physical interdependencies, we want to question whether any social link between both parties exists or can be created or reinforced that could manage these physical interdependencies. Institutions in the sense of Ostrom (1990) may exist that already provide or could establish this social link. But what we are presently even more interested in are loose forms of interaction such as *solidarity.*

Our work is based in the assumption that raising awareness of physical interdependencies *and* social links is a way towards a management of the resource better for the whole system. In this perspective, we think that models are effective tools to reveal and make explicit those situations where the heterogeneity within each group, and the existing relations at the individual level between persons associated with one group or the other is providing opportunities (1) to find institutions to manage the interdependencies, and (2) to find synergies among various perspectives. Meanwhile, we see our modeling work also as a means to achieve a better formalization of solidarity situations and of the solidarity concept itself. This paper reports our progress in this formalization to date.

3 Categories of Solidarity

The capacity to resolve the tensions as water resources become scarce is largely determined by the willingness of stakeholders to share. It is at this point that the concept of solidarity becomes interesting. Solidarity is a loose form of social binding characterized by *latent* reciprocity: the interdependencies need not be explicitly known or understood by those involved. Solidarity is a feeling, rather than a calculated attitude.

On the individual level, solidarity can be defined as "a sense of community between persons who, despite their differences, believe to have the same objective(s) that one has achieved more than the other, from which rises the voluntary obligation to support the other, coupled with the entitlement to support from the other should the situation be reversed" (Hondrich & Koch-Arzberger 1992, p. 14-15). Segall (2005) sees solidarity as a relation between the individual and the collective: "[Social solidarity] comprises the following phenomena:

- Integration – Identification on the part of the individual with the goals and features of the collective (Miller, 1999, p. 26);
- Commitment to the common good – Willingness to forgo self-interest for the sake of the common good (Mason, 1998);
- Empathy – concern for the wellbeing of other members who are less well off than oneself (Cohen and Arato, 1992, p. 472; Mason, 1998);
- Trust – the willingness to suspend suspicion of others, at least until receipt of evidence to the contrary (Misztal, 1996, p. 209; Seligman, 1997, p. 94; Warren, 1999, p. 330)." (Segall 2005, p. 362, the references are his).

Thus, solidarity may be a relation between two individuals or a relation between an individual and the collective that the individual is part of. When considering the (potential) solidarity between urban and rural areas, both relations need to be examined.

In this paper, we focus on solidarity at the individual level, where solidarity involves a voluntary action of person A in support of another person B where A by performing this action knowingly incurs a cost to himself. The definition by Hondrich and Koch-Arzberger is quite precise with regard to the mechanism that leads to solidary action: The sense of community that leads A to support B stems from A's belief that A and B share certain values and objectives. When A furthermore believes to have achieved these objectives more than B, A experiences a *voluntary* obligation to support B, coupled with the *implicit* entitlement to support from B should the situation be reversed. This precision allows us to distinguish the following categories of solidary behavior:

- The action of A is an act of altruism, rather than one of solidarity, when it is motivated *solely* by the belief that this action will contribute to the realization of B's objectives (absence of the shared objectives aspect). *Example*: A is wealthy and donates, for no particular reason other than to do good, to a charity organization that funds B to recover from flood damage.
- The action of A is an act of 'heartfelt solidarity' when it is motivated by the sense of community due to A's belief that B shares some of his own objectives. *Example*:

Village A does not lie in a flood-prone area and its inhabitants nonetheless contribute to a flood damage compensation fund for village B because they believe that those who live in village B aim, just like they do, to preserve their characteristic houses as part of their cultural heritage.

- The action of A is and act of 'opportune solidarity' when A and B have the same objective and A values this objective even more than B (dominance of self-interest). *Example*: Some inhabitants of A contribute to the flood damage compensation fund for village B because they value its cultural heritage even more than the inhabitants of B.
- An action of A is an act of 'calculated solidarity' when A takes this action motivated by his expectation of possible future benefits should he find himself in a situation similar to B (absence of the *latent* reciprocity aspect). *Example*: Village A and B are both situated in a flood-prone area and their inhabitants choose to contribute to the flood damage compensation fund because that will cover their own flood risk.
- The action of A is an act of 'self-interested solidarity' when A's actual motivation to support B does not stem from B's objectives, but from A's own, different objectives. *Example*: Village A does not lie in a flood-prone area and yet some of its inhabitants contribute to the flood damage compensation fund for village B, but they do so merely to increase the fidelity of their customers living in the flood-prone village B.
- An action of A is an act of 'imposed solidarity' when A takes this action as a result of some contract or social rule (absence of the voluntary aspect of the action). *Example*: The people in villages A an B all pay a national income-based tax, the revenues of which are used to compensate people who incur flood damage.

These definitions of solidarity provide categories for *individual* behavioral patterns. In the present stage of our research, we focus on these patterns to study the conditions that lead to different forms of solidarity, and the effect on a group level. In a later stage, we will also consider group solidarity, that is, the degree to which individuals in a group comply with corporate rules in the absence of compensation (Hechter, 1987, p. 39).

4 So-Si-So: A modelling Platform to Analyse Solidarity in Behavioural Patterns

In this section we describe the modelling platform called So-Si-So (Social Simulation of Solidarity) that we have developed to study the simulated behaviour of cognitive agents, situated in physical spaces, and interacting in social spaces.

To be able to distinguish between the different types of solidarity defined in the previous section, the observer must have access to the mental state (beliefs, motivations) of both A and B. Our agent-based models should therefore allow inspection of the motives for agents to take action. The design of So-Si-So agents should otherwise be as simple as possible. This led us to the following choices:

- Agents derive their ultimate motivation for action from values, which they seek to uphold. Some values are intrinsically individualistic (e.g., one's own survival), other values are intrinsically collective (e.g., peace).
- Agents have concerns that define relations between their values and indicators that give them insights on aspects of the physical spaces and/or the social spaces that they have access to.
- Agents acquire factual beliefs about these indicators through their sensors. Being actor-bound, sensors are subjective: different agents may construct different factual beliefs about the same situation.
- Agents become privately concerned when, according to their factual beliefs, one or more of these aspects appear to be in bad shape.
- Agents may also become socially concerned, that is, consider the concerns of other agents situated in their social space(s). Agents will become socially concerned only if they have some inclination towards altruism.
- To address concerns (their own and/or those of other agents), agents evaluate their causal beliefs, looking for actions that they expect to ameliorate the aspects of their environment about which they are concerned.
- For each of their 'active' concerns, agents choose the action (insofar available) that seems best to them.

This agent design permits the modeller to trace for each action taken the values that motivated the agent to perform this action, and therefore also to detect solidary actions (actions that address the concerns of others) and distinguish between different types of solidarity. To demonstrate this, we must define more precisely the way in which So-Si-So agents come to act:

- An *agent* is represented as an 12-tuple *a* = (Val, Con, Sen, Bfact, m, α, Cp, Cs, Cw, Bcaus, budget, Soc) where Val is a set of values, Con a set of potential concerns, Sen a set of sensors, Bfact the set of factual beliefs, m the agent's memory depth (represented as a non-negative integer value), α the agent's inclination towards altruism (represented on the interval [0, 1]), Cp, Cs and Cw sets of, respectively, private, social, and weighted concerns, Bcaus a set causal beliefs, budget the agent's financial resources (represented as a non-negative real value), and Soc the set of social spaces that *a* is part of.
- A *value* is represented as a 3-tuple (idea, type, importance), where idea expresses something that the agent holds dear, for example: 'survival', 'wealth' or 'environment'. The value type is either IND (for individualistic) or COL (for collective). When an agent has multiple values, their relative importance can be represented on the interval of (0, 1].
- A *potential concern* is represented as a 3-tuple (indicator, value, intensity) that expresses the subjective relation that the agent sees between some aspect of the world and what the agent holds dear. When an agent considers multiple indicators to be related to the same value (for example, both quality and quantity of the water resource relate to the value 'ecology'), the relative intensity of their relations with this value can be represented on the (0, 1] interval.
- A *factual belief* is represented as a 3-tuple (tick, indicator, judgement) that expresses how the agent perceives some aspect of its environment at a particular

tick (time step as counted by the model's clock) in the simulation. An agent's factual belief list cannot contain beliefs older than the agent's memory depth. New factual beliefs are added to the agent's factual belief list by the agent's *sensors*, where a sensor typically compares one or several variables in the agent's environment with pre-defined threshold values, and produces a judgement on some scale. In our experiments, we used this scale: AWFUL=-1, BAD=-0.5, NEUTRAL=0, GOOD=+0.5, GREAT=+1, but other scales can be defined, as long as they map onto the same interval of [-1, 1].

- The three lists of *actual concerns* (private, social, and weighted, respectively) contain 2-tuples (concern, state) that express for the current time step the degree (represented by the numeric field state) to which the agent is actually concerned about certain aspects. For example, when in time step 25 an agent has a factual belief (25, 'water quantity', BAD), and the potential concern c = ('water quantity', 'ecology', 1), its private concern will be pc = (c, -0.5). Private concerns are created from the agent's own potential concerns, social concerns are created from other agents' potential concerns. The set of weighted concerns contains the weighted average of an agent's private and social concerns. It is this set of actual concerns that the agent used for taking decisions.
- A *causal belief* is a 3-tuple (action, effects, cost) that expresses the relation that the agent sees between an action, its effects on one or more indicators, and the costs involved. For example, the causal belief ('SELF:produce', {('water quality', BAD), ('income', GOOD)}, 50) represents that the agent believes that its own action 'produce' affects two indicators and will require 50 of the agent's budget units. The list of causal beliefs may be incomplete (thus representing what an agent conceives as possible) and even incorrect (agents may wrongly estimate effectiveness and/or cost). Agents may also have causal beliefs about actions that can be taken by other agents.
- An *action* represents what agents can do. The actual consequences of actions are calculated by the physical space entities and/or social space entities that the agent has access to.

Physical spaces, social spaces and their autonomous evolution (i.e., state changes that are not the immediate result of actions of agents) are defined by the modeller (see the example in section 5). Once physical spaces, social spaces and agents have been initialised, a So-Si-So model iterates through this cycle:

1. *The agents get the opportunity to act.* In the experiments reported in this paper, this happens in a fixed sequence: the order in which the agents were created.
2. *The social spaces coordinate collective actions.* Collective actions are actions that require more than one agent to be effective, so agents can only 'announce' their intention to perform such an action in a social space. When all agents have had the opportunity to act, each social space checks whether the conditions for successful collective action as specified by the modeller (e.g., at least five agents are needed for this action to be effective) are met, and only if that is the case will the consequences of the action be effectuated as changes in the state of certain physical and/or social spaces. In the experiments reported in this paper, this feature has not been used.

3. *The physical spaces evolve*. Although the immediate consequences of agent actions have all been effectuated, the state of the physical spaces may still change autonomously.
4. *The agents evolve.* Similar to physical spaces, certain attributes of agents may also change as a function of time, independently of agent actions.
5. *The model advances to the next time step*: Clock.tick = Clock.tick + 1

A So-Si-So agent *a* comes to act by in a sequence of steps:

1. *Forget*. Agent *a* removes all factual beliefs *fb* from *a*.Bfact for which (Clock.tick – *fb*.tick) > *a*.m.
2. *Perceive*. Agent *a* activates each of its sensors; each sensor in *a*.Sens may add to *a*.Bfact a new factual belief about the state of *a*'s environment.
3. *Appraise*. Agent *a* checks to see whether its factual beliefs activate one or more of its potential concerns: for each factual belief *fb* in *a*.Bfact the agent checks whether it has a potential concern *c* in *a*.Con with *c*.indicator = *fb*.indicator; if that is the case, the agent checks to see whether it already has a private concern *pc* with *pc*.concern = *c*; if not, such a private concern *pc* is created (with *pc*.state initialised to 0) and added to the agent's set of private concerns; the state of the (new) private concern *pc* is then updated as follows:
   pc.state = *pc*.state + *c*.intensity**c*.value.importance**fb*.judgement
4. *Survey socially*. The agent *a* considers the concerns of those agents with which it has a social relation: let *n* be the sum, for each social space *ss* in *a*.Soc, of the number of other agents that are also part of *ss*; then for each social space *ss* in *a*.Soc, the agent *a* checks for each other agent *b* in *ss* whether it has one or more weighted concerns *wc* in *b*.Cw with *wc*.state < 0 and *wc*.concern.value.type = COL. For each such *wc*, agent *a* checks to see whether it already has a social concern *sc* with *sc*.concern = *wc*.concern; if not, such a social concern *sc* is created (with *sc*.state initialised to 0) and added to *a*.Soc; the state of agent *a*'s (new) social concern *sc* is then updated as follows:
   a.*sc*.state = *a*.*sc*.state + *b*.*wc*.state / *n*
   The division by *n* makes that the scale for the state attributes of the agent's private concerns and its social concerns is the same. Note, however, that if some other agent is part of more than one of the agent's social spaces, its concerns will be considered several times. We see this as appropriate, as it reflects that in reality those other agents with multiple relations will also have more influence on the agent.
5. *Weigh up private and social concerns*. Agent *a*'s parameter α is used when the agent merges its private and social concerns into its set of weighted concerns Cw: starting with an empty set, the agent first adds for each of its private concerns *pc* a new weighted concern *wc* with *wc*.state = *pc*.state*(1-α); then for each of its social concerns *sc*, the agent checks to see whether it already has a weighted concern *wc* with *wc*.concern = *sc*.concern; if not, a new weighted concern *wc* is added with *wc*.state initialised to *sc*.state*α; otherwise, the state of the existing weighted concern *wc* is updated as follows:
   wc.state = *wc*.state + *sc*.state*α

6. *Deliberate*. For each of its weighted concerns *wc* (in order of increasing state, so as to deal with the worst problem first) agent *a* checks its causal beliefs looking for actions that are believed to alleviate this concern. More precisely: the agent selects those causal beliefs *b* with *b*.effects containing a 2-tuple *e* = (indicator, impact) with *e*.indicator = *wc*.concern.indicator and *e*.impact > 0 (i.e., GOOD or GREAT). Agent *a* then selects the action that is still affordable (*b*.cost ≤ *a*.budget) and has the highest utility (taking into account its consequences for its other concerns), adds it to its action list, and computes its remaining budget. The utility *u* of an action is calculated as follows: starting with an initial value *u* = 0, the agent checks for each of the action's effects *e* = (indicator, impact) whether it has one or more weighted concerns *wc* with *wc*.indicator = *e*.indicator and *wc*.state < 0, and for each such concern *u* = *u* − *e*.impact**wc*.state. This reflects that *a* only considers the indicators *a* is concerned about (ignoring all other indicators), and will favour those with a (strong) positive effect on indicators whose state is (very) negative.
7. *Act*: the agent performs the actions on its action list. Note that the actual consequences of each action for the agent's own attributes (notably *a*.budget) as defined by the modeller may differ form the anticipated consequences, that is, the effects according to the agent's causal beliefs.

This architecture makes it possible to detect different types of solidarity because it allows scrutinizing an agent's motivation for taking certain actions.

The basic characteristic of an act of solidarity action is that agent *a* takes this action to support agent *b*, and that *a* makes some kind of sacrifice by doing so. In the So-Si-So architecture, agents are aware of the concerns of other agents, but on an aggregate level only. Where *a*'s private concerns *a*.Cp are its own and therefore individual, *a*'s social concerns *a*.Cs are a 'weighted union' of the weighted concerns of a set of other agents. Thus, the 'other agent' *b* in our definitions of solidarity is the aggregate of all agents in *a*'s social spaces, rather than an individual agent.

The terms 'support' and 'sacrifice' both entail a concept of utility: an action of *a* supports *b* when it produces positive utility for *b*, and it constitutes a sacrifice for *a* if it produces negative utility for *a*. By the same type of reasoning, *a* supports *b* by *not* taking an action that would produce negative utility for *b*, even though it would produce positive utility for *a*. The foregone utility for *a* then constitutes *a*'s sacrifice.

While deliberating, agent *a* constructs an action list, selecting (insofar as *a*'s budget permits) for each of its weighted concerns *a*.Cw the action with the highest utility. To see whether a makes a sacrifice, a second action list is constructed according to the same procedure but now based on *a*'s private concerns *a*.Cp. The former (from now on called the 'weighted action list' of *a*) represents what *a* does, the latter (*a*'s 'private action list') represents what *a* would have done if *b* had not existed. Likewise, *a*'s 'social action list' is constructed on the basis of *a*'s social concerns *a*.Cs.

Using these three action lists, we can now define the following types of utility:

- Being based on *a*'s weighted concerns, the utility of an action *x* on *a*'s weighted action list is called the 'weighted utility' of action *x*. The 'private utility' and the 'social utility' of *x* are defined likewise.

- U*soc* cumulates the social utility of actions in *a*'s weighted action list that are *not* in *a*'s private action list PLUS for those actions *x* that are in both lists the 'social utility surplus' defined as the difference (if positive) of social utility of *x* – private utility of *x*.
- U*neg* cumulates the negative private utility of actions in *a*'s weighted action list. Having a negative private utility, such actions will *not* be in *a*'s private action list.
- U*asoc* cumulates the negative social utility for actions in *a*'s weighted action list. When an action *x* has a social utility < 0 and yet occurs in the weighted action list, it must be motivated by private concerns only, hence U*asoc*, which stands for 'anti-social utility'.
- U*opp* cumulates for actions in *a*'s weighted action list that are also in *a*'s private action list (but only those driven by the same concern, with the state of *a*'s private concern being worse than the state of *a*'s social concern) the 'private utility surplus' defined as the difference (if positive) of private utility of *x* – social utility of *x*.
- U*self* cumulates the private utility surplus for the remaining (i.e., not used for U*opp*) actions in *a*'s weighted action list that are also in *a*'s private action list.
- U*foregone* cumulates the private utility of actions in *a*'s private action list that are not in *a*'s weighted action list. Motivated by social concerns, *a* decided not to take these actions, hence the term 'forgone utility'.

As shown in Table 1, these different types of utility allow us to detect and measure four of the six types identified in section 3.

Table 1. Operationalisation of different types of solidarity

Type of solidarity	Detection/measure
altruism	U*soc* + U*neg* + U*foregone*
heartfelt solidarity	The distinction between altruism and heartfelt solidarity can be made in step 4 (*survey socially*). Altruistic actions are precluded by making agent *a* consider only those social concerns that are related to a value that also occurs in *a*.Val. As the present Solid'eau model includes only one collective value, the difference cannot occur.
opportune solidarity	U*opp*
self-interested solidarity	U*self*
calculated solidarity	Not represented in So-Si-So models, as in the present architecture agents cannot anticipate on other agent's actions and therefore not foresee reciprocity.
imposed solidarity	Not represented in So-Si-So models, as presently agents are driven only by their own values. This will change once we have extended the architecture with social norms.

In view of the categories of solidarity we want to test, we did not need to make agents foresighted with respect to expected social behavioural patterns. Agents are myopic and act only upon their knowledge of consequences of potential actions they might perform on their own concerns or on concerns of others. They do not take into account any belief or expectation regarding others' actions like in Conte and Paolucci

(2002) or in Dittrich *et al*. (2003). As noted in Table 1, such sophistication will be needed to detect calculated solidarity.

Agents in So-Si-So compute their perceptions according to their values only, their concerns according to their actual perceptions (and some of their previous perceptions, since agents can keep their factual beliefs for memory depth time steps), and their actions according to their causal beliefs based on actions. Figure 1 below is expanding the Perception-Deliberation-Action cycle from Jacques Ferber (1999) as it is activated in So-Si-So agents. If more cognitive agents are required, the deliberation stage might be more refined, for example through the identification and choice of actions, taking into account more than causal beliefs only.

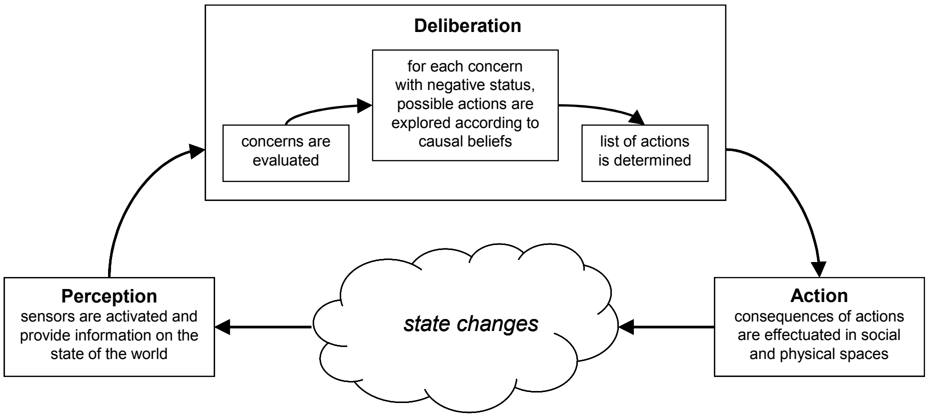


Fig. 1. Extension of Ferber's Perception-Deliberation-Action cycle in So-Si-So

5 A simple Test Model

We tested So-Si-So with what we see as an embryonic version of Solid'eau: a simple decontextualised model of a set of individuals, equipped with the previously described cognitive capacities, located in an urban or a rural area, in which they have a productive activity using units of a shared resource. To facilitate analysis, we use no stochastic parameters. The two areas (modelled as physical spaces) are linked by a water resource. This resource is localized, with a flow from an upstream space to a downstream one, depending on resource availability in the upstream space. As it is the most frequent case, we assume that the rural physical space is located upstream. The connection between the two spaces is described by the two following rules:

- a constant ratio of resource level, φ, disappears from upstream space and flows in downstream space,
- if the resource availability in the upstream space after resource renewal exceeds its maximum capacity, the surplus flows to the downstream part.

The resource in the upstream physical space renews at a fixed rate ρ relative to its maximum capacity.

Agents are localised in either one of the physical spaces. The population is then described by the total number, N, and the ratio in urban area, θ. Urban agents and a part of rural agents (those who have social ties in the urban area) share a social space.

Agents are described by their localisation, the weight α they attach to social concerns, and their thresholds to determine how they assess the information received from their sensors. All agents are driven by three values: ('survival', IND, 1), ('environment', COL, 0.33) and ('wealth', IND, 0.33), the numbers indicating relative importance, with potential concerns ('subsistence', 'survival', 1), ('budget', 'wealth', 1) and ('resource', 'environment', 1).

Possible actions of agents are production at level 0, 1, 2, or 3, where production at level 0 means: do not produce. Production in a rural area consumes πr*level units of resource and generates Yr*level units of budget, respectively πu*level and Yu*level in urban areas. Agents also incur production costs of γ*level, the base production cost γ being equal for urban and rural areas.

All agents have the same causal beliefs:

- ('produce.0', {('subsistence', AWFUL), ('resource', GOOD)}, 0)
- ('produce.1', {('subsistence', GOOD), ('budget', GOOD)}, γ)
- ('produce.2', {('subsistence',GREAT), ('budget',GOOD), ('resource',BAD)}, 2*γ)
- ('produce.3', {('subsistence',GREAT), ('budget',GREAT), ('resource',AWFUL)}, 3*γ)

At each tick, all agents spend the same fixed amount of budget for their cost of living, λ.

Table 2. Parameter setting of the Solid'eau model

Parameter	Value
φ, flow rate	0.2
ρ, resource renewal per tick	0.2
SmaxR, maximum resource level in rural area	100
SmaxU, maximum resource level in urban area	100
N, total population	30
θ, fraction of urban population	0.7
πr, production factor in rural area	3
πu, production factor in urban area	2
Yr, production yield in rural area	200
Yu, production yield in urban area	250
λ, cost of living	50
α, altruism	0.4
β, initial value of agent budget	250
γ, base production cost	50
SB, budget threshold below which sensor yields ('subsistence', BAD)	250
SA, budget threshold below which sensor yields ('survival', AWFUL)	175
BB, budget growth threshold below which sensor yields ('budget', BAD)	0.2
BA, budget growth threshold below which sensor yields ('budget', AWFUL)	0.1
RB, resource threshold below which sensor yields ('resource', BAD)	60
RA, resource threshold below which sensor yields ('resource', AWFUL)	42

The order in which agents can act (and therefore co-determines their actual resource use, given a chosen production level) is fixed for each physical space. All agents have sensors that generate factual beliefs *fb* = (indicator, judgement) for the indicators 'resource' (to assess the resource availability in the physical area where the agent produces), 'subsistence' (to assess whether the agent's budget is sufficient to cover its cost of living), and 'budget' (to assess whether the agent's budget is on average increasing, with a desired percentage per year. These sensors return judgements based on threshold values. In the model we have used so far, all agents to have the same thresholds for the same concern, but these threshold values may be individualised. Table 2 shows the threshold values used, as well as all other parameter values for the model used in the simulations referred to in the next section.

6 Simulation Results

Even though this work is still ongoing, we have been able to verify that the So-Si-So model architecture permits identification and measurement of the various categories and subcategories of solidarity, even with a model with but few options in terms of values and the structure of social spaces. The model's outcomes have been verified through code proofreading and comparison of outcomes with expectations.

Figure 2 and 3 show two scenarios, the only difference being that in the second scenario some of the rural agents do not share the urban-rural social space. As the aim of this paper is to propose a categorisation for solidary behaviour and ways to measure it, the model results are illustrations that hold little surprises. In both scenarios, stimulated by their desire for wealth (a budget increase of at least 10% and preferably 20% or more of their initial budget each time step), the agents who can use the resource first produce at high levels, while those who then see the resource dwindle become concerned about it and produce less. Around tick 30, the least fortunate agents start to produce merely to survive (this causes the high peaks of self-interested utility: production serves both survival and wealth). As more agents become very concerned about the resource, agents who share their social space start foregoing high production (this causes the peaks of foregone utility that first appear around tick 40) and by consequence fall back in wealth. Looking at both the resource level plot and the budget units graph in Fig. 2, it can be seen that gradually more rural agents forego production even when their resource (upstream) would permit it.

As was expected, the first scenario shows more solidary behaviour. In the second scenario, a resource crisis occurs between tick 110 and 125, when more than half of the urban agents run out of budget. In the model used, this makes them become inactive, which represents that the inhabitants leave the region, or—worse—starve. Inactive agents are ignored, so their concerns no longer influence the deliberation of other agents. As a result, the three rural agents who were sensitive to the urban agents' concerns feel less need for solidarity pick up their production.

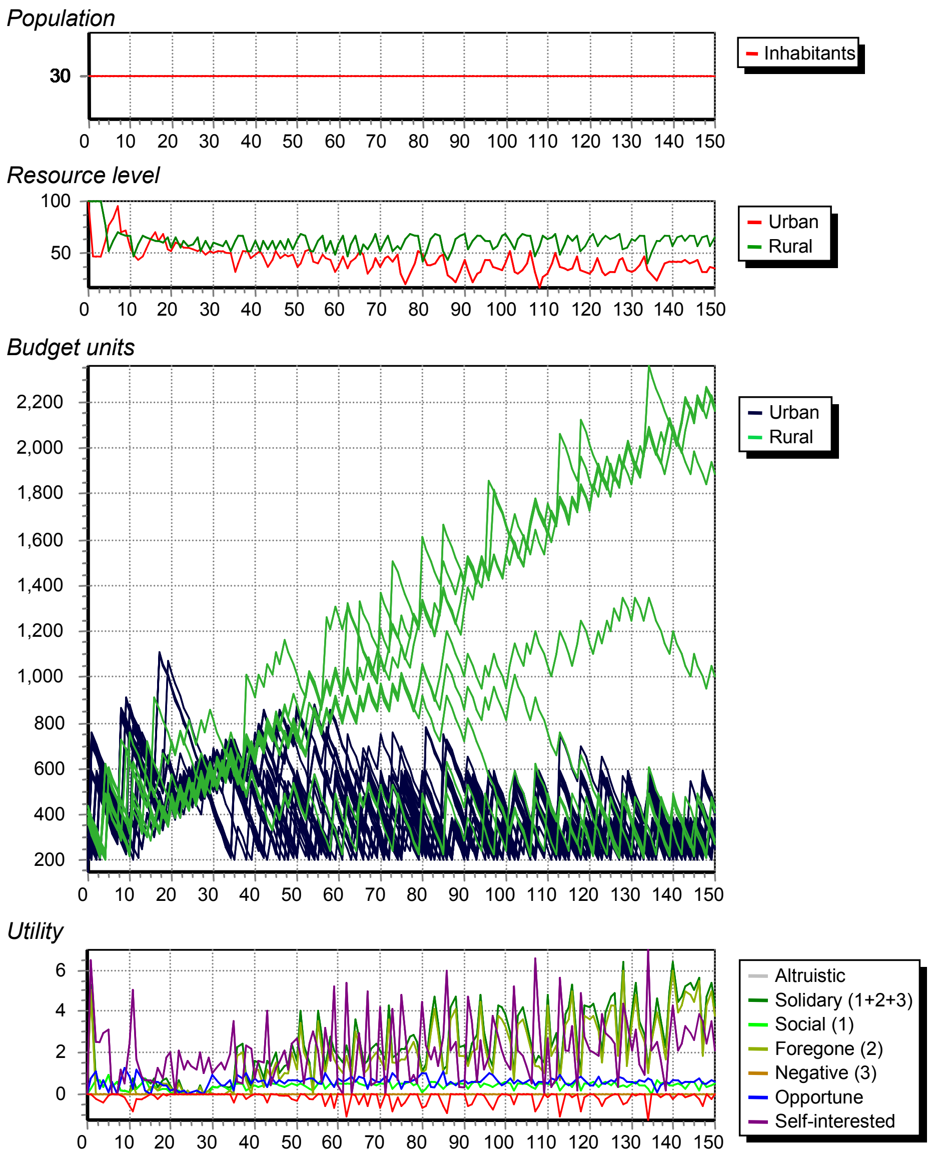


Fig 2. Scenario in which all agents are part of one and the same social space

The steady alternation in Fig. 3 between production at level 0 and level 1 shown by the six rural agents without a social space reflects their own concern with the resource only; they are the last to access the rural resource, and their budget drops only when the three rural agents who do share a social space with the other agents pick up their production.

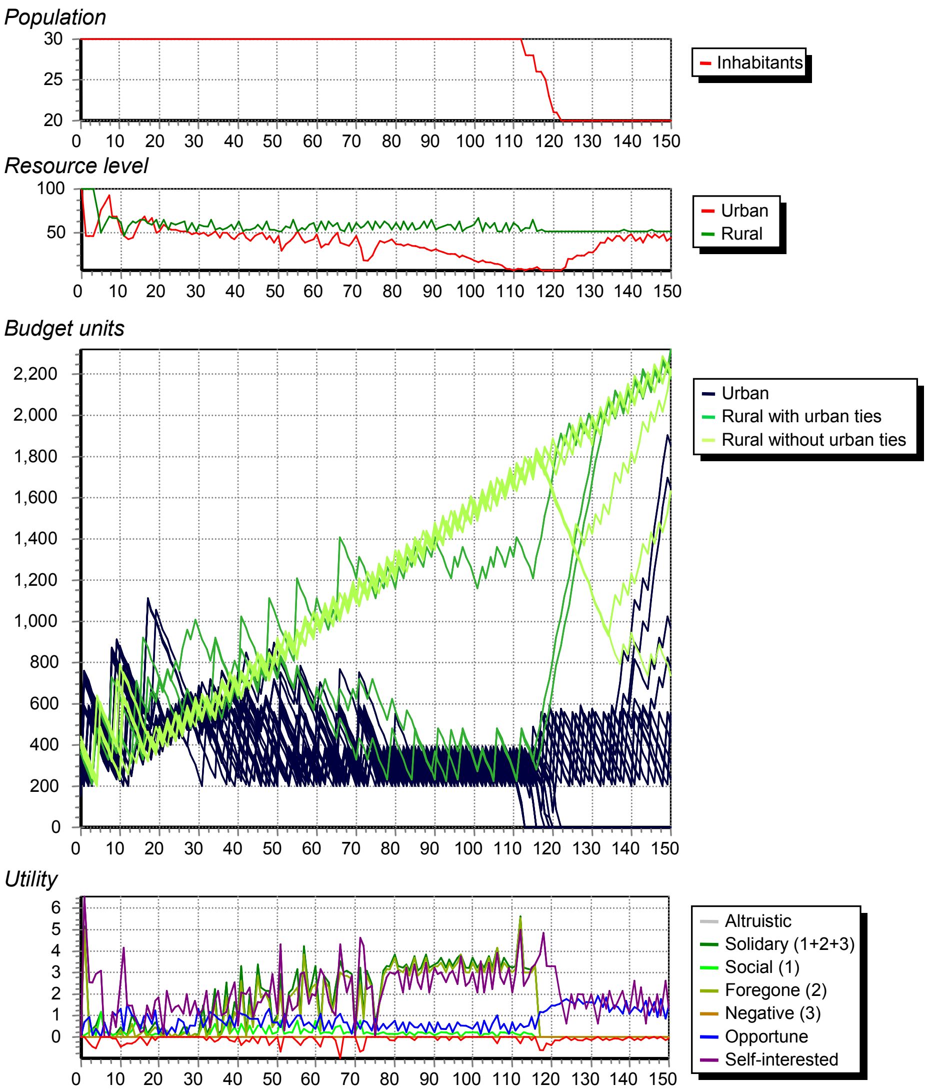


Fig 3. Same scenario, except that 6 of 9 rural agents have no social ties with the other agents

6 Conclusion and perspectives

With this tentative framework to represent solidarity at an individual level we have succeeded in identifying and distinguishing solidary actions in a simple virtual world, and we have paved the way for an evolution towards a more grounded representation of resources and networks of relations between agents (via social spaces) as well as between agents and these resources (via physical spaces).

To achieve this evolution we need to couple this framework with a more structured description of the virtual world, and to include concepts that will permit representation and analysis of other forms of solidarity. For example, to represent imposed solidarity, we will have to include norms.

We aim at incorporating this work in a policy perspective, through the identification of existing social networks which might serve as driving belt for solidarity actions and joint preservation of resources across physical spaces.

References

Cohen, J. L. and Arato, A. (1992). *Civil Society and Political Theory*. Cambridge MA: MIT Press.

Conte, R. and Paolucci M. (2002). *Reputation in Artificial Societies: Social Beliefs for Social Order*. Dordrecht: Kluwer Academic Publishers.

Dittrich, P., Kron, T. and Banzhaf, W. (2003). On the Scalability of Social Order: Modeling the Problem of Double and Multi Contingency Following Luhmann. *Journal of Artificial Societies and Social Simulation* **6**(1), http://jasss.soc.surrey.ac.uk/6/1/3.html

Ferber, J. (1999). *Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence*. Boston, MA: Addison-Wesley Longman.

Hechter, M. (1987). *Principles of group solidarity*. Berkely: University of California Press.

Hondrich, K. O. & Koch-Arzberger, C. (1992). *Solidarität in der modernen Gesellschaft*. Frankfurt am Main: Fischer Verlag.

Mason, A. (1998). 'Solidarity', *The Routledge Encyclopaedia of Philosophy*. New York: Routledge.

Mason, A. (2000). *Community, Solidarity and Belonging: Levels of Community and their Normative Significance*. Cambridge: Cambridge University Press.

Miller, D. (1999). *Principles of Social Justice*. Cambridge, MA: Harvard University Press.

Misztal, B. A. (1996). *Trust in Modern Societies*. Cambridge: Polity Press.

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. New York: Cambridge University Press.

Segall, S. (2005). Political Participation as an Engine of Social Solidarity: A Sceptical View. *Political Studies* **53**: 362–378

Seligman, A. (1997). *The Problem of Trust*. Princeton, NJ: Princeton University Press.

Warren, M. E. (1999). 'Democratic Theory and Trust'. Pages 310–45 in M. E. Warren, *Democracy and Trust*. Cambridge: Cambridge University Press.

Reputation for Innovating Social Networks*

Rosaria Conte[1], Mario Paolucci[1], and Jordi Sabater-Mir[2]

[1] Institute for Cognitive Science and Technology
Via San Martino della Battaglia, 44, Rome, ITALY
{rosaria.conte|mario.paolucci}@istc.cnr.it
[2] Artificial Intelligence Research Institute, Barcelona, SPAIN
jsabater@iiia.csic.es

Abstract. Reputation is a fundamental instrument of partner selection. Developed within the domain of electronic auctions, reputation technology is being been imported into other applications, from social networks to institutional evaluation. Its impact on trust enforcement is uncontroversial and its management is of primary concern for entrepreneurs and other economic operators.

In the present paper, we will shortly report upon simulation-based studies on the role of reputation as a more tolerant form of social capital than familiarity networks. Whereas the latter exclude non-trustworthy partners, reputation is a more inclusive mechanism upon which larger and more dynamic networks are constructed. After the presentation of the theory of reputation developed by the authors in the last decade, a computational system (REPAGE) for forming and exchanging reputation information will be presented and findings from experimental simulations recently run on this system will be resumed. Final remarks and ideas for future works will conclude the paper.

Keywords: Artificial societies, Reputation, Innovation, Social Networks

1 The Problem

In marketplaces, and more generally in social exchange, reputation provides traders and other users with a fundamental instrument of partner selection. Developed within the domain of electronic auctions (like eBay, cf. for a survey [1]), in the last few years reputation technology has been invading other electronic applications, from social networks to institutional evaluation. Its impact on trust enforcement is so uncontroversial, that corporate reputation is counted as an asset, and its management is of primary concern for entrepreneurs and other economic operators [2]. Nowadays, one can make money by assisting people in dealing with, managing, and even refreshing their own reputation[3]. Such a

* This work was partially supported by the European Community under the FP6 programme (*eRep* project, contract number CIT5-028575) and by the Italian Ministry of University and Scientific Research under the FIRB programme (Socrate project, contract number RBNE03Y338)

[3] cf. http://www.reputationdefender.com

far-reaching confidence in reputation probably rests on the assumption that it supports us in the complimentary roles of selecting trustable partners and being selected as such.

Far from discrediting the view of reputation as a trust enhancement mechanism, we would like to enlarge the boundaries of the phenomenon at stake, by pointing to another functionality, namely the enlargement and innovation of social networks.

The rest of the paper will unfold as follows. First, the role of image-based networks in a world where the boundaries of social and trading networks are constantly widened will be questioned. Next, drawing upon the social cognitive model presented in [3], a notion of reputation as a special form of social evaluation will be re-proposed. This notion will be argued to allow for network innovation: on one hand, reputation allows for social evaluation to circulate and complement ones personal experience. On the other, it will be argued to accomplish a most crucial and delicate task, i.e. check and discard misinformation without necessarily discarding the agents responsible for its transmission. In other words, reputation networks will be shown to be more inclusive than image-networks ceteris paribus, and at the same time to help checking the truth-value of the information circulating in the network.

2 Main Claim and Organization of the Paper

The paper is aimed to discuss the view of reputation in the framework presented above. It builds upon the state of the art on reputation theory and technology at the Laboratory of Agent Based Social Simulation (LABSS) of the Institute of Cognitive Science and Technology (ISTC), within the *eRep* project[4] . The starting point will be the results from experimental simulations presented in [4], thanks to the computational system REPAGE, worked out by the authors and presented in [5].

In [4], experiments were meant to show the value added of reputation as a mechanism of partner selection. Results show that an artificial market where agents exchange both image and reputation obtains better results in terms of production quality than a market were agents exchange their own opinions about one another. The reason for such a difference lies in retaliation: as will be argued later on in the present paper, image-based, or familiarity, networks perform more poorly than reputation networks exactly because they induce retaliation.

In the present paper, we will shortly report upon previous findings in order to put forward a more general hypothesis, which seems to be supported by our simulations. Reputation allows for a far more tolerant, gross-grained social selector than image. Hence, whereas shared image forms a selective platform on which familiarity networks that exclude non-trustworthy partners are constructed, reputation is a rather more inclusive mechanism upon which larger and more dynamic networks are constructed. Thanks to it,

[4] http://megatron.iiia.csic.es/eRep/

– candidate (non-confirmed) information may circulate allowing the network to learn new social knowledge,
– the network may innovate, by integrating new partners,
– and put up with errors without discarding the partners that fell prey to them.

In a few words, reputation appears as a more dynamic form of social capital, allowing for social networks to be innovated.

The paper is organized as follows: after the synthetic presentation of the theory of reputation developed by the authors, the REPAGE system will be presented and the experimental simulation recently run by the authors thanks to such a system will be resumed. The findings from that study will be rediscussed in the light of the present hypothesis. Final remarks and ideas for future works will conclude the paper.

3 A Social Cognitive Model of Reputation

In this section we will report on a social cognitive model of reputation presented in [3], where

– the difference between image and reputation has been introduced,
– the different roles agents play when evaluating someone and transmitting this evaluation are analysed,
– the decision processes based upon both image and reputation are examined.

A cognitive process involves symbolic mental representations (such as goals and beliefs) and is effectuated by means of the mental operations that agents perform upon these representations (reasoning, decision-making, etc.). A social cognitive process is a process that involves social beliefs and goals, and that is effectuated by means of the operations that agents perform upon social beliefs and goals (e.g., social reasoning). A belief or a goal is social when it mentions another agent and possibly one or more of his or her mental states (for a discussion of these notions, see [6], [7]).

The social cognitive approach is receiving growing attention within several subfields of the Sciences of the Artificial, in particular intelligent software agents, Multi-Agent Systems, and Artificial Societies. Unlike the "theory of mind" (cf. [8]) approach, this approach aims at modelling and possibly implementing systems acting in a social (whether natural or artificial) environment. The theory of mind focuses upon one aspect, although an important one, of social agency, i.e., social beliefs (knowledge agents have about others).

Here, the approach adopted is aimed at modelling the variety of mental states (including social goals, motivations, obligations) and operations (such as social reasoning and decision-making) necessary for an intelligent social system to act in some domain and influence other agents (social learning, influence, and control).

3.1 Image and Reputation

The social cognitive model is a dynamic approach that considers reputation as the output of a social process of transmission of information. The input to this process is the evaluation that agents directly form about a given agent during interaction or observation. This evaluation will be called here the social image of the agent. An agents reputation is argued to be distinct from, although strictly interrelated with, its image. More precisely, image will be defined as a set of evaluative beliefs about a given target, while reputation will be defined as the process and the effect of transmission of image. As an application of this model, some simple predictions made possible by this conceptualisation will be presented. Furthermore, the decision to accept image will be compared with and distinguished from the decision to acknowledge reputation. Image consists of a set of evaluative beliefs [9] about the characteristics of the target, i.e. it is an assessment of its positive or negative qualities with regard to a norm, a competence, and so on.

Reputation is both the process and the effect of transmission of a target's image. The image relevant for social reputation may concern a subset of the target's characteristics, i.e., its willingness to comply with socially accepted norms and customs. More precisely, reputation is defined to consist of three distinct but interrelated objects:

- a cognitive representation, or more precisely a believed evaluation;
- a population object, i.e., a propagating believed evaluation;
- an objective emergent property at the agent level, i.e., what the agent is believed to be.

In fact, reputation is a highly dynamic phenomenon in two distinct senses: it is subject to change, especially as an effect of corruption, errors, deception, etc.; and it emerges as an effect of a multi-level bidirectional process. In particular, it proceeds from the level of individual cognition to the level of social propagation and from this level back to that of individual cognition again. What is more interesting, once it gets to the population level, it gives rise to a further property at the agent level: agents acquire a bad or good name. Reputation is not only what people think about targets but also what targets are in the eyes of others. From the very moment agents are targeted by the community, want it or not and believe it or not, their lives change: reputation becomes the immaterial, more powerful equivalent of a scarlet letter sewed to their clothes. It is more powerful because it may not even be perceived by those to whom it sticks, and consequently it is out of their control. Reputation is an objective social property that emerges from a propagating cognitive representation, which lacks an identified source, whereas image always requires that at least one evaluator to be identified.

3.2 Reputation and Image As Social Evaluations

According to [9], an evaluation is a hybrid representation. An agent has an evaluation when he or she believes that a given entity is good for, or can achieve, a given goal. An agent has a social evaluation when his or her belief concerns another agent as a means for achieving this goal. A given social evaluation includes three sets of agents:

- a nonempty set E of agents who share the evaluation (evaluators)
- a nonempty set T of evaluation targets
- a nonempty set B of beneficiaries, i.e., the agents sharing the goal with regard to which the elements of T are evaluated.

Often, evaluators and beneficiaries coincide, or at least have nonempty intersection but this is not necessarily the case. A given agent t is a target of a social evaluation when t is believed to be a good/bad means for a given goal of the set of agents B, which may include or not the evaluator. (Social) evaluations may concern physical, mental, and social properties of targets; agents may evaluate a target as to both its capacity and willingness to achieve a shared goal. In particular, more or less explicitly, social evaluations concern the targets' willingness to achieve a goal or interest. Formally, e (with $e \in E$) may evaluate t (where $t \in T$) with regard to a state of the world that is in b 's (with $b \in B$) interest, but of which b may not be aware.

The interest/goal with regard to which t is evaluated may be a distributed or collective advantage. It is an advantage for the individual members who are included in the set B, or it may favour a supra individual entity, which results from interactions among the members of B (for example, if B's members form a team).

It is very easy to find social examples where the three sets coincide: universal norms, such as "Don't commit murder," apply to, benefit, and get evaluated from the whole universe of agents.

There are situations in which beneficiaries, targets, and evaluators are separated, for example, when norms safeguard the interests of a subset of the population. Consider the quality of TV programs during the children's timeshare. Here, we can find three clearly separated sets: children are the beneficiaries, while the adults entrusted with taking care of the children are the evaluators. Of course, here the intersection between B and E still exists, because E may be said to adopt Bs interests. But who are the targets of evaluation? Not all the adults, but the writers of programs and the decision-makers at the broadcast stations. In this case, there is a nonempty intersection between E and T but no full overlap. Also, if the target of evaluation is the broadcaster itself, a supra-individual entity, then the intersection can be considered to be null: $E \cap T = \emptyset$.

To assume that a target t is assigned a given reputation implies assuming that t is believed to be "good" or "bad," but it does not imply sharing either evaluation. Reputation then involves four sets of agents:

- a nonempty set E of agents who share the evaluation
- a nonempty set T of evaluation targets
- a nonempty set B of beneficiaries, i.e., the agents sharing the goal with regard to which the elements of T are evaluated

– a nonempty set M of agents who share the meta-belief that members of E share the evaluation; this is the set of all agents aware of the effect of reputation (as stated above, effect is only one component of it; awareness of the process is not implied).

Often, E can be taken as a subset of M; the evaluators are aware of the effect of evaluation. In most situations, the intersection between the two sets is at least nonempty, but exceptions exist. M in substance is the set of reputation transmitters, or third parties. Third parties share a meta-belief about a given target, whether they share the concerned belief or not. In real matters, agents may play more than one role simultaneously.

3.3 Reputation-Based Decisions

The model presented above focuses on the definition of some critical sets, defining characteristics that we believe to be relevant for reputation. On the basis of our definitions, we will go on from examining the main decision processes undertaken by social agents with regard to image and reputation. To understand the difference between image and reputation, the mental decisions based upon them must be analysed at the following three levels:

– Epistemic: accept the beliefs that form either a given image or acknowledge a given reputation. This implies that a believed evaluation gives rise to ones direct evaluation. Suppose I know that the friend I mostly admire has a good opinion of Mr. Bush. However puzzled by this dissonance-inducing news, I may be convinced by my friend to accept this evaluation and share it.
– Pragmatic Strategic: use image in order to decide whether and how to interact with the target. Once I have my own opinion (perhaps resulting from acceptance of others evaluations) about a target, I will use it to make decisions about my future actions concerning that target. Perhaps, I may abstain from participating in political activity against Mr. Bush.
– Memetic: transmit my (or others) evaluative beliefs about a given target to others. Whether or not I act in conformity with a propagating evaluation, I may decide to spread the news to others. Image and reputation are distinct objects. Both are social in two senses: they concern another agents (the targets) properties (the target's presumed attitude towards socially desirable behaviour), and they may be shared by a multitude of agents. However, the two notions operate at different levels. Image is a belief, namely, an evaluation. Reputation is a meta-belief, i.e., a belief about others' evaluations of the target with regard to a socially desirable behaviour.

The epistemic decision level is grounded upon both image and reputation. An epistemic decision concerns whether to accept a given belief. In the case of image, it concerns evaluations; in the case of reputation, it concerns meta-beliefs (others' evaluations). Both these decisions are relatively independent of one another. To accept a meta-belief does not require that the first-level belief be held to be true, and viceversa: to accept a given image about someone does not imply a belief that that person enjoys the corresponding reputation. To accept/form a given image about a target implies an assessment of the truth value of evaluations concerning the target. In contrast, reputation consists of meta- beliefs about image, i.e., about others' evaluative beliefs concerning the holder.

Conversely, to acknowledge a given reputation does not lead to sharing others' evaluations but rather to the belief that these evaluations are held or circulated by others. To assess the value of such a meta-belief is a rather straightforward operation. For the recipient to be relatively confident about this meta-belief, it is probably sufficient that it be exposed to rumours. In order to understand the difference between image acceptance and reputation acknowledgement, it is necessary to investigate the different roles of image and reputation beliefs in the agents' minds.

But before setting out to do so, a couple of intertwined preliminary conclusions can be suggested. First, reputation is less likely to be falsified than image. Second, the process of transmission, rather than its effect, is prevalent in reputation. In fact, it is more difficult to ascertain whether a given state is true in anyone's mind than in the external world. An external state of the world is more controllable than a mental one. It is relatively difficult to check whether, to what extent, and by whom that state of the world is believed to be true. But the representation of another's mental state is essential for social reasoning, and any clue to such a belief, given a lack of other indications, is better than no information. This easy acceptance of reputation information gives prevalence to the process over the content. Therefore, any study on reputation that concentrates on content only is likely to miss the point completely.

Agents resort to their evaluative beliefs in order to achieve their goals [9]. Evaluations are guidelines for planning; evaluations about other agents are guidelines for social action and social planning. Therefore, the image a given agent has about t will guide its action wrt t, will suggest whether it is convenient to interact with t or not, and will also suggest what type of interaction to establish with t. Of course, image may be conveyed to others in order to guide their actions towards the target in a positive or negative sense. When transmitting its image of t, the agent attempts to influence others' strategic decisions. To do so, the agent must (pretend to) be committed to the evaluation and take responsibility for its truth value before the recipient. Reputation enters direct pragmatic or strategic decisions when it is consistent with image or when no image of the target has been formed. Otherwise, in pragmaticstrategic decisions, reputation is often superseded by image. However, in influencing others' decisions, the opposite pattern occurs: in this case, only reputation considerations apply. Agents tend to influence others' social decisions by transmitting to them information about the target's reputation. Two main reasons explain this inverse pattern:

– agents expect that a general opinion, or at least a general voice, is more credible and acceptable than an individual one

– agents reporting on reputation do not need to commit to its truth value, and do not take responsibility over it; consequently, they may influence others to a lower personal cost.

The memetic decision can be roughly described as the decision to spread reputation. In the case of communication about reputation the communicative action is performed in order to

– obtain the goal that the hearer believes that t is assigned a given reputation by others, rather than by the speaker himself of herself (g2), and to
– obtain the goal that the hearer propagates ts reputation (g4), possibly but not necessarily by having him believe that t is in fact assigned a given reputation (g3).

Whilst g2 is communicative the speaker wants the hearer to believe that the speaker used the language to achieve that effect g4 is not. (Indeed, the speaker usually conceals this intention under the opposite communication: I tell you in confidence, therefore dont spread the news....)

Consequently, communication about reputation is a communication about a meta-belief, i.e., about others mental attitudes. To spread news about someones reputation does not bind the speaker to commit himself to the truth value of the evaluation conveyed but only to the existence of rumours about it. Unlike ordinary sincere communication, only the acceptance of a meta-belief is required in communication about reputation. And unlike ordinary deception (for a definition of the latter, see [10]), communication about reputation implies

– no personal commitment of the speaker with regard to the main content of the information delivered. If speaker reports on ts bad reputation, he is by no means stating that t deserved it; and
– no responsibility with regard to the credibility of (the source of) information (I was told that t is a bad guy).

Two points ought to be considered here. First, the source of the meta-belief is implicit (I was told...). Secondly, the set of agents to whom the belief p is attributed is non-defined (t is ill/well reputed). Of course, the above points do does not mean that communication about reputation is always sincere. Quite on the contrary, one can and does often deceive about others reputation. But to be effective the liar neither commits to the truth of the information transmitted nor takes responsibility with regard to its consequences. If one wants to deceive another about reputation, one should report it as a rumour independent of or even despite ones own beliefs!

4 The Antisocial Effects of Image

The model points to several consequences of image (I) and reputation (R) spreading. Let us examine them with some detail.

First, both I and R spreading are forms of cooperation. Both provide the cognitive matter to informational reciprocity, allowing for material cooperation to take place: agents exchanging shared information about whom they believe to be good and whom they believe to be bad in the group, market, organization or society cooperate at the level of information. By doing so, they allow for material reciprocators, good sellers, norm observers and other good guys to survive and compete with cheaters. Hence, both image and reputation lead to material cooperation.

Secondly, both are expected to lead to social cohesion. Obviously, cheaters may bluff and try to play as informational reciprocators in order to enjoy the benefits of a good image without sustaining the costs of acquiring one. But once bluff is found out, stable social sub-nets are formed by reliable informers who will be sitting there as long as possible. These subnets are more or less what economists and other social scientists call familiarity networks, characterized by reciprocal acquaintance, even benevolence, and trust.

Third, and consequently, both I and R are expected to lead to a reduction in the dimensions of the network of material cooperation or exchange. Acting as selectors, they lead to the initial set of potential relationships to be reduced. Here is where the difference between I and R starts to emerge. I is more selective and R is more inclusive. What is more, unlike R, I spreading reveals the identity of evaluators, or of a subset of them. Shared evaluations make the sources vulnerable, exposing them to possible retaliations. Instead, reported on evaluations protect the identities of evaluators, discouraging or preventing retaliations.

Of course, reported on evaluations provide only candidate evaluations, which often turns to be false and therefore useless. However, one can argue that to find a R disconfirmed is less disruptive than I being disconfirmed. When finding an I received by someone to be wrong, the recipient will face a rather distressing alternative: the source is either misinformed or ill-willed. Either information is unreliable, or the informer's intention is wicked. In any case, the informer cannot be trusted any more, and must be set apart if not punished. Hence, the disruptive effect of image spreading is a function of the amount of informational error and cheating injected into the network. An image-based social network is expected to be rigid, meaning rather sensitive to errors: if a given threshold of error is overcome, the whole system is probably bound to fall apart, and the network will be fatally affected by distrust.

The reason for expecting such a gloomy perspective is complex. For one thing, once recipients of false image have reacted negatively, either getting rid of their bad informers or taking their revenge against them, balance is hardly restored. Mutual defeat will not stop so easily, and retaliation will tend to call for further retaliation in a chain of self-fulfilling prophecies that is usually fatal on both sides. In a stock market, this may even turn into a general collapse.

With reputation, instead, the quality of information received is not necessarily nor immediately tested before being passed on. Misinformation may not be found out so soon, and even when it is finally disclosed, it will not lead the recipient to question the quality of the informer, simply because the latter never committed itself to the truth value of the information conveyed. The reputation network is expected to be more robust than the image-based one, as it puts up with a far larger amount of misinformation without discarding nor punishing the vectors of misinformation, which in fact are not always responsible for such errors.

In the rest of the paper, we will see whether such expectations are met by existing simulation evidence. This was gathered in a study by [4], where our system REPAGE - a REPutation and imAGE tool developed on the grounds of the theory of reputation - was implemented on an agent architecture in order to reproduce an artificial market. In such a setting, buyers were allowed to use either image only (L1 condition) or image plus reputation (L2 condition), and the effects of these two settings were compared in terms of averaged and accumulated quality of products. After a short description of REPAGE, we will turn to show the relevance of these artificial findings to the present view of image and reputation.

5 Repage Model and Architecture

Repage [5] is a computational system based on the theory of reputation presented above [3]. Its architecture includes three main elements, a memory, a set of detectors and the analyzer.

The memory is composed by a set of references to the predicates hold in the main memory of the agent. Predicates are conceptually organized in levels and inter-connected. Each predicate that belongs to one of the main types (including image and reputation) contains a probabilistic evaluation that refers to a certain agent in a specific role. For instance, an agent may have an image of agent T (target) as a seller (role), and a different image of the same agent T as informant. The probabilistic evaluation consist of a probability distribution over the discrete sorted set of labels: Very Bad, Bad, Normal, Good, Very Good. The network of dependences specifies which predicates contribute to the values of others. In this sense, each predicate has a set of precedents and a set of antecedents.

The detectors, inference units specialized in each particular kind of predicate, receive notifications from predicates that change or that appear in the system and use dependencies to recalculate the new values or to populate the memory with new predicates. Each predicate has associated a strength that is function of its antecedents and of the intrinsic properties of each kind of predicate. As a general rule, predicates that resume or aggregate a larger number of predicates will hold a higher strength.

At the first level of the Repage memory we find a set of predicates not evaluated yet by the system. Contracts are agreements on the future interaction between two agents. Their result is represented by a Fulfillment. Communications is information that other agents may convey, and may be related to three different aspects: the image that the informer has about a target, the image that, according to the informer, a third party agent has on the target, and the reputation that the informer has about the target.

In level two we have two kinds of predicates. Valued communication is the subjective evaluation of the communication received that takes into account, for instance, the image the agent may have of the informer as informant. Communications from agents whose credibility is low will not be considered as strong as the ones coming from well reputed informers. An outcome is the agents subjective evaluation of a direct interaction, built up from a fulfillment and a contract. At the third level we find two predicates that are only fed by valued communications. On one hand, a shared voice will hold the information received about the same target and same role coming from communicated reputations. On the other hand, shared evaluation is the equivalent for communicated images and third party images.

Shared voice predicates will finally generate candidate reputation; shared evaluation together with outcomes will generate candidate image. Newly generated candidate reputation and image are usually not strong enough; new communications and new direct interactions will contribute to reinforce them until a threshold, over which they become full-fledged image or reputation. We refer to [5] for a much more detailed presentation. From the point of view of the agent strucuture, integration with the other parts of our deliberative agents is stright-forward. Repage memory links to the main memory of the agent that is fed by its communication and decision making module, and at the same time, this last module, the one that contain all the reasoning procedures uses the predicates generated by Repage to make decisions.

6 Simulation Experiment

In [4] we applied our system REPAGE to a simulation experiment of the simplest setting in which accurate information is a commodity: an agent-based market with instability. The model has been designed with the purpose of providing the simplest possible setting where information is both valuable and scarce. The system must be considered as a proof of concept, not grounded on micro or macro data, but providing an abstract economic metaphor. This simplified approach is largely used in the reputation field (see for example [11]), both on the side of the market design and of the agent design in the study of market with asymmetric information. We follow this approach since our main interest is on the side of agent design, and we must be able to clearly separate complex effect due to agent structure from ones due to market structure.

6.1 Design of the Experiment

The experiment includes only two kind of agents, the buyers and the sellers. All agents perform actions in discrete time units (turns from now on). In a turn, a buyer performs one communication request and one purchase operation. In addition, the buyer answers all the information requests that it receives. Goods are characterized by an utility factor that we interpret as quality (but, given the level of abstraction used, could as well represent other utility factors as quantity, discount, timeliness) with values between 1 and 100.

Sellers are characterized by a constant quality, drawn following a stationary probablilty distribution, and a fixed stock, that is decreased at every purchase; they are essentially reactive, their functional role in the simulation being limited to providing an abstract good of variable quality to the buyers. Sellers exit the simulation when the stock is exhausted and are substituted by a new seller with similar characteristics but with a new identity (and as such, unknown to the buyers). This continuous seller update characterises our model, for example in comparison with recent work as [12], where both sellers and buyers are essentially fixed.

The disappearance of sellers makes information necessary; reliable communication allows for faster discover of the better sellers. This motivates the agents to participate in the information exchange. In a setting with permanent sellers (infinite stock), once all buyers have found a good seller, there is no reason to change and the experiment freezes. With finite stock, even after having found a good seller, buyers, should be prepared to start a new search when the good seller's stock ends.

At the same time, limited stock makes good sellers a scarce resource, and this constitutes a motivation for the agents not to distribute information. One of the interests of the model is in the balance between these two factors.

There are four parameters that describe an experiment: the number of buyers NB, the number of sellers NS, the stock for each seller S, and the distribution of quality among sellers. We defined the two main experimental situations, L1 where there is only exchange of image, and L2 where both image and reputation are used.

6.2 Decision Making Module

In [4], the decision making procedure was shown to play a crucial role in the performance of the whole system. As to sellers, the procedure is quite simple since they sell products required and disappear when the stock gets exhausted. As to buyers, instead, the algorithm is rather more complex. At each turn they must interrogate another buyer, buy something from a seller, and possibly answer a question from another buyer. Each of these actions leads to a number of decisions to be taken.

- Buying. Here the question to be answered is which seller a buyer should turn to. The easiest option would be to pick the seller with the best image, or (in L2) the best reputation if image is not available. A threshold is set for an evaluation (actually, for its center of mass, see [5] for definitions) to be considered good enough and be used for choosing. In addition, a limited chance to explore other sellers is possible, as controlled by the system parameter risk 3. Notice that image has always priority over reputation, since unlike reputation image implies that the evaluation is shared by the user.

- Asking. As in the previous case, the first choice to be made is which agent to be queried, and the decision making procedure is exactly the same as that for choosing a seller, but now agents deal with images and reputations of targets as informers (informer image) rather than sellers. Once decided whom to ask, the question is what to ask. Only two queries are allowed:
  - Q1 - Ask information about a buyer as informer (basically, how honest is buyer X as informer), and
  - Q2 - Ask for some good or bad seller (for instance, who is a good seller, or who is a bad seller). Notice that this second question does not refer to one specific individual, but to the whole body of information that the queried agent may have. This is in order to allow for managing large numbers of sellers, when the probability to choose a target seller that the queried agent has information about would be low. The agent will ask one of these two questions with a probability of 50%. If Q1 is chosen, buyer X as informer would be the least known, i.e., one with less information to build up an image or reputation about.
- Answering. Let agent S be the agent asking the question, R the agent being queried. Agents can lie, either because they are cheaters or because they are retaliating. When a buyer is a cheater, they provide information after having turned its value into the opposite. Retaliation is accomplished by sending inaccurate information (for instance, sending I-dont-know when it has information, or simply giving the opposite value) when R has got a bad image of S as informer. In L1 retaliation is done by sending a I-dont-know message even when R has got information. This avoids possible retaliation from S since a I-dont-know message implies no commitment. If reputation is allowed, (L2) retaliation is accomplished in the same way as if the agent were a liar, except that image is converted into reputation in order to avoid potential retaliations from S. Fear of retaliation leads to sending an image only when agent is certain about evaluation. This is yet another parameter (Strength) allowing the fear of retaliation to be implemented. Notice that if strength is null, there is no fear since any image will be a candidate answer, no matter what its strength is. As strength increases, agents become more conservative, with less image and more reputation circulating in the system.

6.3 Expected Results

Based on the hypthesis that image allows for more retaliation than reputation, we expect the following results to obtain:

- H1 Initial advantage: L2 shows an initial advantage over L1, that is, L2 grows faster.
- H2 Performance: L2 performs better as a whole, that is, the average quality at regime is higher than L1.

Some questions concerning cheating and fairness were also investigated:

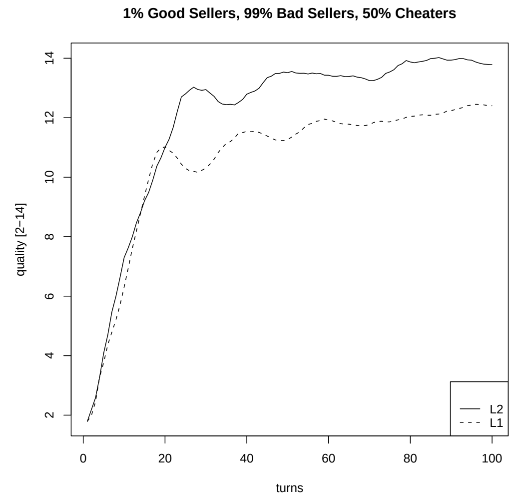


Fig. 1. Accumulated average quality per turn in condition A1 (very few good sellers), 50% informational cheaters, for L1 and L2 agents. L2 agents show better performance even with large amount of false information.

- cheaters advantage: do cheaters effectively reach a significant advantage thanks to their behavior?
- Cheaters' effects: are cheaters always detrimental to the system? In particular, is the performance of the system always decreasing as a function of the number of cheaters?

Simulations to enquire on the relationship between L1 and L2 has been run with the following parameters: with fixed stock (50), number of buyers (25), and number of sellers (100); different values of informational cheaters (percentages of 0%, 25% and 50%); different values of bad sellers, ranging from the extreme case of 1% of good vs 99% of bad sellers (A1), going trough 5% good sellers Vs 95% bad sellers (A2), 10% good sellers vs 90% bad sellers (A3), and finally, to another extreme where we have 50% of good sellers vs 50% of bad sellers (A4). Note that from A1 to A4 the maximum level of quality obtainable increases (from experimental data, we move from a regime maximum quality of about 14 in A1 to nearly full quality in A4). For each one of these conditions and for every situation (L1 and L2) we run 10 simulations. In the figures we present the accumulated average earnings per turn in both situations, L1 and L2. In L1 the amount of useful communications (different from Idontknow) is much lower that in L2, due to the fear of retaliation that governs this situation. In conditions where communication is not important, the difference between the levels disappears.

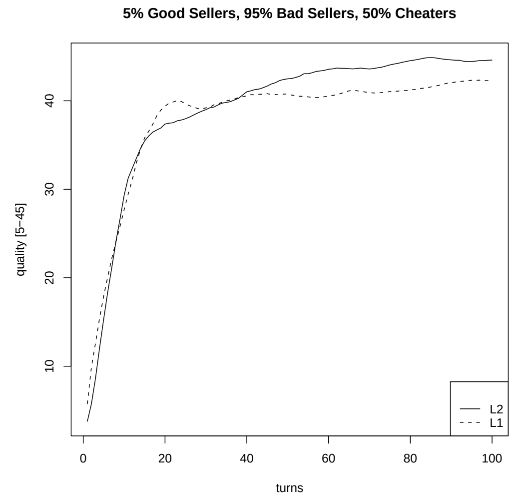


Fig. 2. Accumulated average quality per turn in condition A2 (5% good sellers), 50% informational cheaters, for L1 and L2 agents. The margin of L2 on L1 is reduced.

In the following, we report only the result of the experiment with cheaters, where the difference between L1 and L2 is made more significative by the presence of false information. For a full report, please refer to [4].

6.4 Experiments with Cheaters

We report results of experiments with 50% of informational cheaters in conditions A1, A2, A3 and A4. The large amount of false information produces a bigger impact in situations and conditions where communication is more important. Quality reached in L1 shows almost no decrease with respect to the experiment without cheaters, while L2 quality tends to drop to L1 levels. This shows how the better performance of L2 over L1 is due to the larger amount of information that circulates in L2. In Figure 1, nothwithstanding the large amount of false information, there is still a marked difference between the two levels. Essentially, L2 agents show a better performance in locating the very rare good sellers. The situation starts to change in Figure 2, where the two algorithms are more or less comparable; here, the larger amount of good sellers does not make necessary the subleties of L2. In Figure 3, with an even larger amount of good sellers available, the two algorithms show the same level of performance.

7 Conclusions and Future Work

Results indicate that reputation plus image (as opposed to image only) improves the average quality of products exchanged in the whole system. The value added of reputation is shown under the occurrence of

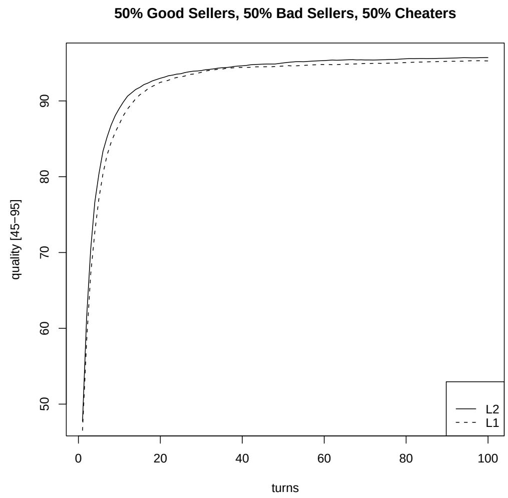


Fig. 3. Accumulated average quality per turn in condition A4 (half good sellers), 50% informational cheaters, for L1 and L2 agents. The two levels are indistinguishable.

- retaliation: personal commitment associated to image transmission exposes the agent to possible retaliation if inaccurate information was sent. Conversely, reputation transmission does not lead to such a consequence, but at the same time provides agents with information that might be useful to select satisfactory partenrs. Future work will concern the effect of cheaters over the whole system in presence of a norm that prescribes agents to tell the truth. The reputation mechanism will turn into a social control artifact aimed to identify and isolate agents that do not follow that norm.
- Communication: There is no reputation without communication. Therefore, scenarios with no or poor communication are irrelevant for the study of reputation. However, in virtual societies with autonomous communicating agents that need to cooperate and are enabled to choose partners, reputation considerably increases the circulation of information and improves the performance of their activities. In our experiments, even when there is no penalty for direct interactions and only one question per turn is allowed, the introduction of reputation improves the average quality per turn. In scenarios where quality is scarce and agents are completely autonomous this mechanism of social control makes the difference.
- Decision making procedure: The decision making model implemented has a decisive impact on the system's performance. In fact, this is where the agent may take advantage of the distinction between image and reputation. In future work, we will elaborate on this distinction, possibly reformulating it in terms of textitmeta decision making, a very promising future line of work to better ground and exploit the image and reputation artefacts.

These results gives us reasons to draw some more general conclusions about the respective role of image and reputation. The antisocial consequence of image spreading seems to be clearly documented in the experiment we have reported upon. But if this is the case, we also find evidence for our argument that social networks based upon image perform more poorly than networks based upon reputation at least when partner selection is a common goal of the network members. An image-network, based upon acquaintanship, if not familiarity, and trusted communication of own evaluations, stimulates retaliation or at least discrimination when informers are found to spread incorrect information. Consequently, such a type of network shows poor robustness against not only deception and cheating, but also errors and rumour.

Conversely, reputation-based networks are more flexible and inclusive, they tolerate errors. Though selecting information before using it, the reputation mechanism does not lead recipients to discard so easily nor, a fortiori, retaliate against bad informers. In such a way, the chain of retaliations is prevented and the consequent lowering of the exchanges' quality is reduced.

Does such a view of reputation point to an account of the evolution of socially desirable behaviour, concurrent with the classical one, based on punishment and strong reciprocity (cf. [13])? Hard to say for the time being. However, this is a fascinating resarch hypothesis for future studies.

References

1. Marmo, S.: L'uso della reputazione nelle applicazioni internet: prudenza o cortesia? l'approccio socio-cognitivo. In: AISC - Terzo Convegno Nazionale di Scienze Cognitive. (2006)
2. Tadelis, S.: What's in a name? reputation as a tradeable asset. The American Economic Review (1999)
3. Conte, R., Paolucci, M.: Reputation in artificial societies: Social beliefs for social order. Kluwer Academic Publishers (2002)
4. Pinyol, I., Paolucci, M., Sabater-Mir, J., Conte, R.: Beyond accuracy. reputation for partner selection with lies and retaliation. In: MABS 07, Eighth International Workshop on Multi-Agent-Based Simulation. (2007)
5. Sabater, J., Paolucci, M., Conte, R.: Repage: Reputation and image among limited autonomous partners. Journal of Artificial Societies and Social Simulation **9**(2) (2006)
6. Conte, R., Castelfranchi, C.: Cognitive Social Action. London: UCL Press (1995)
7. Conte, R.: Social intelligence among autonomous agents. Computational and Mathematical Organization Theory **5** (1999) 202–228
8. Leslie, A.M.: Pretense, autism, and the 'theory of mind' module. Current Directions in Psychological Science **1** (1992) 18–21
9. Miceli, M., Castelfranchi, C. In: The Role of Evaluation in Cognition and Social Interaction. Amsterdam:Benjamins (2000)
10. Castelfranchi, C., Poggi, I.: Bugie, finzioni e sotterfugi. Per una scienza dell'inganno. Carocci Editore, Roma. (1998)
11. Sen, S., Sajja, N.: Robustness of reputation-based trust: boolean case. In: AAMAS '02: Proceedings of the first international joint conference on Autonomous agents and multiagent systems, New York, NY, USA, ACM Press (2002) 288–293

12. Izquierdo, S.S., Izquierdo, L.R.: The impact of quality uncertainty without asymmetric information on market efficiency. Journal of Business Research **60**(8) (August 2007) 858–867
13. Fehr, E., Fischbacher, U., Gächter, S.: Strong reciprocity, human cooperation and the enforcement of social norms. Human Nature **13** (2002) 1–25

The Effect of Social Influence on Market Inequalities in the Motion Picture Industry

Sebastiano A. Delre[1], Thijs L.J. Broekhuizen[2], Wander Jager[1]

[1] Department of Marketing
[2] Department of Strategy and Innovation
Faculty of Economics and Business
University of Groningen
P.O. Box 800, 9700 AV Groningen, the Netherlands
Email addresses: s.a.delre@rug.nl; t.l.j.broekhuizen@rug.nl; w.jager@rug.nl

Abstract. In this paper we investigate the degree to which two social influences, imitation and coordinated consumption, effectuate inequalities in the motion picture industry. We develop an agent-based model based on micro movie visitors' decision-making that generates the observed macro market outcomes. The simulation model makes use of the findings of an empirical survey amongst 1112 cinema visitors. We find that social influences explain market inequalities and that the impact of coordinated consumption on market inequalities is stronger than the impact of imitation.

Keywords: motion picture market, market shares, agent based systems, social influence, imitation, coordinated consumption.

We thank Tammo Bijmolt for his useful comments on previous drafts of the paper. We are very much indebted to Fang Zeng, Maria Rotondi and Annalisa Delre for their assistance in collecting the data.

1 Introduction

Consumers are guided by social and individual needs. Their decisions to purchase or consume a product are influenced by individual values such as self-fulfillment, sense of accomplishment, and self-respect and by social values such as being well-respected, warm relationships with others, etc. [1]. This study investigates how individual and social needs shape the behavior of movie visitors in the cinema market and what the consequences are at the macro level of the market (i.e. distribution of revenues at the box office). Although in the last decade the motion picture market has been object of an increasing number of studies, especially in the marketing field, the large majority of these works have focused on the supply of the industry (for a review see [2]) and, in particular on the effects of marketing efforts on the movies' box office [3-12]. Still, very little is known about how movie goers decide to see a movie, what kind of evaluative criteria they use and, more importantly, how strong the influence of friends, relatives and others is in deciding which movies to visit [13].

Social influences play a dominant role in the motion pictures industry [14-15]. In this study we focus on two types of social influences, namely imitation effect (the influence of other consumers that have already seen the movie) and coordinated consumption effect (the influence of other consumers that have not seen the movie but they are informed about it and they may still want to visit it) and we investigate their separate effects on the distribution of the box office revenues of the movies. Additionally, this study investigates whether the degree of these two types of social influences differs between two types of movies: art house and mainstream movies. The rationale behind this is that movie goers' motivations, attitudes and behaviors substantially differ for these two types of movies [16-18].

With the aim of obtaining useful insights for our agent based model, we empirically investigate the decision making of cinema visitors. Through the use of questionnaires we have collected data from more than one thousand respondents. Basically the data show that the movie goers' decision making and behaviors highly differ across movie types. We first construct a social orientation scale based on the distinction between social and individual motivations and we then find evidence that this scale is higher for visitors of mainstream movies than for visitors of art house movies, that it is negatively correlated with the frequency of attendance and that it is relatively high for genres such as animation and comedy and low for drama and history. We use these insights for the experimental setting of our simulation model and test their effects on market outcomes. We find that market inequalities are explained by both kinds of social influences and that coordinated consumption impacts market inequalities more strongly than imitation does.

2 Social Influence: Imitation and Coordinated Consumption

Social influence has long been recognized as an important force shaping consumer behaviors [18-23]. Social influence may occur before consumption (e.g., when seeking information and/or receiving word-of-mouth), during consumption (e.g., when others are present in consumption contexts), and after consumption (e.g., producing word-of-mouth) [22].

Consumers adjust their behaviors to match the expectations of other people or a reference group [24-25]. Previous research distinguished between informational and normative social influence [26-27]. Informational influence occurs through a process of internalization, where information from others is accepted as evidence about reality (Deutsch and Gerard, 1955). Here, consumers can make more informed and accurate decisions with the help of others [25]. Normative influence is an influence to comply with the expectations of others [28]; consumers can conform to the expectations of others in order to receive rewards or avoid punishments (i.e., utilitarian value), and/or to maintain or enhance their self-image (i.e., value-expressive value) [26][29]. In sum, consumers can gain additional value by attaining three goals: (1) making more informed and accurate decisions, (2) receiving rewards or avoiding punishments, and (3) maintaining or enhancing self-image.

This study focuses on two types of social influence namely *imitation* and *coordinated consumption* [30-31]. Imitation refers to the degree to which consumers are influenced by past behaviors and evaluations of others. It is likely to stimulate visiting behavior, as visitors become aware of the movie and its quality and may engage in imitative behaviors. Imitation is more likely to occur when a greater part of friends have already visited the movie [30]. It is informative in nature[1], as consumers can make more informed and accurate decisions with the help of the recommendations of others. Coordinated consumption refers to the degree to which consumers are influenced by the intended behaviors of other consumers. It positively influences visiting behavior, as consumers are more likely to visit a movie when a greater part of their friends still wants to visit the movie, as they can more easily find friends to accompany them [30]. It is normative in nature, as the consumption of products together is frequently used as a means to strengthen the social bond, and to enhance one's self-image [32]. The strong normative influences in the motion picture industry may partly explain why consumers frequently visit movies together. Normative influences are particularly salient in this industry because the behaviors can be directly observed by friends and relatives [26-29][33] and because movies are hedonic products that are inherently value-expressive.

Apart from the social influences, we also consider the motivations of visitors. Consumer motivations can help explaining the relative strength of social influences on visiting behaviors. Motivation research argues that consumers have individual needs (e.g. learning about self, need for cognition, self-fulfillment, and sense of accomplishment) and social needs (e.g. sense of belonging, warm relationships with others, and experience fun together) that drive their behaviors [34]. This distinction is very useful in explaining the individual's orientation or inclination towards social influence, i.e. people who have strong social needs are concerned with the fulfillment of their social needs and derive more value from it than those who have less strong social needs. Moreover, consumer motivations can also explain the preference for a certain type of movies. For example, an empirical motivational study distinguished between two types of movie visitors: those who seek intellectual curiosity, are interested in human and social issues, value esthetics, and see "film as a form of art", and those who seek excitement, relaxation and want to maintain social relationships [17]. The former visitors, who predominantly visited human and social drama movies, clearly resemble art house movie visitors. The latter visitors, who chiefly visited entertainment movies, can be denoted as mainstream movie visitors.

Based on these insights, we expect that visitors of art house movies have stronger individual motivations than visitors of mainstream movies who are more outer-oriented and have stronger social motivations. We assume that social and individual motivations can explain the degree of social influence, which influences the preference for a movie type (mainstream vs. art house movies).

[1] Imitation may also entail normative aspects (receiving rewards/avoiding punishment, and enhancing self-image) as consumers make norms salient by giving their opinions; however, we believe that the informational aspect of recommendations is dominant.

3 Social versus Individual Motivation: Evidence from a Survey on Movie Visitors

We develop a questionnaire to investigate the degree to which imitation and coordinated consumption occurs, and how these social influences correlate with type of movie, individual and social motivations, and attendance frequencies.

Data have been collected by surveying 1112 cinema visitors that collectively visited 43 movies[2]. We have surveyed visitors that either visited an art house or a mainstream movie [35]. The rationale behind this is that movie goers' motivations, attitudes and behaviors substantially differ for these two kinds of movies [15-18].

Mainstream movies and art house movies differ significantly in a number of ways, but most dominantly in terms of their market share. According to the MPAA[3], art house movies only account for 3 percent of the total box offices. Moreover these types of movies can be distinguished in terms of their association with particular genres, narrative structures and contents [35-36]; the degree of artistic versus commercial qualities [37-38]; the budget size, the participation of movie stars and the occurrence of special effects [35][37]; the number of opening screens [18][39]; and the type of film distributor [8][35][4].

In the existing literature different definitions are used to characterize art house versus mainstream movies. In accordance with prior studies [8][35], we code the movies based on the type of cinema (art house versus mainstream cinema) in which the movie is released. In our data set, examples of art house movies are *Libertine* and *Transamerica*, and examples of mainstream movies are *The DaVinci Code* and *Over the Hedge*.

---

[2] Data have been collected from May until September 2006 in the Netherlands, in China and in Italy. In our analyses, we pooled the data together. Although cinema visitors' behaviors may vary across countries due to cultural differences, this analysis is beyond the scope of this research. We performed additional checks to investigate possible bias effects. The exploratory factor analysis for each country revealed that the same factor structure of individual and social motivations was found in all three countries, meaning that in each country they consider the same distinction between social and individual motivations. Although the mean scores differed across countries for individual motivations ($F(2, 1098) = 21.5$, $p<.001$) and for social motivations ($F(2, 1098) = 45.3$, $p<.001$), in all countries social motivations were stronger than individual motivations. Next the strength of social motivations relative to individual motivations associated with each movie type and movie genre appeared very similar (e.g., in each country social motivations were strongest animation/family and weakest for biography/history). As such, we assume that the possible biases due to cultural differences are rather small, especially because we focus on the link between social and individual motivations with social influences and not on the strength of them.

[3] Theatrical market statistics of the Motion Picture American Association (MPAA). Accessible through http://www.mpaa.org/

[4] In a previous empirical survey [15] we have collected longitudinal data about the kind of consumption of consumers that attended two movies during their complete life cycles. One, *Brothers*, was assumed to be a typical example of art house movie and another one, *The Interpreter*, assumed to be a typical example of mainstream movie. Although the data of this survey brought strong empirical evidence that the kind of consumption highly differed for the two movies, it was quite surprising to see that the analysis of this data did not find any significant correlation between the type of consumption and the time of the consumption. This is the reason why, in this following survey customers' heterogeneity due to timing effects is assumed to be limited and we could submit the questionnaires to movie visitors at any time of the movie life cycle.

Table 1 shows how the respondent profiles (gender, age, education and attendance) and the use of information sources differ for these two types of movies. Subsequent χ^2 tests indicate that females are more likely to go to art house movies, whereas males are more likely to visit mainstream movies ($\chi^2(1)= 7.9$, $p<.001$). Consistent with prior findings [17], our sample indicates that particularly younger people (< 20 years) tend to visit mainstream movies while older people (36 and older) are more likely to visit art house movies ($\chi^2(4)= 70.3$, $p<.001$), and that art house visitors are more highly educated than mainstream movie goers ($\chi^2(2)= 134.5$, $p<.001$). Moreover, art house movie goers also go more regularly than mainstream movie goers (Z-value=3.45, p=.002). Finally, mainstream and art house visitors also differ in their use of and reliance on information sources; compared to art house visitors, mainstream visitors care less about the movie quality before seeing the movie (p=.007), find marketing information sources (ads, trailers and posters) to be more useful (p=.015), and critics' and visitor reviews (magazines, newspapers, and internet) less useful (p<.001). Mainstream movies have greater signaling properties (famous actors and famous directors) and advertise heavily to attract consumers; therefore, mainstream visitors may have a less strong need to engage in search activities.

Table 1. Respondent's profile per movie type

	Mainstream movie sample N=557	Art house movie sample N=535	Difference test to check for dependency
Gender			
Male	259 (45.8%)	204 (37.4%)	$\chi^2(1)= 7.9$, $p<.001$
Female	307 (54.2%)	341 (62.6%)	
Age			
<21	192 (34%)	72 (13%)	$\chi^2(4)= 70.3$, $p<.001$
21-25	123 (22%)	145 (27%)	
26-30	75 (13%)	92 (17%)	
31-35	66 (12%)	67 (12%)	
Education			
Primary school	22 (4%)	5 (1%)	$\chi^2(2)= 134.5$, $p<.001$
High school	184 (32%)	87 (16%)	
Secondary school	86 (15%)	31 (6%)	
College degree	99 (18%)	120 (22%)	
University degree	137 (31%)	301 (55%)	
Number of visits per year	11.6 (SD=15.9)	12.1 (SD=11.8)	Z-value=3.45, p=.002
In general, I want to know something about the movie quality before seeing the movie	3.56 (SD=1.05)	3.75 (SD=1.14)	t-value=2.70, p=.007
Opinions of people who have already seen the movie are useful sources of information	3.54 (SD=1.05)	3.62 (SD=1.10)	t-value=1.58, p<.209

TV ads, trailers, posters, etc. are useful sources of information about movie quality	3.51 (SD=1.14)	3.34 (SD=1.19)	t-value=2.44, p=.015
Reviews on magazines, newspapers, and internet are useful sources of information.	3.49 (SD=1.11)	3.80 (SD=1.05)	t-value=4.69, p<.001

Note: χ^2 values are referring to chi square tests; Z-values are referring to Mann-Whitney tests; t-values are referring to Independent sample t-tests.

The survey contains items to measure imitation and coordinated consumption. Imitation, referring to the degree to which consumers are influenced by other consumers' past behaviors and evaluations, is assessed by asking respondents to report the number of people that recommended them to see the movie. Coordinated consumption, referring to the acquaintances that still want to go to see the movie, is measured by using the number of companions of their current visit. In Table 2 we show the items used to measure imitation and coordinated consumption and how these differ across movie type. It appears that art house movie goers receive fewer messages from their environment than mainstream movie visitors (Z=3.44, p=001), and that coordinated consumption is stronger for mainstream movie visitors as they attend in larger groups (Z=6.88, p<.001).

Table 2. Imitation and coordinated consumption per movie type

	Mainstream movies	Art house movie	Z-value	Significance
Imitation: How many people had recommended you to see this movie?	1.89 (SD=3.27)	1.27 (SD=2.70)	3.44	.001
Coordinated consumption: With how many people did you see the movie?	2.28 (SD=1.57)	1.66 (SD=1.27)	6.88	<.001

In order to shed light on the drivers of the movie goers' decision making we investigate their individual and social motivations. This motivation classification can help explaining which movie type (art house versus mainstream) movie goers visit, and to what degree social influences take place. We test the underlying factor structure by performing exploratory factor analysis. The two factors (individuals and social motivations) found in the dataset explain more than 60% of the variance in the data. Each item loads highly (>.60) on its assigned factor (Table 3).

Table 3. Exploratory factor analysis for social motivations and individual motivations

Item	Social motivations	Individual motivations
I go to see a movie at the cinema especially because I want to have an enjoyable evening with friends/ partners/family members.	.66	
I go to see a movie at the cinema just to spend some time.	.78	
I go to see a movie at the cinema for pure entertainment.	.80	
I go to see a movie at the cinema in order to develop my own idea about a specific issue.		.78
I go to see a movie at the cinema in order to become an expert about movies.		.81
Eigenvalues	1.70	1.36
Variance explained	33.9%	27.1%

Notes: Exploratory factor analysis (N=1101) is performed with Oblimin rotation and based on the eigenvalues criterion. Items are measured using 5 point Likert scales, anchoring at 1=totally disagree to 5=totally agree.

The results of the factor analysis allow us to construe scales for individual motivations and for social motivations based on item means. Table 4 displays the descriptives of these scales per movie type. Although social motivations score higher than individual motivations in both settings, the results clearly demonstrate that art house movie goers have stronger individual motivations relative to mainstream movie goers.

At this point it is possible to obtain a single scale which indicates the *relative* strength of social motivations by considering the strength of social motivations relative to the sum of individual and social motivation: the social *orientation* scale. We investigate how this social orientation scale correlates with imitation, coordinated consumption and the number of visits (Table 5). Correlations are not particularly high, but they are mostly significant. Our social orientation scale is correlated (ρ=.09) with coordinated consumption (i.e., number of companions attending the movie). Moreover, as also expected from the results reported in Table 1, we find a negative correlation between the frequency of attendance and the social orientation scale (ρ=-.15).

Finally, we also check whether the social orientation scale differs across genres[5]. Table 6 ranks the social orientation scale according to the genres of the movies. The analysis reveals significant differences between genres. In particular

[5] The genre of each movie is based on IMDb's classification. Sometimes, multiple genres are assigned to the same movie (e.g., DaVinci Code was coded as *drama*, *mystery* and *thriller*).

animation/family, comedy and adventure/thriller/action/fantasy score high and romance, drama and history/biography score low on social orientation.

Table 4. Mean scores of social versus individual motivations per movie type

	Mainstream movies	Art house movie	t-value	Significance
Social motivations	3.42	3.02	7.22	<.001
Individual motivations	2.35	2.51	-2.79	.005
Social orientation: Social motivations/ (individual + social motivations)	0.55	0.60	6.62	<.001

Note: Constructs are transformed to 5 point Likert scales. 1=weak motivation, 5=strong motivation.

Table 5. Correlations between social orientation, imitation, coordinated consumption, and frequency of attendance

	Social orientation: social motivations/ (individual + social motivations)	Imitation: number of people that had recommended the respondent to go to the movie	Coordinated consumption: number of people that went with the respondent to the movie
Imitation: number of people that had recommended the respondent to go to the movie	-.06		
Coordinated consumption: number of people that went with the respondent to go the movie	.09**	.11**	
Number of visits per year	-.15**	-.04	-.11**

*Note: * sign. at .05; ** sign. at .01 based on one-tailed tests*

Table 6. Mean scores social motivations and individual motivations per movie genre

	Social motivations	Individual motivations	Social orientation: Social /(Individual + Social motivations)
animation/family	3.37	2.25	0.60
comedy	3.38	2.39	0.59
adventure/thriller/action/fantasy	3.33	2.42	0.58
romance	3.15	2.33	0.57
drama	3.15	2.43	0.57
biography/history	2.61	2.52	0.52

Note: Constructs are transformed to 5 point Likert scales. 1=weak motivation, 5=strong motivation. Social orientation ranges from 0 (very weak social orientation to 1 (very strong social orientation).

4 The agent based model

Here below we present the complete simulation model of the motion picture market, and its assumptions. The core of this agent based model is the individual decision-making of movie goers. All agents decide which movie to see at each time step. After agent i is informed about the movies according to (1) or (2), it evaluates the expected utilities of these movies according to (3) and it visits the movie that has the highest expected utility.

$BUZZ_{jt}$ is the buzz of movie j at time t that is generated by the advertising of the movie. In our agent based model it formalizes the probability that agent i is informed about movie j at time t. At time 0, just before the movie is released into the cinema theaters, $BUZZ_{j0}$ depends on the advertisement budget of movie j M_j, and on ω_1 which is a free parameter of the model and it indicates how strong the informative effect of the advertising budget is. After the movie is released, $BUZZ_{jt}$ evolves as specified in (2). $Box_{j,t-1}$ is the box office movie j has obtained at the previous time step and δ_1 is a free parameter. This formalization assumes that $BUZZ_{jt}$ evolves according to the success that the movie j has at the box office. The more the success a movie gains after release, the higher its buzz becomes. Here δ_1 determines how fast the evolution toward the actual box office of the movie is after its release. If δ_1 is very low then agents retain the effects of advertisement budget longer and they are less affected by the results that the movie has at the box office; if δ_1 is very high then agents forget sooner the effects of the initial campaign and they are more affected by the results that the movie has at the box office.

$$BUZZ_{j0} = e^{-\frac{\omega_1}{M_j}} \tag{1}$$

$$BUZZ_{jt} = BUZZ_{j,t-1} + \delta_1 \cdot \left(Box_{j,t-1}/N - BUZZ_{j,t-1}\right) \tag{2}$$

The agent's utility consists of two components: individual utility and social utility (3). Individual utility is based on the fit between the individual preferences and the movie characteristics (4). This fit, $\lfloor 1-(m_j - p_i) \rfloor$, is measured by the distance between preferences of agent i, p_i, and the characteristics of movie j, m_j. Social utility is derived from what other agents do (5). Two concepts are formalized in the social utility: imitation effect a_{jt}, given by the proportion of agents that have already seen movie j (6) and coordinated consumption effect w_{jt}, given by the proportion of agents that are informed about the movie j but have not seen it yet (7). The individual utility increases proportionally to the degree to which the movie characteristics match the preferences of the agent, p_i. The social utility increases linearly when both the effects of coordinated consumption and imitation increase.

$$E\lfloor U_{ijt} \rfloor = \beta_i \cdot x_{jt} + (1-\beta_i) \cdot y_{ij} \tag{3}$$

$$y_{ij} = 1 - | m_j - p_i | \tag{4}$$

$$x_{jt} = a_{jt} + w_{jt} \tag{5}$$

$$a_{jt} = \frac{TotBox_{jt}}{N} \tag{6}$$

$$w_{jt} = BUZZ_{jt} \cdot \left(1 - \frac{TotBox_{jt}}{N}\right) \tag{7}$$

As we assume that an agent cannot visit the same movie more than once, coordinated consumption and imitation are proportions of same market and, hence, the proportions cannot sum up to more than 1. As such, an increase in imitation corresponds with a decrease in coordinated consumption and vice versa. This formalization has a strong shortcoming: because the proportion of agents that have seen the movie (imitation) and the proportion of agents that are informed about the movie but that have not seen it yet (coordinated consumption) are competing proportions of the same population, the social utility function is convex, which is an unrealistic assumption. When, for example, a third of the agents has seen the movie and another third of the agents is informed about the movie but has not seen it yet, the social influence is weaker than when two thirds of the agents have seen the movie or than when two thirds of the agents are informed about the movie but have not seen it yet. To overcome this limitation, we propose a refined formalization where the social utility function is concave (8). In this case, social utility increases at a decreasing rate when both the effects of coordinated consumption and imitation increase. Here c_1 weights the importance of imitation and it determines a minimum level of social influence when w_{jt} is equal to 0 and c_2 weights the importance of coordinated consumption and it determines a minimum level of social influence when a_{jt} is equal to 0.

$$x_{jt} = \frac{c_1 \cdot a_{jt} + c_2 \cdot w_{jt}}{a_{jt} + w_{jt} + c_1 + c_2} \tag{8}$$

The social component and the individual component are weighted by the parameter β_i. This parameter indicates the attitudes of the agents towards the consumption and it measures how strong the social utility is compared to the individual utility. β_i corresponds with the social orientation scale that we have constructed and analyzed in the previous section. Consequently, in the following simulation experiments, settings with high $\overline{\beta}$ formalize markets where movie goers tend to see mainstream movies (strong social orientation) and settings with low $\overline{\beta}$ formalize markets where movie goers tend to see art house movies (weak social orientation).

In order to study how movie revenues are distributed into the market, at the end of each simulation run for each of the *M* movies, we record the number of visitors v_j, compute its market share s_j (9) and we then study the overall market inequality of market shares computing the GINI coefficient g [40]. This can vary from 0 (completely equal market shares for all movies) to 1 (a single movie takes it all) (10).

$$s_j = \frac{v_j}{\sum_{j=1}^{M} v_j} \tag{9}$$

$$g = \frac{\sum_{i=1}^{M} \sum_{j=1}^{M} \left| s_i - s_j \right|}{2 \cdot M \cdot \sum_{k=1}^{M} s_k} \tag{10}$$

The simulation model described above is implemented in a realistic USA cinema market context[6]. Each time step of the simulation corresponds to a week and at each time step new movies are introduced into the market. The model generates 480 movies per year, for 3 years. We select only the 480 movies that enter the market during the second year and we record their complete life cycle at the box office. In this way we avoid initial and final simulation distortions. As such, we only consider the competition of movies that are introduced in the first year and reach the second year, and the complete life cycles of movies introduced in the second year that reach the third year. To make the simulations more realistic, we take into account the famous season effect [41-43]: the number of agents making a decision at each time step given by a probability p_t is proportional to the attendance observed in the real market and the number of movies released each week is also proportional to the attendance. Finally we draw marketing budgets M_j from real data[7] and we set ω_1=50,000,000 and δ_1=0.5. These parameters' values are based on theoretical foundations: the informative effect of advertising is stronger, more persistent and more prevalent than the persuasive effect [3][7]. In our simulation runs, this setting of the parameters makes the majority of the agents aware of the movie before its release

[6] We refer to the Motion Picture Association of America (MPAA), http://www.mpaa.org.

[7] Data have been obtained from http://www.variety.com/, http://www.the-numbers.com, http://www.imdb.com. Marketing expenditures vary linearly from a minimum of $ 7,500 to a maximum of $ 37,000,000.

and lets the advertising effect remain effective for about 4-6 time steps after the release of each movie.

5 Results

In section 5.1 we explore how marketing inequalities change at different levels and for different formalizations of social orientation. In section 5.2 we implement the guidelines of the empirical investigation presented in section 3 and we study their consequences.

5.1 The effects of social orientation and social influences on market outcomes

We begin exploring the outcomes of the model simulating a simple artificial market of 50,000 agents where movies are assumed to have different characteristics (m_j =[0, 1]) and agents have different preferences (p_i =[0, 1]). We vary β_i simulating different markets: from a low $\overline{\beta}$ ($\overline{\beta}$ =[0.0, 0.5]) which implies a weak social orientation to a very high $\overline{\beta}$ ($\overline{\beta}$ =[0.5, 1.0]) which implies a strong social orientation.

The results reported below refer to the 200 movies with the highest market shares. Table 7 shows the variations of the GINI coefficient g for different levels of social orientation. These results clearly show that market inequalities arise because of social influences and increase according to the degree of social orientation included in the decision making of the agents.

However, as mentioned in section 4, the formalization of social utility can be more sophisticated and realistic (8). We investigate the implications of this formalization for imitation and coordinated consumption by setting the model with plausible values (β_i =[0.25, 0.75]; m_j =[0, 1]; ω_1 =50000000 and δ_1 =0.5) and conducting simulation runs for different values of c_1 and c_2. Table 8 shows the variations of the GINI coefficient g for different values of c_1 and c_2. Both the effects of imitation and coordinated consumption increase market inequalities, but it is evident that coordinated consumption has a stronger impact on market inequalities than imitation.

Table 7. GINI coefficient values for different levels of social orientation (from very weak to very strong social orientation)

Social orientation	GINI coefficient
$\overline{\beta}$ =0.25	0.531
$\overline{\beta}$ =0.35	0.553
$\overline{\beta}$ =0.45	0.577

$\overline{\beta}$ =0.55	0.599
$\overline{\beta}$ =0.65	0.619
$\overline{\beta}$ =0.75	0.634

Table 8. GINI coefficient values for different weights for imitation effect and coordinated consumption effect

	c_2 =0.1	c_2 =0.3	c_2 =0.5	c_2 =0.7	c_2 =0.9
c_1 =0.1	0.505	0.522	0.530	0.535	0.538
c_1 =0.3	0.514	0.526	0.533	0.537	0.540
c_1 =0.5	0.518	0.528	0.534	0.538	0.540
c_1 =0.7	0.520	0.529	0.535	0.539	0.541
c_1 =0.9	0.521	0.529	0.535	0.439	0.541

5.2 Micro calibration of the agent based model

The results of the previous section clearly show that social influence matters. In particular, both the definition and the implementation of social orientation and the effects of coordinated consumption and imitation shape the final market outcomes. We have explored the outcomes of the model for an extensive area of the parameter space both for $\overline{\beta}$ and for the weights of the imitation effect and the coordinated consumption effect, c_1 and c_2. The problem social simulation researchers are faced with is that it is difficult to find the parameters that closely represent reality. In our case, we seek to find the values of the parameters of our formalization closely match the actual decision making of the movie goers. Next, we try to understand how these parameters and variables relate to each other. In social simulation, the operation of setting the parameters of the simulation model to the values that most adhere to reality is defined calibration [44]. We decide to use the empirical results of the survey (section 3) in order to conduct a micro calibration of our agent based model and to study the effects on the market outcomes. However, instead of identifying and implementing the precise values of the variables and the relationships among them, we investigate how different strengths of those relationships affect the market outcomes.

First, we investigate the consequences of the relation between social orientation and attendance. In section 3 we have shown that movie visitors that decide more according to their individual preferences tend to visit cinema theaters more often than customers that are more socially oriented. We formalize this correlation deriving a new probability of attendance r_{it} (11) and we substitute this to the previous probability of attendance p_t. We set the model with the previous parameters' setting (β_i=[0.25, 0.75]; m_j=[0, 1]; ω_1=50,000,000 and δ_1=0.5) and compare the results of these simulation runs (Table 9) with the previous ones (Table 8). It is evident that market inequalities are hampered when the frequency of attendance is negatively correlated with social orientation. This effect reduces the values of g of about 1%

which appears to be relatively small compared to the effects of imitation and coordinated consumption.

$$r_{it} = \begin{cases} \overline{\beta} \geq \beta_i \Rightarrow p_t + \left(\overline{\beta} - \beta_i\right) \cdot \left(1 - p_t\right) \\ otherwise \Rightarrow p_t + \left(\overline{\beta} - \beta_i\right) \cdot p_t \end{cases} \tag{11}$$

Table 9. GINI coefficient values when r_{it} is proportional to β_i

	c_2 =0.1	c_2 =0.3	c_2 =0.5	c_2 =0.7	c_2 =0.9
c_1 =0.1	0.502	0.518	0.526	0.531	0.534
c_1 =0.3	0.510	0.522	0.528	0.532	0.535
c_1 =0.5	0.514	0.524	0.530	0.533	0.536
c_1 =0.7	0.516	0.524	0.530	0.534	0.536
c_1 =0.9	0.517	0.525	0.530	0.534	0.536

Second, we investigate the consequences of the relation between social orientation and the movie genre. In section 3 we have shown that movie goers that visit popular movie types like animation/family, comedy, thriller/adventure/action/fantasy tend to report a stronger social orientation than those visiting more niche movie types like drama and biography/history. A simple way to include this correlation in our simulation model is to relate the value of movie characteristics m_j to the social orientation of the agents. In particular, we can assume that niche movie types tend to meet the extremes of the continuum of agents' preferences. This can be formalized in our model moving the value of m_j towards the extremes of the continuum according to the deviation of β_i from $\overline{\beta}$ (12). The parameters' setting for our model is kept the same so that the results of Table 10 are comparable with those presented in Table 8. We notice that the implemented correlation between m_{ij} and β_i slightly decreases market inequalities for low values of c_1 and c_2, and it slightly increases market inequalities for high values of c_1 and c_2. However, also in this case, these effects are minor and very marginal with respect to the effects of imitation and the coordinated consumption.

$$m_{ij} = m_j \pm \left(\beta_{MAX} - \beta_i\right) \cdot m_j \tag{12}$$

Table 10. GINI coefficient values when m_{ij} is proportional to β_i

	$c_2=0.1$	$c_2=0.3$	$c_2=0.5$	$c_2=0.7$	$c_2=0.9$
$c_1=0.1$	0.506	0.523	0.531	0.536	0.539
$c_1=0.3$	0.514	0.527	0.533	0.537	0.540
$c_1=0.5$	0.517	0.528	0.535	0.538	0.539
$c_1=0.7$	0.520	0.530	0.534	0.538	0.541
$c_1=0.9$	0.521	0.530	0.534	0.539	0.542

6 Conclusion

This paper tries to explain the strong market inequalities observed in the motion picture industry by social influence. The methodology used is characterized by a double facet. On the one hand, we develop an empirical study that investigates the social and individual motivations that shape the social influence and hence visitors' behavior. On the other hand, we design a simulation model that makes use of these empirical insights in order to investigate how social influences such as imitation and coordinated consumption determine market inequalities. The empirical survey finds support that the motion picture market is divided into two parts: a segment oriented towards entertainment consumption and a segment oriented towards art consumption. The former segment, in prevalence composed by males, younger and less educated visitors is strongly socially influenced; it mainly visits mainstream movies whose genres usually are comedy, thriller and action; and it does not visit cinema theaters too often. The latter segment, characterized by females, elder and higher educated visitors, is less socially influenced; it visits more art house movies whose genres are more often drama and biography; and it goes more often to the cinema. Obviously caution is needed with interpreting these findings as they represent a strong generalization of the cinema consumption. However, they can easily feed our simulation model furnishing useful insights.

The results of our simulation model show that market inequalities are strongly determined by the segment oriented towards entertainment consumption –which scores high on social orientation– rather than by the segment oriented towards art consumption –which scores low on social orientation. When movie goers perceive cinema as entertainment, their decisions depend more strongly on what other movie goers decide to do. In these cases, the decisions of movie goers converge towards a few movies that obtain an additional advantage due to higher levels of coordinated consumption and imitation. Consequently, these movies more easily become hits and the differences between market shares increases. Further results of our simulation experiments show that the most important positive driver of market inequalities is coordinated consumption. This effect is more prevalent than the positive effect of imitation and it overcomes the negative effect that results from art consumption which has a higher frequency of attendance than entertainment consumption. These results are somewhat contradictory to earlier studies [22][28][33] that find that informational influence (i.e., similar to imitation effect) more strongly influences behaviors than normative influence (i.e., similar to coordinated consumption effect). However, our results can be explained by the strong presence of social influence in the motion picture industry. Visitors frequently visit movies together in order to maintain social relationships and to maintain and improve their self-concept. These normative influences appear to have a strong influence on visitors' behavior. Especially when movies are visiting together, norms become very salient and this stimulates conformity behaviors [25] that lead to convergent behaviors. As a result, the normative influence of coordinated consumption has a stronger influence on market inequalities than the informational influence of imitation. Our results also contribute to the understanding of the peculiar aspects of the motion picture industry. For example, the strong coordinated consumption effect can explain the reasons why big studio producers tend to prefer a platform strategy respect to a sleeper strategy [18]. Big studios prefer to heavily advertise movies before their release in order to convince large groups of movie goers to visit the movie together at the opening weekend. This platform strategy is frequently used, and is likely to be much more effective than introducing the movie with a few cinema screens and then relying on the positive word-of-mouth of the movie that ignites the imitation effect.

References

1. Batra, R., Homer, P.M., Kahle, L.R.: Values, Susceptibility to Normative Influence, and Attribute Importance Weights: A Nomological Analysis. Journal of Consumer Psychology. 11 (2001) 115-128
2. Eliashberg, J., Elberse, A., Leenders, M.A.A.M.: The Motion Picture Industry: Critical Issues in Practice, Current Research and New Research Directions. Marketing Science. 25 (2006) 638-661
3. Elberse, A., Anand, B.: Advertising and Expectations: the Effectiveness of Pre-Release Advertising for Motion Pictures. Harvard Business School Working Paper (2006) Series N. 05-060
4. Zufryden, F.S.: Linking Advertising to Box Office Performance of New Film Releases. Journal of Advertising Research. (1996) July-August, 29-41
5. Prag, J. Casavant, J.: An Empirical Study on the Determinants of Revenues and Marketing Expenditures in the Motion Picture Industry. Journal of Cultural Economics. 18 (1994) 217-235
6. Basuroy, S. Chatterjee, S., Ravid, A.S.: How Critical are Critical Reviews? The Box Office Effects of Film Critics, Star Power, and Budgets. Journal of Marketing. 67 (2003) 103-117
7. Eliashberg, J., Shugan, S.M.: Film Critics: Influencers or Predictors. Journal of Marketing. 61 (1997) 68-78
8. Gemser, G., van Oostrum, M., Leenders, M.A.A.M.: The Impact of Film Reviews on the Box Office Performance of Art House vs Mainstream Motion Pictures. Journal of Cultural Economics. 31 (2007) 343-363
9. Albert, S.: Movie Stars and the Distribution of Financially Successful Films in the Motion Picture Industry. Journal of Cultural Economics. 22 (1998) 249-270

10. De Vany, A.S., Walls, W.D.: Uncertainty in the Movie Industry: does Star Power Reduce the Terror of the Box Office? Journal of Cultural Economics. 23 (1999) 285-318
11. Ravid, S.A.: Information, Blockbusters, and Stars: A Study of the Film Industry. Journal of Business. 72 (1999) 463-492
12. Wallace, W.T., Seigerman, A. Holbrook, M.B.: The Role of Actors and Actresses in the Success of Films: How Much is a Movie Star Worth? Journal of Cultural Economics. 17 (1993) 1-27
13. Wierenga, B.: Motion Picture: Consumers, Channels and Intuition, Marketing Science. 25 (2006) 674-677
14. Liu, Y.: Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenues. Journal of Marketing. 70 (2006) 74-89
15. Delre, S.A. Jager, W., Bijmolt, T.H.A., Janssen, M.A.: Simulating the Motion Picture Market: Why the Hits Take it All? Proceeding of the First World Conference in Social Simulation. 21-25 August. (2006) Kyoto, Japan
16. Gemser, G., De Haas, A.: Determining Value when Nobody Knows Anything: How to Select Winners in the Independent Movie Industry. Working Paper (2006) University of Groningen
17. Möller, K.K.E., Karppinen, P.: Role of Motives and Attributes in Consumer Motion Picture Choice. Journal of Economic Psychology. 4 (1983) 239-262
18. Sawhney, M.S. Eliashberg, J.: A Parsimonious Model for Forecasting Box-Office Revenues of Motion Pictures. Marketing Science. 15 (1996) 113-131
19. Jager, W.: Modelling Consumer Behaviour, PhD thesis. (2000) University of Groningen
20. Bearden, W.O., Netemeyer, R.G., Teel, J.E.: Measurement of Consumer Susceptibility to Interpersonal Influence. Journal of Consumer Research. 15 (1989) 473-481
21. Hoffmann, A.O.I., Jager, W., Von Eije, J.H.: Social Simulation of Stock Markets: Taking it to the Next Level. Journal of Artificial Societies and Social Simulation. 10 (2007) 7
22. Mangleburg, T.F., Doney, P.M., Bristol, T.: Shopping with Friends and Teens' Susceptibility to Peer Influence. Journal of Retailing. 80 (2004) 101-116
23. Terry D.J., Hogg, M.A.: Group Norms and the Attitude-Behavior Relationship: A Role for Group Identification. Personality and Social Psychology Bulletin. 22 (1996) 776-793
24. Sherif, M.: The Psychology of Social Norms. Harper and Row, New York (1936)
25. Cialdini, R.B., Goldstein, N.J.: Social Influence: Compliance and Conformity. Annual Review of Psychology. 55 (2004) 591-621
26. Bearden, W.O., Etzel, M.J.: Reference Group Influence on Product and Brand Purchase Decisions. Journal of Consumer Research. 15(1982) 183-194
27. Deutsch, M., Gerard, H.B.: A Study of Normative and Informational Influence upon Individual Judgment. Journal of Abnormal and Social Psychology. 51 (1955) 629-636
28. Burnkrant, R.E., Cousineau, A.: Informational and Normative Social Influence in Buyer Behavior. Journal of Consumer Research. 2 (1975) 206-215
29. Childers, T.L., Rao, A.R.: The Influence of Familial and Peer-Based Reference Groups on Consumer Decisions. Journal of Consumer Research. 19 (1992) 198-211
30. Hidalgo, C.A.R., Castro, A., Rodriguez-Sickert, C.: The Effect of Social Interactions in the Primary Consumption Life Cycle of Motion Pictures. New Journal of Physics. 8 (2006) 52

31. Tootelian, D.H., Gaedeke, R.M.: The Teen Market: An Exploratory Analysis of Income, Spending and Shopping Patterns. Journal of Consumer Marketing. 9 (1992) 35-44
32. Chao, A., Schor, J.B.: Empirical Tests of Status Consumption: Evidence from Women's Cosmetics. Journal of Economic Psychology. 19 (1998) 107-131
33. Park, C.W., Lessig, P.V.: Students and Housewives: Differences in Susceptibility to Reference Group Influence. Journal of Consumer Research. 4 (1977) 102-110
34. Deci, E.L., Ryan, R.M.: Intrinsic Motivation and Self-Determination in Human Behavior. Plenum New York (1985)
35. Zuckerman, E.W., Kim, T.Y.: The Critical Trade-Off: Identity Assignment and Box Office Success in the Feature Film Industry. Industrial and Corporate Change. 12 (2003) 27–67
36. Bordwell, D., Thompson, K.: Film Art: An Introduction. McGraw-Hill New York (2001)
37. Bagella, M., Becchetti, L.: The Determinants of Motion Pictures Box Office Performance: Evidence from Movies Produced in Italy. Journal of Cultural Economics. 23 (1999) 237-256
38. Baumann, S.: Marketing, Cultural Hierarchy, and the Relevance of Critics: Film in the United States, 1935–1980. Poetics. 30 (2002) 243–262
39. Reinstein, D.A., Snyder, C.M.: The Influence of Expert Reviews on Consumer Demand for Experience Goods: A Case Study of Movie Critics. Journal of Industrial Economics. 53 (2005) 27-51
40. Gilbert, N., Troitzsch, K.G.: Simulation for the Social Scientist. Open University Press Philadelphia (1999)
41. Elberse, A., Eliashberg, J.: Demand and Supply Dynamics for Sequentially Released Products in International Markets: The Case of Motion Pictures. Marketing Science. 22 (2003) 329-354
42. Ainslie, A., Dréze, X., Zufryden, F.: Modeling Movie Life Cycles and Market Share, Marketing Science. 24 (2005) 508-517
43. Vogel, H.L.: Entertainment Industry Economics: A Guide for Financial Analysis. 4th edn. Cambridge University Press Cambridge UK (1998)
44. Janssen, M.A., Ostrom, E.: Empirically based, agent-based models. Ecology and Society. 11 (2006) 37

Primacy Effect with Symmetric Features Propagating in a Population

Emmanuel Dubois[1], Sylvie Huet[1] and Guillaume Deffuant[1]

[1] Cemagref, Laboratoire d'Ingénierie des Systèmes Complexes, 24, rue des Landais, 63 672 Aubière France
{emmanuel.dubois, sylvie.huet, guillaume.deffuant}@cemagref.fr

Abstract. We propose a modification of existing primacy effect models in order to fit better social psychology observations. We keep the main principle, inspired by Festinger theory of cognitive dissonance: individuals tend to ignore features which are slightly incongruent with their global attitude toward an object. Only congruent and strongly incongruent features are thus considered. In the new model, we suppose that the threshold for ignoring incongruent features is linear with the global attitude. This hypothesis is inspired by recent social psychology studies about "attitude strength" being able to explain primacy effect. Moreover, and also in agreement with the social psychology literature, this new model exhibits a primacy effect when congruent and incongruent features are of equal importance (the previous model does not). We also study the conditions in which, interactions among individuals enhance the propagation of primacy effect.

Keywords: Primacy effect, attitude model, individual based model.

1 Introduction

It is generally agreed that attitudes exert selective effects at various stages of information processing [8]: information may be filtered by the individuals, i.e. they ignore it. Here, "attitude is a psychological tendency that is expressed by evaluating a particular entity with some degree of favour or disfavour" [8]. Festinger, in his theory of cognitive dissonance [9], proposes some mechanisms for this selection: people seek out information that supports their attitudes and avoid information that challenges them, in order to minimize their cognitive dissonance. Such selection mechanisms can lead to an attitude formation which depends on the order of information delivery. This is typically the case of the primacy effect: individuals define their attitude with the first few bits of information they catch, and then only consider information which reinforces this attitude.

More recent researches [20, 21] show that when individuals have a low motivation to process information, a recency effect or an exhaustive processing of information occur, when individuals have a high motivation to process, a primacy effect is observed.

Several researches in social modelling include links between attitude and information transmission [1][10][12][14][16]. The belief-adjustment model [22] focuses on the impact of information presentation on attitude. It is particularly devoted to the recency effect and doesn't take into account the motivation. The bounded confidence models [23][24][17][18] implement reception and emission filters on attitudes or opinions. In the model of innovation diffusion of [4], bounded confidence dynamics are coupled with information propagation. However, none of these models focuses specifically on the primacy effect.

Recently, Huet and Deffuant proposed a simple agent based model of individual primacy effect [5][6][11][12], which abstracts from the researches in social psychology. The model is very simple: it supposes that individuals filter incongruent features if their importance is below a given fixed threshold. They show that, for some choice of positive and negative features, the model exhibits primacy effects: depending on the order of feature reception, the agents have different final attitudes. Moreover, in some cases, the number of agents which show primacy effect is significantly increased and in other cases significantly decreased when agents interact with each other.

Discussing this model with social psychologists, we realised that, despite its interesting first results, it does not fit experiments on the primacy effect. Indeed, these experiments show that the primacy effect can take place even when congruent and incongruent features are of equal importance. Unfortunately, the model shows a primacy effect only if the features of one sign are of higher importance, and their value is well chosen relatively to the threshold. Clearly these constraints are too high. It seems that the primacy effect takes place much more easily in real subjects.

In this paper, we suppose that the threshold for filtering incongruent features is not fixed, but evolves linearly with the absolute value of the global attitude. We study the conditions in which this model exhibits a primacy effect, and show that it can take place even when positive and negative features are of equal importance. Therefore this new model fits better the social psychology literature.

Moreover, we study how the primacy effect propagates when individuals interact, and receive randomly the features from a media. We observe that, when the features are asymmetric, interactions increase the number of individuals showing a primacy effect.

First, we describe the individual dynamics in details, and show why it shows primacy effect even with symmetric features. Second, we study the effect of interactions, and analyse when there is a propagation of the primacy effect. Finally, we conclude and point out some next steps of this research.

2 Individual dynamics: a filtering threshold linear with the global attitude

In this section, we consider an isolated individual: he receives an object's description from a media and doesn't exchange with others. Firstly, we detail the individual dynamic. Then we analyze the filtering conditions of information. Finally, we determine proportion of individuals subjected to primacy effect in a population of isolated, or non-interacting, individuals

2.1 The model

The model is very similar to the one studied in [6] and [11]. The basis of our modelling is dissonance theory [9] on the one hand, and Allport's work on rumour diffusion [1] on the other hand. We notice that individuals avoid incongruent information, and, keep only important information.

We consider an individual which is told by a media about the features of an object. We define this object by a set of features $= (1,2..., \ )$, which are associated with positive or negative real values $\left({}_{1},..., \ ,..., \ \right)$ with $\in \Re$. An individual can have a partial view of the object, in which case it has a real value for some features and nil for others. To simplify we use feature instead of feature value.

An individual is characterised by:

- : an initial attitude (suppose the same for all individuals in the following).
- : a subset of containing the features currently retained by the individual (empty at the beginning).
- $= \ + \sum_{\in}$: global attitude about the object (related to information integration theory of Anderson [3]).

The model is based on the congruency principle. A feature is congruent when it has the same sign as the individual's global attitude to the object, incongruent otherwise. This means, when ≥ 0, feature is said congruent for individual , and incongruent otherwise. The dynamics of the model have three main aspects:

1. **Exposure to feature values:** we suppose that, at each time step, a media sends a randomly chosen feature to the individual.
2. **Selective retaining**: The dynamics of filtering are determined by an incongruence threshold Θ. Being told about feature , the individual will react as follows:
   - If is congruent → "retain the feature".
   - If is incongruent:
     - if $\left| \ \right| > \Theta$ → "retain the feature";
     - otherwise "ignore the feature".

   Here, "retain the feature" means that is added to (if does not include yet), "ignore the feature" means that the feature is filtered (not added to).
3. **Computation of attitude**: an individual computes its global attitude each time it retains a new feature.

In [6] [11], Θ is a constant positive number. In [5], Θ is dynamic depending on the current attitude value plus a constant. Both cases imply that the primacy effect appears only when the features are asymmetric. Indeed, we need for instance that positive features are filtered when incongruent, whereas negative features are not. To achieve this, positive features should be lower than Θ, and negative features higher (in absolute value). See [5],[6], [11] for details.

These constraints seem much too strong, because psychological studies (see [2][15] classics) show that the primacy effect can also occurs when the opposite features have the same absolute value.

In order to weaken the constraints on the primacy effect, we suppose now that threshold Θ is linear with the current value of the global attitude with factor . Therefore, we have:

$$= \alpha \left| \quad \right| \tag{1}$$

One can interpret parameter as the tendency of individuals to privilege their own attitude. When is very small, individuals consider all features, because the threshold is very small. On the contrary, when is high, individuals rapidly only consider features that fit their own view. α varies from 0 to 1. Moreover, we argue in the discussion that may be related to recent psychological studies. In the following study, we consider a very simple case where α has the same value for all individuals.

This new model significantly weakens the conditions on the primacy effect. The global attitude can grow in absolute value as long as congruent features are retained, which increases linearly the threshold, and reinforces the tendency to ignore incongruent features (and consequently to show primacy effect). Therefore, as we shall verify analytically, the primacy effect can take place, even with symmetric features.

2.2 Analysis of the conditions for filtering positive or negative features

To simplify the analysis, we suppose that all positive features have the same value $_{+} = 1$, and all negative features the same value $_{-} = -\beta$ ($\beta \geq 0$). We get symmetric features when $\beta = 1$. We are now studying the conditions which lead an individual to filter the features of one sign or the other.

Suppose that an individual retained positive features and $_{-}$ negative features. In these conditions, we have:

$$= \; _{+} - \beta \; _{-} + \tag{2}$$

The conditions in which negative features will be filtered are:

- the global attitude is positive, which implies that negative features are incongruent;
- the threshold is higher than β.

These conditions can be written as follows:

$$_{+} - \beta \; _{-} + \quad \geq \frac{\beta}{\alpha} \tag{3}$$

The same derivation can be made about the conditions for filtering positive features. It yields:

$$_{+} - \beta \, _{-} + \quad \leq -\frac{1}{\alpha} \qquad (4)$$

When = 0 and β 1 In this case formulae (3) and (4) can be simplified. Both become equal to $1/\alpha$ if respectively consider $_{-}$ = 0 and = 0. Fig. 1 shows the possible states for = 0 and β 1. We note that in this case, the figure is totally symmetric.

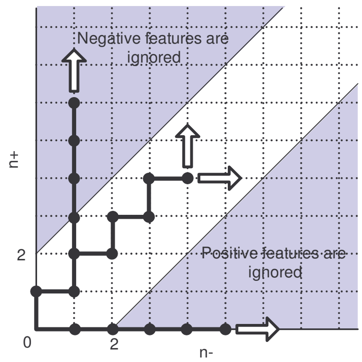


Fig. 1. States of an individual regarding the number of negative (x-axis) or positive features (on y-axis) for $\beta = 1$, $\alpha = 0.5$ and = 0. In the white area, there is no primacy effect. In the grey areas, either positive or negative features are ignored, which implies a primacy effect. Black dots and lines represent three different examples of individual trajectories.

When an individual goes from the non-filtering zone to a filtering zone, it cannot come back into the non-filtering zone. This means that once individuals are in a filtering zone, they retain only congruent features, and filter all incongruent ones. In other words, they exhibit the primacy effect. Their final attitude is designed by the first features they receive before they begin to filter a type of feature (positive or negative).

On the contrary, individuals staying in the non-filtering zone retain any feature, negative or positive, they receive. They do not show the primacy effect. The width of this zone depends on α. However it is easy, due to consecutive reception of several same sign features for example, to go out of this zone and fall into a filtering zone. Thus, when the number of positive and negative features, randomly diffused by the media, grows infinitely, the probability that an individual remains in the non-filtering zone, i.e. to retain all the features, tends to zero.

Let's consider now a case of asymmetric features. Fig. 2 shows the filtering and non-filtering zones when negative features are 4 times more important than positive features ($\beta = 4$). We notice on figure 2 that the negative filtering zone is smaller than the positive filtering zone. The fact the negative features have a higher value creates an asymmetry in their favour.

Let's consider now the effect of a strictly positive value of the initial attitude . Fig. 3 shows the individuals states for = 4. As on figure 2, we notice on figure 3 an asymmetric behaviour of the individual model. The positive filtering zone is smaller than the negative one. The initial positive attitude favours the positive feature.

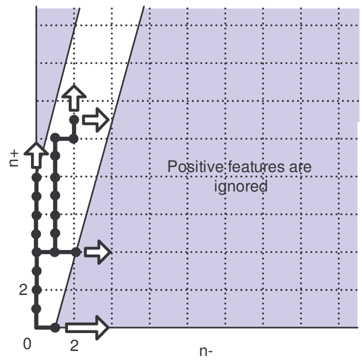


Fig. 2. States of an individual regarding the number of negative (x-axis) or positive features (on y-axis) for $\beta = 4$, $\alpha = 0.5$ and = 0. We notice the negative filtering zone is smaller than the positive filtering zone. The fact the negative features are of higher importance than the positive features creates an asymmetry favouring negative features.

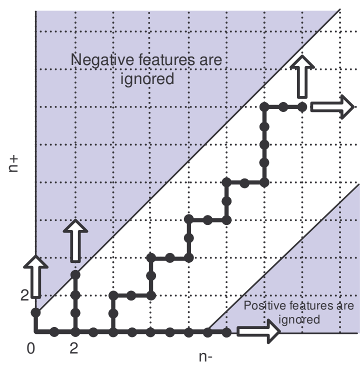


Fig. 3. States of an individual regarding the number of negative (x-axis) or positive features (on y-axis) for $\beta = 1$, $\alpha = 0.2$ and = 4. We notice here a strong asymmetry favouring positive features.

2.3 Probability to show primacy effect in simple examples

We consider simple examples in which we suppose that the object includes 4 features. As we have 4 features with two possible values (two positive values and two negative ones), 6 orders of feature presentation by the media are possible.

We first consider the case of β 1, = 0. Table 1 shows the final list of retained features for the different orders of presentation, and different values of . The last line is the probability that an individual shows the primacy effect, when all orders of presentation have the same probability (which is our hypothesis when features are only diffused by the media). This probability is the same for a unique individual or a population of such individuals when they do not interact each other, i.e. when they are isolated.

Consider the presentation order ++-- and = 0.5, as an example (first line third column table 1). We are following our individual state in time. Firstly, it receives its attitude goes from 0 to 1 and its list of retained features becomes ={+}. Secondly it also receives which is also congruent since its current attitude is positive, so its attitude becomes 2, and L={++}. Then it is exposed to which is incongruent. Then, Θ 0.5 × 2 1 thus the feature -1 is ignored since it is not higher than Θ The attitude of the agent is 2, with L={++}. The last feature is also filtered since it is incongruent; the individual shows the primacy effect.

Table 1. Final list of retained features for an isolated individual with = 0 and β = 1. Grey cells indicate a primacy effect.

	α		
Order of presentation	[0;0,5[	[0,5;1[	1
++--	++--	++	++
--++	--++	--	--
+--+	+--+	+--+	++
-++-	-++-	-++-	--
-+-+	-+-+	-+-+	--
+-+-	+-+-	+-+-	++
Primacy effect proba.	0	1/3	1

Following the same method, we can determine the final state of a population composed of isolated individuals for β = 4. Results are presented in table 2.

Table 2. Final state of an isolated individual = 0 and β = 4

	α				
	[0;1/8[	[1/8;1/7[	[1/7;1/4[	[1/4;1/3[	[1/3;1]
Primacy effect proba.	0	1/6	1/2	2/3	5/6

We can also determine the population final state for = 4 and β = 1. The results are presented in table 3.

Table 3. Final state of a non-interacting population of individuals for = 4 and β = 1

	α			
	[0;1/6[	[1/6;1/5[	[1/5;1/4[	[1/4;1]
Primacy effect proba.	0	1/6	2/3	1

After having studied the individual case and the case of a non-interacting population, we are now comparing the latter to the case of a population of interacting individuals.

3 Interacting individuals

We suppose now that individuals communicate with each other. Thus, they can receive a given feature from a peer or from the media. We firstly define the interaction model. Then we simulate simple examples in order to compare the interacting population dynamic with isolated individuals. To simulate the interactions, the model is implemented in the Java language.

3.1 The model of interaction

For sake of simplicity, we suppose that each individual can communicate with any other. The principle is that, at each time step and for times (considering a population of individuals), we choose a pair of individuals at random, and the first communicates to the second one of its retained features at random. More precisely, at each time step, the algorithm is as follows:

N times repeat:

- *Media diffusion:* choose individual k at random with probability f, choose feature m at random in the object, send feature m to individual k.
- *Interactions:* choose couple of individuals (i,j) at random; i tells j about one of its retained features, chosen at random.

3.2 Comparing interacting population with isolated individuals

We consider the parameter values of section 2.3, with an object of 4 features (2 positive and 2 negative) and the different values of and β.

We performed simulations with an interacting population of 1000 agents. Each result is obtained with 100 replicas. For a replica, we perform 200 iterations. This duration is largely sufficient to insure that all individuals have been exposed to all features. The delivery frequency by the media is 0.1.

We then compare the probability of isolated individual to show primacy effect with the proportion of individuals showing the primacy effect in the interacting population.

Symmetric case (g = 0, β = 1)

We performed simulations with the relevant values of = {0.1; 0.5; 1} chosen from table 1. Indeed with these values, we get all possible situations. Fig. 4 shows the results.

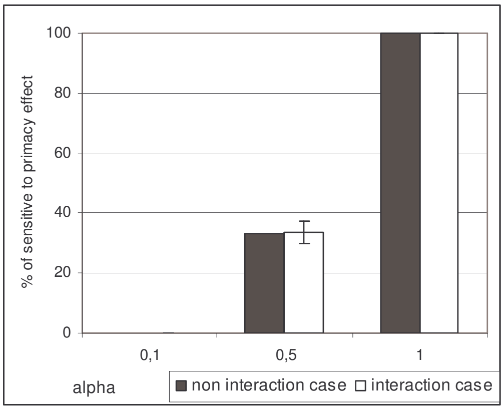


Fig. 4. Comparison between the interacting and the isolated individuals for = 0, β = 1. The error bar takes the min and max.

We observe on Fig. 4 that the proportion of primacy effect in an interacting population is the same as for isolated individuals. This result can easily be understood. Indeed, in the model, the situation is totally symmetric between positive and negative features. Therefore, the number of negative filtering is the same as the number of positive filtering, which implies that in the interactions within the population, the probability to communicate a negative feature is the same as the probability to transmit a positive feature. Therefore, the probability of receiving the different orders of features is the same with or without discussions.

We can make the assumption that when introducing asymmetry in this model, the interactions in the population will modify the probability of primacy effect.

Asymmetric case due to a higher value of negative features (g = 0, β = 4)

We first investigate an asymmetry due to β 4. We chose relevant values of = {0.1; 0.13; 0.2; 0.3; 0.4} (see table 2). The results are shown on Fig. 5. As expected, it shows a clear increase of the primacy effect by the interactions (more particularly for α = 0.3). The reason is that the first individuals which receive negative features filter positive ones afterwards, and then communicate only negative features while the individuals receiving firstly positive features do not filter negative ones and communicate about both features. Therefore, among the individuals, the discussions are more likely to be about negative features, which reinforces the primacy effect. For some values of α, the asymmetry is enhanced, in which case the effect of the discussions is more significant.

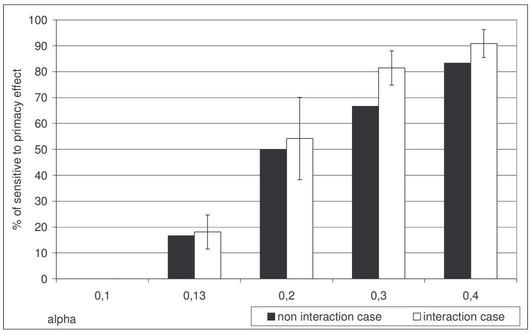


Fig. 5. Comparison between the interaction and the non interaction case for = 0 and an object described by 4 features $\beta = 4$. Error bars take the min and max.

Asymmetric case due to the initial value of attitude ($g = 4$, $\beta = 1$)

For this final case, we considered the relevant values of = {0.1; 0.18; 0.2; 0.3} (see table 3). The results are shown on Fig. 6.

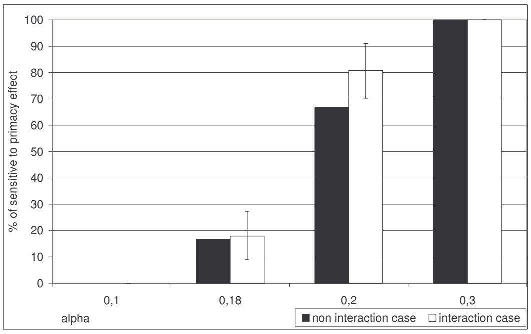


Fig. 6. Comparison between the interaction and the isolated individual primacy effect for = 4, $\beta = 1$. The error bars take the min and max.

Again, for some values of α, we observe a reinforcement of the primacy effect by the interactions. Indeed, in this case, the positive value of favours an initial filtering of negative features. Therefore, in the discussions, the negative features are less likely than the positive ones. This increases the tendency we observed for isolated individuals to filter negative features.

4 Discussion, conclusion

The first contribution of this paper is a new model of primacy effect. In this new model, primacy effect takes place also when positive and negative features are of the same importance. To do so, the individual incongruence filtering threshold evolves linearly with the global current attitude weighted by α. Thus, the new model fits better recent social psychology experiments.

Indeed is a coefficient of taking into account the current attitude in filtering threshold determination. It can be related to a motivation degree to process the information [19] [21] since it modulates the threshold between 0 and the current attitude of the individual. When is low, individual have a low resistance to incongruent features and retain almost all transmitted features. On the contrary, when is very high, resistance and filtering become systematic and often generate the primacy effect.

Nevertheless this interpretation of as a motivation degree is only a hypothesis at this stage, because interpreting experiments of [21] in the framework of our model is not direct. Indeed, the experiments use two groups of same sign features. For example, they present to subjects a positive feature group and next a negative feature group. In our model, we suppose that all features are separated and we consider all the possible presentation orders. Therefore, other experiments would be necessary to confirm or falsify our hypothesis about .

The second contribution of this paper is the comparison between isolated and interacting individuals. In the line of the results we obtained with previous models, we found that interactions can increase significantly the primacy effect. More particularly, in the new model, we underlined the necessity of an asymmetry between positive and negative features to obtain such an impact of interactions on the primacy effect. We observed that the asymmetry can also be due to an a priori bias toward one or the other type of features (the value of parameter), even though the object's features have equal importance. We intend to provide more complete results in the future, probably through the design of an aggregated model.

To conclude, we would like to stress the experiments in social psychology that this work suggests. Besides already mentioned experiments related to the interpretation of , we would need to measure the attitude after each feature presentation, to further evaluate the hypotheses of the model. Finally, it would be interesting to investigate the impact of interactions on the primacy effect.

Acknowledgments. We warmly thank Diemo Urbig for his crucial suggestions about the model.

5 References

1. Allport, G.W., Postman, L.: The psychology of rumor. Russel & Russell, Inc., (1947), 1965.
2. Asch, E.S.: Forming Impressions of Personality. Journal of Abnormal and Social Psychology, Quarterly (1946), 41:258-290.
3. Anderson, N. H.: Integration theory and attitude change. Psychological Review, (1971) 171–206.
4. Deffuant, G., Huet, S. and Amblard, F.: An individual-based model of innovation diffusion mixing social value and individual payoff dynamics. American Journal of Sociology 110-4, The university of Chicago press January (2005), 1041-1069.

5. Deffuant, G. and Huet, S.: Propagation effect of filtering incongruent information. Journal of Business Research Elsevier (2007), 60, 8, 816-825.
6. Deffuant, G. and Huet, S.: Collective Reinforcement of First Impression Bias. First World Congress on Social Simulation, Kyoto (Japan), August (2006).
7. Deffuant, G., Weisbuch, G., Amblard, F., Faure, T.: Simple is beautiful... and necessary. Journal of Artificial Societies and Social Simulation, (2003) 6(1).
8. Eagly, A.H.; Chaiken, S.: The psychology of attitudes. Thomson/Wadsworth; 1993, (1998), 800 pages
9. Festinger, L.: A Theory of Cognitive Dissonance. Stanford Univ. Press (1957).
10. Galam, S.: Modelling rumors: the no plane Pentagon French hoax case. Physica A., Elsevier (2003), vol. 320., pp 571-580.
11. Huet, S. and Deffuant, G.: When do interactions increase or decrease primacy effects? Proceedings of Model-2-Model, Marseille (2007), 19 pages.
12. Huet, S. and Deffuant, G.: Effets d'un filtre cognitif sur la diffusion d'information. (2006), Rabat.
13. Jacoby J., Morrin M., Jaccard J., Gurhan Z., Kuss A., Maheswaran D.: Mapping Attitude Formation as a Function of Information Input: Online Processing Models of Attitude Formation. Journal of Consumer Psychology, Lawrence Erlbaum Associates (2002) 12(1), 21-34.
14. Lawson, G., and Butts, C.T.: Information Transmission Through Human Informants: Simulation, CASOS'04 (2004).
15. Miller, N. and Campbell, D.T.: Recency and primacy in persuasion as a function of the timing of speeches and measurements. Jour. Abnormal Social Psychology, APA (1959) 59, 1-9.
16. Tsimring, L.S., and Huerta, R.: Modeling of Contact Tracing in Social Networks. Physica A, Elsevier (2003) 325, pp. 33-39.
17. Urbig, D.: Attitude Dynamics With Limited Verbalisation Capabilities. Journal of Artificial Societies and Social Simulation (JASSS). (2003) vol. 6 no. 1.
18. Urbig, D. and Malitz R.: Dynamics of Structured Attitude and Opinions Third Conference of the European Social Simulation Association Proceedings, September 5-8, Koblenz, Germany, (2005) 206-212.
19. Haugtvedt, C. P., Wegener, D. T.: Message order effects in persuasion: An attitude strength perspective. Journal of Consumer Research 21, The University Of Chicago Press (1994) 205-218.
20. Petty, R. E., Cacioppo, J.T.: Communication and persuasion: Central and peripheral routes to attitude change. NewYork: Springer-Verlag. (1986).
21. Petty R. E., Tormala, Z. L., Hawkins, C., Wegener, D. T.: Motivation to Think and Order Effects in Persuasion: The Moderating Role of Chunking. Pers Soc Psychol Bull 27, Society for Pers. and Soc. Psy. (2001), 332.
22. Hogarth R. M., Einhorn H. J.: Order effects in belief updating: The belief-adjustment model. Cognitive Psychology 24, Elsevier (1992) 1-55.
23. Deffuant, G., Neau, D., Amblard, F., Weisbuch, G.: Mixing beliefs among interacting agents", Advances in Complex Systems, 3, 87-98, (2001).
24. Hegselmann, R. and Krause, U.: Opinion Dynamics and Bounded Confidence Models, Analysis and Simulation. Journal of Artificial Societies and Social Simulation, (2002), vol.5, 3. http://jasss.soc.surrey.ac.uk/5/3/2.html.

Market Dimensionality and the Proliferation of Small-scale Firms

César García-Díaz

Faculty of Economics and Business, University of Groningen (The Netherlands)
c.e.garcia.diaz@rug.nl

Arjen van Witteloostuijn
Faculty of Applied Economics, University of Antwerp (Belgium)
Arjen.vanWitteloostuijn@ua.ac.be

Gábor Péli
Faculty of Economics and Business, University of Groningen (The Netherlands)
g.peli@rug.nl

Abstract. We build an agent-based computational model to study how market structure evolution, in terms of the increasing diversity of available product options in an *n*-dimensional space of product attributes, affects the performance of firm strategies (i.e. being a large-scale or a small-scale firm). Our aim is to study how the degree of space heterogeneity affects firm viability when different levels of scale advantage characterize firm structure. Such an increasing heterogeneity is associated with the increasing number of dimensions in the so-called resource space. A resource space with higher number of dimensions implies a market where there is higher product diversity. We study the effect on firm performance through the realized profit/cost ratio of every time period. We build a novel approach to measure the evolution of the number of dimensions in the product space, based on the assumptions that i) not all the possibilities of a single dimension (attribute) are active at a given point of time, and ii) product attributes or characteristics in a new dimension might start to emerge when the previously established dimensions have not been fully developed yet. This motivates us to think that the number of dimensions may be a "fraction" of the Euclidean dimensions. We use the concept of "similarity dimension" in order to measure those fractional dimensions in the resource space. Market starts with only one product characteristic option (dimension = 0) and evolves until a maximum of dimension = 2. We confirm that increasing dimensionality gives a differential advantage to small-scale firms. However, this advantage comes from the "opening up" of new product characteristics mostly generated by large-scale firms. We also find that large-scale firms may also benefit from increasing dimensionality, whenever it allows only a small degree of differentiation without weakening the power of scale advantages.

1. Introduction

The study of spatial features of consumer distributions (i.e., resource spaces) and its effects on the shaping of social structures has seen a development of interesting ideas in the field of organizational sociology in the recent years. Specifically, the subfield of Organizational Ecology (Hannan and Freeman 1989; Carroll and Hannan 2000) has generated a set of interesting studies and fruitful theories about the effects of resource heterogeneity in market structures and organizational form viability (Caroll 1985; Carroll and Hannan 2000; Boone et al. 2002), and the effect of resource spatial distributions on markets (Boone and Witteloostuijn 2004; Witteloostuijn and Boone 2006). Moreover, the fact that the evolution of the number of dimensions of such spaces might generate market structural changes have also drawn special attention in the literature. For instance, Péli and Nooteboom (1999) use a geometrical (sphere-packing) model to demonstrate that the increase in the number of dimensions opens up new spots for specialist organizations. However, Péli and Bruggeman (2006) demonstrate that the benefits a specific organization gets might be non-monotonic respect to dimensionality change. That is, it might be the case that a decrease in dimensionality influences positively a specialist, while affecting negatively a generalist organization.

Although the study of the effects of spatial features reveals an interesting direction for the study of social structures (Freeman 1983, McPherson 1983, 2004; Carroll et al. 2002), the study of increasing dimensionality remains still very abstract and disconnected from empirical attempts. With such aim in mind, we build an agent-based model to study how market structure evolution, in term of the increasing diversity of product options in an *n*-dimensional space of product characteristics, affects the performance of firm strategies (i.e. being a large-scale or a small-scale firm). Such an increasing diversity is associated with the increasing number of dimensions in the resource space.

We study effects on firm strategy performance through both profits and market "density" (the number of firms in the market, as defined by organizational ecologists). Our computational model offers three contributions i) To offer an economic ground, based on profit-seeking behavior, of the effects of organizational form and location in the resource space, ii) To understand how dimensionality of the consumer distribution in the product space affects firm performance; and iii) To offer a framework where dimensionality changes endogenously.

We build a novel approach to measure the evolution of the number of dimensions in the product space, based on the assumptions that i) not all the possibilities of a dimension are active at a given point of time, ii) product attributes in a new dimension might start to emerge when the previously established dimensions have not been fully developed yet. This motivates us to think that the number of dimensions may be a "fraction" of the Euclidean dimensions in a space. We use concepts derived from fractal geometry in order to measure those fractional dimensions in the product space (Mandelbrot 1983). In our model we set a market that starts with only one product characteristic combination (dimension = 0) and evolves up to maximum of m^2 (dimension = 2).

In order to analyze our results, we mainly consider two separate treatments to simulated data. We aim i) To measure the number of firms per firm type in the market

(market density) over time; and ii) To measure the relationship between profits and dimensionality. We find that increasing dimensionality has a *positive relationship* with large-scale firm performance *at a decreasing rate*; and a slightly positive relationship with that of small-scale firms.

This paper is divided as follows: we present a brief theoretical background that link to our objectives in this paper; then we present the computational model built for such purpose, analysis of results and concluding remarks.

2. Theoretical background

Organizational ecologists have posited that the scope of environmental resources where firm operate, along with its change dynamics, is a determinant of organizational form performance (Hannan and Freeman 1983, 1989). Such set of environmental resources might be the distribution of consumer purchasing power over a set of preferences, usually mapped onto a space of socio-economic characteristics (the so-called Blau spaces; see McPherson 1983, 2004). Such distribution is usually named "resource space" by organizational ecologists. Organizational Ecology (OE) theories have developed fragments where the fate of typical organizational forms is explored. OE usually explores effects on founding and mortality rates of generalists (this is, firms that target a wide set of consumer preferences) and specialists (firms that "specialize" their offering by targeting a narrow set of consumer preferences). For instance, niche-width theory (Hannan et al. 2003) states that specialist organizations benefit from a population of a narrow set of consumers preferences, but are very weak if the environmental pattern changes due to specific market conditions. On the other hand, generalist spread their risk over a wider set of consumer preferences, but at the expense of a decrease in average "fitness", which is by-product of handling a larger environmental complexity (Hannan and Freeman 1977; Hannan et al. 2003).

On the other hand, resource-partitioning theory (Carroll 1985; Carroll and Hannan 2000; Carroll et al. 2002) deals with a different effect: the strategic location of generalist and specialist in a unimodal resource space characterized by enough taste heterogeneity. Moreover, Péli and Nooteboom (1999) argue that the increase in the number of dimensions of the space (this is, adding new features to the already existing set of consumer preferences) generates difficulties to generalists in dealing with the increased resource heterogeneity, causing the release of space in favor of specialists (Carroll et al. 2002). Péli and Nooteboom (1999) demonstrate through geometrical concepts that the increasing number of dimensions makes it difficult to generalists to cover the whole space. They assume that generalists's catchment area of consumers is a hypersphere (e.g., a hypersphere in a two-dimensional space is a circle). The total space covered by all the hyperspheres declines as the number of dimensions increases, leaving empty spots suitable for specialist entry.

Although empirical studies reveal the effect of resource heterogeneity on market structures (Boone et al. 2002), the change in the number of features of the resource space is less understood and far from theories of economic behavior. Next we present a model that may give light to empirical studies in implementing the idea of dimensionality change and reduce the gap in putting this set of ideas into a more economic framework.

We develop a computational model of the evolution of the number of dimensions in a space and its effect on firm performance. We use the term "resource space" to represent the consumer distribution on a *n*-dimensional product characteristics space (Lancaster 1966). We don't have consumer preferences coinciding exactly with product characteristics. We assume that consumers are re-allocated in the product space depending on firms' influence, which allows for the evolution of space dimensionality. This is grounded in terms of how firms try to understand consumer behavior, help to develop new market segments, deploy advertisement campaigns and ultimately contribute to the shaping of consumer preferences over time (Basmann 1954, Zinam 1974, Lachaab et al. 2006).[1] In a very simplified way, we assume firm target new spots in the product space (differentiated enough with current positions), invest in opening a demand for such position and consequently generate a re-distribution of the whole demand in such a product space.

The OE's terms of generalism and specialism usually refer to the width of firm's niche. In our computational model, niche width changes over time (contracting or expanding) according to current market conditions. In order to give a permanent label to every participating firm in the market, we endow firms with scale economies but classify them as having limited incentives to grow (small-scale) or being benefited from unit cost reductions of large production (large-scale).[2]

With our computational model we demonstrate how, from a set of micro-level rules of profit-seeking behavior, the effect of increasing dimensionality opens up opportunities to small-scale firms as suggested by Péli and Nooteboom (1999) in a geometrical model. We also show that the effect of increasing dimensionality on large-scale firm is *nonmonotonic*, which means that large-scale firms increase their performance up to some (fractional) dimension, before observing a subsequent decrease. This means that large-scale firms are also benefited from dimensionality increase up to some extent (the point where a small product differentiation results beneficial without undermining scale effects). In next section, we present the main features of the agent-based model.

3. The Model

3.1 The resource space

As mentioned in the previous section, we name the "resource space" as the *n*-(Euclidean) dimensional arrangement of product characteristics, along which consumers are distributed. Specifically, we deal with a two-dimensional space of product characteristics, where every combined value of two different attributes distinguishes a

[1] Organizational ecologists have also speculated about the way specialists unlock new tastes at the periphery of the market. See Carroll and Hannan (1995).

[2] Scale economies is an important driver of firm growth in resource-partitioning theory, where the effect of dimensionality has been considered (Carroll et. al 2002).

point of the space. Assuming that each attribute has m possible different values, the two-dimensional space has $m \times m$ possible combinations, each one representing a distinctive product variety (Dawid et al. 2001). As mentioned above, in OE, the term "resource space" usually refers to the n-dimensional arrangement of socio-demographic characteristics. These types of space renderings are referred to as Blau spaces (honoring Peter Blau. See McPherson (2004)), For instance, Boone et al. (2002) established four relevant dimensions in the study of resource space heterogeneity in the Dutch newspaper industry: age, education, political affiliation and geographical location. Organizational ecologists have generally assumed that each point of the space of socio-demographic characteristics defines a consumer taste (Carroll et al. 2002).

The space of product attributes has also been considered in economic theories of consumer behavior (Lancaster 1966) and also has used it as an operationalization of consumer tastes (Péli and Nooteboom 1999). However, here we depart from such one-to-one correspondence and assume that firms exert product differentiation through the opening of new product combinations that ultimately persuades consumers to move across the product space until they stabilize in some specific product combination (when the space has been fully covered and a stable market for each point of the product characteristics space has been activated). We consider that a space with only two dimensions (supported by our fractional measure) provides enough dynamics in order to explore the dimensionality effect and a convenient visualization of results.

We usually assume that markets start with one firm at the beginning of the industry, but use different starting populations when necessary. Since a new product attribute might appear in the market even when the possibilities for the other attribute have not been exhausted, it implies that one new Euclidean dimension might start to emerge when the previous ones are still underdeveloped (this is, when not all their positions have become active). This argument leads us to consider an alternative way to measure dimensionality, which we explain next.

3.2 Dimensionality computation

Summarizing the above, we assume that a) the market possibly starts with one or very few product varieties, b) not all the possible product varieties are active all the time and, c) varieties belonging to one of the dimensions might start to emerge when the values from the other dimensions have not been exhausted yet. This is, the second dimension might start to be developed by firms even when the first dimension has not been fully exploited.

Under such a context, we think that a measure of a "fractional" dimension may be convenient. We adapt concepts from fractal geometry (Mandelbrot 1983) that have already found applications in bio-ecology (Haskell et al. 2002, Olff and Ritchie 2002) to measure resource space heterogeneity. We use the concept of "similarity dimension" (Mandelbrot 1983), which assumes the space divided in *N hypercubes* with identical area r^D, where r is the edge of the hypercube and *DIM* is the corresponding Euclidean dimension.

If the total space is normalized to have an area equal to 1, we have that the total area might me represented by $Nr^{DIM} = 1$. Solving for *DIM* we get $DIM = \ln(1/N)/\ln(r)$. In our model, where we consider a two-dimensional resource space of size *m* x *m*, if we assume that total area is equal to 1, we get $Nr^2 = 1$, where $N = |m \times m| = m^2$. Note that in a two-dimensional rendering, hypercubes are just squares with edge equal to *r*. Solving for *r* we get $r = (1/m^2)^{1/2} = 1/m$. If $N(t)$ corresponds to the number of active cells at time t, the dimension of the space at time t is computed as:

$$DIM(t) = \frac{\ln(1/N(t))}{\ln(1/m)} = \frac{\ln(N(t))}{\ln(m)} \tag{1}$$

One property of the similarity dimension is that its integer values coincide with the Euclidean dimension values. In our model, if only one dimension is fully operating in the market, while the other is absent, there are only m active cells, so that $DIM = \ln(m)/\ln(m) = 1$. If the values of all possible combinations in two dimensions are active, there are *m* x *m* active cells, so that $DIM = \ln(m \times m)/\ln(m) = 2$. We present the following example: let us assume that $m = 25$; so the space has 25 x 25 possible preferences. An example of resource space with dimension equal to 1 appears when a complete row of the space is active (see figure 1). Higher usage of the space leads to a higher dimension. See also figures 2 and 3.

3.3 Demand distribution

We adopt calibrated values from previous related works (García-Díaz and Witteloostuin 2006). For instance, we assume that market remains constant an equal to $Q_T = 5500$ consumers. As the number of active positions increase, the total market size Q_T is simply distributed among the active cells in the market.

Since the dimensionality problem is closely to the resource-partitioning context (Carroll 1985; Carroll and Hannan 2000; Péli and Nooteboom 1999), we explore two cases: and a) when the resource space evolves to a uniform distribution among the active positions (a flat space), and b) when the resource space evolves to a configuration of a mainstream product preference surrounded by a scarcer periphery (a unimodal space). For a justification of the use of uneven distributed spaces, the reader can refer to Carroll et al. (2002). In their work, the authors illustrate through a review of several empirical studies the presence of unimodal spaces in partitioned markets. In either of the two cases, we assume that at $t = 0$ there is only one active position at the center of the resource space, which contains Q_T.

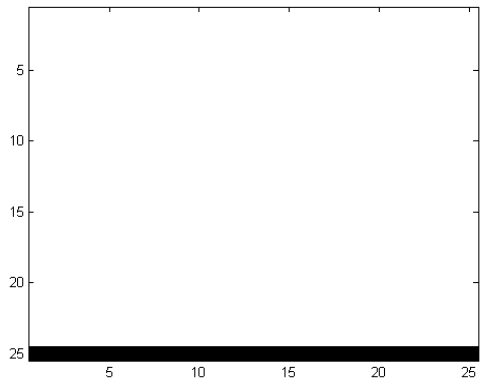


Fig. 1. Resource space with m = 25, dimension = 1 (Active cells in black)

As the number of active cells increase due to the firm actions, Q_T is re-distributed along the set of current active positions. Let us represent the set of active positions at time t by $\Omega(t)$. For the flat space, the re-distribution is trivial and is made just by assigning a demand $D_i = Q_T/\Omega(t)$, $i \in \Omega(t)$ to each active cell. For the unimodal resource space case, the potential market at cell i is computed as $D_i = p_i(t)Q_T$, where

$$p_i(t) = \frac{\frac{1}{d_i(t)+1}}{\sum_{k:k\in\Omega(t)} \frac{1}{d_k(t)+1}}, \tag{2}$$

and

$$d_{i_X,i_Y}(t) = \left\|(i_X, i_Y) - (x_o, y_o)\right\| = \left[(i_X - x_o)^2 + (i_Y - y_o)^2\right]^{1/2} \tag{3}$$

Being (i_X, i_Y) the X-Y coordinates of taste i, and (x_o, y_o) the location of the initial active taste (the market center) at $t = 0$. This implies that opening a position far from the market center (the mainstream position) and finding consumers for it is more difficult that opening near the mainstream taste. In other words, *the marginal cost of convincing one consumer to buy a product with attributes different from the mainstream characteristics increases with "distance" to such mainstream characteristics*. It can be shown that the two dimensional space evolves to a unimodal space with abundant central and scarce peripheral resources.

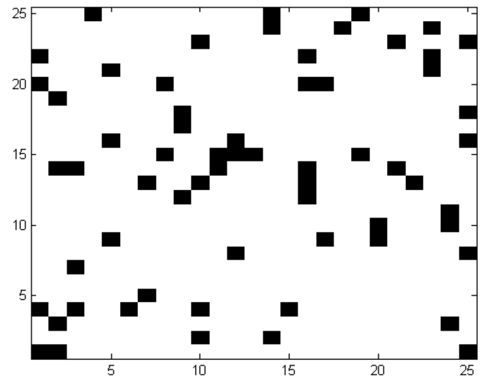


Fig. 2. Resource space with m = 25, dimension = 1.29 (Active cells in black)

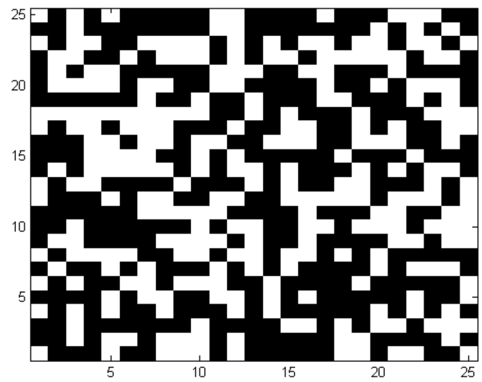


Fig. 3. Resource space with m = 25, dimension = 1.81 (Active cells in black)

3.4 Market entry

In order to model firm entry to the market, we used a mechanism that is line with previous empirical studies regarding organizational founding (Hannan and Carroll 1992; Barron 1999): the density-dependent mechanism Entry is modeled as an stochastic process using a varying entry rate $\lambda(t)$, which is a function of the so-called market "density" (that is, the number of firms in the market, as organizational ecologists define it). Technically, entry is obtained drawing a number from a negative binomial distribution with probability of success equal to $K/(K+\lambda(t))$, where $1/K$ is interpreted as an "overdispersion parameter" and generally set to 0.5 in organizational founding models (see Lee and Harrison 2001, Harrison 2004 for details). The arrival rate is represented by $\lambda(t) = \exp[\delta_0 + \delta_1 N(t) + \delta_2 N(t)^2]$, where $N(t)$ is density at time t, with $t = 1,2,\ldots,T$. Parameter values for δ_0, δ_1 and δ_2 are derived from previous density-dependence models (Cf. Lee and Harrison 2001, Harrison 2004) and calibrated to monthly events. Parameter values and mechanism behavior are presented in table A1 and figure A1 in the Appendix, respectively. If it is not mentioned otherwise, we assume that the model starts with one single firm.

3.5 Firm behavior

The cost function for a firm has two components, one related to production cost, $C^i_{PROD}(Q_i,t)$, and the other related to niche-width expansion costs, $C^i_{NW}(t)$. For firm i at time t, total costs at time t are represented by:

$$C^i(Q_i,t) = C^i_{PROD}(Q_i,t) + C^i_{NW}(t) \tag{4}$$

The production function corresponds to a Cobb-Douglas function.[3] We consider a cost function with two production components, F and V. The cost of each unit F is W_F and the cost of each unit V is W_V. Total production costs are represented as:

$$C^i_{PROD}(Q_i,t) = W_F F_i + W_V V_i(t) \tag{5}$$

Production Q_i is calculated as:

$$Q_i(t) = AF_i^{\alpha} V_i(t)^{\beta} \tag{6}$$

A corresponds to a scale parameter and $\alpha + \beta > 1$, which is needed to produce positive scale economies. Parameters of the LRAC (the long-run average cost curve)[4] are calibrated to produce a minimum average cost for the whole industry equal to 1 when Q is maximum (α, β, W_F, W_V and A). Parameter values are taken from previous simulation work (García-Díaz and Witteloostuijn 2006) and are summarized in the Appendix. We assume the existence of two firm types: large-scale and small-scale firms. A firm has equal probability to select either type. The amount of units V are computed according to the solution of the following optimization problem (for firm i):

$$\begin{aligned} \min \quad & C^i_{PROD}(t) \\ s.a. \quad & Q_i(t) = AF_i^{\alpha} V_i(t)^{\beta} \end{aligned} \tag{7}$$

[3] A Cobb-Douglas function (Mas-Colell et al. 1995) is a typical type of production function used in microeconomic theory to specify production level as a function of two or more production factors.

[4] A LRAC curve results from obtaining the set of points where the average production cost is minimum (i.e., the set of the most efficient points in term of average unit cost), when the two production factors are allowed to vary. In a Cobb-Douglas function, $\alpha + \beta > 1$ implies the existence of scale economies.

Figure 4 shows cost behaviour (i.e. two different short-run average costs, *SRAC*[5]) under two different capacities: $SRAC_2$ reflects a firm that has larger capacity (a large-scale firm), respect to $SRAC_1$ (a small-scale firm), and thus a higher potential scale advantage. It is important to mention that, in order to avoid firms starting abnormally large, we constrain entrants to produce a maximum capacity, which corresponds to the intersection point between the *SRAC* and the *LRAC*. Such constrain also forces firm to explore the value range of their more feasible operation zone: the down sloping part of the average cost curve.

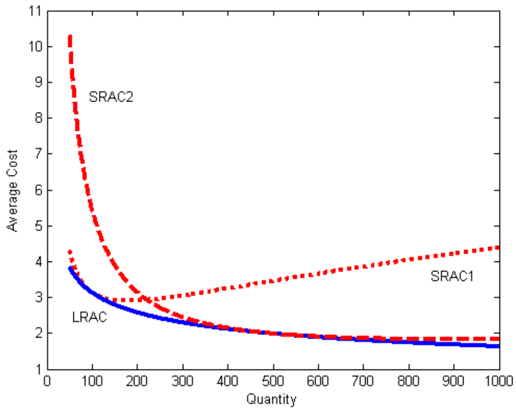


Fig. 4. . Long-run average cost (LRAC) curve and two examples of short-run average (SRAC) cost curves

For convenience, we name the "scale" value as the quantity at which the firm's *SRAC* and the *LRAC* intersect. These values are set so that the cost production efficiency points of the two *SRAC*s are located at the two opposite extremes of the quantity axis. For instance, the small-scale value is set arbitrarily at $Q_{SC} = 10$ and the large-scale at $Q_{LC} = Q_T/2$. Variations of the value Q_{SC} is subject to sensitivity analysis, while the value Q_{LC} is kept fixed. This is done that way because simulation outcomes might change due to the relative scale difference only. Thus, the scale distance of the two firm types is changed when necessary (changing Q_{SC}) in order to explore some properties of the results under the sensitivity analysis.

The second cost component refers to niche width costs. We define the constant *NWCost* as the niche-width cost coefficient, and the niche-width related costs as NWC times the Euclidean distance between the niche limits:

$$C^i_{NW}(t) = NWCost\left\| w_i^{\,u}(t) - w_i^{\,l}(t) \right\|, \tag{8}$$

where $w_i^l(t)$, $w_i^u(t)$ the upper and lower niche limits of firm *i*. The reader can notice that a firm occupying only one position does not incur in niche costs. Since every ex-

[5] A *SRAC* curve reveals the behavior of the average production cost curve when one of the factor is kept fixed, while the other one is allowed to vary according to desired production level. The *LRAC* curve is the envelope function of all the *SRAC* possibilities, as it can be thought of by looking at figure 4.

isting firm in the market has at least one position, each of them would need to incur in a “default operation cost” (regardless of the firm type) so its consideration is not relevant and would just generate a re-calibration of the parameters of interest. Thus, this cost is just excluded from the simulation model for convenience. We rather focus on the relative difference of niche costs among firms. Thus, niche-width costs are those in which a firm incurs as a consequence of operating in more than one market segment.

As argued before, we assumed niches to be rectangular. The fact that firm expansion is either horizontal or vertical corresponds to a random selection. However, a firm does decide to expand if expansion brings about a profit increase. Also, the fact that niches are rectangular implicitly make us assume that a firm has a dominant taste on one dimension, along which it expands and generates varieties in a second dimension (see figure 5).

3.6 Selecting an entry cell

Each cell has a demand function. Every active cell (i,j) has an associated price, $P_{i,j}(t) = P_o - B\Sigma_j O_{i,j}(t)$, where $O_{i,j}(t)$ is the amount produced by firm i in cell j. P_o is set as the highest point of the calibrated *LRAC* curve ($P_o = 11.7435$), while B is set so that a price equal to zero clears the whole market ($B = P_o/Q_T = 0.0021$). Firms stay in the market as they are able to maintain nonnegative profits. Firms also receive an endowment upon entry, which is a multiplier of their fixed costs. This means that large-scale firms receive more endowment than small-scale firms. Since the density-dependent mechanism is calibrated to generate monthly entries, we make sure that each firm receives endowment enough to cover their fixed costs for twelve time periods (one year) with no sales.[6]

Upon entry, a firm decides to either enter an a) occupied cell, or b) to open a new one:

a) *Entering an occupied cell*. Among the active cells, the firm builds a probability function of the unserved demand. If there was no unserved demand at time t-1, the firm treats every active position as an alternative with equal probability. A Random number is drawn from this distribution and a position is selected. Formally, at time t firms build a probability distribution from the sales percentage per cell at time t-1 ($SPC_{j,t-1}$), which is computed as the total sells a position j divided by D_j. Bottom line, firms assemble a (discrete) probability distribution $Pent_{j,t}$, $j \in \Omega(t\text{-}1)$ as follows:

[6] Although not reported here, we also experimented with endowments of 6 months and 24 months. We did not observe any significant qualitative change in the final results.

$$Pent_{j,t} = \begin{cases} 1/\Omega(t-1) & if \sum_{k\in\Omega(t-1)}(1-SPC_{k,t-1}) = 0 \\ \dfrac{(1-SPC_{j,t-1})D_{j,t-1}}{\sum_{k\in\Omega(t-1)}(1-SPC_{k,t-1})D_{k,t-1}} & Otherwise \end{cases} \quad \textbf{(9)}$$

This indicates that, among the active cells, a firm is more likely to enter that one reflecting a larger potential (unserved) amount of consumers. Let us name j^*, $j^* \in \Omega(t)$, as the position randomly selected according to the probability function mentioned above. Then, if a firm decides to enter an active cell, it enters the position j^* with the following quantity offering Q^*:

$$Q^* \equiv \max_{q_1, q_2, \ldots, q_{T_r}} P_{j^*}(Q^{-i}_{j^*,t-1} + q_i)q_i - C(q_i, t) \quad \textbf{(10)}$$

Where Tr is the number of scenarios the firm executes (Tr is set to 10), P is price and $Q^{-i}_{j^*,t-1}$ is the quantity sold by the other firms (all but firm i) in the previous time period, at position j^*.

b) *Opening a new cell*. The second option is to open a new position, which would generate a redistribution of demand if the taste position were effectively opened by the firm. First of all, the firm has to decide among the inactive cells which one to open. We assume that firms are more able to open those taste positions adjacent to the active ones. There is justification for that: the reasoning is that firms try to take advantage of the existent positions by attempting to pull consumers with closely similar tastes and to bring up a new market segment with minimum (but enough) differentiation. That is, a firm takes advantage of the current taste similarity but influences consumers to slightly refine their product preference in order to generate sufficient differentiation from current "market segments".[7] This in fact generates a new market segment. A firm chooses randomly from a uniform distribution an adjacent inactive position k^* from the set of adjacent-to-active cells $\varphi(t)$. In such cases, a firm has to pay a one-time extra cost *NewPos* (the cost of opening a new market segment). If more than one firm happens to enter the same inactive taste position, all the entrants

[7] There is a theoretical framework in Organizational Ecology that uses the concept of "engagement" (Hannan et al. 2007). In a taste space, engagement is the effort of an organization to "materialize" an offering to a taste where such an organization has observed a market opportunity (through, for example, investment in production capacity). Other arguments in line with endogenous sunk costs effects (Sutton 1991) argue that a firm commits itself to advertisement or R&D expenses in order to influence consumer behavior through the increase of his willingness to pay and setting up of brand recognition. Although our model is not directly related to these concepts, we do assume that demand distribution in a product space is affected by firms' offerings.

split the cost $NewPos$ equally. A firm considers to enter an unoccupied cell $k^* \in \varphi(t)$ if:

$$\left\{ \max_{q_1,\ldots,q_{T_r}} P_{k^*}(q_i)q_i - C(q_i,t) \right\} - \left\{ P_{j^*}(Q_{j^*,t-1}^{-i} + Q^*)Q^* - C(Q^*,t) \right\} > NewPos \qquad \textbf{(11)}$$

$$with\ k^* \in \varphi(t),\ j \in \Omega(t)$$

That is, *a firm decides to pay an extra cost to open a new cell if such alternative results in higher expected profits than those of entering an already occupied cell*.[8] We name "innovators" to those firms that open new cells. Note that firms also assume that the (expected) aggregated new production level of an associated position is the sold amount in last period. With last iteration prices and sells, firms build market expectations for next round. Recall we assume firms build a series of scenarios Tr = 10 for each option a) and b), generating a quantity from a uniform distribution. From each set of scenarios, firm chooses the amount that maximizes its utility. In summary, firms evaluate between a) entering an already active (occupied) cell and b) paying an extra cost to open a new cell and "pull" consumers to a new market segment, away from current competition. Then the firm compares the two alternatives through their expected profits and makes a decision based on the option that brings higher expected profits.

We warn the reader that the above computations for expected profits for entry at an active cell are still missing an additional explanation: if the position's total expected production at a given cell is less than the total available demand, each firm may assume that it could sell everything it produces. However, how does a firm compute profits if the total expected production surpasses the total available demand at a given cell? How much sells does a firm expect to receive in such scenario? We answer this question in the next section.

3.7 Market competition and profit calculation

Once the firms have set their production quantities and target positions, competition starts and new profit calculations take place. Since firms choose quantities while seeking for profits, competition resembles somehow what in Industrial Organization is called Cournot competition.[9] As mentioned above, a crucial point is to determine how

[8] Somebody might ask what if there is the option to differ the *NewPos* cost to some number z of periods, so firms better calculate expected profits in the horizon of z periods and check if the return is higher that *NewPos*. Such procedure would just imply a re-calibration of the model, but without loss of generality we prefer to keep it as simple as possible, as it has been presented here.

[9] Cournot competition is usually referred to as a one-shot game theoretical model where firms choose quantities in order to maximize their profits, in the presence of a down sloping demand curve (Tirole 1988).

firms split quantities when there is overproduction and they target the same cell(s). There two cases worth analyzing in this section: a) when aggregated production in a given cell (result of adding up the production levels of all participating firms in such a cell) do not surpass total demand at that cell, and b) when aggregated production surpasses total available demand at that cell.

In the first case (case a), we assume that firms sell everything they produced, since there will be anyway a portion of unserved consumers. Thus, total realized sells equal total production. In the second case (case b), when aggregated production exceeds total available demand at a given cell, sells are split according to firm's contribution to total production.[10] Firms also take into account this information in order to set their production quantities for next time period.

3.8 The organizational niche

If successful in the market competition, a survivor firm (i.e., with nonnegative cumulative profits) might look for further expansion to other points of the resource space, as we explain later. An incumbent firm decides to expand either horizontally or vertically, which implies that niches adopt a rectangular form. We can interpret this as if firms had a main consumer feature over which they develop their business, while they generate product variations in the other (complementary) feature. This is a modeling convenience but also has some sensible arguments to support it. First, avoiding expansion in any direction we also avoid potential issues like having non-convex niches (Hannan et a. 2003) and simplify the rules for niche expansion, which can be only in the direction of the firm's main feature. The possibility of having "rectangular" niches is well supported by the sociological literature (Freeman 1983, McPherson 1983). In marketing, for instance, metrics generalization (i.e., the Minkowski-metric) in optimal product positioning in an attribute space may show that firm's catchment area may be "square-like" (Albers 1979).[11] See figure 5.

[10] The argument to do this is to assign sales participation according to some measure of firm's "scale advantage". In reality, firms with higher scale advantage are usually larger, with larger supply networks and consequently better prepared logistically to place their products. In our model, such scale is reflected in the production capacity of the firm. Some other measures have been used in the literature to measure such scale advantage. For instance, Carroll and Swaminathan (2000) and Dobrev and Carroll (2003) use distance measure as a way to quantify it.

[11] Péli and Nooteboom (1999) also argue that, if the number of dimensions is low, a reasonable assumption might be to have rectangular niches, as proposed by McPherson (1983).

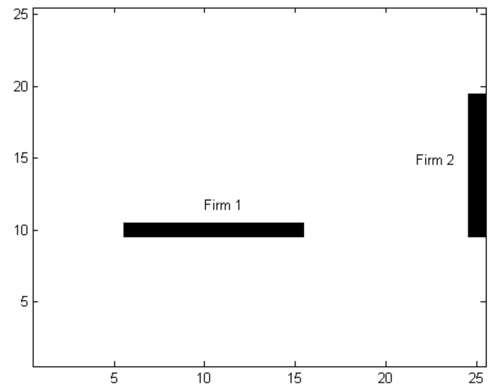


Fig. 5. Two different firms and their niches

3.9 New production level and eventual niche reduction

After transactions have taken place and firms have computed their profits, they set a new production level for the current niche for next time period. In a similar fashion to market entry, evaluate T_r trials to re-adjust their quantities, but incorporating to the analysis last iteration prices and aggregated sold amounts. Projected quantities are drawn from a uniform probability distribution and evaluated. To avoid unrealistic jumps in quantities, we force firms to take the next production level as the average of those projected quantities and the last sold amount. Firms set those quantities that result in higher expected profits for next iteration (with respect to what was observed in last transaction).

If firms are able to expand, they might have been able to pull back as well. Since the space distribution is being transformed by the simultaneous interaction of all the firms, all of them evaluate if dropping the upper or lower taste position of its niche brings incremental profits. This is especially valid for large-scale firms, which may enjoy high production levels with a low number of active cells, but as the market develops, they might prefer to contract since maintaining a spread demand with a very heterogeneous set of consumers might prove to be costly. On the other hand, before expansion or reduction procedures, at time t firms adjust production for time t+1 according to other firms' realized sells, which constitute a basis for next time period quantity estimations. Assuming that firm's k niche positions are ordered from the lowest (position 1) to the upper (position k), firm j computes expected incremental profits at the lowest position if its niche according to the procedure described in Table 1. In the same fashion, the firm computes the expected profits at position k. Then the firm computes the two values and decides to drop the position that generates higher incremental profits. If reduction is not attractive to a firm (because dropping positions lowers its profits), the firm considers expanding.

We also assume that niche reduction may not turn an active position into an inactive one. This means that once a position is opened, no other firm has to pay to open it again, even it such position remains totally unserved due to the niche reduction procedure.

3.10 Niche expansion

Those firms that decided no to reduce their niche are candidates for expansion.[12] Niche expansion evaluates a niche's upper and lower adjacent cells to its current niche, which can also be to either occupied or empty cells. We treat next these two cases separately:

a) Expansion to an already active cell

In order to expand to an already active cell, firms evaluate T_r trials and choose the one with highest expected incremental profits. Each trial consists of selecting randomly (from a uniform distribution.) a quantity to produce in the newly targeted cell. Let us assume a firm attempts expansion to an active adjacent cell. Then, a firm takes into account the additional cost for adding a new position (*NWCost*), the new production cost that includes the newly added demand, and the additional revenue. For instance, if a firm attempts expansion to the upper adjacent position j, it checks if this operation brings positive incremental profits and consider it for expansion. Incremental profits $\Delta\pi_{i,t}$ for a single trial are computed as follows:

$$\Delta\pi_{i,t} = \qquad \textbf{(12)}$$
$$P_j(Q^{-i}_{j,t-1} + q^u)q^u - NWCost - [C(qsold_{i,t-1} + q^u) - C(qsold_{i,t-1})]$$

Terms $Q^{-i}_{j,t-1}$, q^u and $qsold_{i,t-1}$ represent position j's total sold amount at t-1, expected quantity at niche's upper adjacent position and firm i's total sold amount at t-1, respectively.

b) Expansion to an inactive active cell

If the cell is inactive, the procedure is more complex since expansion produces a redistribution of the demand of the whole active space. Again, a firm attempts T_r trials in order to decide which way to expand and with which amount. Qualitatively, in every trial a firm proceeds as follows:

i) The firm builds an "expected space". Based on it computes incremental sales. That is, sales from current scenario (the current space) minus potential sales with the expected space plus the attempted new position.

ii) The firm computes incremental costs (the additional production cost, the additional niche width *NWCost* and the cost of opening a new position *NewPos*).

---

[12] The reader might ask why firms do not compare niche reduction to niche expansion directly, and chooses the one that serves highest profits. The reason by which we have firms evaluating those options sequentially is because preliminary computer experimentation confirms that the profits function is convex respect to quantity, so the decision between reducing and expanding niche is exclusive, as we have modeled it here.

iii) If incremental sales surpass incremental costs, the firm is willing to consider the position as a candidate for expansion.

Firms evaluate the net effect and compute incremental profits for the lower and upper adjacent cells. Firms decide to move towards the cell where incremental profits are positive and larger. A cell is always discarded if bring negative incremental profits.

```
x1 ← lowest_niche_position_row;
x2 ← lowest_niche_position_column;

// CostSaving is computed as -{C(new estimated production) -
C(last production)}

CostSaving ← -(CostCalculation(sum([firm(j).q])-firm(j).q(1))-
CostCalculation(sum([firm(j).q])));

//TotalQ is the total expected amount in (x1,x2); QuantityPer-
Cell(x1,x2) is the available demand at (x1,x2); Q(x1,x2) is
the total sold volume at (x1,x2) in last iteration. The field
qsold is firm's sold amount

TotalQ = Q(x1,x2)- firm(j).qsold(1) + firm(j).q(1);

if TotalQ ≤ QuantityPerCell(x1,x2)

   ExpMarginProfit ← max(0, NWCost + CostSaving - max(0,(A -
B*TotalQ))*firm(j).q(1));

else

   QProp ← firm(j).q(1)/TotalQ;

   ExpMarginProfit ← max(0, NWCost + CostSaving - max(0,(A -
B*TotalQ))*QProp*QuantityPerCell(x1,x2));

end
```

Table 1. Pseudo-code for firm's niche reduction

4. Experimental design

We set a resource space with 100 different positions (this is, $m = 10$). We run every simulation for 300 time periods. We deal with two parameters in our simulation trials: the cost of expanding *NWCost*, and the cost of opening a new taste position *NewPos*. To determine the value range of *NWCost*, we experimented with the model to determine which value will still give positive profits to a large-scale firm when expanded up to the boundaries of the resource space. On the other hand, we determine the value range for *NewPos* by taking into account that a small-scale firm would be able to open a new position and still find positive profits, at least initially. That is, *NewPos* should not exceed the initial endowment of a small-scale firm. Those exploration resulted in $NWCost \in [0, 850]$ and $NewPos \in [0, 400]$.

We first consider a scenario with costless expansion ($NewCost = NewPos = 0$). Among the positive set of values, we "sweep" the range of *NWCost* using 200, 400, 600 and 800. Likewise, we use 100, 200, 300 and 400 for *NewPos*. All of this gives a total of 4x4 = 16 simulation combinations of parameter values. We run 30 simulation runs for every combination. We perform the same procedure in two scenarios: a) when the consumer distribution evolves onto a flat space, so the demand is equally distributed, and b) when the consumer distribution evolves onto a peaked space, where a dominant product attribute combination dominates. A simulation run plot showing the shape of the peaked resource space can be seen in the appendix (figure A2). Thus, the total number of simulation runs is 1x30x2 + 16x30x2 = 1020.

5. Findings

5.1 Costless expansion

This scenario reveals that small-scale firms are unable to proliferate when dominant large-scale can freely expand without incurring in any scope diseconomies, as showed in figure 6. This figure illustrates the aggregated data for the average profit/cost ratio. Very similar patterns where obtain in both flat and peaked spaces. We also observed the average profit / cost ratio per firm of the last transaction (i.e., last iteration). As observed in the figures, an initial dimensionality increase makes large-scale firms able to benefit from some degree of product differentiation. Further differentiation makes total demand split among a number of players, which reduces the large-scale profit / cost ratio, but without getting necessarily worse than small-scale ratio.

Figures 7 revealed the behavior of the large-scale and small-scale populations over time, illustrating the difficulty perceived by the small-scale to proliferate. It is an example It is apparent to see that the decline in profit / cost ratio for the large-scale firms is due to the fact that there are more large-scale firms competing and sharing the market. Specifically, a higher entry rate generates a larger presence of large-scale firms and thus higher incentives to differentiate, producing larger product diversity and an increase in dimensionality.

Additional experiments with different entry rates (e.g., a stochastic entry process with a constant rate) confirm this result. For instance, an extreme case is a realization of one entry according to a Bernoulli process with probability $p = 0.01$ show that not necessarily the profit / cost ratio decreases for the large-scale firms, since entry is very unlikely and incumbents only attempt to differentiate enough to increase profitability. We also experimented with different starting populations (other than 1) and, again, observed that the large-scale ratio can be increasing with the number of dimensions, while the small-scale population gradually disappears over time. Figure 8 shows s simulation run in a peaked space with a Bernoulli entry process with probability equal to 0.01.

Proposition 1: a) In a costless expansion scenario, large-scale firm take over the whole market, b) average large-scale profit/cost ratio behavior increases with dimensionality up to some level of product differentiation but declines as long as entry becomes more intensive, c) average large-scale profit/cost ratio appears to be always higher than the small-scale ratio.

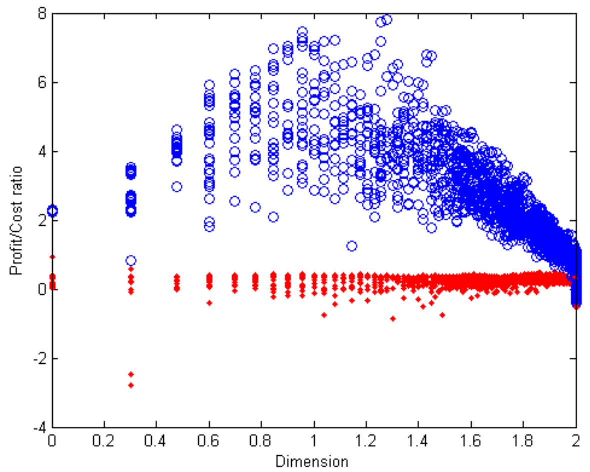


Fig. 6. **Costless expansion / flat space (blue circles = large-scale, red dots = small-scale)**

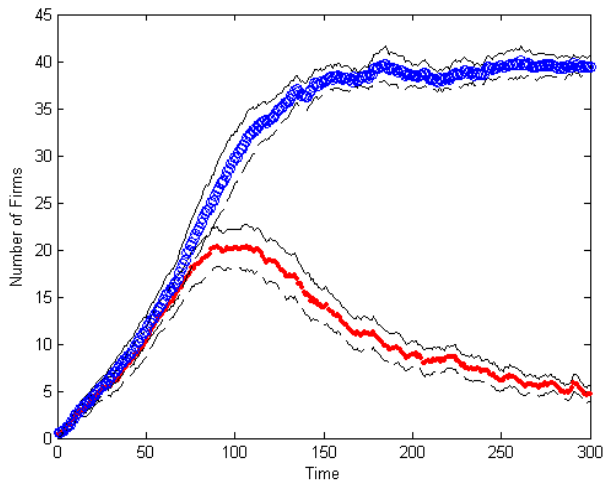


Fig. 7. Costless expansion / flat space. Averages are represented by blue circles (large-scale) and red dots (small-scale). Black solid and dashed lines indicate confidence intervals at 95%

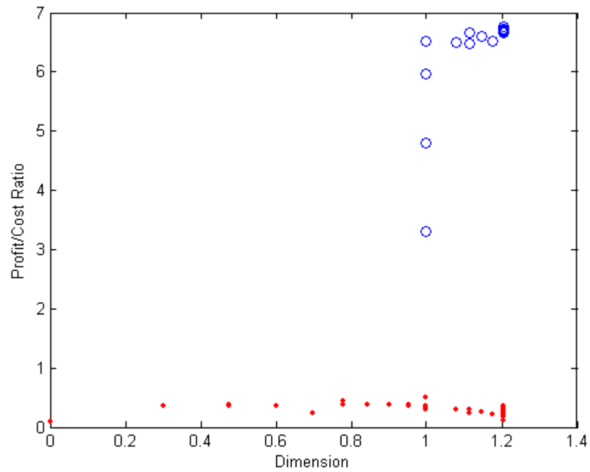


Fig. 8. Costless expansion / peaked space / Bernoulli entry with p = 0.01.

5.2 Costly expansion

Inclusion of expansion costs leads to a series robust results across the different tested parameters, resource space shapes (either flat or peaked) and different entry mechanisms. Under the entry mechanism explain in section 3, it is observed a proliferation of small-scale firms over time. Figure 9 reveals such pattern. Experiments with the other parameter values reported behaviors closely similar to the one presented in figure 9. Also, a nonmonotonic behavior of the profit/cost ratio for large-scale firms was observed. Again, large-scale firms benefit from some degree of product differentiation (i.e., at low dimensionality), but at higher dimensionality decreases even below that of the small-scale firms. Figures 10, 11 and 12 reflect such pattern of results. Alike the costless scenario, the nonmonotonic behavior of the large-scale profit/cost ratio was always observed, even with different entry mechanisms: constant entry with Bernoulli process and low probability of success (e.g., 0.1, 0.05) and at different values and different starting populations (e.g., 40, 50).

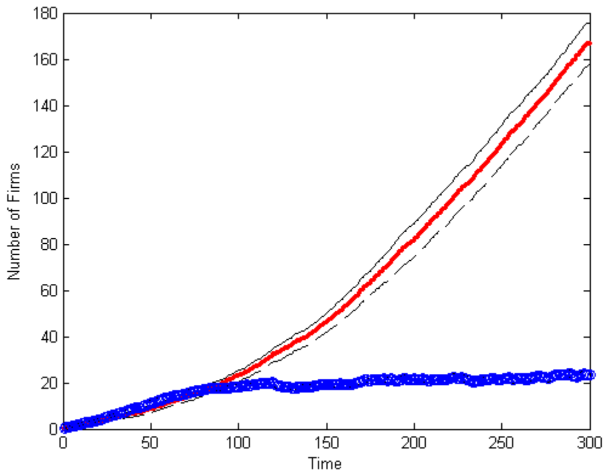


Fig. 9. Population evolution in peaked space (blue circles = large-scale, red dots = small-scale), NWCost = 200; NewPos = 100. Black solid and dashed lines indicate confidence intervals at 95%

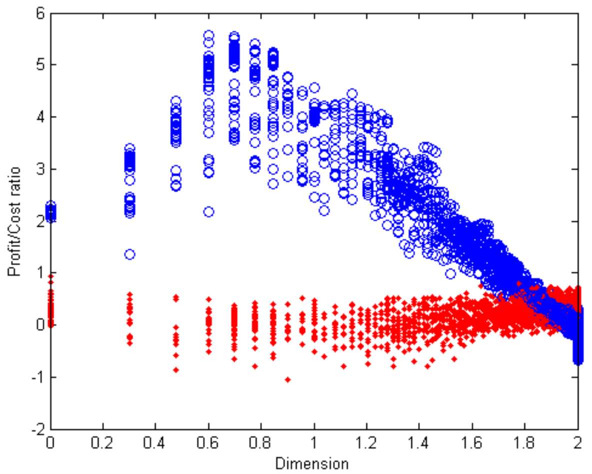


Fig. 10. Average profit/cost ratio in peaked space (blue circles = large-scale, red dots = small-scale), NWCost = 200; NewPos = 100

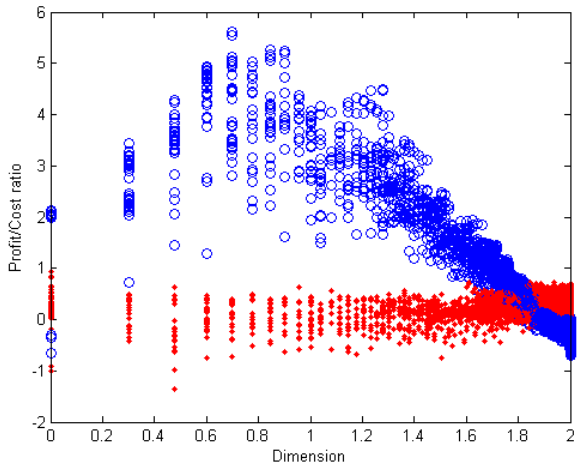


Fig. 11. Average profit/cost ratio in peaked space (blue circles = large-scale, red dots = small-scale), NWCost = 200; NewPos = 400

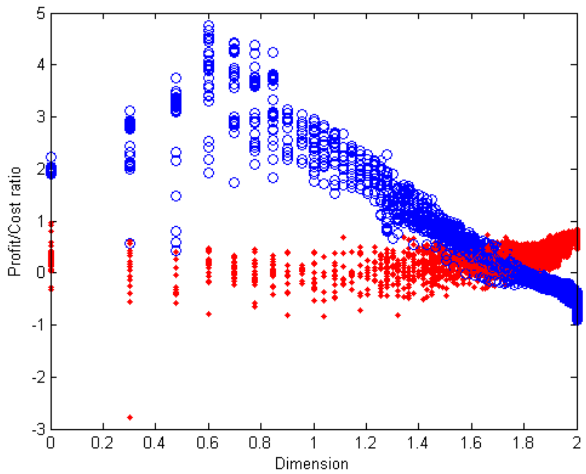


Fig. 12. Average profit/cost ratio in peaked space (blue circles = large-scale, red dots = small-scale), NWCost = 600, NewPos = 100

One more interpretation of why the large-scale type has an average profit/cost ratio more attractive than the small-scale one, is due to the fact that large-scale firms are able to price discriminate: if a large-scale firm attempts to open a new position, it may enjoy the benefit of having a small set of consumers charged at a higher price. That justifies the expansion of large-scale firms at low dimensionality. However, those benefits decline to a point lower than the small-scale ratio, since additional niche operational costs become prohibited.

Proposition 2: a) Small-scale profit/cost ratio is likely to be higher than the large-scale one as dimensionality increases when expansion is penalized with niche costs and position opening costs, b) small-scale firms are able to proliferate in resource spots where scale dominance does not compensate cost expansion.

5.3 Who are the innovators?

The last question we investigate is, which firm type is the major innovator force (i.e., those who open new positions in the space) over time? Is it the large-scale or the small-scale type? Is it influenced by any factor like cost of expansion, resource space type or the relative difference between niche and new opening costs?

In the costly expansion, we observed similar patterns for all the parameter combinations and the two types of resource space. In the costly expansion, all innovators are the large-scale firms. The average behavior under the costless expansion is observed in figure 13: many small-scale firms become innovators; however, the cumulative number of small-scale innovators never surpasses that of the large-scale firms.

This result complements what was mentioned above, regarding the interpretation of the decline of the profit/cost ratio for the large-scale firms: when the population of incumbent firms grows, the incentives for product differentiation increase as a way of differentiation from competition. It is also clear that the increase in expansion costs leave the small-scale unable to benefit from any position opening. Figure 14 reveals the innovation pattern for the costly expansion case.

However, it is important to recall that these results are dependent on what we have called small-scale and large-scale. Further sensitivity analyses that involve increases in the scale advantage of small-scale firms tend to modify this result. That is, taking advantage of the robustness of previous results, and using a peaked space with *NWCost* = 200 and *NewPos* = 100, we varied the small-scale from $Q_{SC} = 10$ to higher values (50, 70, 75, 80, 100, 250) and found that i) their incentive to innovate becomes higher after the value range of [70 80], ii) their market density figures sharply decline. When their scale advantage surpasses the value range of [70 80], to name them "small-scale" is inappropriate, and to talk about a real "proliferation" may be out of context. After the value of 250, these firms resemble similar behavior to the large-scale firms with respect to dimensionality, but are no longer "small-scale". Under a costless expansion scenario, increases in the values of Q_{SC} generated inconclusive results that deserve a further separate experimental design.. However, those results never proved to alter the profit/cost behavior previously observed in the costless expansion scenario.

Proposition 3: When expansion is costly, and the difference in scale between the large and the small is big, only large-scale firms have incentives for opening new positions as a mean for product differentiation. The collective effect of innovation of large-scale firms generates enough product diversity to stimulate small-scale entry. Small-scale firms tend to increase their incentives to innovate as their scale advantage increase.

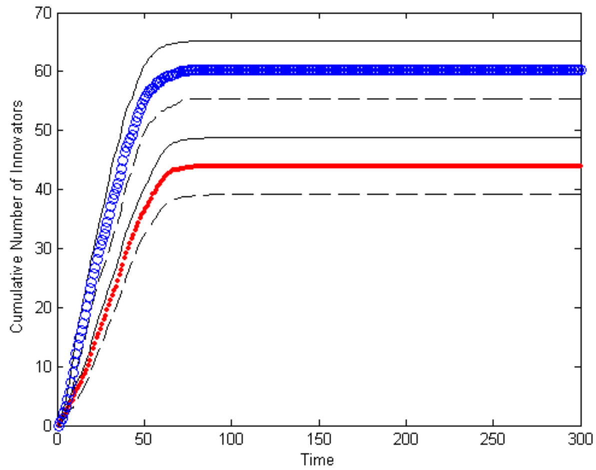


Fig. 13. Average cumulative number of innovators per type under costless expansion in peaked space (blue circles = large-scale), $Q_{SC} = 10$

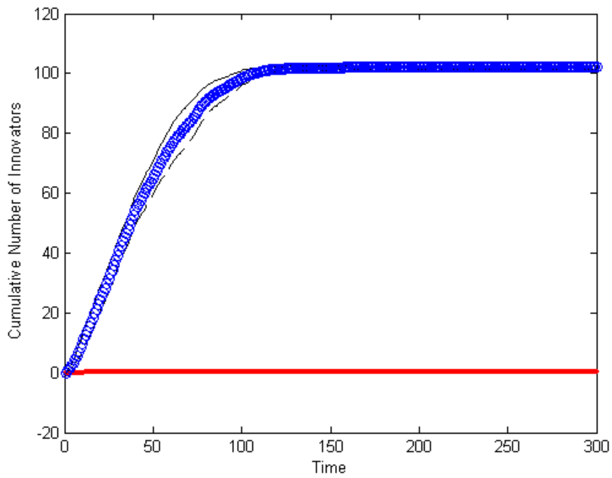


Fig. 14. Average cumulative number of innovators per type under costly expansion in peaked space (blue circles = large-scale). NWCost = 600, NewPos = 100, $Q_{SC} = 10$.

5. Concluding remarks

We have presented and agent-based modeling approach of a market where the active product features of the space change over time. We have investigated the effect of increasing dimensionality on small-scale and large-scale firms. The model illustrates i) an approach to account for dimensionality in economic markets, ii) a consistent set of findings related to previous research that complement other modeling approaches (Péli and Nooteboom 1999, Péli and Bruggeman 2006). Our model also makes endogenous the dimensionality change, and dependent on firm-level interaction. It resembles a spatial repeated Cournot competition where expectations are updated according to previous market transactions.

The model also reveals that i) the advantages of increasing dimensionality for small-scale firms are not due to their improved strategy or capabilities, but to the fact that large-scale firms suffer more the spreading of consumers across the resource space; ii) large-scale firms are able to enjoy a small dimensionality increase, which softens competition with other large-scale firms, while allowing a small degree of differentiation. Such small increase pushes up their profit/cost ratio. However, a further dimensionality increase weakens the scale advantage and provides better chances for non scaled-based competitors. A graphical summary of results is presented in figure 15.

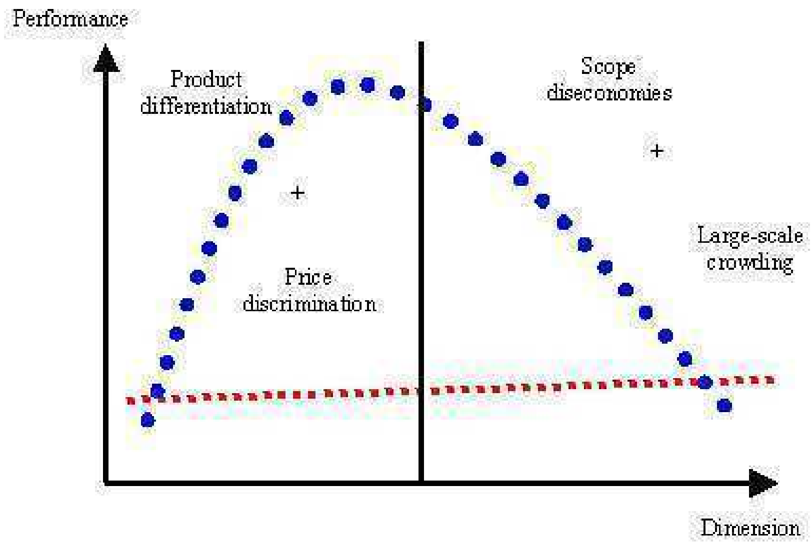


Fig. 15. Graphical representation of propositions (blue circles = large-scale; red dots = small-scale)

Appendix

δ_o	-0.9849
δ_1	0.02
δ_2	-0.00005

Table A1. Parameter values used for density-dependent model

α	0.7
β	0.7
W_V	4.1520
W_F	8.3040
A	1

Table A2. Parameter values used for LRAC curve

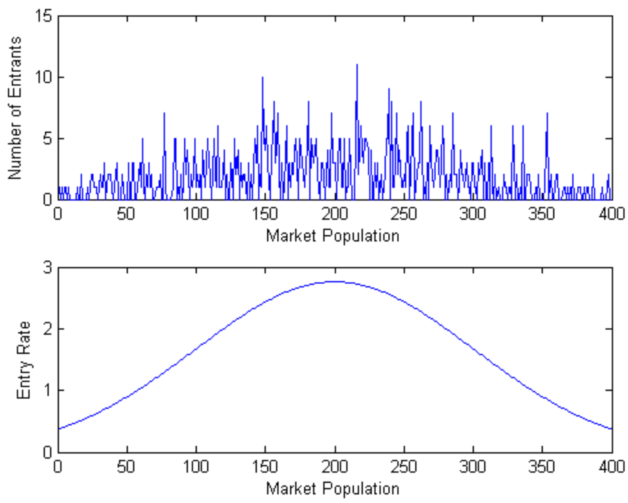


Figure A1. Example of density-dependent entry mechanism

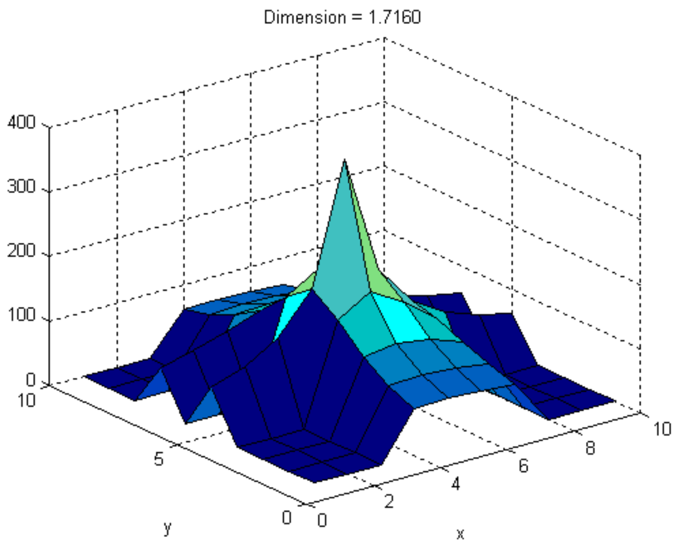


Figure A2. Peaked resource space

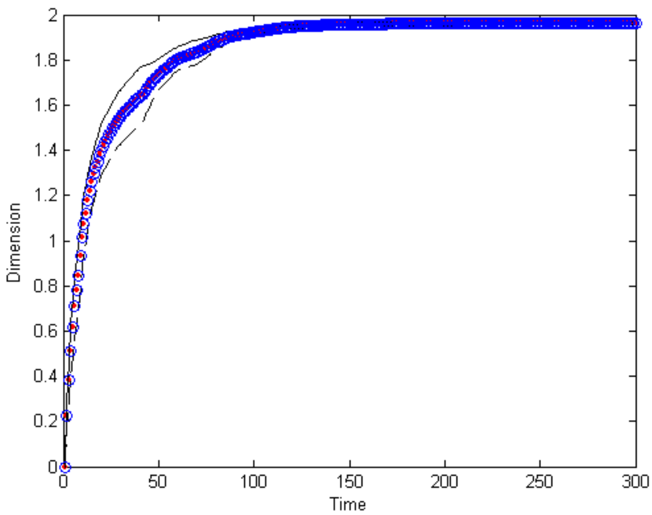


Figure A3. Evolution of dimensionality

References

Albers S (1979). An extended algorithm for optimal product positioning. *European Journal of Operations Research* 3: 222-231.

Basmann R (1954). A theory of demand with variable consumer preferences. *Econometrica* 24: 47-58.

Boone C, Carroll G, Witteloostuijin, A. van (2002). Environmental resource distributions and market partitioning: Dutch daily newspaper organizations from 1968 to 1994. *American Sociological Review* 67: 408- 31.

Boone C, Wittteloostuijn A van (2004) A unified theory of market partitioning: an integration of resource-partitioning and sunk cost theories. *Industrial and Corporate Change* 13: 701-725

Carroll G (1985). Concentration and specialization: Dynamics of niche width in populations of organizations. *American Journal of Sociology* 90: 1262-1283.

Carroll G, Hannan M (1995). Resource-partitioning. In Carroll G, Hannan M (Eds.) *Organizations in Industry: Strategy, Structure and Selection*. NY: Oxford University Press, pp. 215-221.

Carroll G, Hannan M (2000). *The Demography of Corporation and Industries*. Princeton University Press: Princeton.

Carroll G, Dobrev S, Swaminathan A (2002) Organizational processes of resource partitioning. In: Kramen R, Staw B (eds) Research in organizational behavior, New York: JAI/Elsevier. pp 1-40.

Carroll G, Swaminathan A (2000). Why the microbrewery movement? Organizational dynamics of resource partitioning in the U.S. industry. *American Journal of Sociology* 106(3): 715-762.

Dawid H, Reinmann M, Bullerheimer B (2001). To innovate or not to innovate. *IEEE Transactions on Evolutionary* Computation 5: 471-481.

Dobrev S, Carroll G (2003). Size (and competition) among organizations: Modeling scale-based selection among automobile producers in four major countries, 1885-1991. *Strategic Management Journal* 24(6): 541-558.

Freeman L (1983). Spheres, cubes and boxes: graph dimensionality and network structure. *Social Networks* 5: 139–156.

García-Díaz C, Witteloostuijn, A van (2006). Co-evolutionary market dynamics in a peaked resource space. In Bruun C. (Ed.). *Advances in Artificial Economics - Lecture Notes in Mathematics and Economic Systems* 584: 93-104. Berlin: Springer.

Hannan M, Carroll G (1992). *Dynamics of Organizational Populations*. Oxford, England: Oxford University Press.

Hannan M, Freeman J (1977). The population ecology of organizations. *American Journal of Sociology* 82: 929-964.

Hannan M, Freeman J. (1989). *Organizational Ecology*. Harvard University Press.

Hannan M, Carroll G, Pólos L(2003) The organizational niche. *Sociological Theory* 21 (4): 309-340.

Hannan M, Carroll G, Pólos L(2007). *The Logics of Organization Theory: Social Codes and Ecologies*. Princeton NJ: Princeton University Press.

Haskell JP, Ritchie M, Olff H (2002). Fractal geometry predicts varying body size scaling relationships for mammal and bird home ranges. *Nature* 418: 527-430.

Harrison JR (2004) Models of growth in organizational ecology: a simulation assessment. *Industrial and Corporate Change* 13: 243-261.

Lachaab M, Ansari A, Jedidi K, Trabelsi A (2006). Modeling preference evolution in discrete choice models: a Bayesian state-space approach. *Quantitative Marketing and Economics* 4: 57-81.

Lancaster K (1966). A new approach to consumer theory. *Journal of Political Economy* 74: 132-157.

Lee J, Harrison R (2001). Innovation and industry bifurcation: the evolution of R&D strategy. *Industrial and Corporate Change* 10(1): 115-149.

Mandelbrot B (1983). *The Fractal Geometry of Nature*. W. H. Freeman, New York.

McPherson JM (1983). An ecology of affiliation. *American Sociological Review* 48: 519-532.

McPherson JM (2004). A Blau space primer: prolegomenon to an ecology of affiliation. *Industrial and Corporate Change* 13, 263–280.

Olff H, Ritchie M (2002). Fragmented nature: consequences for diversity. *Landscape and Urban Planning* 58: 83-92.

Péli G, Bruggeman J (2006). Networks embedded in n-dimensional space: The impact of dimensionality change. *Social Networks* 28: 449-453.

Péli G, Nooteboom B (1999). Market partitioning and the geometry of the resource space. *American Journal of Sociology* 104: 1132-1153.

Sutton J (1991). *Sunk Costs and Market Structure: Price Competition, Advertising and the Evolution of Concentration.* Cambridge MA: MIT Press.

Tirole J (1988). *The Theory of Industrial Organization*. MIT Press: Cambridge.

Swaminathan A, Carroll G (1995). Beer brewers. In Carroll G and Hannan M. (eds.) *Organizations in Industry: Strategy, Structure and Selection*, Oxford University Press. pp. 223-243.

Witteloostuijn A van, Boone C (2006) A resource-based theory of market structure and organizational form. *Academy of Management Review* 31: 409-426.

Zinam O (1974). Role of consumer preferences in an economic system. *American Journal of Economics and Sociology* 33: 337-350.

Growing *qawms*: A case-based declarative model of Afghan power structures

Armando Geller and Scott Moss

Centre for Policy Modelling, Manchester Metropolitan University Business School
Aytoun Building, Aytoun Street, Manchester, M1 3GH, United Kingdom
{armando, scott}@cfpm.org

Abstract. By means of evidence-based and declarative social simulation we grow *qawms*. *Qawms* denote solidarity networks in Afghanistan. They are dynamic social modules that contribute to the establishment of social structure. The study of *qawms* lends insight into the structural and processual dynamics of Afghan society. In particular we concentrate on the evolution of power structures. A computational model is presented whose ontology is based on a notion of power structures traceable in contemporary conflicts. The model's agent behaviour, however, is informed by qualitative data derived from case studies on Afghanistan and, in terms of cognition, by the conception of endorsements. Our preliminary findings suggest – in accordance with existing case-studies – that actors are deemed if they are isolated, but perform strong if they are socially embedded and act according to the principles of neo-patrimonial behaviour.

Keywords: Afghanistan, anomie, critical realism, declarative modelling, evidence-based social simulation, neo-patrimonialism, power structures.

1 Introduction

We computationally grow *qawms*. *Qawms* denote (opportunistic) solidarity networks in Afghanistan. They can be perceived as dynamic modules based on actor interconnectedness which generate social order. Because *qawms* permeate all different spheres of Afghan social life, their analysis allows multidimensional insight into a conflict-torn society. We are particularly interested in the formation of power structures amongst Afghan stakeholders, for whom *qawms* are a means to acquire, maintain and increase power. It is in this respect the notion of neo-patrimonialism is pivotal to our understanding of power. Although neo-patrimonialism has been identified as an important organisational principle in contemporary conflicts (cf. [1], [3], [21], [22]), it has been routinely ignored in model-based approaches to conflict analysis.

The methodological approach we apply is evidence-based and declarative multi-agent social simulation. It is evidence-based because agent behaviour is informed by qualitative data derived from case studies and interviews conducted in Afghanistan; it is declarative because of a descriptive – in contrast to imperative – implementation of the evidence-based agent behaviour.

There are three reasons such a research design has been chosen. i) Statistical data is either scarce or unreliable in the context of contemporary conflicts, constraining the scope of applicable methodological tools. ii) A variety of social phenomena exhibit characteristics which can be labelled as complex, a condition in which agent behaviour and social interaction combine to generate macro level outcomes that firstly could not be predicted from knowledge of the behaviour and nature of interactions alone and secondly result in sporadic volatile episodes the timing, magnitude, duration and outcomes of which are themselves unpredictable [2]. In the Afghan context, the interactions within *qawms* determine a pattern of actions that could be described as episodic clusters of aggressive activity or even extreme violence. Such *qawm* level behaviour leads to interactions amongst the *qawm* that cause episodic conflicts of unpredictable magnitude, duration and outcome. It is the virtue of agent-based social simulation to analyse such phenomena [5], [8], [10], [14], [17]. iii) Declarative programming supports the use of mnemonic terms resembling natural language terms in which stakeholders describe behaviour, social environments and social interactions.

Our research design is developed against the backdrop of critical realism, a meta-theory compatible with complexity [10]. Social science in the tradition of critical realism focuses on a context-sensitive approach (see also [4]) to agency and structure, the interplay of which leads to emergent phenomena, underlining the generative paradigm of computational social science [26]. The *entrée* for the object of investigation is provided by a well-informed, but intuitional model that serves as a real definition and thus specification to the context-relevant aspects of agency and structure [19] (see sec. 2). Once properly defined, the intuitional model needs to be qualitatively validated to ensure the descriptive accuracy of the computationally implemented processes (agency) and architectures (structure). A construct valid model is expected to yield strong results which can be cross-validated against real world (statistical) data (if available).

This study is expected to be of general interest as a research design is presented that allows for systematic and dynamic but context-sensitive analysis of social phenomena under statistical data scarcity, as for example in conflict-torn societies.

2 Anomie, neo-patrimonialism and Afghanistan

2.1 Power in a deranged order

Power structures are anthropogenic [20]. In conflict regions such as Afghanistan, Chechnya or the Democratic Republic of Congo the anthropogeneity of power has been shown in a variety of studies [3], [21], [22].

Sofsky [28] argued that conflict societies are societies *sui generis*. They function according to their own social laws and are structurally and processually disjointed from societies lacking a comparable degree of organised violence. In conflict-torn societies virtually everything goes. This can be illustrated by the concept of *anomie*.

Anomie is the situation in which the upper and lower normative boundaries for the aspirations of members of a society are thrown awry [7] [11]. An anomic situation emerges when the means to attain a specific goal, such as accumulation of wealth or power, run out of social control [13]. Accordingly, in a space emptied of restricting norms, i.e. an anomie, virtually everything goes – also the creation of power structures to one's own ideas and interests.

Anomic spaces are political spaces lacking strong modern institutions, such as the state's monopoly on organised violence, stability of the law and protection of property rights. In these circumstances only highly adaptive stakeholders prevail. The socio-structural outcomes of this organisational process are manifold and so are the adopted means that serve one's interests.

In contemporary conflict societies this outcome is neo-patrimonialism (cf. [12], [22], [10]). Weber [30] understands patrimonial power as power based on authority, suppressed subjects and paid military organisations, by virtue of which the extent of a ruler's arbitrary power as well as grace and mercy increases. Stakeholders interested in gaining power in contemporary conflict settings have to act neo-patrimonially to accumulate and redistribute monetary and material as well as social resources. The range of related activities is broad and includes corruption, clientelism, patronage, nepotism, praebendism and so forth (cf. [12]).

2.2 A model of *qawm*

Although twenty-seven years of conflict accentuated two important factors in Afghan society, namely ethnicity and religion, the traditional organisational principle of the *qawm* rested sound [23], [25], [27]. Less mentioned, however, is a general decline of norms and values in Afghan society leading to a Hobbesian form of society [29]. Today's Afghanistan is an *anomie*.

The causes for this development are complex, but nevertheless directly linked to the Jihad of 1979 to 1989. Although trends of neo-patrimonial politics are already recognisable in the very beginning of the Jihad – and are indeed a characteristic of Afghan politics throughout history –, the war's proper goals started to mutate with its increasing duration, from throwing out the Soviets to personal enrichment and personal aspiration for power. Some of the adopted means of warfare have been traditional, such as organised violence, intrigue, alliance formation and dissolution; others have been imported, such as religious extremism, radicalisation of ethnicity, corruption and narcotics.

The *qawm* is a dominating feature of Afghan society [23]. Mousavi [18] refers to it as a complex interpersonal network of political, social, economic, military, and cultural relations. Afghan social structure does not take the form of a unified hierarchy and nor does an individual *qawm*. However, each *qawm* has a *primus inter pares* who competes with other *primi inter pares* as well as with *qawm*-internal rivals for manifold reasons [24].

Figure 1 depicts an informed intuitional and idealtypical representation of an Afghan solidarity network or *qawm*. Our notion of *qawm*-based social structure has been informed either by the literature cited above or by data collected by ourselves.

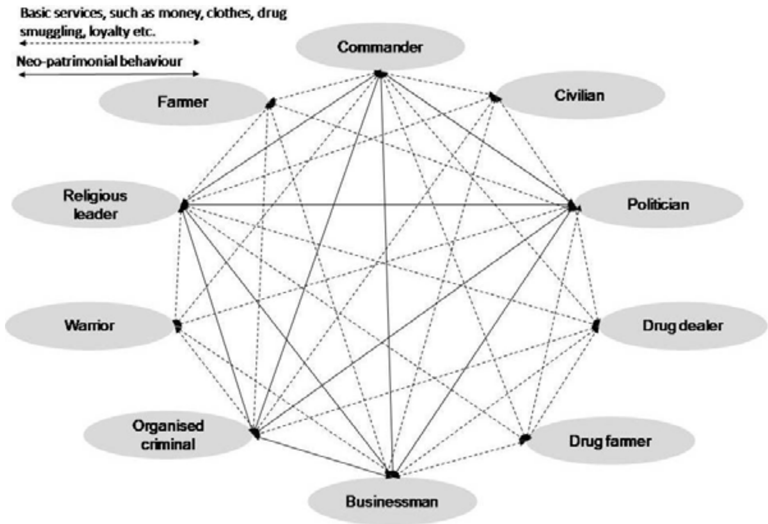


Fig. 1. A case-study informed intuitional model of a *qawm*.

The *qawm*-model consists of ten actor types: politicians, religious leaders, commanders (meritocratic title for a militia leader), businessmen, warriors, civilians, farmers, drug farmers, organised criminals and drug dealers. An important abstraction from reality is that in our model each actor has its distinct role, whereas in reality actors may incorporate a variety of roles. For example a commander can be a (military) commander, a politician and a drug lord at the same time, for which General Mohammed Daud is a good example, a former Deputy Minister of Interior Affairs, Governor of Badakhshan, favourite commander of Ahmad Shah Massud and drug lord. We proxy individual role pluralism by mutual interdependence, i.e. each actor has virtues another actor may be in need of and vice versa, leading to mutual coorporation and interdependence. This, of course, is also a common pattern in reality, where there is no clear distinction between role incorporation and cooperation.

The following examples explain the *qawm*-model in terms of agency. If a politician is in need of military protection, he approaches a commander. In return, a commander receives political appreciation by mere cooperation with a politician. If a businessman wants to be awarded an official construction contract by the government, he relies on a politician's political connections. In return, the politician receives a monetary provision, for example bribes. If a politician wants beneficial publicity, he asks a religious leader for support. The religious leader, in return, becomes perceived as a religious authority. If a warrior seeks protection and subsistence for his family, he lends his services to a commander, who, in return, provides him with weapons, clothes, food and/or money. If an organised criminal wants to carry drugs, he relies on the transport business of a businessman who, in return, receives a share of the drugs sold. If a drug farmer needs protection for his poppy fields, he affiliates with a commander, who, in return, receives a tithe on the drugs sold to a local drug dealer. Our model represents this behaviour.

3 Modelling *qawms*

3.1 Evidence-based and declarative modelling

The evidence on which we base our model is largely qualitative and drawn from case-studies (cf. [1], [9], [18], [23], [24], [25], [27], [29]) or data collected by ourselves during semi-structured interviews with Afghan urban elites between May 2006 and January 2007. Interviewees were motivated to reflect on power. The main reason we rely on rich qualitative data is the lack of reliable statistical data in Afghanistan and other comparable contexts.

Evidence denotes information about the target system that allows to develop a representative model *of* reality. This information stems from case-studies, empirically tested theories *and* interviews with experts and stakeholders. A model is evidence-based if the rules according to which agents behave are directly derived and reified from this information. This presupposes that the data makes concrete declarations of how an actor behaves in a particular social situation. The triangulation of the information sources is vital for the model's validity.

Unlike other simulation approaches, evidence-based modelling pursues construct validity. It is important that the modelled processes and structures resemble the processes and structures identified in the target system. Agent-based models are more than mere input–output models. As demanded by critical realism, they direct a researcher's focus on internal processes (agency) and structures and allow for the analysis of them.

Results are more valid if an evidence-based social simulation's output can be cross-validated [17] and not only "validated" by circumstantial evidence. There are three strategies: i) If models generate statistical output, this output is statistically analysed and the resulting significant signatures are compared with statistical signatures generated from target system data. If the model yields numerical output but no statistical target system data is available, then validation must rely on qualitative data. In this case, validation must ii) either seek systematic structural and processual similarities between the model and the target system or iii) find circumstantial evidence in the target system that can also be found in the simulation. Except for civilian casualties data only ii) and iii) can be applied to our model.

A program is declarative if there are a set of statements on a database, rules have a set of conditions which are statements with some values left open as variables, and consequents exists which are another set of statements. When all of the statements in the conditions of a rule are matched by statements on the database, then the variables are given their specific values from the database statements and the consequent statements are added to the database. When a set of conditions is satisfied and a rule fires (i.e. puts its consequents on the database), then the state of the environment as represented by the database is changed and perhaps other rules will now be able to fire and so on until all rules have fired and no further matches of conditions can be found on the database. The sequence of rules that will fire and the particular values of their variables' instantiations are determined only as the program is running. The sequence of actions represents the process of agent behaviour and leads in each case to a new state of the environment. If all agents are implemented declaratively, then they will be changing the state of the environment for one another and the pattern of rules and therefore actions of all of the agents taken together will be influenced by one another.

The outcomes for the model as a whole are, in these circumstances, impossible to predict with any exactitude. Frequently, such models exhibit the sort of episodic volatility associated in the first section with complexity. The same effect can be achieved by other means, but declarative representations of agents have a number of virtues in terms of ease of development as new evidence becomes available and in terms of yielding comprehensible outputs stored as statements on the databases.

3.2 Endorsements

As we have implied above, power relations are interactions between at least two actors. The computational implementation of these interactions must be based on certain grounds. This can be knowledge an actor has about another actor; it can also be experiences an actor has made in the past with his environment. Endorsements are a "natural" way of computationally implementing reasoning about this knowledge or experience. As cognitive-behavioural modules they represent within the computational model what has been described in section 2.2 as agency aspects of the *qawm*, i.e. interaction, but need further specification (see below).

Endorsements were introduced by Cohen [6] as a device for resolving conflicts in rule-based expert systems (cf. also [16]). Endorsements can be used to describe cognitive trajectories aimed at achieving information and preferential clarity over an agent or object from the perspective of the endorsing agent himself. We use endorsements exactly in this sense, namely to capture a process of reasoning about preferences and the establishment of a preferential ordering (cf. [15], [17]).

Because endorsements capture an agent's (the endorser) reasoning process about another agent (the endorsee), the information collected by the endorser cannot be identical with the endorsee himself. It is more precise to state that during the endorsement process the endorser's endorsement scheme is projected onto the endorsee. If a commander endorses for example a businessman, he has no base to rate if the businessman is a better Muslim if he is a Sunni or a Shia. But the individual endorsement scheme tells the commander how important it is for himself that the businessman is Sunni or Shia. If this is done for every of the endorsee's attributes, the so called overall endorsement value (E) for the endorsee can be calculated. E allows the endorser to select the preferred endorsee among the endorsees.

The process of endorsing an agent is embedded in an agent's environment, i.e. his neighbouring agents. The endorsement process allows an agent to find the agent most appropriate to *him* – he does not seek the most successful or most reliable of all agents. This implies that the chosen agent may not be preferable to differently embedded agents.

The main advantage in applying the idea of endorsements lies in the fact that they allow for combining the efficiency properties of numerical measures with the richness and subtleties of non-numerical measures of interest or belief.

Based on our interview data and on secondary data, particularly [1], we had to develop an endorsement scheme for an idealtypical "Afghan agent", which, in principle, had to answer the following questions: When is an agent powerful? How does a powerful agent behave? How does an agent behave towards a powerful agent? Azoy [1] argues that authority depends on *hisiyat*, character, and *e'tibar*, credit. The analysis of our interview data supports this view. Logically, *hisiyat* and *e'tibar* are the dimensions in which an "Afghan agent" reasons about another agent. *Hisiyat* is related to the social embeddedness of an actor. An actor has character if he is of particular kin, religion and/or neighbourhood and can, in case of cultural pattern matching, be trusted. An actor who is creditworthy and has political support disposes of *e'tibar*. *E'tibar* has to do with meritocracy and reliability.

Hisiyat and *e'tibar* can be straightforwardly operationalised. *Hisiyat* are generally intangible endorsements and are attributed at the beginning of the simulation, such as ethnicity, religion or kin. *E'tibar* are generally dynamic endorsements which change their values during the simulation, such as payment or success.

3.3 Model specifications

The simulation is spatially based on a 2D-grid topology, the dimension of which is 30 × 30 cells. Each cell can be inhabited by one agent. The total number of cells are distributed into 4 ethnic regions (Pashtun 40%, Tajik 25%, Uzbek 25%, Hazara 10%).[1] There are ten agent types and a total number of agents of 132: 3 politicians, 3 religious leaders, 3 businessman, 3 organised criminals, 4 commanders, 6 drug dealers, 10 drug farmers, 20 warriors, 30 farmers and 50 civilians. In the beginning of the simulation, each agent is assigned an ethnicity (proportional to the land indices), a religion (Sunni or Shia), a number of kinspersons and a (Moore) neighbourhood. The majority of the land is rural. There is only one city. Rural areas are rather homogeneous in terms of ethnicity and religion, whereas the city is a "multicultural" space. Some agents belong only to rural spaces, such as farmers (drug farmers), some only to the city, such as organised criminals. However, we do not model geographical representation as such and our agents are spatially static.

The overarching model architecture is simple. Everything that has to do with cognition is implemented in Jess, the rest is implemented in Java. For example the decision if a commander wants to invest money with a businessman is coded in Jess. Money transaction and control over assets, by contrast, are coded in Java. Besides the arguments put forth in section 3.1 such an architecture makes efficient use of computational resources. The model makes use of the Repast libraries.

4 Results

We discuss three categories of results. First we describe a commander's "life". Subsequently we refer to the sequence of a single simulation run. Whereas the first two categories address problems of agency, the final category deals with networks and thus with structural aspects of the simulation results.

[1] Numbers and proportions are vague approximations as the size of the Afghan population, something between 20 and 25 million people, is an unreliable and politicised figure.

All data presented is drawn from the same representative simulation run with the abovementioned agent configurations.

4.1 A commander's "life"

A commander behaves according to the rules listed in table 1. The names of the rules are self-explanatory. Other agents have comparable sets of rules.

Table 1. A selection of the set of rules according to which a commander behaves (implemented in Jess). "h" denotes *hisyiat*, "e" denotes *e'tibar*.

Rule
default-daily-payment-commander-to-warrior [e]
commander-endorses-warrior-as-reliable [e, h]
commander-endorses-warrior-as-unreliable [e, h]
commander-endorses-warrior [e, h]
commander-endorses-businessman [e, h]
commander-endorses-politician [e, h]
commander-asserts-trustworthiness-affiliation-with-politician [h]
commander-sends-message-to-answer-politician-protection-request [e, h]
commander-endorses-religious-leaders [h]
commander-sends-message-to-best-endorsed-religious-leader [h]
commander-sends-message-to-answer-religious-leader-spiritual-leader-request [h]
commander-sends-message-to-answer-businessman-protection-request [e]
commander-sends-message-to-accept-businessman-protection-request [e]
commander-invests-money [e]

During the setup of the simulation (time step (tick) 0), a commander (actually commander-1) checks how many warriors he can afford on the basis of his assets. He scans his neighbourhood for warriors, kin and spatial neighbours. He endorses warriors, selects a number of them, offers to hire them and pays those he hires.

In tick 1, the commander endorses several warriors as unreliable and at the same time endorses a number of agents for same religion and ethnicity. He keeps on collecting warriors and pays their salary on a daily basis. He also endorses a politician, a businessman and a religious leader.

In tick 2, the commander mainly collects warriors and pays them. In principal this goes on as long as he has money to pay them.

In tick 3, the commander again endorses a businessman and a religious leader but is not able to affiliate with either of them.

Although commander number 1 is a particular boring commander – at least in this simulation run – we can see clearly how he gradually builds up his neighbourhood, namely by constantly watching the area surrounding him for possible affiliations. If he finds a suitable agent he endorses him and decides on the basis of varying facts – for example ethnicity, religion, kin, reliabilty, trustworthiness – if he wants to affiliate with him. It becomes also obvious that the ability to accumulate and redistribute resources is vital. Commander number 1 will have a hard time, as he is lacking affiliations with businessmen and thus will soon have no more money to pay his warriors on a regular basis. This will characterise him as untrustworthy. As a consequence he will have even less affiliations leading to a “death” in isolation. As a neo-patrimonial agent commander number 1 failed.

4.2 The sequence of events

We stated above (section 3.1) that the sequence of actions represents the process of agent behaviour, the totality of which represents the course of the simulation and thus the model dynamics. Let us consider the same simulation run, but this time from a macro perspective. Figure 2 represents the number of rules that fired per tick for each type of agent (tick 0 to 20). In total 19983 rules fired during these 20 ticks.

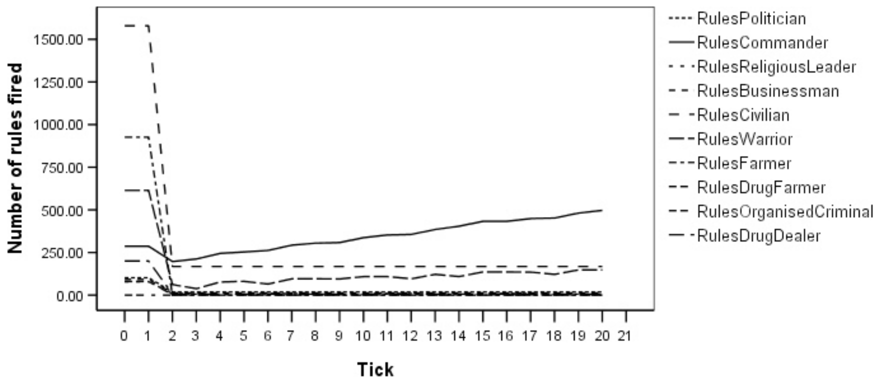


Fig. 2. Number of rules that fired per tick for each type of agent.

In general, figure 2 indicates that the agents, and thus the model, are in constant flux. Between tick 0 and 2 a sharp decline in the number of rules that fire is recognisable. This is not only due to a high number of rules that fire during the model setup in tick 0, but it also mirrors an organisational process. Agents intensively endorse each other and select with whom they want to interact in an initial organisational phase.

After the decline comes to an end in tick 2, two interesting developments occur. On one hand, commanders and warriors start getting very active, while on the other hand the rest of the agents remain in a constant state of action. The frequency of this constant activity is between 18 fired rules per tick for politicians and 0 fired rules per tick for farmers. This result appears to be sensitive to the initial agent distribution and needs to be further explored in upcoming model versions.

Per contra, increasing activity among commanders and warriors indicates, from an agency point of view, the establishment of a dense network among these two types of interacting agents and a cooperation takeoff. This might be a consequence of the model’s initialisation as well as of the number of agents in the model. Naturally, the likelihood of mutual interaction increases with an increasing population density. The effects of both, model setup and population density on mutual cooperation must therefore be analysed in more detail. We expect that other types of agents, for example politicians or religious leaders with civilians, experience similar takeoffs.

4.3 *Qawms* as neo-patrimonial networks

After having discussed the model output in relation to individual and aggregated agency, let us turn now to structural output, i.e. networks. Again, we analyse the same simulation run. Figure 3 depicts a simple network visualisation for our model output. Agents affiliated with each other are linked via a line.

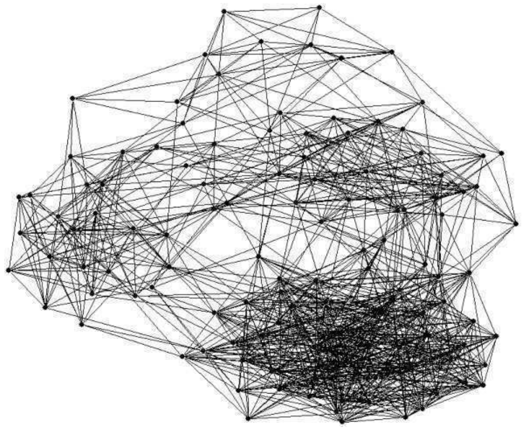


Fig. 3. Network of agents exhibiting three distinct clusters.

Three distinct but nevertheless interconnected clusters of agents are visible in the network. Each cluster consists of a variety of agent types. Agents assumed to be more powerful then others, i.e. politicians, commanders, religious leaders and organised criminals, are prevalent in the two more dense clusters on the right hand side of figure 3. The third cluster on the left consists only of civilians, farmers and warriors. However, there is one exception to this finding, namely the lonely commander number 1 discussed in section 4.1. He can be found rather isolated on the top of the network not embedded in one of the three clusters. It is possible that real centres of power emerge in highly populated areas exhibiting a large variety of agent roles.

The reasons for the evolution of this network of clustered affiliations are manifold: agents affiliate because they share the same ethnicity or religion or because they have established a business relationship or because they seek protection with a commander. But in general, the clusters can be perceived as emergent properties of agent neo-patrimonial behaviour as reified by our agent rules. The model generates data of the sort we expected (see section 2.2) and its output can therefore be considered as representations of *qawms*.

5 Conclusions

This paper is about growing *qawms* computationally. It includes meta-theoretical, contextual and social modelling considerations which have been exemplified by reasoning about critical realism as meta-theory for social simulation, by the introduction of a case-study informed intuitive model of *qawms* and by a conceptual as well as implementational discussion of an evidence-based and declarative social simulation model and its output. We consider the proposed proceeding to be a feasible and promising general research design for evidence-based social simulation.

Because we cannot provide stringent cross-validational results at this stage of research, our conclusions are preliminary. Nonetheless, section four reports interesting results which indicate processual and structural homology with the target system. In particular, we find neo-patrimonial behaviour is a highly socialised strategy to gain power. Isolated actors are doomed to fail – even if they have the personal potential of becoming powerful. Moreover, neo-patrimonial behaviour requires constant activity. Actors must be continuously aware of potential chances to affiliate with other actors for their own good. This can result in a mutual takeoff-process. Both of these features of the model output are found on the ground in Afghanistan where a strongman is always embedded and where there are constant changes of alliances in hope of positional gains.

The sequence of events referred to in section 4.2 produced a dynamic pattern that resembles the dynamics prevalent in conflict-torn societies, where all-out war is the exception, but constant noise and hard to forecast volatility is the rule [10]. This applies to Afghanistan as well.

Finally, we succeeded in simulating an artificial social structure that resembles *qawms*. Mutual interaction among different types of actors leads to the emergence of a complex organisational structure which contains a number of centres of power which themselves consist of a number of political, economical and military stakeholders cooperating in a limited way. Fuchs [9] reports similar structures in a network analysis of Afghan regional leaders based on qualitative data.

The model presented suggests possible insights into the complexities and evolution of Afghan power structures. It also suggests directions of further qualitative, ethnographic research regarding individual actors, such as spatial movement and conflict behaviour. Finally, the emergent network must undergo an in-depth social network analysis.

Acknowledgements. We would like to thank Sayyed Askar Mousavi for advice in the data collection process, Zemaray Hakimi for his translation and faciliator skills and two anonymous reviewers for their helpful comments. We gratefully acknowledge Shah Jamal Alam's, Ruth Meyer's and Bogdan Werth's modelling support.

References

1. Azoy, Whitney G.: Buzkashi. Game and Power in Afghanistan. 2nd edn. Waveland Press, Long Grove (2003)
2. Bak, P.: How Nature Works: The Science of Self Organized Criticality. Oxford University Press, Oxford (1997)
3. Bayart, J.-F., Ellis, S., Hibou, B.: The Criminalization of the State in Africa. Indiana University Press, Bloomington Indianapolis (1999)
4. Boero, R., Squazzoni, F.: Does Empirical Embeddedness Matter? Methodological Issues on Agent-Based Models for Analytical Social Science. J. of Artificial Societies and Social Simulation. 4 (2005) http://jasss.soc.surrey.ac.uk/8/4/6.html

5. Cederman, L.-E.: Modeling the Size of Wars: From Billiard Balls to Sandpiles. Amer. Pol. Sc. Rev. 1 (2003) 135–150
6. Cohen, P.R.: Heuristic Reasoning About Uncertainty: An Artificial Intelligence Approach. Pitman Advanced Publishing Program, Boston London Melbourne (1985)
7. Durkheim, E.: Suicide. A Study in Sociology. The Free Press, New York [1897] (1951)
8. Epstein, J.: Modeling Civil Violence: An Agent-Based Computational Approach. Proceedings of the U.S. National Academy of Sciences. 3 (2002) 7243–7250
9. Fuchs, C.: Machtverhältnisse in Afghanistan. Netzwerkanalyse des Beziehungssystems regionaler Führer [Power Structures in Afghanistan. A Network Analysis Among Regional Leaders.]. 1992–2004/05. Unpublished M.A. thesis. University of Zurich (2005)
10. Geller, A.: Macht, Ressourcen und Gewalt: Zur Komplexität zeitgenössischer Konflikte, Eine agenten-basierte Modellierung [Power, Resources, and Violence: The Complexity of Contemporary Conflicts, An Agent-based Model]. vdf, Zurich (2006)
11. Marks, S.R.: Durkheim's Theory of Anomie. Amer. J. Soc. 2 (1974) 329–363
12. Médard, J.-F.: L'État patrimonialisé [The Patrimonialised State]. Politique Africaine. 39 (1990) 25–36
13. Merton, R.K.: Social Structure and Anomie. Amer. Soc. Rev. 5 (1938) 672–682
14. Moss, S.: Game theory: Limitations and an Alternative. J. of Artificial Societies and Social Simulation. 2 (2001) http://jasss.soc.surrey.ac.uk/4/2/2.html
15. Moss, S.: Canonical Tasks, Environments and Models for Social Simulation. Computational and Mathematical Organization Theory. 3 (2000) 249–275
16. Moss, S.: Critical Incident Management: An empirically Derived Computational Model. J. of Artificial Societies and Social Simulation. 4 (1998) http://jasss.soc.surrey.ac.uk/1/4/1.html
17. Moss, S., Edmonds, B.: Sociology and Simulation: Statistical and Qualitative Cross-validation. Amer. J. Socio. 4 (2005) 1095–1131
18. Mousavi, S.A.: The Hazaras of Afghanistan, An Historical, Cultural, Economic and Political Study. St. Martin's Press, New York (1997)
19. Outhwaite, W.: New Philosophies of Social Science. Realism, Hermeneutics and Critical Theory. Macmillan Education, London (1987)
20. Popitz, H.: Phänomene der Macht [Phenomena of Power]. 2nd edn. Mohr Siebeck, Tübingen (1992)
21. Reno, W.: Warlord Politics and African States. Lynne Rienner Publishers, Boulder London (1998)
22. Richards, P.: Fighting for the Rain Forest. War, Youth and Resources in Sierra Leone. The International African Institute; James Currey, Oxford; Heinemann, Portsmouth (1996)
23. Roy, O.: Afghanistan: From Holy War to Civil War. The Darwin Press, Princeton (1995)
24. Roy, O.: The New Political Elite of Afghanistan. In: Weiner, M., Banuazizi, A. (eds.): The Politics of Social Transformation in Afghanistan, Iran and Pakistan. Syracuse University Press, Syracuse (1994) 72–100
25. Roy, O.: Islam and Resistance in Afghanistan. 2nd edn. Cambridge University Press, Cambridge (1990)
26. Sayer, A.: Realism and Social Science. Sage, London Thousand Oaks New Delhi (2000)
27. Schetter, C.: Ethnoscapes, National Territorialisation, and the Afghan War. Geopolitics. 1 (2005) 50–75
28. Sofsky, W.: Zeiten des Schreckens. Amok, Terror, Krieg [Times of Terror. Amok, Terror, War.]. S. Fischer, Frankfurt a.M. (2002)
29. Tarzi, S.M.: Afghanistan in 1992: A Hobbesian State of Nature. Asian Survey. 2 (1993) 165–174
30. Weber, M.:. Wirtschaft und Gesellschaft. Grundriss der verstehenden Soziologie [Economy and Society: An Outline of Interpretive Sociology]. 5th edn. J.C.B. Mohr (Paul Siebeck), Tübingen [1921] (1980)

Prior knowledge vs. constructed knowledge: what impact on learning?

Widad Guechtouli[1]

[1] GREQAM, Paul Cézanne University, 15/19 Allée Claude Forbin, 13627, Aix-en-Provence, France
Widad.Guechtouli@univ-cezanne.fr

Abstract. The aim of this paper is to model the process of learning within a social network and to compare the level of learning in two different situations: one situation where individuals know others competencies as a given data and interact on this basis; and one situation where individuals know nothing about others competencies but rather build this knowledge over time, according to their past interactions. For this purpose, we build an agent-based model and model these two scenarios of simulations. Results are partly analyzed using network analysis and show that in the second type of simulations agents are able to identify the most competent agents in the network and increase their competencies. Results also show that learning is easier when there is no prior knowledge of others' competencies. Otherwise, agents deal with a congestion effect that slows down the learning process.

Keywords: Learning, knowledge, network, agent-based model

1 Introduction:

Knowledge became very important for organizations since knowledge-based economy appeared. Nowadays, it represents a crucial asset that every organization should take care of, just like every other asset and yet, in a quite different way [1]. Knowledge is intangible and therefore is not easy to capitalize within an organization, or share between a set of individuals. Knowledge is acquired through a learning process where we can distinguish two levels: individual learning and social learning [2], [3], and [4].

First, let us start by giving a brief definition of individual learning. We choose to follow Dibiaggio's definition [5] who defines this process as a means to reach a goal, solve a problem or answer a question. He considers that it is related to the difference between the knowledge one already has, and the necessary knowledge required to answer a question. In this paper, we will consider that individual learning occurs when an individual increases his/her competency. Salomon and Perkins [6] state that, in reality, it is not possible to consider individual learning as an isolated process, as it is always related to a social context with social norms and influences. This leads us to a rather social perspective of learning. Zimmermann [4] gives the following definition: « Social learning corresponds to a situation where agents or individuals are able to modify their behavior, state, opinion or other factor, on the basis of information derived from the observation of their neighbors (Bala and Goyal, 1998) or more generally from the observation of these agents' behavior and performances». Learning is thus a socially constructed phenomenon [8].

The aim of this paper is to compare learning in two different situations: one situation where individuals know others competencies as a given data and interact on this basis; and one situation where individuals know nothing about others competencies but rather build this knowledge over time, according to their past interactions. For this purpose, we will build an agent-based model and model these two scenarios of simulations:

- Simulations with prior knowledge: simulations where agents know each-other competencies;
- Simulations without prior knowledge: simulations where agents acquire this knowledge through their interactions.

2 Description of the model

2.1 The agents

We have a population composed of 110 agents. Each agent is characterized by the following features:

- Knowledge vector: this vector is composed of 100 knowledge concerning 100 different subjects.

Subjects →	1	2	…	99	100
Knowledge vector →	1	0	…	1	1

Fig.1. Example 1: an agent has the knowledge concerning subjects 1, 99 and 100 but knows nothing about subject 2.

- Competency: defined as the number of subjects an agent knows about.
- Memory: where an agent stocks information about past interactions (name of agents previously met and the answer given by each one of them).
- Availability: defined by the number of questions that an agent is allowed to answer per time-step.
- Tolerance threshold: defined as the number of unanswered questions that an agent is willing to accept from another agent, before deciding not to ask him anymore.

According to these features, every agent is potentially a knowledge-seeker or a knowledge-provider, or both, according to his competency:

In terms of answering questions and providing knowledge, we consider that the population of agents is divided in two parts: priority knowledge-providers (*pkp*) and secondary knowledge-providers (*skp*). The members of the former have knowledge about the 100 subjects of an agent's knowledge vector; they have a competency equal to 100. Whereas the latter members have a competency equal to or higher than a competency threshold (CompMin) defined as the minimal competency required in order to have the ability for answering questions. This threshold is equal to 75[1].

In terms of asking for knowledge, each agent that has a competency smaller than 100 is a knowledge-seeker. This includes *skp* as well.

The initial structure of the population is the following:

- 1 agent with an initial competency equal to 100;
- 9 agents with an initial competency equal to 75;
- 100 agents with an initial competency equal to 0.

2.2 Interacting

- An interaction is defined by a question asked by agent *a* to agent *b*, and answer given by agent *b* to agent *a*.
- Each agent can only ask one question per time-step.
- An agent asks a question about a subject it knows nothing about.
- An agent answers a question if it has the specific knowledge asked for and if it is available; otherwise, it will ignore the question.

2.3 Learning process

Each time an agent gets an answer to a question; it raises its knowledge of that particular subject to 1, and won't ask questions about this subject anymore. Following example 1, an agent increases her knowledge of subject 2, as shown in fig. 2.

	1	2	…	99	100
Subjects →	1	2	…	99	100
Knowledge vector →	1	1	…	1	1

Fig.2. An agent learns and acquires knowledge about subject 2

2.4 Choosing a knowledge-provider

- Simulations with prior knowledge: agents choose the most competent agent in the population.
- Simulations without prior knowledge: to choose a knowledge-provider, an agent will base its decision on the performance of each knowledge provider

[1] We led simulations for several values for CompMin and 75 is the value where the highest numbers of agents increase their individual competencies.

towards it. The performance of agent *j* towards agent *i* at time-step *t+1* ($perf_{ij}^{t+1}$) is calculated as follows:

$$perf_{ij}^{t+1} = \alpha \ perf_{ij}^{t} + (1-\alpha)\, nbAnswers_{ij}^{t+1} \qquad (1)$$

Where: $\alpha = 0.2$[2] and: $nbAnswers_{ij}^{t+1}$: Number of answers given by agent *j* to agent *i* at time-step *t+1*.

At the first time-step, knowledge-providers are selected randomly. From the second time-step on, each knowledge-seeker will select the agent with the highest performance towards it.

An agent leaves the community once it has no more agents to ask questions to. We run simulations until no more questions are asked.

2.5 Parameters of simulation

- Availability: we will make this parameter vary between 1 and 10 questions per time-step.
- Tolerance threshold: will also vary between 1 and 10 unanswered questions per agent.

2.6 Indicators

We will observe the following:

- Number of priority knowledge-providers: this indicator will show how many agents were able to raise their competencies to the highest level (100).
- Mean learning of agents who left the community: measured by the mean competency of these agents at the end of the simulations. This indicator will let us know the level of learning reached by some agents, before leaving the community.

With the several variations of the simulation parameters, we have 100 different scenarios of simulation. We run each scenario 30 times, and the results presented are the mean results of the 30 iterations of each scenario.

3 Results of simulations

3.1 Simulations with prior knowledge:

[2] We led simulations for values of α between 0 and 1 and α = 0.2 is the value where the highest number of agents increases their competencies.

Number of Priority Knowledge-Providers. The following figure shows the number of *pkp* at the end of the simulations. We can see that this number reaches 110 agents (i.e. all the agents of the network) as soon as the smallest value of agents' availability and for a tolerance threshold equal to or higher than 7.

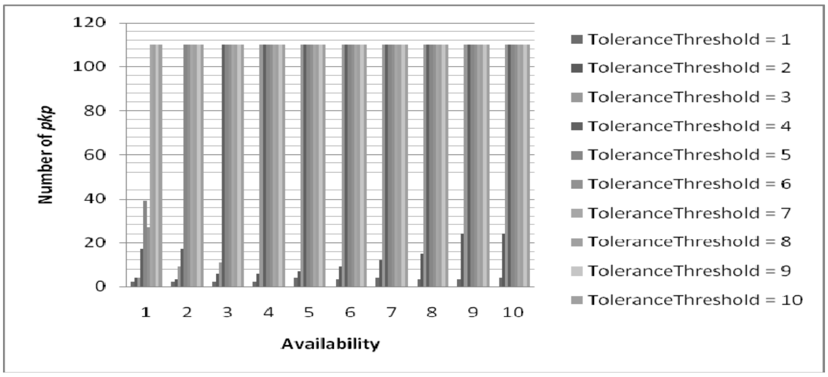


Fig. 1. Number of *pkp* at the end of simulations with prior knowledge

The same scenario is observed for other values of availability. All agents become *pkp* when certain equilibrium is reaches between values of knowledge-providers' availability and knowledge-seekers' tolerance threshold. These values are:

Table 2 Minimal tolerance threshold required for each value of availability in order to have a maximal number of *pkp*

Availability	1	2	3	4	5	6	7	8	9	10
Tolerance threshold	7	5	4	2	2	2	2	2	2	2

For all values smaller than the values presented in table 2, the number of *pkp* at the end of the simulations is quite small. Thus, the learning process is poorer.

Mean Learning of Agents who Left the Community. With this indicator, we can have an idea of the number of agents that could increase their competencies to a level equal to or higher than CompMin (75). Thus, we are able to know how much did the knowledge-seekers (i.e. agent that had an initial competency equal to 0) learn before leaving the community. Results are shown in fig.2

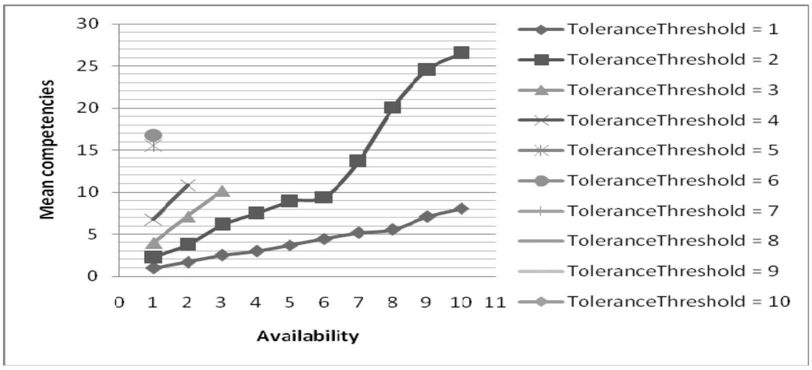


Fig. 2. Mean competencies for knowledge-seekers that left the community in simulations with prior knowlegde

This figure shows the mean competencies of leaving agents for several values of availability. This competency is not observable for some values of availability and tolerance threshold. These values match the values of these two parameters where all of the agents were able to become *pkp*. From this figure, we can see that, before leaving the community, knowledge-providers did not learn that much. The maximal mean competency barely reaches 26.

To sum up the results presented above, it is obvious that the learning process depends on agents' availability and tolerance. But it is likely that it depends more on knowledge-providers' availability than on knowledge-seekers' tolerance. This could be explained by the fact that, as all agents know each-others' competencies and always ask the most competent agents first, there is a congestion effect. Thus, access to knowledge is not possible for every agent, and therefore, some of them leave the community with a quite low level of competencies. Agents' tolerance should help for better learning, but this parameter remains at a secondary level, whereas the most important one is knowledge-providers' availability. Let us see now the results obtained by simulations without prior knowledge about others' competencies.

3.2 Simulations where agents know nothing about others' competencies:

Number of Priority Knowledge-Providers. Next figure shows the number of *pkp* at the end of simulations without prior knowledge about others' competencies. Here, for all values of tolerance threshold, all agents are able to become priority knowledge-providers as soon as knowledge-providers' availability is equal to 4. Below this value, the number of *pkp* at the end of the simulations is quite low. However, it remains higher than the number of *pkp* observed at the end of the previous set of simulations.

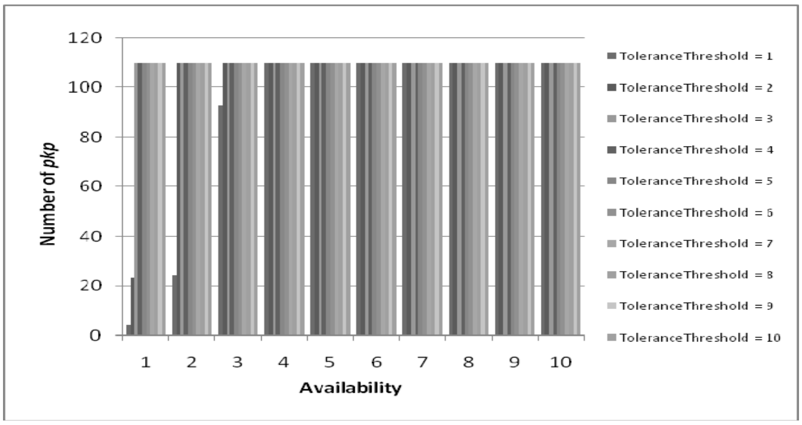


Fig.3. Number of *pkp* at the end of simulations without prior knowledge

Mean Learning of Agents who Left the Community. Figure 4 shows the competencies reached on average by initial knowledge-seekers before leaving the community. We can see that the level of mean competencies appears for points that correspond to: (availability = 1, tolerance threshold = 1), (availability = 1, tolerance threshold = 2), (availability = 2, tolerance threshold = 1) and (availability = 3, tolerance threshold = 1). These are the only combinations of availability and tolerance values where not all the agents were able to become *pkp*. Nevertheless, the level of competencies reached on average is quite high if compared with simulations where agents had prior knowledge about each-others' competencies. It reaches 65 for (availability = 3, tolerance threshold = 1).

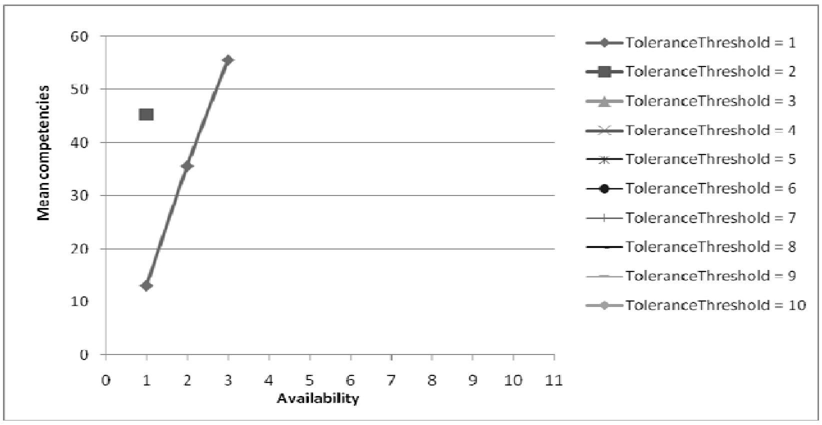


Fig.4. Mean competencies for knowledge-seekers that left the community in simulations without prior knowlegde

From figure 4, it is even clearer than in simulations with prior knowledge, that knowledge-providers' availability has more influence on knowledge-seekers' learning, than knowledge-seekers' tolerance. And we can see from both figures 3 and 4 that agents learn more when they don't have prior knowledge about each-others' competencies, they if they did. In such a situation, the congestion effect observed in the first set of simulations tends to disappear. Thus, agents have an easier access to knowledge.

4 Results discussion:

4.1 The congestion effect:

Results from simulations above stress that agents increase more their competencies when they don't have prior knowledge about each-others' competencies, than when they do. This can be explained by the strong congestion effect observed in the first set of simulations. In the summary of the first set of results, we suggested that agents' tolerance should relax the congestion effect. Let us see if that is true. To do so, we use network analysis.

We choose to use Freeman's degree centrality measures on the adjacency matrix derived from simulations where the congestion effect is the strongest, i.e. where knowledge-providers' availability is equal to 1. We will compare results when knowledge-seekers' tolerance is equal to 1 and when it is equal to 10 and see whether the congestion effect is weaker or not. Results are as follows:

Table 3. In-degree centrality when availability = 1 and tolerance threshold = 1

	In-Degree
Mean	12.327
Standard deviation	37.951

From the table above, we can see that the in-degree centrality (which represents the number of questions received by a knowledge-provider) is quite unequally spread among the knowledge-providers in the community. In fact, at the end of the simulations, the community is composed of the 10 initial knowledge-providers and the only one knowledge-seeker that was able to become a *pkp.* These 11 agents get between 109 and 208 questions, whereas the 99 other agents (knowledge-seekers) left the community without being able to increase their competencies enough to answer questions. This indicates that almost all initial knowledge-seekers asked the same set of agents, causing a congestion effect and making access to knowledge not easy.

The in-degree network centralization gives a description of the population as a whole. It expresses "the degree of inequality or variance in our network as a percentage of that of a perfect star network of the same size" [8]. Here, it is quite low: 1.812 % and we will now compare it with results of simulations without prior knowledge.

Table 4. In-degree centrality when availability = 1 and tolerance threshold = 10

	In-Degree
Mean	125.336
Standard deviation	358.361

In this set of simulations, the mean number of questions asked is larger than in the first case. It reaches around 121 questions whereas it was equal to 12 when agents' tolerance threshold was equal to 1. This makes sense because agents stay longer in the community and ask more questions as they are more tolerant towards knowledge-providers. However, we can see that in-degree centrality is even more unequally spread among knowledge-providers, the standard deviation reaches around 358. This means that most agents still ask the same ones, and this can be verified by looking at the in-degree centrality for each agent. Again, at the end of the simulations, the population counts 11 knowledge-providers (the 10 initial ones and the only one knowledge-seeker that became a *pkp*). These 11 agents received from 1096 to 1198 questions throughout the simulations, whereas the rest of the agents (99 knowledge-seekers) didn't get any question because they didn't increase enough their competencies.

If we look at the in-degree network centralization, we will notice that increasing agents' tolerance increases this measure. The network in-degree centralization was equal to 1.812% when agents' tolerance was equal to 1, and it shifts to 9.111% when agents' centralization is equal to 10. This is quite interesting because one would expect that if agents were more tolerant, then they would learn more. In fact, according to our results, it seems that if agents were more tolerant, the congestion effect would get stronger as there are more agents waiting for an answer without necessarily having it at the end of the simulations.

4.2 The core of the population:

The core of the population is composed of the knowledge-providers with the highest in-degree centrality, i.e. agents that received the largest number of questions at the end of the simulations. We will use results from simulations without prior knowledge about others agents' competencies. We wish to compare these agents to the most competent agents in the community and see whether knowledge-seekers were able to identify knowledge-providers and ask them questions, or not.

Results from simulations without prior knowledge show that agents with the highest in-degree centrality are always the 10 initial knowledge-providers, and this is always the case, no matter the values of availability and tolerance. We can show this result in the next figures.

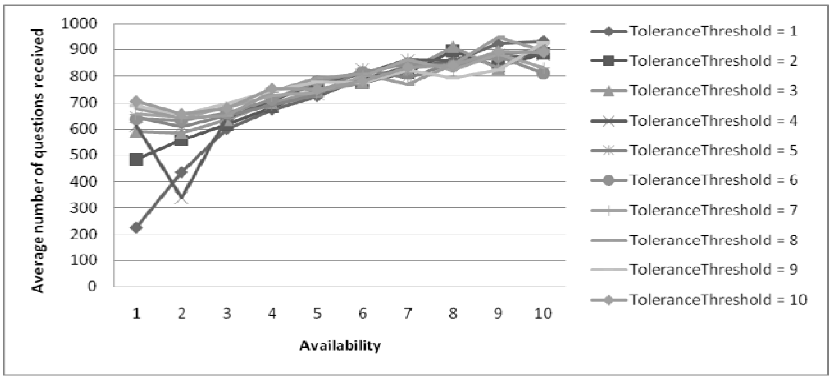


Fig.5. Mean number of questions received by an initial knowledge-provider

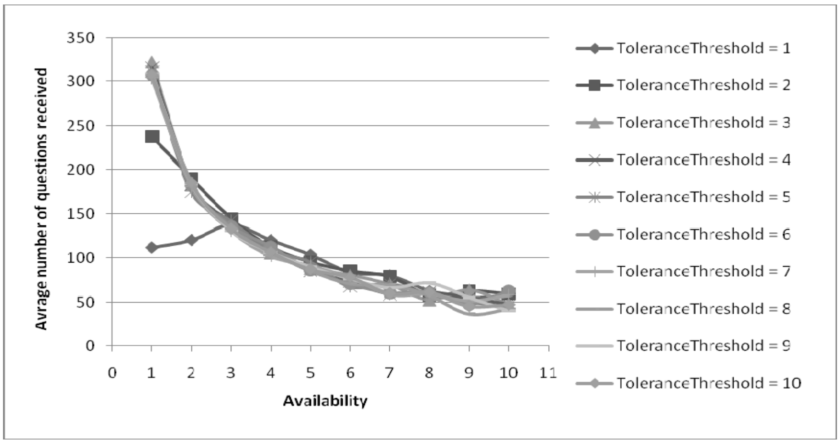


Fig.6. Mean number of questions received by an initial knowledge-seeker

From the two previous figures, we can see that the average number of questions received by an initial knowledge-providers is almost twice larger than the average nulber of questions received by an initial knowledge-seeker who managed to become a knowledge-provider, when the values of availability and tolerance are quite small. The gap bewteen these two average numbers gets even bigger as the knowledge-providers' availability increases. As agents are more available, knowledge-providers gets more and more questions, whereas initial knowledge-seekers don't.

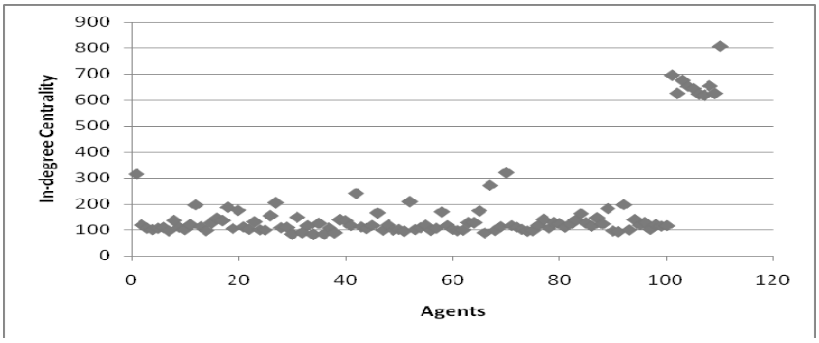


Fig. 7 In-degree centrality when availability = 4 and tolerance = 1

From figure 7, we can see that the core of the population is clearly visible. These agents are the most competent ones (as given by the initial structure of the population) and they also happen to be the ones with the highes in-degree centrality. This shows that knowledge-seekers were able to identify, through their inetractions, the most comepetent agents in the population. Learning is quite good (all agents were able to become *pkp* in most simualtions) given a certain balance between konwledge-providers' availability and knowledge-seekers' tolerance. This balance is summarized in the folowing table:

Table 5 Minimal tolerance threshold required for all values of availability in order to have a maximal number of *pkp*

Availability	1	2	3	4	5	6	7	8	9	10
Tolerance threshold	3	2	2	1	1	1	1	1	1	1

5 Conclusion:

Let us compare the equilibrium to be reached in simulations with prior knowledge and in simulations without prior knowledge, in order for all agents to become priority knowledge-providers. Table 6 shows this equilibrium in both types of simulations.

Table 6. Equilibrium required between availability and tolerance threshold in both types of simulations

	Availability	1	2	3	4	5	6	7	8	9	10
Tolerance threshold	Simulations with prior knowledge	7	5	4	2	2	2	2	2	2	2
	Simulations without prior knowledge	3	2	2	1	1	1	1	1	1	1

It seems that the equilibrium needed between these two parameters is easier to reach when agents have no knowledge about others' competencies and build this knowledge over time, through their interactions. In fact, in this type of simulations, all knowledge-seekers become *pkp* as soon as knowledge-providers' availability is equal to 1, and tolerance threshold equal to 7.

When agents have prior knowledge about others' competencies, knowledge-seekers need to have a tolerance threshold equal to 7, in order to be able to become *pkp*. This scenario of simulations can be thought of as a situation where agents interact through a network where there is an established hierarchy, and where agents have to follow some strict rules of interaction. These rules make them interact with a very specific set of individuals (in our model, agents had to interact with the most competent ones). This looks like a rather formal network.

The scenario of simulations without prior knowledge about other agents' competencies would rather represent an informal network. The most competent agents in the community are not explicitly nominated. Thus, there is no explicit hierarchy or strict interaction rules. Agents have to interact and learn with one another and eventually learn by increasing their individual competencies. According to this interpretation of the results presented above, it seems that the equilibrium needed between availability and tolerance is easier to reach in an informal network than in a formal one. Thus, learning in an informal network should be easier than in a formal one.

We can compare this result with the work of Cataldo, Carley and Argote [9] where the authors studied knowledge transfer in a social network with a hierarchical structure, and in a network with a fully connected structure. Results showed that the hierarchical structure was for more restrictive in terms of knowledge transfer, than the fully connected structure.

What we call here a formal network, given by the simulations with prior knowledge about other competencies, has a star structure, very hierarchical, with the *pkp* agents in the core of the graph and the rest of the agents situated at the periphery (cf. appendix A). Whereas what we call an informal network has a fully connected structure (cf. appendix B). The results of our simulations are consistent with the work of the previous others and show that the network structure has a strong effect on knowledge transfer. Besides, results show that, in terms of knowledge transfer and individual learning, it is better if agents have no prior knowledge about others' competencies and rather learn to know the other agents through their past interactions. This means that agents are able to coordinate their actions in a way and have access to knowledge without communicating or sharing information. To do so, they only need to consider their own experiences and do not need to communicate or exchange information about that.

This model presents a number of limitations. Future developments of the model should include a new parameter that would summarize the balance needed between agents' availability and tolerance and include both parameters. We did not address the eventual decrease of agents' knowledge. For instance, when an agent does not receive a question about a particular knowledge, it forgets it. This is something that we are currently working on and should appear in a forthcoming paper. In addition, one should think about new parameters to model the formal/informal feature of a network.

We could be able to know in a more precise way the role played by the nature of the network (formal or informal) in the learning process of the member of the network. And then, maybe we'll be able to confirm the results that these simulations are suggesting. Despite these limitations, this study provided an insight on how networks structure affects knowledge transfer among their members, and therefore the impact of such a feature on the learning process that takes place within networks.

References

1. Foray D.: L'économie de la connaissance, éditions La Découverte, (2000).
2. Bala V. and Goyal S.: Learning from Neighbours, Review of Economic Studies, 65, 595-621 (1998).
3. Leroy, F. : L'apprentissage organisationnel : revue de littérature critique". 7ème Conférence Internationale de Management Stratégique, Louvain (1998).
4. Zimmermann J.-B. : Social networks and economic dynamics, In: Bourgine P. and Nadal J.-P. (eds): Cognitive Economics: an interdisciplinary approach, pp 399-416, Springer (2004)
5. Dibiaggio L. : Information, connaissance et organisation, *Ph.D. Thesis*, Université de Nice (1998).
6. Salomon G. and Perkins D. N.: Individual and Social Aspects of Learning, Review of Research in Education, Volume ,23 ,P. David Pearson and Ali Iran-Nejad, editors (1998).
7. Cook S.D.N., Yanow D. : Culture and organizational learning, In: Cohen M.D. and Sproull L.S. (Eds), Organizational learning, Sage, 430-459 (1996).
8. Hanneman, R. A. and Riddle M.: Introduction to social network methods, Riverside, CA: University of California, Riverside (2005).
9. Cataldo, M., Carley K. M. and Argote L.: The effect of personnel selection schemes on knowledge transfer (2001).

Appendix A:

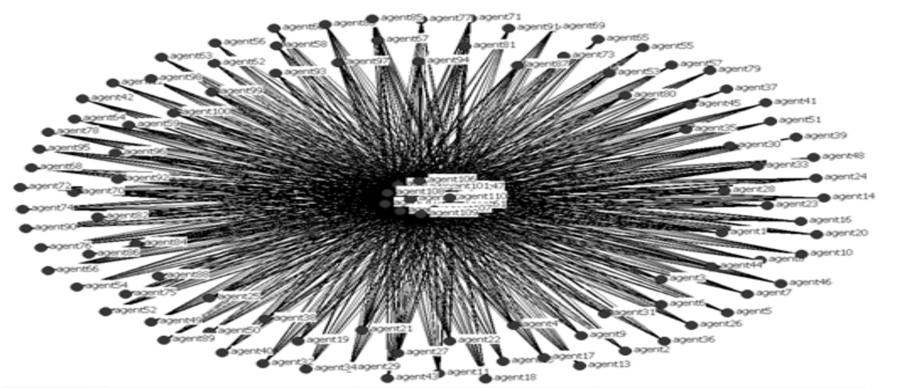


Fig. 8 Network structure in simulations with prior knowledge about others' competencies when agents' availability is 1 and agents' tolerance threshold is 10

Appendix B:

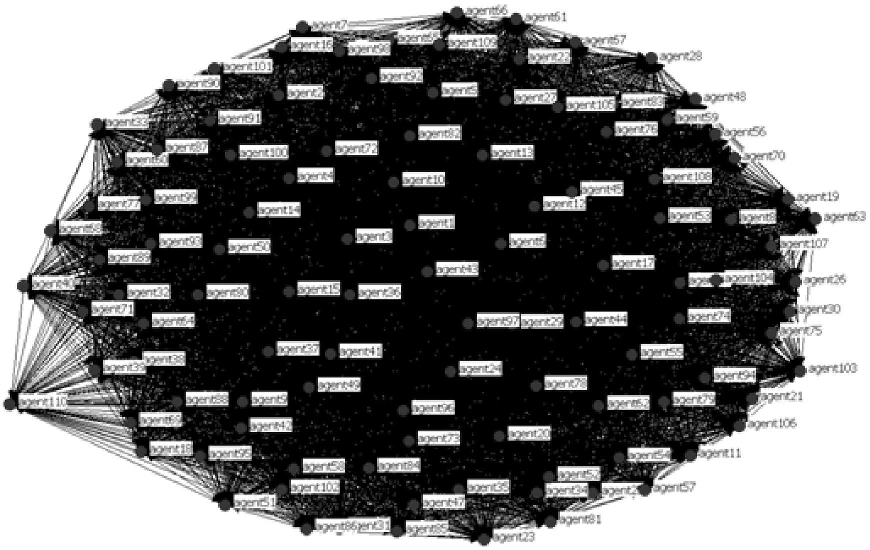


Fig. 9 Network structure in simulations without prior knowledge about others' competencies when agents' availability is 1 and agents' tolerance threshold is 10

Historical Simulation: A Study of Civil Service Examinations, Family Line, and Cultural Capital in China

Setsuya Kurahashi[1] and Takao Terano[2]

[1] Graduate School of Systems Management, University of Tsukuba
3-29-1 Otsuka, Bunkyo, Tokyo, Japan
kurahashi@gssm.otsuka.tsukuba.ac.jp

[2] Computational Intelligence and Systems Science, Tokyo Institute of Technology
4259 Nagatsuda-Cho, Midori-ku, Yokohama, Japan
terano@dis.titech.ac.jp

Abstract. In this study, we investigate what would happen in a Chinese historical family line. We analyzed a particular family line, which had so many successful candidates, who passed the very tough examinations of Chinese government officials over 500 years long. First, we studied the genealogical records 'Zokufu' in China. Second, based on the study, we implemented an agent-based model with the family line network as an adjacency matrix, the personal profile data as an attribution matrix. Third, using "inverse simulation" technique, we optimized the agent-based model in order to fit the simulation profiles to the real profile data. From the intensive experiments, we have found that both grandfather and mother have a profound impact within a family to 1) transmit cultural capital to children, and 2) maintain the system of the norm system of the family. We conclude that advanced agent-based models are able to contribute to discover new knowledge in the fields of historical sciences.

1 Introduction

It is more than 30 years since Pierre Bourdieu introduced the structure of reproduction in relation to cultural capital and education. He introduced the system of the norm (Habitus) within a family which reproduces cultural capital and plays a critical role in the selection of social stratification [1]. Furthermore, he referred to the civil service examination [3] which was used as the selection system for government officials in former days in China and also indicated the role played by cultural capital in the selection mechanism of examinations, when he pointed out the importance of examinations in the French education system [2].

However, in modern societies, the traditional concept of the family is forced to change in various areas because of the changes in the social system and local

[3] The term "Civil Service Examination" in the historical science field means the very tough examination for government officials of higher classes in China. The examination system runs for about 1,300 years.

society through the advance of globalization. Under these circumstances, it is getting more important to know the function of the family as the fundamental element of the social system [7].

There have been many changes in sociological approaches to family study. Now, there are approaches based on historical demography which interprets the time-series changes of family composition / population size, approaches based on network logic which borrows from the analysis method of the social network, and also traditional approaches based on the comparative institution, historical sociology and exchanges.

As mentioned above, the method of approach changes gradually to computable sociology in the field of family study. In this study, we construct an ABM based on the viewpoint of historical demography and the social network, and analyze the family system of a particular Chinese family line over a period of about 500 years. We then clarify the system of the norm which is maintained by the family, through a simulation of time-series changes of the attribution of family members, and then by inverse simulation.

2 Related work

Study of the family has been promoted from various angles: sociology, historical science, anthropology and biology. We begin with the sociology of the family.

2.1 Analysis of the Family System

The sociology of the family commences from a fundamental sociological concept and assumption. In other words, the sociology of the family consists of the concept of reward, benefit, and exchange, and the assumption that regards human being as a rational existence. The sociology of the family and its analysis is specified from the notion that the family is one of the social systems. This analysis is an approach to the family from the point of a system based on the fundamental concept that the system is the body which organizes and regulates the norms/ rules / customs of the human being,

Where the data is arranged in time series order, it is an historical comparison. Where the data is arranged by country or by area, it is an international comparison. "Clan, Caste and Club" by Hsu is regarded as a typical comparison analysis [10]. Hsu analyzes the influence of relationships within the family, based on the framework of the inheritance system as family typology.

2.2 Historical Sociology

Study of the traditional family shows clearly that the line maintains its continuity and formality through various strategies such as traditional practices within the family [6].

In the theory of family changes, change is examined from the point of the family structure and its scale, from the stem family system to the conjugal family system [14].

Historical demography has shown the change of population and the social conditions of the past through the empirical approach of drawing on parish records in Europe on historical trends of families/households. The footsteps of individuals over the past hundreds of years can be reconstructed from such data as the historical records of individual religious affiliation/ official family registry in Japan which recorded the details of names, ages, relationships to head of household, births, deaths, marriages and moving destinations of members of each house [8]. The data from genealogical records is similar to the above records.

2.3 Family Demography

It is common in the study of family demography that the major subjects of study are births, deaths, movement and marriages and analysis is done of family events throughout the cohort by year of birth [9].

2.4 Archeology and Cultural Anthropology

There are other branches of study which have many contact points with study of the family: (1)Archeology which looks at ancient monuments/ remains / relics and studies human cultures of the past, (2)Cultural Anthropology which compares such matters as ways of life, languages /common practice /ways of thinking and tries to find the rules which are common to humanity.

Especially in family anthropology looked at by, for example, E.Tod, there is research into family systems throughout the world and matters relating to cultures and social systems are considered anthropologically. Ethology finds the germination of paternity in the behaviors of chimpanzees and studies the creation of society and the organization of kindred of early peoples [16].

2.5 Theory of Cultural Capital

In "La Distinction" Pierre Bourdieu[1] defined cultural capital as the coming together of tangible and intangible property related to culture in the broad sense of the term. He classified it into the following three categories;

a) The variety of knowledge, accomplishment, expertise, liking and sensitivity that each individual has accumulated through his/her family circumstances and school education (physicalized cultural capital).
b) Cultural property which can be held in the form of materials such as books/ pictures/ tools/ machines (cultural capital which can be held).
c) Qualifications and titles from the education system and examinations (institutionalized cultural capital).

According to Pierre Bourdieu, education institutions do not monopolize the production of cultural capital, itself, although they have the ability to change inheritance cultural capital to educational status capital by monopolizing the issuance of certificates. This shows that such accomplishment may have an advantage for seemingly fair examinations of the implicit norms of the ruling class, and that culture and Habitus are actually invisible standards of selection. Habitus means the system of norms as the system of all tendencies, which produces a manner of action and sense, specific to certain classes / groups. This cultural capital is handed on to the next generation within the family.

2.6 Theory of Social Capital

The importance of social reliance is recognized by the study results of social capital and social system, described by Patnum [12][13]. The norm of reciprocation is maintained by imitation, socialization and forcing by sanction. Reciprocation is classified into the following two categories; "Balanced Reciprocation", which is to exchange reciprocally with a specified person. "Generalized Reciprocation", which benefits in a one-sided way and lacks balance at a particular time, but it is to be returned in the future. The norm of generalized reciprocation is a very productive component of social capital and is related to the close network of social exchanges. It can be said that family lines not only maintain within a family the system of the norm which reproduces the norm of reciprocation, but they also make use of marriage which implies the system of the norm between families that have produced many excellent descendents. This study intends to examine the above assumption.

2.7 Agent Approach

In addition to the above mentioned sociological approaches, studies of family and history have also been carried out by social simulation employing agent technology. The appearance of mating was reproduced in SugarScape, Epstein by the interactions between the birth rate of the agent and the population density [5]. The mating of agents was phrased as the network of family line and the relationship of the agents by blood was shown. Timothy A. Kohler and other scholars applied agent simulation to archeology in the Village Project and brought out the connections between change of vegetation, migration and change of population [11]. Furthermore, Cathy A.Small analyzed the influence of the marriage system in Polynesian society by employing a simulation model [15]. As mentioned above, use of the agent model is common in sociology and anthropology.

These studies have shed light on social structures and historical matters. Most of them are done from the analytical point of view. However, the agent approach focuses attention on the active aspect of history and has reproduced the change of matters by employing a simulation model. However, these simulation models are built based on historical facts which are already known and have been analyzed. This study tries to bring out undiscovered historical facts and structures by employing the above agent models and the inverse simulation method which estimates such acts and variables that may meet the historical facts.

3 History of Civil Service Examinations

In the past in China, there were examinations for the recruitment of government officials which were institutionalized as the civil service entrance examinations in the Tang Era. The golden age of the examinations was during the Sung Era, when politicians who passed the examinations displayed great abilities and reached the heights of politics. The examinations comprised a provincial examination, a metropolitan examination and a palace examination, with an entrance examination for school in each prefecture as a preliminary step. Successful candidates who passed the final step of the palace examination were called "chin-shih". As there was no way to make money other than passing the examinations over the hundreds of years from the start of the examination in the 6th or 7th centuries, many people tried hard to pass the examination. As a result, the competition heated up and the environment surrounding individuals became more important than the ability bestowed upon individuals in order to beat the competition. If an individual had the same abilities as others, then being rich became advantageous, it was better to be from an educated family and to live in an urban area with more advanced culture rather than being poor, from unlearned parents and living in the country. This phenomenon factored in the progress of the skewed distribution of culture and wealth. When European civilization surged into China at the ending of the Ching dynasty, it became more important to have an education in such things as natural science, experimentation and industrial art. The civil service examination system was finally brought to an end in 1904 by the Ching Government.

As mentioned above, it was a qualifying examination system for high-ranking officials which had been implemented for more than one thousand three hundred years. There were the following titles awarded as qualifications by the department of literature:

- chu-jen for those who passed the local examination.
- sheng-yuan for the status of student at a prefecture school, as a preliminary step.
- kung-sheng and chien-sheng for those who were recommended as central students.

According to Benjamin A. Elman [4], where the state emphasized the examination system for the production of loyal bureaucrats to the state, the candidates regarded the system as the most authoritative method for achieving their own personal success through successfully passing the said examinations. However, it took a huge investment in time, effort and training in order to achieve such success. Such candidates set their family, clan and lineage as the strategic targets of the social reproduction of their community. In the Ming and Ching eras, the school education system accepted only those candidates who already had a good command of the official language and were literate in classical Chinese. It was the responsibility of each house to obtain and maintain these elite positions as the "House of the Bureaucrat" at the initial stage of educating a son and preparing for his entering government service. Further more, Elman points out that it is possible for seemingly ordinary candidates to have achieved academic success because they had bureaucrats among their close relatives or affinities to the same lineage. Under these circumstances, the examination system proved successful in that it created elite families as areas of cultural reproduction, and it guaranteed that the right background gained superiority for a successful future social and political career and that the candidates came from a family which had the tradition of learning classics and spoke the official language.

However, there was a tendency that certain family lines could produce more successful candidates even from among these elite families, and there was a big difference between each family line. In China, records of family trees had been made from old times and kept as genealogical records: "Zokufu". Zokufu refers to records relating to family tree and lineage. It is a paternal record from the primogenitor and includes name, birth year, year of death, antemortem achievement, wife's name, number of children, place of residence and other information for each family member.

In this study, we used the Zokufu of the Y Family in the Ming and Ching Eras. The Zokufu mainly consists two parts: the "sekei" which shows generally the family tree, and the "sehyo" which records the details of the profile of each member. Each example is shown in Fig.1.

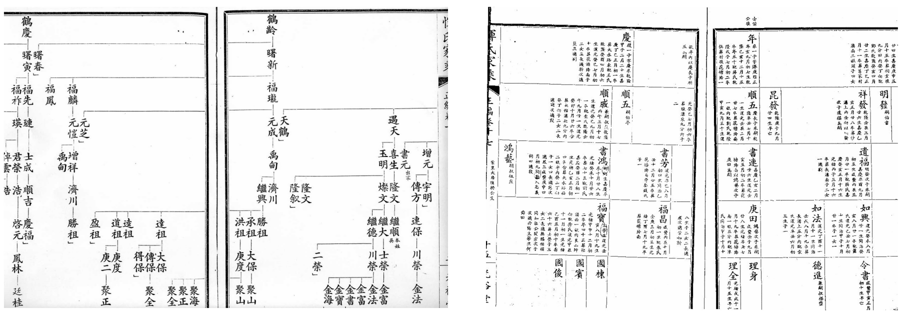


Fig. 1. Family Tree(left) and Personal Profile(right)

Changzhou, Jiangsu, the home of the Y Family, is located in the Gangnam region, which produced the highest number of successful examination candidates, who ranked 1st or 2nd throughout the country in the Ming and Ching eras. It was clear that most of these candidates were from certain families, and there were twenty seven families which kept producing chin-shih and chu-jen for the period of more than five generations during the Ming and Ching eras. Among these families, the Y Family was one of the typical cases, it produced twenty two successful candidates for the period of more than twelve generations. By analysis based on agent simulation, we began to know why so many such successful candidates were produced from the same family, by employing the Zokufu of the Y Family.

4 Agent Simulation of Zokufu

We prepared the types of data as available for the simulation from the "sekei" and "sehyo" Zokufu data. The sekei data show the relationship between father and son, and we were able to prepare the adjacency matrix from this data. The Zokufu of the Y Family contains data for a total of 1237 persons. It makes the adjacency matrix of 1237×1237 which shows the relationship between parent and child, and "0, 1" represents this relationship. In the same manner, we prepared the attribution matrix of each person from the sehyo data. The attribution involves chin-shih, chu-jen, kung-sheng, sheng-yuan, chien-sheng, chuan-na, merchant, painter, poet, qualified status of the examinations of wife's family home / daughter's married family and others. Each of these elements is represented by 0, 1 Fig.2. We can reproduce the family tree with attribution from the above two matrixes and implement the simulation based on this family tree. Each member of the family tree is grouped by birth year and tallied as successful candidate by cohort.

Outline of the Agent Simulation as follows:

- Each agent can transmit cultural capital from parent to child, from grandfather to grandson, from great-grandfather to great-grandson, by face to face along the family tree shown by the adjacency matrix.
- There are two categories of cultural capital: knowledge cultural capital and art cultural capital.
- Where there is a successful examination candidate on the mother's side of the family, his cultural capital is transmitted from parent to child in the same manner.
- Children have by birth the character of knowledge and art.
- The degree of a child's cultural capital depends on the synergetic effect of the character of the child and the cultural capital transmitted by others. However, only knowledge cultural capital affects success in the examinations and art cultural capital does not directly affect the rate of success in the examinations.

The agents can take the above mentioned actions. At the same time, they have parameters that decide each pattern of action. The parameters are in common with all the agents, which are as follows:

- Who is the transmitter? (father, grandfather, great-grandfather)
- Degree of effect on individual cultural capital (rate of transmission from father and others).
- Degree of effect by education (the increasing rate of the effects by cultural capital and by education of character).
- Mode of transmission of cultural capital (how knowledge cultural capital and art cultural capital are transmitted).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
2		0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
3			0	0	0	0	0	1	0	0	0	0	0	0	0	0
4				0	0	0	0	0	0	0	0	0	0	0	0	0
5					0	0	0	0	1	1	1	1	0	0	0	0
6						0	0	0	0	0	0	0	1	1	1	1
7							0	0	0	0	0	0	0	0	0	0
8								0	0	0	0	0	0	0	0	0
9									0	0	0	0	0	0	0	0
10										0	0	0	0	0	0	0
11											0	0	0	0	0	0
12												0	0	0	0	0
13													0	0	0	0
14														0	0	0
15															0	0
16																0

	Shins	Kyoj	Kous	Seii	Taig	Enou	Gunk	Riin	Saik	Sais	Dghts
1	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	1	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0
6	1	0	0	0	0	0	0	0	0	0	0
7	0	0	0	1	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	1	0	0	0	0	1	0	0
14	0	0	0	1	0	0	0	0	0	0	0
15	0	0	0	0	1	0	0	0	0	1	0
16	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	1	0	0	0	1	0	0
19	0	0	0	0	0	0	0	0	0	1	0
20	0	0	1	0	0	0	0	0	0	0	0

Fig. 2. Adjacency matrix(left), Attribution matrix(right)

– Degree of effect of the mother's side of the family (transmission rate of cultural Capital.

5 Inverse Simulation

The agent simulator model and code are shown in Fig.3, Fig.4. As shown in this chart, cultural capital is transmitted to children by the system of norms and the parameter which characterizes the norms that is transmitted along with the family line. The agent simulations are plurals implemented at the same time by way of the rule. The profile information of all the agents, which appears as a result, is compared with the actual profile information based on the attribution data prepared by the sehyo. These profile data are employed after tallying by cohort. The objective function sets the error of mean square of this simulator profile information and actual data profile information. The objective function is as follows:

$$min : CohortFitness = \sum_{i=1}^{n} \sum_{j=1}^{m} (c_{ij} - sc_{ij})^2, \tag{1}$$

where n : the number of cohort, m : the number of caltural capital, c_{ij} : caltural capital degree, sc_{ij} : simulated caltural capital degree. We select the better models by way of tournament with the value of the objective function by each generation and produce agent models that have the next generation parameter after the process of crossover and mutation. As a result, we can obtain an agent model that indicates results similar to the actual profile information. By analyzing the parameters of this model, we can estimate the strategies of the family lines which produced many successful examination candidates.

6 Experimental Results and Consideration

At the beginning, we implemented simulation experiments for which parameters were arbitrarily provided. However, we knew it was difficult to find a family

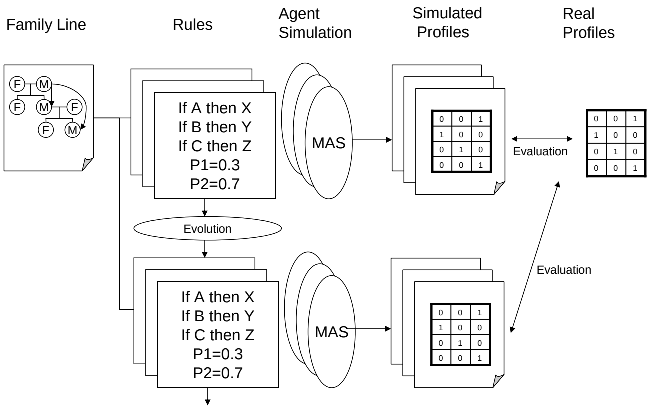


Fig. 3. Inverse simulation model

that could keep producing the high level of cultural capital to be able to have successful examination candidates. This is why there are too many parameters, variables, action patterns to be set for discovering suitable combinations. Therefore, as a second step, we implemented inverse simulations under the following conditions: Selection by tournament, crossover rate: 0.8, mutation rate: 0.05, number of models: 100, number of generations: 100. As the result of the experiments, we obtained the following parameters after a lapse of one hundred generations(Tabel. 1).

Table 1. The result of the experiments

Parameter	The result of Inverse Simulation
People who transmit cultural capital to child	Grandfather only
Effect by education : Effect by cultural capital	1 : 1
Effect by mother	100%, if maternal grandfather is a successful candidate.
Crossover rate between knowledge and art cultural capital	20%
Transmission method of cultural capital	Both cultural capitals of knowledge and art are transmitted to the child from parents.

From the above parameters, the following five patterns are prepared for people who transmit cultural capital to a child. Among them, only No.4 (from grandfather) is selected. This may show that the influence of the grandfather is

```
Inverse-Simulation(realData)
  set parameters and rules of each society to random
  for each society in the world
    Create-Society(parameters, rules)
  end for
  while generation < maxGeneration
    for each society in the world
      Simulate-Society(parameters, rules)
      fitness <- fitness-function(realData)
    end for
    SelectRecombinate-Society(fitness)
  end while
  optimumParameters-and-rules <-
            parameters and rules of society[maxFitness]
  return optimumParameters-and-rules
end
```

Fig. 4. Inverse simulation code

stronger than that of the father. 1) from father, 2) from father and grandfather, 3) from father, grandfather and great-grand father, 4) from grandfather only, 5) from great-grand father only. Based on these results, we implemented a statistical analysis by family line data and attribution data, and examined the effect of the data. The results are as follows:

- where grandfather is a successful examination candidate, father is a non-successful candidate and grandson is a successful candidate, and where grandfather is a non- successful candidate, father is a non-successful candidate and grandson is a successful candidate, the odds ratio is 3.4 between the above two cases. The effect on grandson is 99% screened and shows a significant difference where only grandfather is a successful candidate.
- where father on mother's side of the family is a successful examination candidate and son is a successful candidate, and where father on mother's side of the family is a non-successful candidate and son is a successful candidate, the odds ratio is14.2 between the above two cases. The effect on son is 99% screened and shows a significant difference where only father on mother's side of the family is a successful candidate.

It is contrary to what we expected that the result of the experiments shows that the effect on the child from the grandfather is stronger than from the father in transmission of cultural capital. Furthermore, we find that the transmission of cultural capital is made from the mother's side of the family to children of the married family. The result of such screenings has been verified by statistical analysis based on these findings of agent simulations. This shows that the doubling-up of three generations has had a big effect on education under the big family system in China. It shows that such a norm system can help to rebuild the family fortunes when there is marriage with girls from "good families". It is very interesting result that the effect on children by grandfather is stronger, than by father and that the effect by mother is important, considering the characteristics of the family line which kept producing the successful examination candidates. It may be possible that these facts are the custom which had been transmitted as the norms through generations. Further, the following are found as the characteristics of this particular family;

- Knowledge cultural capital and art cultural capital of parents are transmitted equally to child.
- 20% of each cultural capital affects the other.

The above matters are assumed by the family as the transmission method of cultural capital. It can be said that the following are some of the factors which resulted in keeping producing the successful examination candidates:

- Transmission of artistic ability is made to descendants, as well as of knowledge.
- It is more important that such education is provided within the family to make the most of this ability, if the child has strong artistic abilities, even if knowledge is inferior.

The result of transmission function of caltural capital is the following,

$$cl_k^c = r(cl_k^p \cdot pc_k^p) + (1-r)(cl_a^p \cdot pc_a^p), \tag{2}$$
$$cl_a^c = r(cl_a^p \cdot pc_a^p) + (1-r)(cl_k^p \cdot pc_k^p), \tag{3}$$

where cl_j^i : i's caltural capital about j, pc_j^i : i's personal characteristics about j, i : c=child, p=parent, j : k=knowledge, a=art, r : crossing rate of cultural capital. This may imply that there is a strong bond in the exchanges between artists and intellectuals, and in the relationships between brothers and sisters, which can be seen in the present society.

7 Summary and Challenges

In this study, by employing agent technology, we analyzed the family line for a family which produced many successful examination candidates for a period of five hundred years. We implemented an inverse simulation based on a multi-agent model which expresses the family line network and the personal profile data as an adjacency matrix and as an attribution matrix, respectively, and sets the real profile data as an objective function. As a result, we found that grandfather and mother had a strong effect on the child for the transmission of cultural capital within the family. This was verified by statistical analysis. With this model, we showed the possibilities that an agent based model could contribute to new discoveries of facts in the fields of historical science and sociology. We also showed the possibility of historical simulation by ABM.

As challenges for the future, we need to research the changes of family motto which can be obtained by inverse simulation through analyzing by generation and by branch of family, not using the method employed in this study, to all the family lines without exception. In addition, we plan to construct a model that can simulate the influence of the daughter's married family and of artists such as painters, and then to refine the models.

References

1. P. Bourdieu. *Distinction: A Social Critique of the Judgment of Taste*. Harvard University Press, 1987.
2. P. Bourdieu. *Reproduction in Education, Society and Culture*. Sage Pubns, 1990.
3. P. J. Carrington, J. Scott, and S. Wasserman. *Models and Methods in Social Network Analysis*. Cambridge University Press, 2005.
4. B. A. Elman. Political, social, and cultural reproduction via civil service examinations in late imperial china. *The Journal of Asian Studies*, 50(1):7–28, 1991.
5. J. M. Epstein and R. Axtell. *Growing Artificial Societies*. The MIT Press, 1996.
6. K. Fujii. *Historical Sociology of The Family and Kinship (in Japanese)*. Tosui Shobou, 1997.
7. V. Gayle, D. Berridge, and R. Davies. Young people's entry into higher education: quantifying influential factors. *Oxford Review of Education*, 28(1):5–20, 2002.
8. A. Hayami. *The world of Historical Demography (in Japanese)*. Iwanami Shoten, 1997.
9. K. Hiroshima. *Eurasian Project on Population and Family History Working Paper Series, 4(23)*. International Research Center for Japanese Studies, 1998.
10. F. L. K. Hsu. *Clan, Caste, and Club*. Van Nostrand, 1963.
11. T. A. Kohler, J. Kresl, C. Van, W. E. Carr, and R. H. Wilshusen. Be there then: A modeling approach to settlement determinants and spatial efficiency among late ancestral pueblo populations of the mesa verge region, u.s. southwest. In T. A. Kohler and G. J. Gumerman, editors, *Dynamics in Human and Primate Societies*, pages 145–178. Oxford University Press, 2000.
12. R. Patnum. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton University Press, 1993.
13. R. Patnum. *Bowling Alone: The Collapse and Revival of American Community*. Simon & Schuster, 2000.
14. H. Shimizu. The approaches of family change (in japanese). In H. Nonoyama and H. Watanabe, editors, *Beginning The Sociology of The Family*, pages 42–68. Bunka Shobou Hakubun Sha, 1999.
15. C. A. Small. The political impact of marriage in a virtual polynesian society. In T. A. Kohler and G. J. Gumerman, editors, *Dynamics in Human and Primate Societies*, pages 225–249. Oxford University Press, 2000.
16. E. Tod. *L'invention de l'Europe(in Japanese)*. Fujiwara Shoten, 1996.
17. S. Wasserman and K. Faust. *Social Network Analysis, Methods and Applications*. Cambridge University Press, 1994.

A Multi-Agent System to Model the Labor Market: Simulating a New Job Contract Introduction

Zach Lewkovicz and Jean-Daniel Kant

Computer Science Laboratory (LIP6), Université Paris 6, Paris, France
zach.lewkovicz@lip6.fr, jean-daniel.kant@lip6.fr

Abstract. The aim of this work is to design a Multi-Agent System (MAS) simulation to model the French labor market. We departed from an economic model proposed by Cahuc and Carcillo to model the introduction of a new job contract into the labor market. We designed a specific methodology to convert this equation-based model to an agent-based model, and calibrated our MAS to reproduce the data found in the economic simulations. As we observed the same tendencies found in the former one, a new dimension emerged from the agent-based simulation: an increase of oscillations for the characteristic rates, revealing an increase of precariousness (job instability) due to the new type of contract. These encouraging results lead us to pursue into that direction, where several extensions of our model can be proposed, including the move to a large-scale simulation framework.

1 Introduction

In this paper, we aim to build a computational economic model that imitates the conditions of the French labor market in order to study the possible changes in the market, due to the introduction of a new job contract. We intend to take advantage judiciously of the tools provided by MAS, i.e. systems made of artificial agents that are autonomous and in strong interaction with each other [1], in order to translate an economic model into an agent-based one. This approach is common in the Agent-based Computational Economics (ACE) community [2]. The ACE community captures economic changes and developments, and translates them into dynamic computational models. In these models, entities (agents) interact with each other and with the artificial environment. The meaning of computational entity is a collection of data and behavioral methods.

This transposition from a pure economic model to a multi-agent system was conceived in two phases. We started by identifying the main actors in the model. Each actor has his/her own main role, and his/her own behaviors (methods). These behaviors introduce the dynamics to the simulation. We have afterwards integrated the calculation mechanisms of the economic model, which consist the decision processes of the agents.

The implementation of the multi-agent mechanisms is sometimes straightforward, like when we encode the individual decision process, which helps the agents to choose their future action. However, some variables, which are computed by equations in the classical economic approach, can be directly set in the MAS simulation from the results of the agents' interactions. This is the case, for instance, of the unemployment rate: in the MAS simulation, depending on the decisions of workers and employers, jobs are created or destroyed, positions are filled or persons are fired, and the unemployment rate fluctuates accordingly during the simulation. Hence, one central issue when we design our MAS model is to decide whether a particular variable can be derived from simulation or had to be computed via equations.

In the long run, the goal of this work is to provide a useful and reliable tool to political deciders. A tool that will allow, through the agent-based simulation, to evaluate and predict the efficiency of particular economic policies for the labor market.

2 Model Features

2.1 The Economic Model

The economic model we adopted was proposed by Cahuc and Carcillo [3] in order to model the introduction of a new job contract (CNE) into the French labor market. Coming from a microeconomic approach of Labor Economics [4], they used several systems of equations to describe the evolution of productivity, unemployment rate, expected utility, decision-making, etc. There are three types of job contracts in the model: *CDD* (short and fixed duration), *CDI* (no duration limit), *CNE* (no duration limit but a trial period of 2 years). The model is divided into two parts: one before the introduction of the new contract and the second just after this introduction. For each part, an independent system of equations defines the thresholds for the various behaviors in the market.

In the first part of the simulation two types of contracts exist, namely CDD and CDI. CDD is a contract of limited duration. After two years this contract has to be either destroyed or transformed into a CDI contract. The CDI is a contract of unlimited duration. As the productivity of employees change over time, companies prefer to hire with the CDD contract, which is less binding. An upper limit is introduced to the possible number of hires with CDD per company (the case in the real French labor market).

A worker may loose his/her job in two cases. After two years of work with a CDD contract, the company assesses the productivity of its employee: if this productivity is insufficient, the contract is not transformed into a CDI contract and the person looses his/her job. In the second case, employees working with a CDI contract, are evaluated every month, and can be fired upon each evaluation. However the company must justify this decision. The productivity thresholds for each type of contract are computed separately.

In the second part of the simulation the CDD contract is replaced with the CNE contract. During the first two years, the company is allowed to fire an employee at the end of every month upon evaluation, without having to justify its decision. In addition to these evaluations, a general evaluation is made at the end of two years: if the contract is not destroyed, it is transformed into a CDI contract. Hires are made only with CNE contract.

The model defines several thresholds, which take part in the decisional mechanisms in the simulation. Employers decide whether to fire employees or not, concerning their productivity, and also decide whether to open vacancies or not. Employees decide whether to look for a job or not to participate in the labor market.

2.2 Main MAS features

When using MAS, one intends to study a specific problem and therefore chooses a particular architecture for the model. Nevertheless, certain characteristics are common to all MAS [1]. Let us briefly review these main features while referring also to our model:

The Agents All existing ACE models of the Labor Market use Worker agents and Employer agents. Some introduce also a Government agent [5]. The number of agents in the simulation can be an important factor. In [6] we find 100 agents and in [5] we find 200, while [7] uses 24 agents: 12 workers and 12 employers. Many researchers want to have the possibility of obtaining "zero unemployment" ([5], [7]), which – in their models – implies to have the same number of workers and employers.
Our simulation allows setting the number of agents freely, just before it is launched. There is no obligation of having the same number of worker agents and employer agents. Moreover, the user can choose the maximal number of jobs that a single company (employer agent) will be able to open. In our simulation we set the number of workers to 400.

Heterogeneity In the real labor market it is obvious that actors, let's say workers, are not homogeneous. One can consider psychological aspects of people, like the level of their education, their history in the labor market or their cognitive abilities. Heterogeneity of agents in ACE simulations is used to study specific aspects of the labor market. For instance, [7] introduces heterogeneity by using the memory (the history of interactions) of agents. In that way a persistent relationship between two agents can be studied. In other models, agents have various reservation wages [6] or skill endowments [5].
In our model agent's heterogeneity is introduced in several aspects: each agent has its own working-site history, productivity rate and "well-being" (the expected utility thresholds with which it reasons).

Goals of the agents We can often see that agents do not have an explicit goal. In ACE models the agents may have a reasoning process that allows them to maximize their profit, for example by choosing their next action according to a

profit matrix [7]. In [5], agents imitate a winning strategy without knowing what the meaning of the term ”winning strategy” is.
As in our model agents live and reason constantly, their permanent goal is to improve their situation, by choosing a higher expected utility state.

Representations of information The information in the simulations has two main characteristics.

- A *complete* information environment, where everybody knows or may know everything.
- A *perfect* information environment, where the information, which the participants have, is a 100% exact.

We can find different approaches dealing with information in the simulations. In [5] an assumption of complete and perfect information is made. To compute the investment rate of each agent in his own professional training, the average profits of all other agents is taken in account. This method may facilitate calculations and may also simplify implementation, but it is less realistic when talking about decision making in the real labor market. Incomplete and perfect information is used in [7], where the agents play a strategy game over a profit matrix of a ”prisoner's dilemma” type. As both agents choose their strategy simultaneously they ignore the adversary's strategy, the type of information is incomplete. Then again, as both agents know the matrix, and gain exactly what they should, the information is also perfect in this case. [8] uses incomplete information when agents do not know (or have access) to all the vacancies in the labor market. A searching mechanism for the agents is implemented in the simulation, but this mechanism is quite demanding (from a cognitive point of view), so that agents cannot discover all vacancies in the market. Uncertainty is introduced by [5] in form of probability. Shocks hit sectors in the labor market, which oblige employers to close down their companies. Neither companies nor employees can predict these shocks, in this case an imperfect information assumption is made.
In our model the nature of information varies according to the types of agents (see section 3 for more details on agents in our system). *Person* agents (i.e. job seekers or employees) are somewhere between complete and incomplete information. On one hand, they use the rate of unemployment to calculate their chances and their expected utility of finding a job, a process that uses complete information. This process is not completely unrealistic, as persons do read the newspaper and can get a general impression of the unemployment situation in the labor market. If a person sees that the unemployment rate is high, he/she may think that his/her chances of finding a job are too small and so, give up and leave the labor market. On the other hand, the *Person* agent cannot know nor estimate the duration of time required to be matched with a job offer: this incertitude is a clear case of incomplete information. This same kind of analysis is valid for *Company* agents (employers) as well.
As all reasoning processes in the simulation take place simultaneously for all agents, we can say that the information that agents have is imperfect. At the

moment an agent decides about his future (taking in account, let's say a certain unemployment rate), it is possible that many other agents have gone through the same process and have decided to change their state. So the agent has really a quite imperfect image of the situation in the labor market.

Interactions between agents The ACE community uses message sending to implement communication and interactions between agents, as it is often done in MAS. The protocols for this messages sending are not always defined explicitly in the simulations, although we can get a general idea about their structure from principles like "first applied, first hired" or "an agent can not apply to more than one job" [7]. Sometimes the order of message sending is crucial to the results obtained, and sometimes it has no importance, as in the Gale-Shapley algorithm [7].
We also use message sending to make the agents interact, like transmitting information to each other. However, in our model, the type of communication is closer to reality, as it takes place simultaneously and asynchronously: each agent makes its decisions independently of other agents and communicates information whenever it decides to do so.

The environment The most complex environment is *dynamic, stochastic, inaccessible, non-deterministic, non-markovian* and *continuous*, alas, this is the case of the labor market. To be able to deal with this complexity ACE researchers use various simplifying hypothesis. The environment in previous ACE experiments is usually *markovian, deterministic, static* and *discrete*.
In our simulation, the environment is *dynamic* – the unemployment rate can change every round, *non-markovian* – the current productivity of an agent does not depend on its previous productivities, *inaccessible* – agents cannot know the decisions of other agents concerning their future, *deterministic* – the actions of agents are executed fully and *discrete* – each agent has its own life cycle.

Everyday life in the company Everyday life in the company is usually treated implicitly in ACE models. From the moment when a person has been hired, his/her wage and his/her productivity stay constant over time. Exogenous events may occur in the labor market, like economic shocks [5] or technological progresses [8]. In [7] interactions take place explicitly in the company, where the employees and the employers constantly play a game of strategies. This game is played over a profit matrix by choosing a cooperation strategy or a defection strategy.

Our simulation does not include a bilateral relationship in the company. A worker performs with certain productivity and it is the employer who decides whether to keep or to fire this worker. The wage of the worker is constant over the time that he/she works, as his/her expected utility. The separation between a worker and an employer can be initiated only by the latter one that means that en employee cannot quit.

3 Our Model's Agentification

The agents participating in the model are:

- The *Person* agent, who produces while working and gets paid a salary. It can be employed, unemployed or not participating in the labor market.
- The *Company* agent, who hires and fires *Person* agents. It can also decide whether to open a job or to keep it closed.
- The *Matching* agent, who represents the environment of the labor market. It matches unemployed *Person* agents with vacant jobs.
- The *Government* agent, who sets economic policies in the labor market.

The *Person* agent (worker) The *Person* agent can be in one of *three states*:

worker it fills a job, produces and gets paid.
unemployed it has no job, but it is looking for one. In this case it gets unemployment benefits.
non-participant it does not fill a job neither does it look for one. In this case it gets social welfare allowances.

The *Person* agent has a capital which is the sum of money it earns by filling a job or the allowances it gets. It is characterized by the productivity with which it performs its job. This productivity represents the quantity of work and the investment of the agent in the job it fills [3]. This productivity takes its values randomly from a probability distribution every month. We can sum up the *Person* agent's behaviors in the state transition diagram depicted in Figure 1.

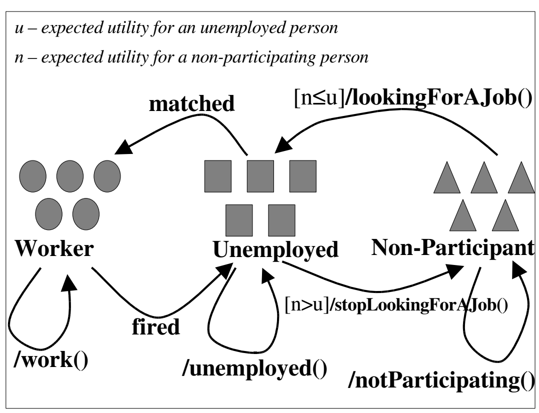


Fig. 1. States of the *Person* agent

The *Person* agent uses its decision mechanism in two situations: in the state of unemployment and in the state of non-participation. The transition between these two states is due to a reasoning process, where it decides which state will be more profitable economically. Thus, if the expected utility u for an unemployed person is greater than the expected utility n for a non participating person, the *Person* agent shifts from non-participant to unemployed state; otherwise, we have the opposite shift. In order to compute these expected utilities, the agent has to solve a system of four equations (a separate system exists for each model). For more details see [3].

The *Company* agent (employer) In the labor market the *Company* agents offer vacancies, decide whether to hire a person or not and eventually decide whether to keep or to fire this person. *Two states* exist for a job:

vacant the *Company* looks for a *Person* to fill the job.
filled the job is populated, and produces a profit.

A *Company* agent can decide not to look for a person to fill a vacancy if it thinks its chances of succeeding are too low, and that it will loose more money looking for an employee than keeping the job closed. If such decision is made after firing a worker, we will say that the job is destroyed. In this model a *Company* agent has a size, which defines the number of jobs it can offer.

The *Company* agent uses its decision mechanism in the beginning of each cycle and after each firing. If a firing takes place or vacancies that have not been published exist in the company, the agent decides whether to open these jobs or not. This decision is made according to a threshold C computed for the profitability of a job. The decision mechanism also intervenes after each productivity report, when evaluating the employee and deciding about his/her future in the company. Figure 2 summarizes the *Company* agent's behaviors.

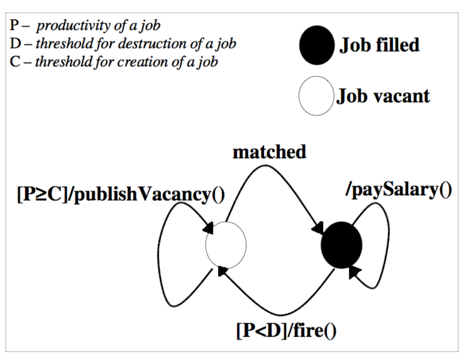


Fig. 2. States of the *Company* agent

The *Matching* agent In order to integrate the environment of the labor market, we introduce a *Matching* agent. The role of this agent is to manage the lists containing the states of the agents and the states of the jobs in the labor market. This agent also computes the matching rate $m(\theta)$ for each round that defines how many vacancies and unemployed persons will be coupled per unit of time. Following [3], we use the same matching function: $m(\theta) = m_0\theta^{-\eta}$, where m_0 and η are constants in [0,1] and θ is the tightness of the labor market.

The fact of having simultaneously unemployed persons and vacancies in the real labor market, could explain the necessity of having such a function. As we know, looking for a job consumes time and money. Furthermore it is possible that the offer and demand of different types of jobs do not coincide. We use the matching function to integrate this notion of delay and frictions.

The *Government* agent The main functionality of this agent is to set the economic policies in the labor market. The agent calibrates the model and is responsible for setting the policy wanted, for example which contracts will be available (CDD, CDI, CNE ...). This agent decides the amount of annual salaries for the different types of contracts, unemployment benefits, social welfare allowances etc. It also computes the benefits paid to fired workers.

4 Results

4.1 The simulation

The simulation takes place in two parts, each part corresponds to different types of contracts. In the beginning we can set the number of agents (*Persons*, *Companies*) and the size of *Companies* that will take part in the simulation. We can also modify the level of salaries in the model, the amount of the unemployment benefits and the level of social welfare allowances. The new thresholds for the decision mechanisms are computed automatically by the system. The simulation is then launched.

The agents are created and start to interact with the environment of the labor market and with each other. The *Person* agents decide constantly whether to look for a job or not to participate in the labor market, while the *Company* agents decide whether to open jobs or not. The user can observe the unemployment rate, the vacancy rate, the participation rate, the tightness of the labor market and the messages sent between the agents. The user then decides on what moment to introduce the new contract and the simulation adopts automatically the new model.

We implemented our system using the Cougaar Platform [9], because it is an open source agent-based architecture that emphasizes cognitive agents, groups and organizations. Moreover, it fully supports large-scale applications, and we plan in the future to design a large-scale multi-agent system for the labor market, a market that is made of several millions of agents in the real world. The peculiarity of Cougaar is that it is fully asynchronous; the agents are total autonomous entities and are difficult to synchronize. The drawback of this, when we perform a simulation and compute some rates, like unemployment, vacancy, etc., is that we get sometimes high peaks in the values, higher than what we find in reality, (Cf. Figures 3- 5 below) that are partly due to synchronization delays in the system. We plan to work on that issue in the very next step.

4.2 Observed tendencies

The results we found in our simulations resemble to the results found in the economic model: our MAS/ACE model reproduces the same average behaviors in the labor market. We find the same tendencies while looking at the rates of unemployment, participation etc.

For instance, if we take the evolution of the *unemployment rate* (percentage of unemployed persons within the labor force population) depicted in Figure 3, after the introduction of CNE contracts, there is a temporary drop in the unemployment rate, due to the brief rise in the vacancy rate (percentage of vacant jobs) and the decrease of the labor force. After this brief drop, the market's average unemployment rate stabilizes on a higher level than before, since it is easier to fire a person with CNE. Overall, with CNE, we have a transition from 7.1% of unemployment to 16% on average.

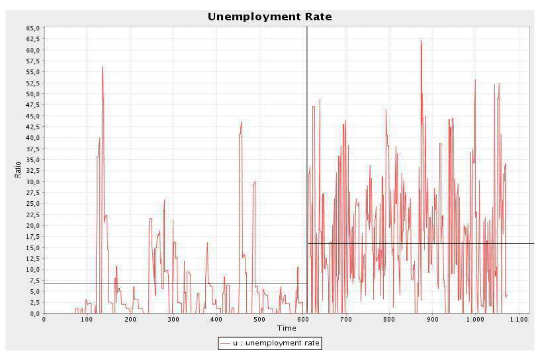


Fig. 3. Unemployment rate through time. Vertical bar shows the time when CNE contract is introduced, horizontal bar is the mean for each period.

Similarly, with the shift to CNE, the vacancy rate (percentage of vacant jobs) is increasing, as shown in Figure 4. The explanation to this phenomenon is given by Cahuc and Carcillo [3]. The companies decide to open new jobs, because their expected value is higher if filled than if vacant. Regarding individual persons, when fired, the expected utility tends to be higher in the non-participation then in the unemployment state, so they do not look for a new job and the number of vacancies stays high. In our simulations, we measured a shift from 21% to

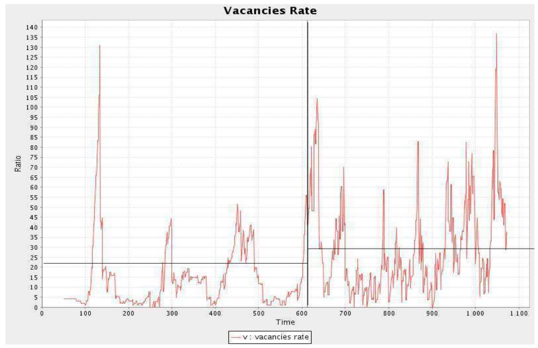


Fig. 4. Vacancy rate through time. Vertical bar shows the time when CNE contract is introduced, horizontal bar is the mean for each period.

29% in the vacancy average rate. Finally, as shown in the Figure 5, persons have a higher tendency to consider the non-participation state with higher expected utility than the unemployment state. More and more workers, who have just lost their jobs, choose to abandon the attempt of looking for a new job because they consider the market and the jobs in it to be extremely unstable. The fall of the working population rate drops from almost 90% to 64%, on average.

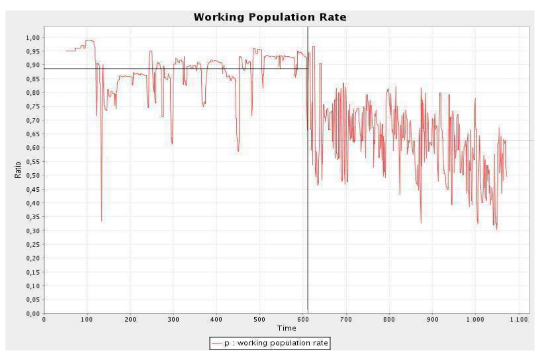


Fig. 5. Working population rate through time. Vertical bar shows the time when CNE contract is introduced, horizontal bar is the mean for each period.

4.3 The impact on the well-being

The novelty in our approach comes from the fact that MAS simulations add the possibility of taking into account more realistically the interactions between agents. Although in the economic model we can compute the unemployment rate at any given period, we cannot observe the interactions that led to this rate – our MAS simulation allows that. For instance, in Figures 3-5, we can see the increasing frequency and also the increasing deviations from the mean after the introduction of CNE contracts. This phenomenon clearly shows that the well-being of persons decreases: the frequency of these oscillations shows frequent shifts from work to unemployment or non-participation, and vice-versa. Thus it measures the precariousness (as job instability) of persons and the volatility of the labor market. All three rates (unemployment, vacancy, working population) show an increasing frequency of oscillations, after the introduction of the CNE contract.

5 Conclusion and future directions

In this study, we translated a pure economic model into a MAS simulation. We reproduced the same results as in the economic model and added a supplementary dimension to the results obtained. Our model allows exploiting better the tools provided by MAS. By using these tools intelligibly, one can achieve finer and richer results. For example, all the reasoning processes take place explicitly at the agents level and implicitly in the aggregated labor market level. In our study, the precariousness induced by the CNE contract clearly emerges, and can be measured, as a dramatic increase of oscillations for all the rates: unemployment, vacancy, and working population.

Contrary to most approaches in economics, we adopted a bottom-up approach. This approach is based on a higher granularity of the model. Most economic models, including models that deal with microeconomic questions, use to average behaviors of agents when reasoning about durations or decision-making. When using an agent-based simulation, we are obliged to look into the agent, and thus to define its architecture. Many types of architectures are possible. In the future, we plan to model the agent's cognition more profoundly, in order to tackle more adequately the human decision processes involved in the labor market. For such a purpose, we could use and adapt our cognitive architecture CODAGE [10] we already applied for an experimental financial market. The whole idea is to "humanize" more the agents, to provide them with finer and more precise reasoning and decision abilities, something that will hopefully bring them closer to real human agents in the labor market.

We have not so far exploited all the possibilities provided by MAS, and several improvements will help to bring the model closer to reality in the future. First, we could diversify the agents a little more. By diversity we mean setting a field of expertise for each agent. In that case, an agent will be hired only for jobs, which suit its professional profile. This enrichment will make the matching mechanism more realistic and add frictions to the labor market environment.

Second, in the real world, the matching process is not perfect and straightforward. Even if an employee meets with an employer, a matching does not necessarily take place. A selection or sorting phase may be integrated into both sides. An employee may wave away a job offer because he/she finds the work conditions not satisfying enough. An employer may refuse a candidate because of a bad interview, a bad vita or simply because he/she prefers hiring someone else.

Third, the meeting between a worker and an employer may be considered as a bargaining interaction as well, where both sides negotiate the wage, the work conditions and so on. This first meeting may end in separation or in hiring. By allowing agents to negotiate, the interactions in the simulation will depend more on the labor market situation. If the unemployment rate is low an employee should have more power when negotiating his/her work conditions then when there are a lot of unemployed persons and not many vacancies.

Finally, our model allows an employer to fire an employee, but the opposite case where an employee decides to quit is not permitted. This situation is biased and a future version of our model should allow an employee to quit his/her job as well.

References

1. Weiss, G.: Multiagent Systems A modern approach to distributed artificial intelligence. MIT Press (1999)
2. Tesfatsion, L., Judd, K., eds.: Handbook of Computational Economics, 2: Agent-Based Computational Economics. North-Holland (2006)
3. Cahuc, P., Carcillo, S.: The Shortcomings of a Partial Release of Employment Protection Laws: The Case of the 2005 French Reform. IMF Working Paper No. 06/301 (January 2007)
4. Cahuc, P., Zylberberg, A.: Labor Economics. MIT Press (2004)
5. Neugart, M.: Labor market policy evaluation with ACE. In: Agent-Based Models for Economic Policy Design (ACEPOL05), Bielefeld (June 2005)
6. Neugart, M.: Endogenous matching functions: An agent-based computational approach. In: Industry and Labor Dynamics: The Agent-Based Computational Economics Approach - Proceedings of the Wild@Ace 2003 Conference, Torino (October 2003) 90–106
7. Pingle, M., Tesfatsion, L.S.: Evolution of worker-employer networks and behaviors under alternative unemployment benefits: An agent-based computational study. In: Industry and Labor Dynamics: The Agent-Based Computational Economics Approach, Proceedings of the Wild@Ace 2003 Conference, Torino (October 2003) 129–163
8. Gabriele, R.: Labor market dynamics and institutions: an evolutionary approach. LEM Papers Series from Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies, Pisa, Italy (2002)
9. Cougaar: An Open Source Agent Architecture. http://www.cougaar.org/
10. Kant, J.D., Thiriot, S.: Modeling one human decision maker with a multi-agent system: the codage approach. In: AAMAS '06: Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems, New York, NY, USA, ACM Press (2006) 50–57

Testing Marshallian and Walrasian Instability with an Agent Based Model

Marta Posada, Cesáreo Hernández, and Adolfo López-Paredes

INSISOC, Universidad de Valladolid,
E.T.S. de Ingenieros Industriales, Paseo del Cauce s/n, 47011 Valladolid, Spain,
{posada,cesareo,adolfo}@insisoc.org

Abstract. Do the Marshallian and the Walrasian models have a firm foundation on micromotives, or are they just macro abstractions that we could dispense of in Microeconomics? Previous evidence from experimental economics with humans-subjects in continuous double auction markets shows that, if the supply is downward-sloping or the demand is upward-sloping, the Marshallian stability model captures the observed phenomena but the Walrasian does not. But of course in human subjects experiments, the basic question of what was the human agents' behaviour was remains open. We build an artificial agent based model to show that the path of convergence needs more intelligence than the Zero-Intelligent framework. We confirm with artificial intelligent agents the results from experimental economics at a low cost and an increase in the reliability of the experiment. Not only the existence but the computation of the stability is relevant. These results cast a shadow on the interest of both instability concepts in Economic Theory.

1 Scope and Related Work

Multiagent-based simulation (MABS) has become a popular tool to theorize about distributed but interacting phenomena as markets and economic activity are. Economics is a Social Science. Being social inherits complexity and being a science calls for experimentation. Historical records are not enough to test economic theories. Happily enough Experimental Economics with humans has provided a replicable Lab to suggest new solutions to economic problems and to test standing models and theories. But human behaviour in the experiments is not directly controllable and the question of what the agents' behaviour is remains open. Artificial Economics, mainly MABS, has broadened the scope of Experimental Economics, allowing the experimenter to check alternative individual behaviour. In this sense MABS is a "killer application" of Economic Theory.

In this paper we consider alternative individual agents' behaviour in a Continuous Double Auction (CDA) market to examine which of the two concepts of instability (Walrasian and Marshallian) emerge from these alternative behaviours and their practical and theoretical relevance.

The Walrasian and Marshalian market adjustment. Two classical concepts of market stability exist in the literature. One is based on the analysis introduced by Alfred Marshall, traditionally called "Marshallian stability" and the second was introduced by Leon Walras, traditionally called "Walrasian stability". The two concepts of instability (Walrasian and Marshallian) stem from two different models of market adjustment. The Marshallian model views volume as adjusting in response to the difference between demand price and supply price at that volume. The Walrasian model views price as changing in response to excess demand at that price.

If there is an awkward question one has to suffer when teaching economics, that is who does the job of the so called market adjustment? In contrast with the usual microeconomic models, these processes of adjustment do not represent any optimization of the economic agents' behaviour, but just differential equations brought out from nowhere. As Nicholson [7, chapter 19] states in his well known microeconomics book: *"in the last instance this speculation on the adjustment mechanism hardly makes any sense, because neither the Walrasian nor the Marshalian adjustment reflect the real behaviour of the economic agents"*. Being virtual models taken from the mechanical analogy of vector fields, they can be compatible with different micro behaviours of the participant agents. Nevertheless it is interesting to asses the relevance of both concepts of instability, and to which extent they emerge from individual agents' behaviour.

Learning and perfect competition, an outstanding puzzle in microeconomics. In general equilibrium theory it was undoubtedly the imposition of high standards of mathematical rigour that led to the almost exclusive concern with the existence and characterization of equilibria rather than the adjustment process that lead to them. In the nineties there was an upsurge in interest in determining how economic agents might learn their way into equilibria in general.

But these works have been focused on the question of learning about the parameters of the model (as captured by some Walrasian or not equilibrium model). They have concentrated on proving the existence of such an equilibrium and perhaps some evolutionary mechanism that will lead to the fixed point equilibrium, under the forceful assumption of rational expectations.

Markets are information gathering tools that will allow many reasonable and realistic heuristic learning-decision models from agents yet achieving market convergence and equilibrium. Assuming that learning is primary focused towards finding the true parameters, which may describe the macroscopic market behaviour as a "virtual" entity, is just totally unnecessary and even unrealistic. Why should the agents have a "virtual correct" macro model in mind and try to capture the parameters of the model? Their search may be local and no a super agent such as the Walrasian auctioneer is needed for a market clearing-price as the evidence form Experimental Economics with humans, and recently with artificial agents, has proved

The equilibrium puzzle. If we accept the existence of a virtual auctioneer that computes prices in a centralized way (Walrasian adjustment) or the existence of a Marshallian path to the equilibrium, one may ask two questions. To which extent are they efficient in computational terms? Under what circumstances this adjustment will lead to the expected market equilibrium? The first question has been answered recently in a great paper by Axtell [1]. A decentralized system with agents, who have local information and are engaged in k-lateral exchange, has much better computational complexity than Walrasian equilibria have.

The second question is the main objective of the paper. With this aim in mind, and following Smith [13], we single out the three dimensions that are essential in the design of any market experiment: the environment (E) (agent endowments and values, resources and knowledge), the agent' behaviour (A) and the institution (I) (it is both the exchange rules and the way the contracts are closed, and the information network). The institution we have considered is a CDA, since it has been extensively used in experimental economics and in works related to this one (Posada [10], Posada *et al.* [11, 12]). The environment dimension is essential to analyze the convergence properties to the clearing price and to check for Marshallian or Walrasian instability. In environments where the supply is positively sloped and the demand is negatively sloped (as in Figure 1.a), there is only one equilibrium and it is both Walrasian and Marshallian stable (MS/WS). The controversy arises when a priori there are several points that should be Marshallian (Walrasian) stable (unstable). This happens when the supply is downward-sloping (when the price decreases the quantity supplied increases) or when the demand is upward-sloping (when the price decreases the quantity demanded decreases). For example, in Figure 1.b where the supply is downward-sloping, if the demand cuts the supply from above (below), the equilibrium is Walrasian unstable (MU) and Marshallian stable (MS) (Walrasian stable (WS) and Marshallian unstable (MU)). The situation is the same in Figure 1.c where the demand is upward-sloping since an upward-sloping demand is a mirror image of the downward-sloping demand.

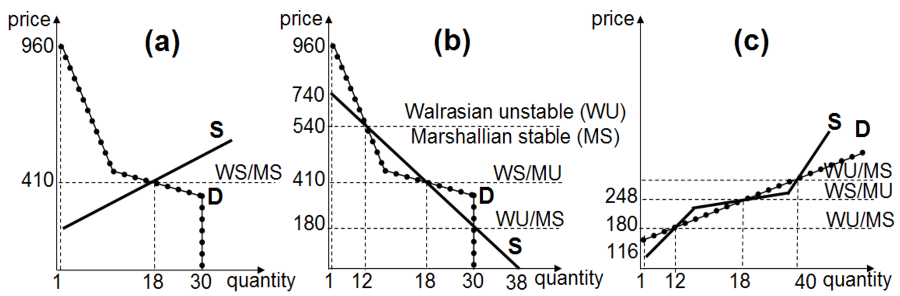


Fig. 1. Environments: (a) standard; (b) downward-sloping supply; (b) upward-sloping demand

Economic theory suggests two situations that lead to such unusual shapes. One is related to income effects (the famous Giffen good for an upward-sloping demand and the labour-leisure trade-off for a downward-sloping supply). The second one is related to externalities. On the supply side, downward-sloping supplies are thought to be produced by increasing efficiency due to expanding industrial scale. On the demand side, a similarly externality can produce the upward-sloping demand curves. The externality in this case is due to fads (preferences such as desires to mimic others' behaviour). In this paper we consider the second one, the use of externalities to create both a downward-sloping supply and an upward-sloping demand.

In the standard environment where the Walrasian stable equilibrium coincides with the Marshallian stable equilibrium, human-subjects experiments confirm the theoretical predictions. Do the experiments with artificial agents lead to the expected equilibrium? Brewer *et al.* [2] examine Zero Intelligence (ZI) agents (developed by Gode and Sunder [6] and Sunder [14]) trading under standard symmetric environments. Furthermore Posada [10] and Posada *et al.* [11, 12] examine other agents' behaviours (GD developed by Gjerstad and Dickhaut [5], ZIP developed by Cliff and Bruten [3], K the winner in Santa Fe tournament in 1993) under standard symmetric and asymmetric environments.

The main conclusions are that the institutional design (I) matters and so does the agents' intelligence (A). That perfect competition is compatible with strategic agent's behaviour. Mapping different arrangements of (ExA) in a CDA market, they showed that agents' behaviour influences the convergence to the clearing price equilibrium. They obtained that under asymmetric environments the clearing price equilibrium is not achieved in homogeneous ZI populations, the price convergence is achieved in GD/ZIP populations (GD agents takes less time to exchange and to learn that ZIP agents) and the presence of K agents in the population affects negatively both price convergence. Computation of the equilibrium is affected by intelligence, but the concepts of instability are not questioned.

However human-subject experiments under a downward-sloping supply environment (Plott and George [8]) and under a upward-sloping demand environment (Plott and Smith [9]) found that in a continuous double auction market the Marshallian stability model captures the observed phenomena whereas the Walrasian model does not. In this paper we replicate and generalise with artificial agents the results of Plott's experiments. We wonder if these results depend on a particular agents' behaviour because nothing is said about this issue.

This paper addresses four questions: Can both stable and unstable equilibria be observed? If markets do exhibit instability, which of the two classical models will lead to the right equilibria prediction? How robust are the results against alternative learning agents and against alternative environments? In view of these results, what is the interest of both instability concepts for policy focused modelling and simulation?

2 The Model

The main features of our model are described using the triple (IxExA): institution, environment and agents' behaviour. The model has been programmed in

SDML, a declarative model language developed by the Centre for Policy Modelling at the Metropolitan Manchester University (http: //cfpm.org).

2.1 The Institution (I)

Under the standard continuous double auction (CDA) rules, any trader can send (or accept) an order at any time during the trading period. A new bid/ask has to improve previous pre-existing bid/ask. A trade occurs when a new ask (a) is made that is less than a pre-existing bid (bc), or when a new bid (b) is made that is greater than a pre-existing ask (ac). The trading is equal to that of the pre-existing bid/ask, whose acceptance is triggered automatically by the new entry. Pseudo code of CDA protocol (but note that in SDML it has been programmed using rules)

```
1. {Initialize the initial bid/ask: ac=∞, bc=0 when a period starts
   or a deal takes place}
2. {Several situations might arise during a round}
   Repeat
      When a seller-agent submits an ask a
         if a ≥ ac then a is an invalid ask
         if bc<a<ac then ac is update to a
         if a ≤ bc then this seller-agent makes a deal at bc goto
         1
      When a buyer-agent submits a bid b
         if b ≤ bc then b is an invalid bid
         if bc<b<ac then bc is update to b
         if b ≥ ac then this buyer-agent makes a deal at ac goto 1
   until no bids/asks are submitted
```

2.2 The Environment (E)

We consider two alternative environments: a downward-sloping supply (Figure 1b) and a upward-sloping demand (Figure 1c). Since they are mirror images, we test both the replicability and the robustness of the results.

Forward falling supply (Figure 1b). In our simulations we have used the same valuations which were used in Plott and George [8]'s human-subject experiments.

In the demand side, there are six buyers (each one with six units). There are two buyers of each type. Each buyer of a given type has identical reserve price given by [960, 600, 440, 350, 330, 0], [880, 640, 410, 390, 310, 0] and [800, 720, 410, 390, 290, 0], respectively. Buyers know their reserve price with certainty. In the supply side, there are six sellers. Each seller is uncertain about his marginal costs because they depend on their own output and the output of all other sellers. The externality implies that, as market volume increases, the marginal cost decreases even though the individual seller's marginal cost increases with an increase in his own volume. Each seller has eight units to trade. There are two sellers of each type (a, b, c). Each seller of a given type has identical marginal costs. In Table 1 we show the marginal costs of type-a sellers. As unit increases, the marginal cost value increases. Note as volume of others increases, the marginal cost value decreases. Table 2 lists the equilibria according to both, Walrasian and Marshallian theories in the downward-sloping supply environment.

Table 1. Valuations of type-agent. Marginal Costs of type-a sellers

Volume sold by others							
units	0	1	2	3	4	...	18
1st	820	790	760	730	700	...	130
2nd	900	870	840	810	780	...	210
⋮	⋮	⋮	⋮	⋮	⋮	⋱	⋮
8th	1380	1350	1320	1290	1260	...	690

Table 2. Equilibria in downward-sloping supply environment

price	quantity	Marshall	Walras
500-540	12	stable	unstable
380-410	18	unstable	stable
140-180	30	stable	unstable

Upward-sloping demand (Figure 1c). In our simulations we have used the same valuations which were used in Plott and Smith [9]'s human-subject experiments.

In the supply side, there are six sellers (each one with ten units). Sellers know their marginal cost with certainty. The aggregated supply is [308 306 304 302 300 298 296 294 292 290 288 286 284 282 280 278 276 274 272 270 268 266 264 262 260 258 256 254 252 250 248 246 244 242 240 238 236 234 232 230 228 226 220 212 204 196 188 180 172 164 156 148 144 140 136 132 128 124 120 116].

In the demand side, there are six buyers. Each buyer is uncertain about his reserve prices because they are influenced by their own buying behaviour and the buying behaviour of other buyers. The externality implies that, as market volume increases, the reserve price increases even though the individual buyer's reserve price decreases with an increase in his own volume. Each buyer has eight units to trade. There are two buyers of each type (a, b, c). Each buyer of a given type has identical reserve prices. In Table 3 we show the reserve prices of type-a buyers. Table 4 lists the equilibria according to both, Walrasian and Marshallian theories in the downward-sloping supply environment.

Table 3. Valuations of type-agent. Reserve prices of type-a buyers

Volume bought by others							
units	0	1	2	3	4	...	18
1st	120	140	146	152	158	...	272
2nd	108	124	130	136	142	...	256
⋮	⋮	⋮	⋮	⋮	⋮	⋱	⋮
8th	12	28	34	40	46	...	160

Table 4. Equilibria in upward-sloping demand environment

price	quantity	Marshall	Walras
176	12	stable	unstable
248	18	unstable	stable
273	40	stable	unstable

2.3 Agents' Behaviour (A)

In the downward-sloping supply environment described above, sellers are uncertain about their marginal costs because they depend on their output and the output of all the other sellers. In the upward-sloping demand environment buyers are uncertain about their reserve prices because they are influenced by their own buying behaviour and the buying behaviour of other buyers.

The first decision taken by each agent who is uncertain about his valuations is to estimate the volume traded in the market. Agents form expectations on the volume traded by others using their own past experience. They update their volume expectations about the volume traded by others ($\hat{q}_n$) according to the actual volume observed in the market (q_n). In particular, the agent (indexed in i) uses the following simple updating rule:

$$\hat{q}_{n+1} = (1 - \lambda)\hat{q}_n + \lambda q_n \,. \tag{1}$$

The λ learning rate measures the responsiveness of agents' volume estimation to new data (a memory weighting factor). If $\lambda = 1$ the agents believe that the quantity traded by others in the next period will be equal to the traded volumes observed in the current period. Thus sellers will under-estimate their marginal cost and buyers will over-estimate their reserve prices. If $\lambda = 0$ the agents do not use the information generated in the market to improve the estimation of the initial traded volume (myopic behaviour).

Once an agent estimates his valuation (marginal cost or reserve price), he faces the following three decisions: How much should he order? When should he submit an order? When should he accept an outstanding order? To take these decisions each agent only knows his own valuations, which are private, and the information generated in the market.

How much should he order? We try two alternative bidding strategies: ZI (Gode and Sunder [6]) and GD (Gjerstad and Dickhaut [5]). Each bidding strategy has different answers to this question. For space reasons we refer the reader to these references for the details of the bidding strategies.

Each ZI agent chooses his order randomly between his valuation and the best ask outstanding in the market.

Each GD agent chooses the order that maximizes his expected surplus, defined as the product of the gain from trade and the probability for an order to be accepted.

$$\text{Sellers: } \max \prod_a (\text{price-marginal cost}), \tag{2}$$

$$\text{Buyers: } \max \prod_b (\text{reserve price-price}). \tag{3}$$

GD sellers estimate the probability $\prod_a$ calculating a belief function $q(a)$ using the history of the recent market activity (AAG accepted ask greater than a, BG accepted bid and ask greater than a and RAL rejected ask less than a) to calculate the probability for an ask at price a to be accepted. Interpolation is used for values at which no orders are registered.

$$\hat{q}(a) = \frac{AAG(a) + BG(a)}{AAG(a) + BG(a) + RAL(a)} \,. \tag{4}$$

When should he submit an order? When an agent is active at a given time step, he may submit an order (a new or a replacement an open order). The agents have a constant activation probability of 25%. Of course, orders must be also in agreement with the spread reduction rule of the institution.

When should he accept an outstanding order? A seller accepts the current bid if his ask (submitted or not) is equal to or greater than the current bid. A buyer accepts the current ask if his bid (submitted or not) is equal to or less than the current ask.

3 Results

Note that agents have two basic learning tasks: learning about the externality (on the supply side or on the demand side) and learning how to bid/ask. What is the performance of humans, ZI and GD agents in these two tasks? Following we discuss the meaning of the results of both volume traded and prices.

Total traded volume. We find in our experiment a market volume of 12 which is Marshallian stable and Walrasian unstable, in every period and for any population under any environment as it happens in Plott's human-subject experiments [8, 9].

Price convergence. On the contrary the emerging price patterns towards convergence of humans, ZI and GD agents are different. Figure 2 shows the price distributions of the transactions in 30 runs of a downward-sloping supply environment, each one with ten trading periods and 100 rounds per period. We can see that the transaction prices in homogeneous ZI populations are more volatile than the transaction prices in homogeneous GD populations, where the price convergence to the Marshallian stable equilibrium is very clear. The results are similar under an upward-sloping demand environment as we can see in Figure 3. Since the downward-sloping supply environment is a mirror image of the upward-sloping demand, we confirm the replicability and the robustness of our results.

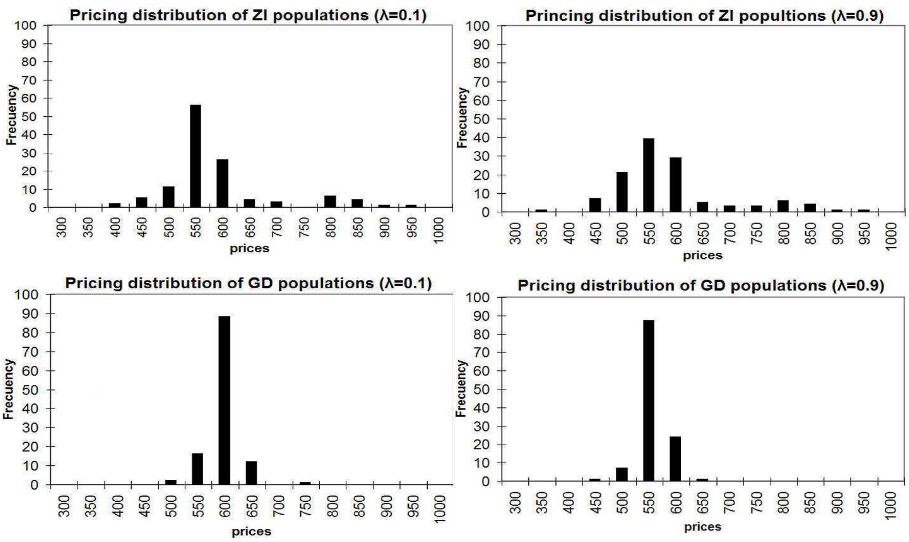


Fig. 2. Pricing distribution of both ZI and GD homogeneous populations under extreme learning rates ($\lambda = 0.1$ and $\lambda = 0.9$) in a downward-sloping supply environment

Although the distribution of the transaction prices in homogeneous ZI populations is more flat, there is a false convergence appearance. In Figure 4 which shows the time series of transactions, we observe that there is no price convergence to the equilibrium price in ZI homogeneous populations.

The learning rate plays an important role in GD populations.Figure 4 shows the time series of transaction of a GD population in a downward-sloping supply. We have represented with a discontinue line the price range (540-500) which is Marshallian stable and Walrasian unstable equilibrium. Here we observe that the transaction prices remain slightly over 540, with a positive bias, when $\lambda = 0.1$ (myopic behaviour) as it happens with human subjects experiment [8]. The transaction prices are unbiased and they cluster around the theoretical equilibrium of (540-500) when $\lambda = 0.9$.

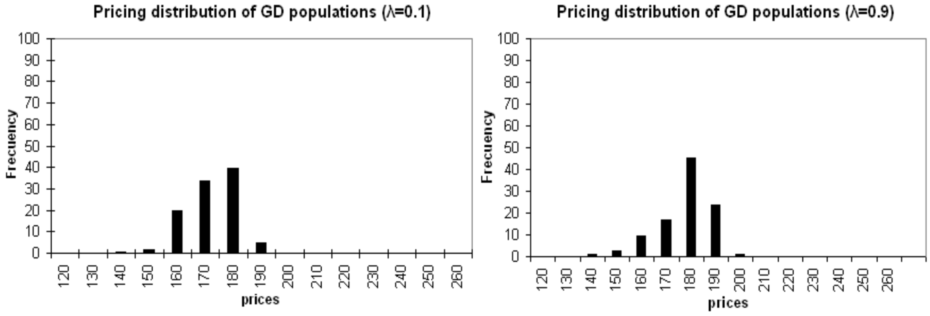


Fig. 3. Pricing distribution of GD homogeneous populations under extreme learning rates ($\lambda = 0.1$ and $\lambda = 0.9$) in a upward-sloping demand environment

GD sellers may have negative profits as a result of their wrong estimation of the volume sold by others. GD sellers with a low learning rate make more mistakes than GD sellers with a high learning rate. Do humans make mistakes when they estimate the volume sold by others as GD sellers with a low learning rate make? Are these mistakes the reason for the similar equilibrium price range of humans and artificial intelligent agents with a low learning rate? The source for the bias in the equilibrium convergence with human agents remains an open question.

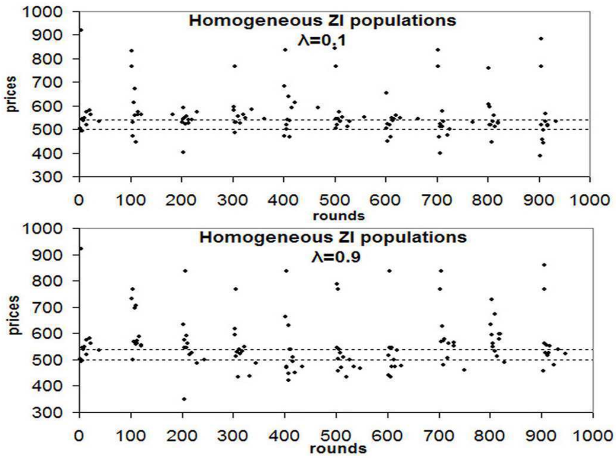


Fig. 4. Price dynamics for homogeneous ZI populations under a downward-sloping supply

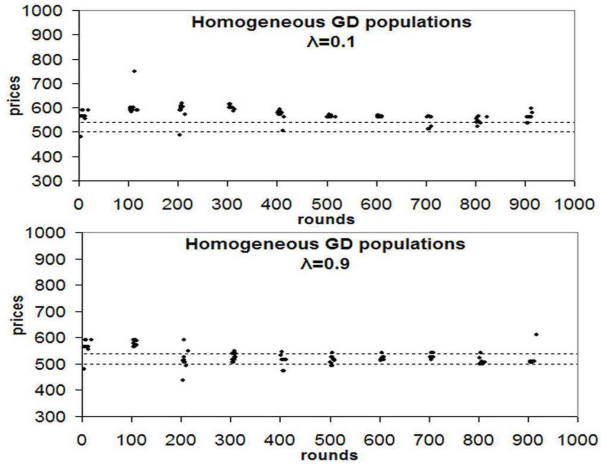


Fig. 5. Price dynamics for homogeneous GD populations under a downward-sloping supply

4 Conclusions

We have shown, as in previous works, that ABM simulations we can test the results from Experimental Economics with humans that have enlightened Economics for the last two decades. In this paper we have clarified one of the most unpleasant questions one can face when teaching Economics: *who does the job of the so called market adjustment?* We have found answers to the following questions: Can both stable and unstable equilibria be observed? If markets do exhibit instability, which of the two classical models will lead to the right equilibria prediction? How robust are the results against alternative learning agents? In view of these results, what is the interest of both instability concepts for policy focused modelling and simulation?

The micro behaviour of the artificial agents leads to a Marshallian stability model but not to a Walrasian equilibrium. We confirm that the emerging price pattern from with artificial GD agents is compatible with the Marshallian model as it happens in the Plott's human-subject experiments [8, 9]. This answers the first two questions.

Marshallian stability is achieved and it is a valid instability criteria of the market when a stable equilibrium exists. Whereas Walrasian is not. These results are not robust against alternative agents learning models. In a CDA market populated by ZI agents, market efficiency and the right estimation of the traded volume is achieved, but there is no price convergence to equilibrium.

As for the last question, we share the following comment by Axtell [1] to fix up our conclusion: *In the end we advocate not the jettisoning of this useful abstraction* (Walrasian equilibrium) *but merely its circumspect use whenever focused on questions for which it has limited ability to adjudicate an appropriate* *answer: distributional issues and actual prices.a direct consequence of the results described above is to at least cast a pale on the utility of such analysis,* (Walrasian or Marshallian) *if not vitiate them altogether.*

Since we have the instruments for the analysis of markets from the agents behaviour (microbehaviuor) what should be claimed is a proper instruction in MABS as a prerequisite for research in Economic Theory.

Acknowledgements

This work has received financial support from the Spanish MEC, n^o 2005-05676 and from Junta de Castilla y León n^o VA029B06. We acknowledge the useful comments of the two referees.

References

1. Axtell, R.: The complexity of exchange. Economic Journal 115 (2005) 193–210
2. Brewer, P. J., Huang, M., Nelson, B., Plott, C. R.: On the Behavioral Foundations of the Law of Supply and Demand: Human Convergence and Robot Randomness. Experimental Economics 5 (2002) 179–208
3. Cliff, D., Bruten, J.: Zero is not enough: On the lower limit of agent intelligence for continuous double auction markets. HP Laboratories, Tech. Rep. HPL-97-141 (1997)
4. Gjerstad, S.: The competitive market paradox. Tech. Rep 1180 of the Institute for Research in the Behavioral, Economic, and Management Sciences of Purdue University (2006)
5. Gjerstad, S., Dickhaut, J.: Price formation in double auctions. Games and Economic Behaviour **22** (1998) 1–29
6. Gode, D., Sunder, S.: Allocative efficiency of market with zero-intelligent traders: Market as a partial substitute for individual rationality. J. of Political Economy **101** (1993) 119–137
7. Nicholson, W: Microeconomic Theory. Dryden Press (7th ed) (1997)
8. Plott, C.R., George, G.: Marsahallian vs Walrasian stability in an experimental market. The Economic Journal 102(1992) 437–460
9. Plott, C.R., Smith, J.: Instability of Equilibria in Experimental Markets: Upward-Sloping Demands, Externalities, and Fad-Like Incentives. Southern Economic Journal 65(1999) 405–426
10. Posada, M.: Strategic Software Agents in Continuous Double Auction under Dynamic Environments. Lecture Notes in Computer Science **4224** (2006) 1223–1233
11. Posada, M., Hernández, C., López, A.: Learning in a continuous double auction market. Artificial Economics - Lecture Notes in Economics and Mathematical Systems **564** (2006) 41–51
12. Posada, M., Hernández, C., López, A.: Strategic behaviour in a Continuous Double Auction Market. Lecture Notes in Economics and Mathematical Systems **584** (2006) 31–43
13. Theory, Experiment and Economics. Journal of Economic Perspectives, Winter 1989) 783–801
14. Sunder, S.: Markets as Artifacts. In M. Augier and J. March (eds.): Models of a Man: essays in memory of Herbert A. Simon. Cambridge: MIT Press (2004)

Modelling endogenous rule changes in an institutional context: The ADICO sequence

Alex Smajgl[1], Luis Izquierdo[2], and Marco Huigen[3]

[1] CSIRO Sustainable Ecosystems, University Drive,
Townsville QLD 4810, Australia, alex.smajgl@csiro.au
[2] Universidad de Burgos, Edificio la Milanera, Burgos, Spain
[3] Universität Hohenheim, Stuttgart, Germany

Abstract. Agent-based modeling is increasingly used to simulate systems that involve social dynamics. Changes to constraints humans face trigger often also modifications at the scale of the social group. Including such adaptation dynamics would allow modeling 'pure' emergence. This paper discusses such an approach in an institutional framework and develops a sequence that allows modeling endogenous rule changes. Parts of this sequence are translated in a NetLogo KISS implementation and simulation results are provided at the end of this paper.

1 Introduction

A major advantage of agent-based models (ABM) in social science is the exploitation of the agents' heterogeneity. Several models using this aspect of ABM display modelled patterns that emerge on the system level. The emergence of a collective social pattern is not visible by looking at the individual characteristics only and result from a multitude of socially heterogenous interactions. However, it may be claimed that this emergence is already baked-in by the parameterization of the agents, their often static rules and their world. A 'pure' form of emergence is only then possible if agents are able to come up dynamically with new actions for themselves and are allowed to dynamically formulate rules on the system level to guide their own and their peers' actions. Research on institutional change regarding common pool resources require such a dynamism of 'pure' emergence.

Ostrom [8] defines institutions as "the shared concepts used by humans in repetitive situations organized by rules, norms, and strategies" [9]. Commons [4] identifies a similar link between the collective and the consequence for individual behaviour when he defines institutions as "collective action in constraint, liberation or expansion of individual action". Ostrom [8] distinguishes between rules and norms by emphasising that norms do not have a punishment component while rules have.

Stepping towards rule change dynamics, Buchanan and Yoon [3:1] state the core question behind the existence and dynamics of rules: "Why do persons choose rules that seem to constrain or limit their choices?" Bromley [2] rejects the interpretation of rules as solely constraining mechanisms and refers to the enabling properties of rules, similar to Commons [4]. While most researchers refer to rules as a constraining mechanism, i.e. North [7], Bromley (2006) points out that, for instance, a non-smoking rule constraints the group of smokers but it enables the group of non-smokers. The existence of rules is therefore based on the expected consequences resulting from the trade-off between groups of individuals that are enabled and a second group that is constraint.

Smajgl and Larson [11] discuss institutional change as an effect of changes in economic, social, technological or other institutional variables. Two examples shall clarify institutional causality: Economic decisions such as the liberalisation of markets have a multitude of flow-on effects that include new incentives for potential producers to enter the market and incentives for existing producers to create market barriers protecting them from new competition. Therefore, rules have to be established to regulate how market access is granted and what kind of fees can be charged by whom. Similarly, new technologies create new strategies for individuals, such as the internet providing an effective network for data exchange. Such new strategies for individuals go along with various flow-on effects that might require new rules for how to use the new technologies. A new rule itself is able to create requirements to modify existing rules in order to increase consistency of legislation. Such *institutional ripple effects* [10, 11] are based on a complexity of interactions of individuals, their epistemological background and their goals.

Essential for the linkage between the epistemological level of individuals, their interaction and the social level on which rule changes emerge is the critical incentive perceived by individuals within the social system. Following Bromley [2] the change is based on the constraining and the enabling property of a rule. Therefore, the consequence of a (new or modified) rule expected by individuals defines an incentive that triggers an investment (in time or other resources) to provide assistance to rule changing processes or to hinder them.

Given such a principle mechanism of social systems with enabling and constraining properties its implementation in models simulating social systems is still in its infancy. The few publications that have focussed on this aspect of social simulation [1, 6] show the difficulties in developing a generic approach for implementing the dynamics of rule changes in social systems.

In this paper we analyse necessary elements to model endogenous institutional rule change by (1) identifying structural rule components from a modelling perspective, (2) structuring an agents' knowledge-base to determine rule change, (3) propose a sequence protocol, and (4) developing a kiss model that demonstrates work in progress for modelling endogenous rule change dynamics.

2 From rule structure to rule dynamics

While Ostrom [justifiably] uses the term strategy we prefer the term (agents') action in this paper. An agents' strategy implies an agent with an endogenous reasoning capacity of selecting a set of actions to pursue a long-term goal, while in this paper we equip our agents solely with one time-step objectives.

The difference between actions an individual can choose and the rules perceived by an individual is pivotal. While an action can be "Pump 20 l/min" and another one "Pump 100 l/min" a rule has several additional building blocks. An example of a rule is: *No individual is allowed to pump more than 50 l/min at all times from the aquifer or else he gets fined by local police*.

Crawford and Ostrom [5] define a *Grammar for institutions* and provide a general structure of rules. These are built from the elements of *attributes*, *deontic*, *aim*, *conditions* and *or else*. Their example is (see Crawford and Ostrom 1995: 584)

Attribute A	Deontic D	Aim I	Conditions C	or else O
"All villagers	must not	let their animals trample the irrigation channels	at all times	Or else the village who owns the live-stock will be levied a fine."

One of the structural differences between an agent's action and an institutional rule is that the action just entails the elements A-I-C while a rule demands all five elements. Rule dynamics implies that at least one of these five elements changes. For instance, the attribute might change while the rest of the rule remains unchanged.

As said, one of the major objectives of this paper is to explore the necessary building blocks to model institutional rule change. We identify two major mechanisms that lead to institutional rule change:

a. Innovation mechanism: agents develop new interactions allowing them to deal with their own resources in a more efficient way.
b. System-level rule mechanism: agents formulate new rules on the system level allowing the community to deal with the resources better.

An example of the innovation mechanism is the introduction of a market for tradable water quotas without observing such a system somewhere else. In order to allow new rules to emerge the (Aim) I field has to be split up into two aspects, the verb and the object. For instance, 'to trade' as a verb might be part of the knowledge of all agents but might not be linked to 'water', which defines the object of the Aim(I). These two subfield (I-verb and I-object) expand the potential variety in creating new rules and allow for a rule that defines, for instance the need for a quota (entitlement) for the use of irrigation water and subsequent the rule regarding trading these entitlements.

For a institutional rule change of the second type the agents 'discover' in a situation when individual payoff is below expected payoff that the cause of their disappointing returns is an externality on the system level. A rule change on the system level dealing with this externality might improve their own payoff. Hence, here the dissatisfaction of an agent leads to an evaluation of the *effectiveness* of the system, while the former mechanism results from improving the *efficiency* of their resource use.

3 Agent Architecture

The approach developed here outlines an endogenous solution to rule changes in an agent based model. The main model components are summarised in Figure 1:

- a population of agents, with each agent endowed with
  - a set of actions she can potentially use,
  - a set of goals,
  - a set of motivations, and
  - a knowledge base that maps the impact of own actions and other agents actions against the agent's goals
- a given set of context variables such as environmental assets, and
- a set of institutional rules.

Agents are assumed to be heterogeneous according to their attributes defined in action set, goals, motivation and knowledge. Additionally, the endowment or access to context variables can vary depending on the context this structure is implemented in. This latter point can be defined by access to specific infrastructure, such as fishing boats, or political influence.

This approach to model rule change is based on perceived incentives, which requires a definition of goals and motivation. A **goal** is a desired state of one or more system attributes and agent attributes at time t (infinite) linked to an agent. This means that an agent can aim for achieving individual goals, such as profit or covering nutritional needs, or system goals, such as maintaining fish stocks as a sustainable option for future generations.

Motivation is the context-dependant (time, variable state, agent state) mapping of possible actions against their contribution to agent goals. The **knowledge base** defines the memory of an agent and records effectiveness regarding chosen actions and goal achievement. It has to be a combination of reinforcement learning capturing the payoff response from changes in own actions and observation based learning, such as fictitious play. This need is based on the trigger for institutional change. An agent who does not achieve its goals although trialling various actions must have to ability to triangulate other agents' actions with own actions and own payoff. Technically, this step of identifying negative (or positive) externalities flags discontent with a specific combination of activities. This means an agent must have to ability to look back and to rationalise what changed on the system that triggered the decreasing payoff.

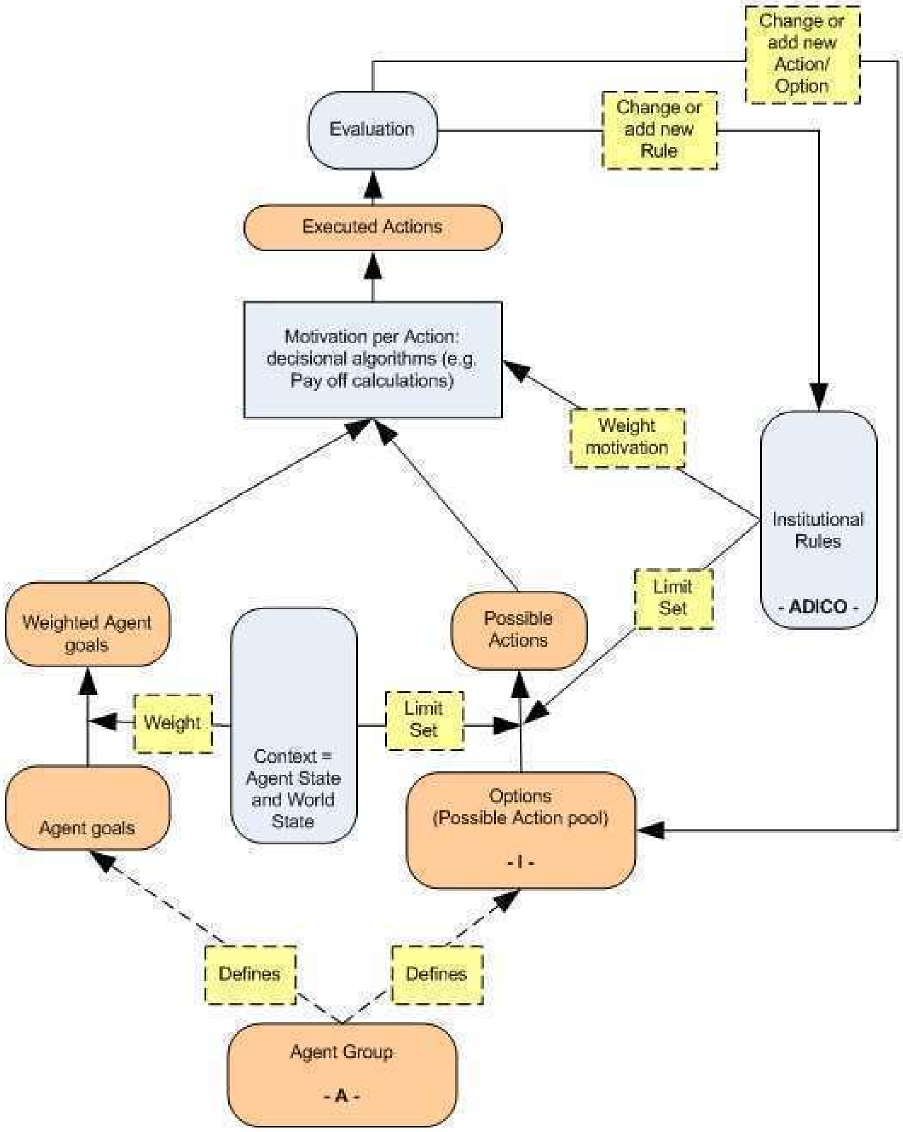


Figure 1. Flow diagram with model components

Therefore, the motivation that triggers the action choice of an agent can be defined as

$$M_i = f(A_i, A_j, G_i, KB_i, IR) \quad \text{with} \quad \begin{array}{l} IR : \text{Institutional rules} \\ KB_i : \text{Knowledge base of agent } i \\ G_i : \text{Goal set of agent } i \\ A_i : \text{Action set of agent } i \\ payoff_i \in G_i : \text{Reward of agent } i \end{array}$$

The motivation of agent *i* to choose a certain action to achieve a perceived goal is influenced by the **institutional rule**, which applies to all agents in the system. Institutional rules are defined as a combination of five fields ADICO, following the Crawford and Ostrom approach.

As mentioned above, institutional change and emergence of new rules requires disaggregating the Aim (I) into two parts. The aim consists of a verb and an object, which have to e separated to allow new combinations of actions. Firstly, a verb (I-V) is identified, such as to go fishing or to trade something. Secondly, the object (I-O) is defined in order to identify the tool or mechanism, such as fishing boat or a market. This allows an innovation to emerge such as trading water quotas, which is based on the pre-existing knowledge about 'trading' as a possible action (verb) and the specific market applied to an object.

It is assumed that if each individual achieves her goal no incentive would exist to question *IR*. Depending on the individual's role in the social system she can be endowed by more or less power in the rule changing process. Additionally it is important that individuals can base their motivation also on system goals such as the conservation of fish stocks.

4 Sequence for Institutional Rule Change

Given the model components, the sequence starts with updating goals, motivations and list of possible actions. The principle relation of model components is mapped in Figure 1. Starting with the initialisation of agent group

1.1 Load goal(s)
1.2 Load possible actions
1.3 Define Motivation, i.e. calculate payoff for each possible action
1.4 Choose action with highest payoff and execute the action

Figure 1 shows that the next step is the evaluation in which the agent compares the expected payoff with the realised payoff. Realising the action leads to the next steps:

2.1 Evaluation: Compare actual payoff with expected payoff

2.2 IF $payoff_{i,t} = f\left(A_i^t, Context\right) \geq payoff_{i,t,\text{expected}}$ GOTO 1.1 (outer arrow leading back to options)

2.3 ELSE $payoff_{i,t} = f\left(A_i^t, Context\right) < payoff_{i,t,\text{expected}}$ update payoff function:

$$payoff_{i,t} = f\left(A_i^t, A_j^t, Context\right)$$

2.4 Load observable strategies of other agents into A_j^t

2.5 Calculate costs / effort of new elements for own strategy

2.6 IF elements feasible GOTO 1.1 ELSE GOTO 3.1

In step 2.3 the option for a rule change is prepared. The agent learns about externalities of other agents' action and if the observable actions cannot be implemented by the agent herself (i.e. due to resource constraints) changing the incentives for all agents remains the only options.

3.1 Rule formulation step:

3.1.1 Read knowledge base, identify unachieved goal and check belief matrix for action-object combination with extreme values

3.1.2 Reduce matrix to goal-action and goal-object combinations with same signs

3.1.3 Take most extreme value

3.1.4 Check sign of fitness of goal-verb-object combination and determine *deontic D* (i.e. if big nets leads to highest expected reduction in own profit deontic would be 'must not')

3.1.5 Define *aim I_verb* by reading related action (i.e. 'fish' or 'use')

3.1.6 Define *aim I_object* by reading related instrument (i.e. 'fishing boat')

3.1.7 Define *attribute A* by reading agent properties employing observed stratagies that have negative impact on own payoff

3.1.8 Define *condition C* based on given pre-defined list

3.1.9 Define *or else O* based on pre-defined list

3.1 Rule formulation step:

3.1.1 Read knowledge base, identify unachieved goal and check belief matrix for action-object combination with extreme values

3.1.2 Reduce matrix to goal-action and goal-object combinations with same signs

3.1.3 Take most extreme value

3.1.4 Check sign of fitness of goal-verb-object combination and determine *deontic D* (i.e. if big nets leads to highest expected reduction in own profit deontic would be 'must not')

3.1.5 Define *aim I_verb* by reading related action (i.e. 'fish' or 'use')

3.1.6 Define *aim I_object* by reading related instrument (i.e. 'fishing boat')

3.1.7 Define *attribute A* by reading agent properties employing observed stratagies that have negative impact on own payoff

3.1.8 Define *condition C* based on given pre-defined list

3.1.9 Define *or else O* based on pre-defined list

An example for a fishing context would be that a fisherman learns about decreasing catch and starts observing other agents actions. The fisherman realises that the other agents started fishing with big boats, which is not part of the agent's own action choice. Protecting her own payoff the agent thinks about a rule change and assembles a new rule:

A: Detect attributes by agent types	**All fishermen**
D: Given list for D	must not
I (Verb): Detect actions as I (Verb)	**USE**
I (Object): Detect actions as I (Object)	**BIG NETS**
C: Given list for C	at all times everywhere
O: Given list for O	or else fined $1,000.

This step is followed by a fictitious play in which the agent calculates the impact of a potential rule change.

4.1 Start fictitious play in which agent determines expectation on potential new rule (ADICO combination); IF $payoff_{i,fictitiousplay} > payoff_{i,t-1}$ WAIT until all agents that passed 3 evaluated their ADICO combination; GOTO 4.2 ELSE EXIT

4.2 Diffuse new knowledge in given social structure (i.e. increased fitness of a new rule)

4.3 If new knowledge GOTO 4.1 ELSE increase points IF positive expectations ELSE decrease fitness

4.3.1 Read *I* field

4.3.2 Read goals (individual goals such as profit and potential system goals such as conservation of fish stocks for future generations)

4.3.3 If action (*I*) - Goal combination has extreme negative value THEN support

4.3.4 IF I-G combination has extreme positive value THEN do not support

4.3.5 IF $payoff_{i,fictitiousplay} > payoff_{i,t-1}$ increase fitness of rule ELSA decrease fitness

In case the agent expects the rule change to provide a higher payoff the agent will attempt introducing or changing the rule. This process depends certain properties of the agents (i.e. power) and properties of the system (i.e. network structure and what kind of 'voting'). In order to remain generic the next step would be to diffuse the new idea through the social network in order to form a coalition.

5 Formation of coalitions (new attribute within agent properties) depending on pro_new_rule or contra_new_rule

Depending on certain system properties the negotiations between the two (or more) groups takes place (i.e. in the parliament, on the streets, on the battlefield).

6 Negotiation

6.1 Specify decision making function for any given context (i.e. democratic process 1 stage or 2 stages, corruption based), this can include system goals such as sustainability or employment in goal set of representatives (power nodes)

6.2 Calculate for each coalition the relevant fitness (i.e. votes, votes*power, payment)

6.3 Depending on institutional structure activate new rule

Such a simple sequence resembles one of the possible processes of rules change in the real world. While the steps are kept generic, the contextualisation can lead to very complex implementations, such as voting and lobbying processes under step 6 that might include investment in confirming expectations. This paper implements endogenous rule change in the simplest format. The following section presents a KISS approach that shortcuts several of the six steps describes above in order to focus on the main mechanism that models rule change in an agent-based model.

5 KISS Implementation: EndoRules

5.1 The conceptual model

In this section we explain the conceptual model that **EndoRules** implements. The information provided here should suffice to re-implement the same conceptual model in any platform. The reader may also want to consider following the explanation of the model using it at the same time. For this purpose, we have uploaded an applet of **EndoRules** at http://www.luis.izquierdo.name/models/endoRules. **EndoRules** has been implemented in NetLogo 3.1.4 (Wilensky, 1999). We use bold red italicised Arial font to denote ***parameter*** names.

In **EndoRules** there is a population of ***num-agents*** agents. Events occur in discrete time-steps, which can be interpreted as successive periods of time (e.g. weeks) during which the agents exploit a common resource. In each time-step, every agent must decide the *intensity* with which she will exploit the resource; this individual *intensity*, which can be interpreted as e.g. the number of hours spent exploiting the resource in that time period, is a positive integer no greater than ***max-intensity***. Thus, in each time-step, every agent obtains a certain *payoff* –or benefit– which depends on both her individual *intensity*, and the total intensity with which the resource is exploited by the whole community of agents (i.e. *total-intensity*, which is the sum of all individual *intensities*). In general, the state of the resource at the beginning of a certain time-step may depend on the exploitation carried out in previous time-steps. As we show below, our model accounts for this but, for the sake of simplicity in this paper, we will assume that the resource is fully renewed every time-step.

A valid metaphor for this model could be the exploitation of a certain water body (e.g. a lake or a sea) by a community of fishermen who decide how many hours they spend fishing during a month. The monthly catch that each one of them obtains depends on the length of time that the particular fisherman spends fishing during that month, and also on the total time that the whole community of fishermen spends fishing that month. As mentioned above, generally the catch may also depend on the intensity with which fishing was conducted in previous time periods, but here we assume that the resource is renewed every time-step.

Specifically, the *total-payoff* that the common resource yields in a certain time-step is a quadratic function of the *total-intensity* (i.e. the intensity of the exploitation carried out by the whole community of agents):

total-payoff = *available-resource* · *total-intensity* · (***num-agents*** · ***max-intensity*** – *total-intensity*)

The variable *available-resource*, which denotes the percentage of resource available, is set to 1 at the beginning of every time-step (i.e. the resource is fully renewed every time-step). The *total-payoff* is distributed among the agents proportionally to their individual *intensity*. Under these conditions, agents have incentives to overexploit the common resource. More precisely, it can be shown that the *total-payoff* reaches its maximum value when the *total-intensity* equals ***num-agents*** · (***max-intensity*** / 2), whereas the unique symmetric Nash equilibrium implies that every agent's individual *intensity* is equal to ***max-intensity*** · ***num-agents*** / (***num-agents*** + 1). Hence, for any number of agents greater than 1, the symmetric

Nash equilibrium is suboptimal, and the greater the number of agents, the more inefficient the Nash equilibrium. These results are sketched in Figure 2.

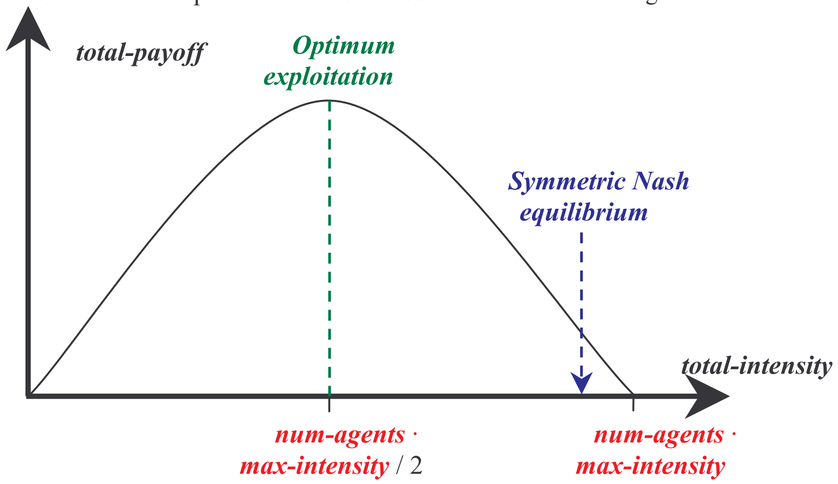


Figure 2. Total payoff (*total-payoff*) that the common resource yields in a certain time-step as a function of the intensity of the exploitation carried out by the whole community of agents (*total-intensity*).

Every agent has a set of possible *intensities* that she is entitled to select. This set, called *possible-intensities*, may be different for each agent and it generally changes over time. At the beginning of a simulation run, each agent's set of *possible-intensities* is a random subset (of random size) of the set of integers comprised in the interval [0 , ***max-intensity***]. The set of *possible-intensities* for each agent may expand or contract over the course of a simulation run due to the emergence of rules which stipulate that certain levels of *intensities* are allowed or forbidden for everyone. An example of an allowing rule would be "All agents may use level of intensity equal to 3 from now on or else a penalty has to be paid", while a forbidding rule would be, for example: "All agents must not use level of intensity equal to 45 from now on or else…". Rules only differ in the verb ("may" or "must not") and in the specific level of *intensity* (e.g. "3") they contain. For simplicity reason, the penalty is assumed to be constant and monitoring is assumed to be 100% effective (uncertainty equals nil).

Rules emerge endogenously in the sense that they are proposed by agents (see below). Only one rule can be imposed over the population in each time-step, and every agent will necessarily follow such approved rule (i.e. agents modify their set of *possible-intensities* accordingly). When there are several rules that have been proposed in the same time-step, the selection of the specific rule to be implemented depends on the number of supporters each of them has. (One single agent may support as many rules as she wishes.) Specifically, the probability of a certain rule being accepted is proportional to the number of agents that support it. The proposals of rules are valid only for the time-step when they originated, i.e. the list of proposed rules is cleared at the beginning of every time-step.

The following subsections describe the remaining details necessary to implement our model in any platform, i.e. the agents' algorithms (a) to select one *intensity* among their set of *possible-intensities*, (b) to decide whether to propose a rule or not, (c) to formulate a rule, and (d) to decide whether to support or not a particular rule. We explain each of these in turn.

Agents' algorithm to select a level of intensity

Agents satisfy in the sense that they keep exploiting the resource with the same *intensity* as in the previous time-step if they obtained a *payoff* that they consider good enough (i.e. no less than their ***aspiration-threshold***, which is the same for every agent) and such level of intensity is still allowed (i.e. it is still a member of the agent's set of *possible-intensities*). If one of these two conditions fails and there are levels of *intensity* in the agent's set of *possible-intensities* that she has never tried out before, then she chooses one of these unexplored *intensities* at random. Finally, if one of the two conditions mentioned above fails and the agent has previously tried out every *intensity* in her set of *possible-intensities*, then she selects the one that provided her with the highest payoff the last time she used it[1]. This algorithm is sketched in Figure 3

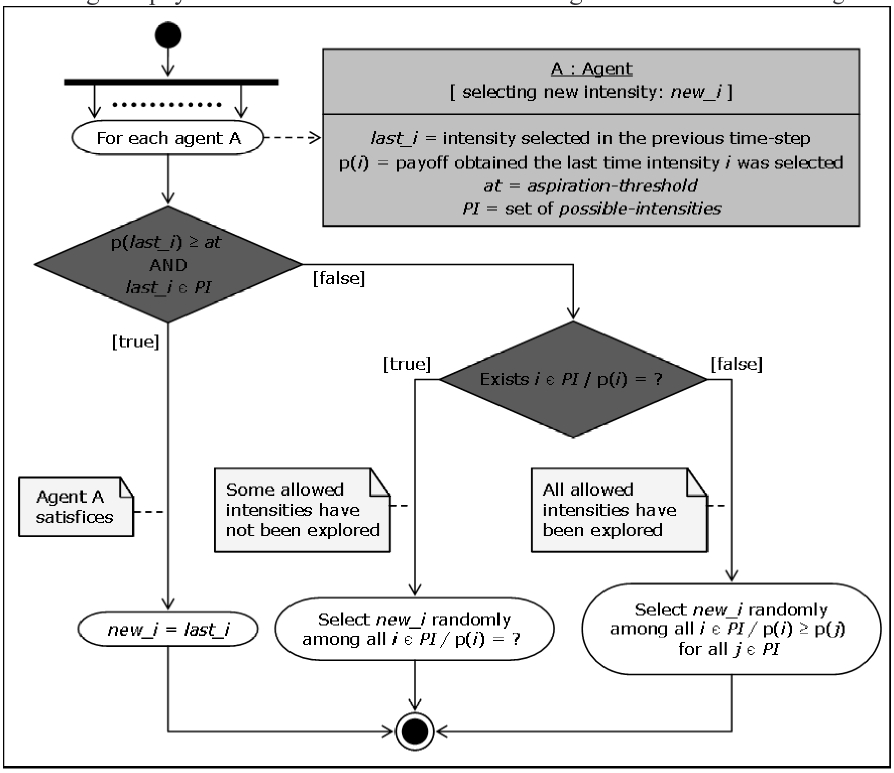


Figure 3. UML Activity diagram of the algorithm agents follow to select the level of intensity to exploit the resource.

[1] In case of tie, the agent will choose randomly among the subset of *possible-intensities* that provided her with the highest *payoff* the last time she used them.

Agents' algorithm to decide whether to propose a rule or not

Agents decide whether to propose a rule right after having received their *payoff*. An agent will propose a rule if she has already explored all the *intensities* in her set of *possible-intensities* and the *payoff* she just obtained is below ***aspiration-threshold*** (i.e. the agent is not satisficing) and also below ***payoff-threshold-to-formulate-rule*** (the value of which is the same for every agent).

Agents' algorithm to formulate a rule

The formulation of new rules is completely random, i.e. both the verb ("may" or "must not") and the specific level of *intensity* in the new rule are randomly selected. The set of *intensities* the agent may choose from when formulating a rule is the set of integers in the interval [0 , ***max-intensity***].

Agents' algorithm to decide whether to support or not a particular rule

Agents decide whether to support a rule or not randomly.

Summary of Parameters

num-agents: Number of agents in the population.
max-intensity: Maximum level of *intensity* allowed.
aspiration-threshold: minimum *payoff* to keep the same level of *intensity* as in the previous time-step (assuming such *intensity* is still allowed).
payoff-threshold-to-formulate-rule: *payoff* below which agents decide to propose a new rule (assuming that all *intensities* in their set of *possible-intensities* have already been tried out before).

The value of ***aspiration-threshold*** and ***payoff-threshold-to-formulate-rule*** can be modified at run-time, with immediate effect on the model. This enables the user to closely interact with the model by observing the impact of changing these values during the course of one single run.

Summary of Scheduling

The following is a summary of the sequence of activities that take place in every time-step:

1. Every agent selects her intensity of exploitation for the period (see Figure 3).
2. The resource is exploited and agents collect their payoff (see Figure 2).
3. Agents decide whether they want to propose a rule or not.
4. Those agents who decided to propose a rule in step 3 formulate the new rule.
5. The number of supporters for each proposed rule is calculated; one of the new rules is imposed over the population of agents.
6. The state of the resource is updated (in this paper, the resource is renewed).

5.2 Displays

EndoRules provides various displays which are shown in Figure 4. Some of these displays are time-series plots showing the historical evolution of the value of a particular variable throughout time (e.g. the plots on the bottom left corner, showing maximum, minimum, and average values of *intensities* and *payoffs* in every time-step), whereas others refer only to the last time-step (e.g. the large square in the middle and the grey monitors on the right hand side).

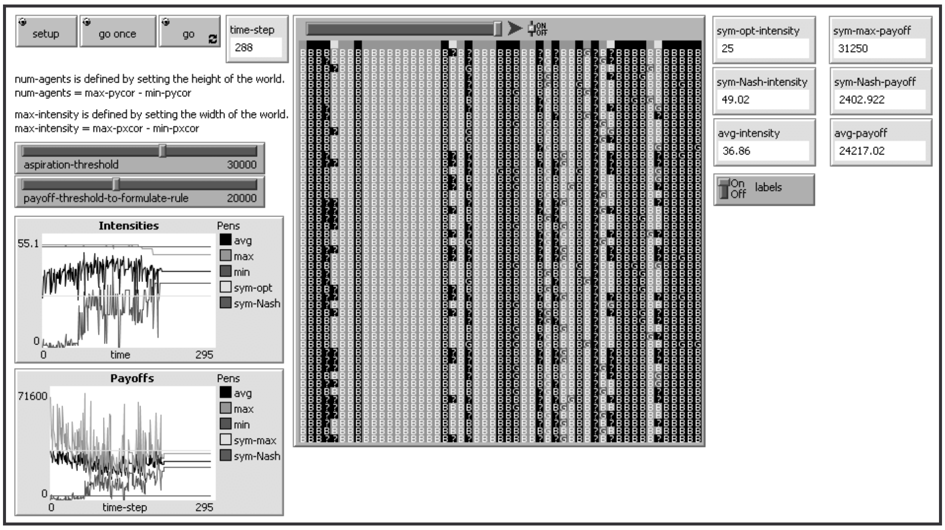


Figure 4. Snapshot of a simulation run in EndoRules.

The large square in the middle of the interface is a representation of the whole set of *intensities* for every agent. Each cell in the 2-dimensional grid represents a certain *intensity* for a particular agent. Each column denotes a certain level of *intensity*: the first column corresponds to *intensity* equal to 0, the second column is for *intensity* equal to 1, and so on until the last column, which corresponds to *intensity* equal to ***max-intensity***. Agents are represented in rows, but the first row is reserved for the institutional environment (see below). The legend of the colours is the following:

- Cells coloured in green indicate allowed *intensities* (*intensities* contained in the agent's set of *possible-intensities*).
- Cells coloured in black indicate forbidden *intensities* (*intensities* that are not contained in the agent's set of *possible-intensities*).
- Cells coloured in red indicate the intensity selected by the agent in the current time-step.

The first row is a representation of the institutional environment, i.e. the rules that are in place in the current time-step. The *intensities* for which there is no (allowing or forbidding) rule in place are coloured in yellow. Green indicates that there is a rule in place that enables agents to use such level of intensity, whereas black indicates that such intensity is currently forbidden.

When the switch named ***labels*** is on, every cell corresponding to every agent displays a label in white text:

- Label "?" indicates that the agent has never used such level of intensity.

- Label "G" indicates that the payoff the agent obtained the last time she used such level of intensity is greater than or equal to ***payoff-threshold-to-formulate-rule***.
- Label "B" indicates that the payoff the agent obtained the last time she used such level of intensity is below ***payoff-threshold-to-formulate-rule***.

5.3 Preliminary Results

For many parameterisations this model will end up in an absorbing state, i.e. a state such that no agent will change her *intensity*. As an example, every simulation run where ***payoff-threshold-to-formulate-rule*** does not exceed the maximum *total-payoff* divided by ***num-agents*** will necessarily end up in an absorbing state. This is so because under such conditions there exists an absorbing state –the symmetric optimal exploitation of the resource– which can be reached from any non-absorbing state in a finite number of time-steps. Some simulation runs where the mentioned condition does not hold may also end up in an absorbing state if some agents satisfice.

While the previous analytical results are indeed useful, they do not tell us anything about the efficiency and equity of the absorbing states where simulations eventually lock in, or about how long it may take to reach such states. To investigate these questions, we have run many simulation runs for different values of ***aspiration-threshold*** and ***payoff-threshold-to-formulate-rule***.

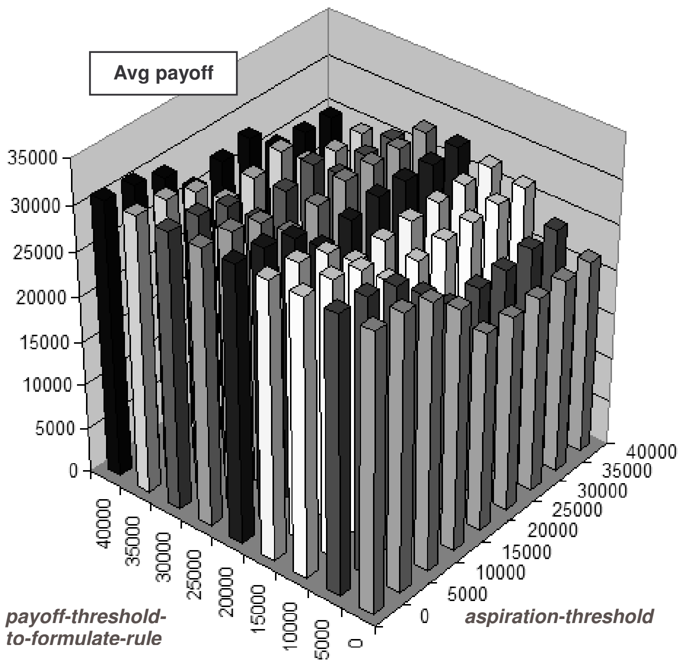


Figure 5. Average payoff obtained by the population at time-step 500 for different values of *aspiration-threshold* and *payoff-threshold-to-formulate-rule*. The value plotted in each column has been calculated averaging over 500 runs.

Figure 5 shows that the average payoff obtained by the population is not very sensitive to either parameter. The highest values are achieved for low aspiration thresholds. This is so because initially agents select their intensities at random, and doing so leads to very efficient levels of exploitation. When aspiration thresholds are low, these (random) efficient levels of exploitation are preserved. If aspiration levels are higher, then some agents will look for better strategies, and this leads to slightly more inefficient (but much more equitable, as we show below) levels of exploitation. The impact of **payoff-threshold-to-formulate-rule** is negligible for low levels of **aspiration-threshold** (since agents who satisfice do not propose new rules); for higher levels of **aspiration-threshold**, efficiency is highest when **payoff-threshold-to-formulate-rule** is below but close to the symmetric optimum level of exploitation, as one would expect.

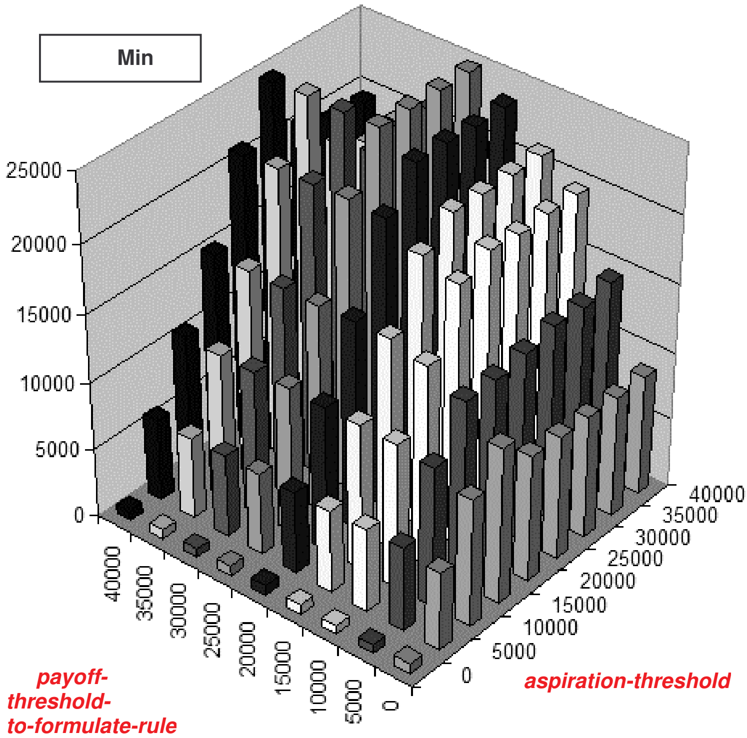


Figure 6. Minimum payoff obtained in the population of agents at time-step 500, for different values of *aspiration-threshold* and *payoff-threshold-to-formulate-rule*. The value plotted in each column has been calculated averaging over 500 runs.

Figure 6 shows that efficiency is only one part of the story. Parameterisations that led to similar levels of efficiency imply dramatically different distributions of the *total-payoff* among the agents. Thus, it is clear that the **aspiration-threshold** greatly influences equity, even though it hardly affected efficiency (particularly for low values). Similarly, **payoff-threshold-to-formulate-rule** has a greater effect on equity than on efficiency, particularly for high values of **aspiration-threshold** (remember that agents who satisfice do not propose rules). Again, the most equitable and efficient results are achieved when ***payoff-threshold-to-formulate-rule*** is below but close to the symmetric optimum level of exploitation, as one would expect.

6 Summary

This paper highlights one of the important gaps of ABM in the context of simulating social systems and offers a first step in finding a solution for endogenising rule change dynamics. This is critical in simulating 'real emergence' in social systems. The paper is accompanied with the EndoRules model that implements most of the sequence developed in this paper. It implements endogenous rule change in the simplest form in order to develop a base model for future improvements that add elements such as punishment, monitoring, negotiations including agent properties such as power and variations in system goals. The model design is still hypothetical and reduced to the critical steps of changes in institutional rules. Due to this focus the KISS approach allows analysing thresholds regarding agents' satisfaction and agents' investment in rule change processes. Testing the generic approach is various contexts will be endeavoured in future steps.

Acknowledgments. The authors wish to thanks Alessio Sebastiano Delre, Panomsak Promburom, and Scott Heckbert for their discussions during the eABM Workshop 2006. We are also grateful for the funding we received from the CSIRO Water for a Healthy Country flagship.

References

1. Axelrod R. Building New Political Actors: A model for the emergence of new political actors. In: Axelrod R (ed.), The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration. New Jersey: Princeton University Press. 1995: 121-144
2. Bromley D. Sufficient Reason: Volitional Pragmatism and the Meaning of Economic Institutions, Princeton University Press. 2006
3. Buchanan JM, Yoon YJ. Generalized Increasing Returns, Euler's Theorem, and Competitive Equilibrium. History of Political Economy 1999; 31:511-523
4. Commons JR. Institutional Economics, Madison, WI: University of Wisconsin Press. 1961
5. Crawford SE, Ostrom E. A Grammar of Institutions. American Political Science Review 1995; 89:582-600
6. Janssen MA. An immune system perspective on ecosystem management 532. Conservation Ecology 2001; 5 (1):13
7. North DC. Institutions, Institutional Change and Economic Performance 734, Cambridge: Cambridge University Press. 1990

8. Ostrom E. Governing the Commons: the Evolution of Institutions for Collective Action
776, Cambridge: Cambridge University Press. 1990
9. Ostrom E. Institutional Rational Choice: An assessment of the Institutional Analysis and Development Framework. In: Sabatier PA (ed.), Theories of the Policy Process. Boulder, CO: Westview Press. 1999: 35-71
10. Smajgl, A. Advancing the capacity to simulate institutional change. The First World Congress on Social Simulation, August, 21-25, 2006, Kyoto, Japan. 2006. (GENERIC)
Ref Type: Conference Proceeding
11. Smajgl A, Larson S. Institutional Dynamics and Natural Resource Management. In: Smajgl A, Larson S (eds.), Sustainable Resource Use: Institutional Dynamics and Economics. London: EarthScan. 2007: 3-19

Representing beliefs as associative networks to simulate the diffusion of innovations

Samuel Thiriot[1,2] and Jean-Daniel Kant[1]

[1] Computer Science Laboratory - University of Paris 6 (LIP6), France
[2] France Telecom R&D
thiriot@poleia.lip6.fr, Jean-Daniel.Kant@lip6.fr

Abstract. A lot of agent-based models were built to study diffusion of innovations. In most of these models, beliefs of individuals about the innovation were not represented at all, or in an highly simplified way. In this paper, we argue that representing beliefs could help to tackle problematics identified for diffusion of innovations, like misunderstanding of information, which can lead to diffusion failure, or diffusion of linked inventions. We propose a formalization of beliefs and messages as associative networks. This representation allows to study the social representations of innovations and to validate diffusion models against real data. It could also make models usable to analyze diffusion prior to product launch. Our approach is illustrated by a simulation of iPod™diffusion.

1 Why representing beliefs ?

Diffusion of innovations is an interdisciplinary field that studies "the spread of new ideas, opinions, or products troughout a society" [1]. Rogers defines diffusion "as the process by which an innovation is communicated through certain channels over time among the members of a social system" [2, p. 11].

Several models were built to study diffusion of innovations, including multi-agent based simulations, with different purposes. *Explicative models* aim to reach a better understanding of how individual interactions make collective dynamics appear. A great part of these models studies the decision/judgment level (adoption, opinion, perceived utility[3], payoff [4], attitude, etc.). For instance, in the threshold model (e.g. [5]), social pressure makes individuals influenced by opinions of their neighbours. Several models also include the beliefs level, that is what individuals trust for a given object (one use "belief" rather than "knowledge" because these beliefs can be false or subjective). It is the case of models focused on informational cascades (see [6] for a review) or in the consumat approach [7]. In these models, beliefs are represented as single values or as a vector of values, and rarely aim to be matched against data collected on the field.

Predictive models aim to produce an estimation of the future diffusion rate of an innovation. The well-known model, and the most used in industry, is the Bass aggregative model [8]. It includes parameters for adoption due to media messages, adoption due to interpersonal communication and an index of market potential for the new product. It permits to reproduct the classical S-curve of cumulated adoption.

Despite of the large amount of litterature about diffusion of innovation, there still remain several problems that are not studied. The first lack resides in *explicative power*. Rogers [2] underlines that models are not able to explain innovation failures (sometimes due to misunderstanding of what innovations are or to incompatibility with beliefs or values). Rogers also remarks that most of the said "innovations" launched in markets are in fact incremental products. In this case people already understand what the innovation is, how it works, so the diffusion becomes quicker. Such processes cannot be modelled without representing beliefs of the population about innovations. The second lack is about *predictive power*. The Bass model can predict the future adoption rate of an innovation only after its launch, based on the adoption data from innovators and early adopters. But at this time, costs are already engaged (for building the product, for communication, etc.). Obvisouly, the predictive interest of the model is highly lowered. So, firms use less formal methods to test new concepts, like interviews or focus groups, which provide some insights on subjective perception and expectations about the innovation. Here again, it seems that modellers cannot avoid to represent beliefs.

Our main concern is to be able to tackle real-world cases. In this paper we study how a modeller can represent individual beliefs in an agent-based simulation. For such a simulation, we need a model for knowledge representation that is complex enough to be explicative and representative, but also simple enough to make its parameters' settings and data collection possible. We illustrate this approach with the simulation of iPod™ diffusion using beliefs collected across forums.

2 Model

2.1 Beliefs as associative links

Individual Associative Network

The concept of associative network has been widely used in social sciences and artificial intelligence to model beliefs: bayesian networks, causal networks, social representations represented as proximity networks, etc. A marketing methodology called the Means-end Chains Theory (MCT) [9] proposes to formalize the perception of products as cognitive chains linking concrete attributes to perceived consequences for the individual and satisfaction of his values. As shown by the MCT, associative networks are relevant to represent the beliefs about products (an example is provided in figure 1). These chains can be retrieved by semi-directed interviews, surveys or stastistical data analysis. Messages like advertisement or consumer reviews can also be represented as chains [10], as shown in figure 3.

Associative networks permit to represent several kinds of knowledge. We categorize knowledge as private, concrete or subjective[1]. The subjective part of

[1] This taxinomy follows the one provided by Audenaert and Steenkamp's studies on means-end chains theory [11], and the discussion in the field of consumer value [12],

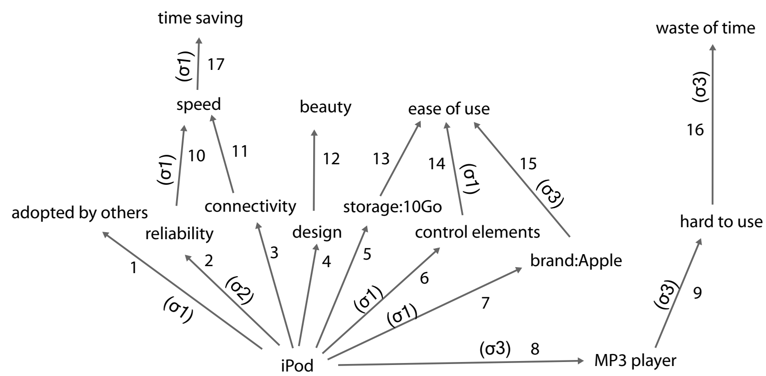


Fig. 1. Exemple of Individual Associative Network (IAN) retrieved by interview for iPod^{TM}(study on what people like or dislike for this product). To improve lisibility only useful supports are provided. σ_1 represents the support "personal experience", σ_2 means "indirect experience" and σ_3 "no credibility"

information is about the innovation itself, like product attributes (links 2-7) and perceived functional consequences of the product (e.g. 11,13). This kind of information is received or retrieved by individuals through mass medias or interpersonnal communication. The private part of beliefs are about individuals themselves. These beliefs are more stable for an individual accross time [13]. For instance, the belief "speed → time saving" is used for all technological innovations. Private beliefs can be heuristics, like "high price → high quality". Private beliefs are provided as initial data by the modeller based on the population segmentation. The last kind of beliefs is about abstract judgments and is built by the individual itself based on its local information, as "product adopted by others". This knowledge is represented in agents by simple computational rules held by each agent.

From the modeller viewpoint, concepts in the model are a finite set $\mathcal{C}$, which is created based on data collection or expert hypothesis. Sometimes two or more concepts are incompatible: an agent cannot trust both of them in the same time for the same social object. As in theory of evidence, we define frames of exclusivity called $\Theta^{\mathcal{X}}$, with $\mathcal{X} \subset \mathcal{C}$. Some exemples of frames are: (solid, breakable), (good connectivity, bad connectivity).
Formally, we define knowledge as directed associations between concepts. Mathematically, a belief is a binary relation in $\mathcal{C}^2$. $C_1\ b_\sigma^{a,t}\ C_2$ is the conviction held by an agent $a \in \mathcal{A}$ at time t that two concepts (C_1 and C_2) $\in \mathcal{C}^2$ are associated with a given support $\sigma \in \Sigma$. The support represents the confidence of the agent on this belief (more details on support are provided below). In this model, existence of a link represents belief. No link means ignorance. Disbelief is modelled as the

which concludes that perceived value depends both on the intrisic product properties and on the subjective perception of consumers belief in the opposite concept. Each individual possesses his own set of beliefs; we name this set an Individual Assocative Network (IAN).

Some concepts are considered as object of interest by the agents $\mathcal{A}$ (agents will speak about them, they want to understand them, they can take decisions about these concepts). We use a psychosocial term [14] to design these objects of common interest: these concepts are *social objects* $\mathcal{O} \subset \mathcal{C}$. When we model the diffusion of innovations, social objects are innovations. A set of beliefs about a social object o forms the *representation* $R_o^{a,t}$ of this object. This representation is the subgraph rooted in the social object. If a representation is shared between several agents, it becomes a *social representation* in the social psychology meaning, noted $SR_o^{\mathcal{X},t}$ with $o \in \mathcal{O}$, $\mathcal{X} \subset A$.

Beliefs Revision

Insights about persuasive communication are provided by social psychology [14]. Persuasivness of a communication depends on properties of the source like credibility, expertise, self-interest, structure of argumentation, messages order, etc. No formal model exists to compute the total persuasiveness of a communication based on these parameters. However, several formalisms are available to represent beliefs and their strength, mainly with probabilities or belief functions (see [15] for a comparative review). But, all of these models are normative and lead to results incompatible with observable evidence. They would require us to include quantitative valuation of beliefs (as probabilities or belief masses), which would make the model harder to validate, less representative and harder to manipulate. So, we developped a solution based only on the qualitative properties of beliefs.

The sources of informations are perceived as more or less credible by individuals. Broadly speaking, personal experience is stronger than other advices, themselves stronger than advertisement. We define a set Σ that contains several levels of support (in other words: credibility, certainty, revisability, strength). Each source of information is categorized by the agents in one of these levels. Levels are defined operationally to fit observations from the population and the needs of the model. Currently we work with the following levels: *no credibility* is used for information from advertisement, *plausible* is used for advice from someone, *indirect experience* represents feedback of someone based on its personal experience. *Personal experience* represents the strongest level for beliefs acquired by the agent direct experience.

We assume that a stronger source erases the previous advice, because the new source is considered to be more credible. In some cases, however, it is possible for a strong belief (acquired by direct experience) to be modified by new weakly-supported information, because individuals accept to revise old beliefs, comply with social consensus, can be convinced by a good argumentation or another reason. That's why we choose to model belief revision based on probabilities of revision between support categories $p(revise|\sigma_{old}, \sigma_{new})$. We built this

	no credibility	plausible	indirect experience	personal experience
no credibility	0.9	0	0	0
plausible	1	0.9	0.01	0.001
indirect experience	1	1	0.9	0.001
personal experience	1	1	1	0.9

Table 1. probability of revising a belief based on the support level of the previous belief σ_{old} (top) and on the support level of the new information σ_{new} (left column)

function (table 1) based on qualitative observations. A weak support has a low probability to modify a stronger support. However in long term, this probability becomes higher and higher, leading to invalidate old beliefs. This model is easier to validate than a quantitative representation of strength.

Retrieving from memory

We need to be able to retrieve the representation of a social object contained in an IAN. Retrieving a representation is an spreading activation process: start from the social object, then browse all the links connected to this node to build the representation $R_o^{a,t}$. We assume an activation propagation inspired by evidence networks: *the activation strength of a concept for an object is the strength of the weakest link in the chain that links the social object to this concept.* When activation follows a link, activation is filtered by the belief strength. For instance in figure 1, activation of concept "time saving" for the social object "iPod" is "indirect experience", which is the lowest support in the chain (σ_2). If a node receives several levels of activation from its parents, the stronger activation is kept (MAX-activation, which is also an OR logical interpretation). In the example of figure 1,"ease of use" has a support of "personal experience". As a result, the *activated representation* contains the beliefs activated and their support.
In the particular case of incompatible beliefs, the activation process only keeps the strongest belief. For instance in figure 1, the frame of exclusivity $\theta_{complexity} = \{\text{ease of use, hard to use}\}$ forbids both of these concepts to be trusted at the same time. The spreading activation process sets a low activation to "hard to use" and an higher to "ease of use", so only the last one will be included in the activated representation.

2.2 Communication

As shown before by the agent-based modelling community, the social structure has a huge impact on the system dynamics (e.g. [16]). A model of communication, that it too simple - like random meeting or cellular automata - doesn't seem adequate. As a consequence, we detail here explicitely the channels that

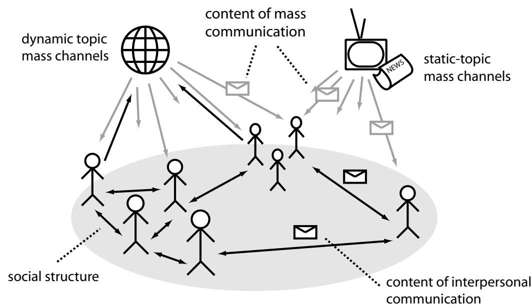


Fig. 2. Communication in a real population

support communication, the structure of messages themselves, and the topics (social objects) agents are talking about.

A channel is a support of communication that transmits information from an information source to an audience. Historically mass media were controlled by firms for persuasive communication, while interpersonal channels were only used for uninterested communication. Today individuals' reviews through specialized websites or forums could challenge traditional mass media, and interpersonal communication begins to be modified by individuals who are paid to propagate positive recommandations. To take this evolution into account, we propose to categorize channels based on their audience size and the determination of topics (table 2). An unidirectional channel will always have a static topic (because the information source communicates about the object of its choice) while bidirectionnal channels allow interactive choice of topic. Modelling interactive topics implies modelling information research, and not only passive information reception.

	big audience	small audience
interactive topic choice	forums, search on internet	face-to-face
static topic	press, advertisement, direct experience	weblogs

Table 2. taxinomy of channels

A *mass channel* is connected to a great number of agents. The agent exposure defines its probability to receive messages through this channel. An *interpersonal channel* represents the fact that two individuals can exchange information with a given exposure parameter (the probability for the agents to meet). A static-topic channel will only transmit passively messages, so the topic is determined by the information source. An interactive-topic choice channel asks both agents which topics they want to discuss about (the salient social objects set of each agent) and pick up randomly a social object in the union of the two sets.

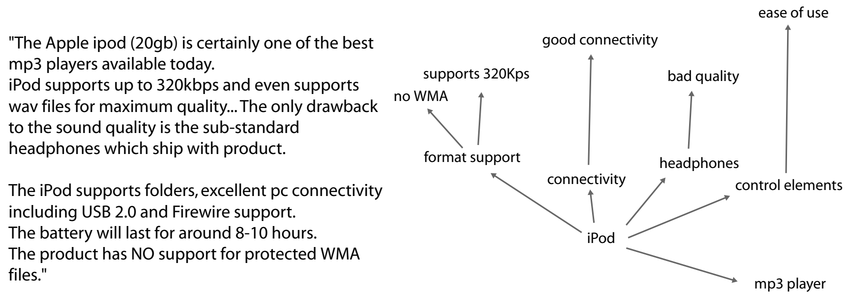


Fig. 3. Transcription of a consumer advice retreived on a website (left) as a TAN (right)

Messages

Each transmission of information (either from mass media or interpersonal) is a message. A message is intended to transmit information[2] about a social object. A message is sent by a sender over a channel; audience will be determined by the channel itself. A communication campaign is composed of several messages broadcasted on channels during a given period.
The content of a message is a *transmissible associative network* (TAN), which is made of associative links (see figure 3 for example). A TAN typically embodies only a representation of a single social object. Sometimes - especially in the case of co-branding - the network can include several social objects and their associated representations. A TAN transmitted by an extrinsic information source is provided by the modeller. A TAN from an intrinsic information source is dynamically built by this agent.

2.3 Agents

A consumer agent represents a unit of adoption. It embodies a belief base, a list of currently salient social objects and is linked to an agent profile. An agent profile contains the default exposure to mass channels, background knowledge, subjective production of knowledge. It also contains functions to evaluate attractivity and decide adoption. It can also include some rules to create the subjective knowledge based on local information. For instance, the fact that others have adopted a product (belief number 1 in figure 1) is modelled by a threshold on the observed relashionships that possess the product.
The definition of the agent's behavior is out of the scope of this paper. The modeller can implement whatever models he wants based on the internal representation of beliefs, which provides both beliefs and their strength. Several

[2] This definition of a message is voluntarily simplified to fit the frame of this paper. A message, especially an advertisement, also embodies some non-semantical components.

models exist to describe attitude formation or adoption based on beliefs, as the Theory of Planned Behaviour [17], the Fishbein model or any multi-criteria model.
As an example we currently use the behavior process represented in figure 4. We designed multicritera functions to compute attractivity and adoption, which take into consideration the support of beliefs.

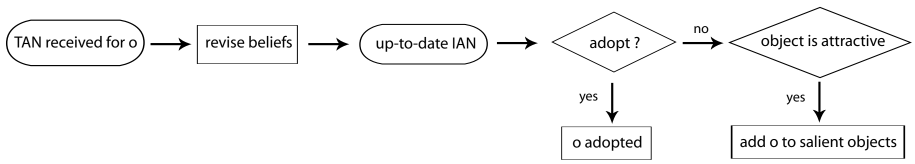


Fig. 4. Exemple of decision process

Based on these three functions, the following process appears, which is compliant with existing models of buying steps or adoption process [2]. (1) first the agent becomes aware of the innovation, and receives prior information. (2) if the information is attractive enough, the agent decides to look for it (3) if the agent thinks it has enough information, it decides to adopt or not (4) using the product, it receives more information by usage and participates to word-of-mouth.

3 Application to iPod$^{\mathrm{TM}}$

This simulation is provided as an illustration of our approach. The model is implemented with the Repast Framework (http://repast.sourceforge.net/). In this discrete-time simulation, each step represents one week.

3.1 Data collection

We retreived data from the published means-end chains analysis of iPod$^{\mathrm{TM}}$ [18] and from statistical analysis of reviews provided by consumers on specialized websites. This data is used to determine the content of interpersonal messages and to insert background knowledge into agents. Associative networks permit to represent background knowledge. For instance in figure 1, the links (9,16) represent fears of late majority about technology: it's hard to use and leads to waste of time.
We identified the following static-topic mass channels: TV advertisement, generalist and specialized press, experience with the product. We set exposure to each media from general statistics published about TV ads exposure, press reading, etc. We used as a social structure a small-world graph (a regular lattice with shortcuts as proposed in [19]). The exposure level to social interactions is retreived from a study [20] about word-of-mouth, which quantifies on average 15 word-of-mouth episodes per week.

3.2 Agent profiles

We adopt the classical segmentation used in diffusion of innovations. *Innovators* like what is new, fun. They enjoy to spent time to learn how an innovation works. They are able to understand technological terms. They read specialized press nearly ones a week. They are more impulsive than others, and can adopt an innovation as soon as it is available. They easily speak about innovations. They like to be alone to possess new things, and an innovation already possessed by others loose its interest. *Early adopters* sometimes read specialized press. They like new thinks, they carefully study available information before buying. Individuals from *early majority* like to be on-trend, with new products. They already have a good knowledge about technology, but like to have feedback from first adopters before buying. *Late majority* don't cares about the novelty of a product. They focus on the utilitarian aspect, don't like to loose time to learn new techologies. As part of their background knowledge, they believe that technological innovations are hard to use (as represented by beliefs 9,16 in fig 1). They consider a piece of information as true only if it comes from someone else with direct experience. *Laggards* have a low exposure to press, and retreive most of their information from interpersonal communication.

3.3 Simulation

Graph 5 shows the output of the model. Awareness starts before adoption due to annoucement information transmitted about iPod$^{\mathrm{TM}}$. Because an announcement is only transmitted in specialized press, mainly innovators and early adopters are aware of the product and can propage word-of-mouth around them. Then the product is launched, with information in generalist press and TV advertisement. All the population becomes aware of the product and can adopt it. Early majority requires indirect information from previous users or independant reviews to adopt. Late majority needs indirect feedback to adopt. The last curve in this figure shows that the diffusion is made quicker if another media (here: internet) permits to retrieve others advices quicker than face-to-face communication; this media is highly efficient because it permits to determine interactively topics and to retrieve credible information.

3.4 Observations

How to improve diffusion ? In this model advertisement on its own doesn't lead to adoption, but can make the product salient in individual minds and provoque adoption or word-of-mouth. The best idea to make diffusion quicker is to facilitate word-of-mouth, which is required to persuade late majority and laggards to adopt. A good timing, and attractive information, is require to stimulate word-of-mouth. If new information is sent when individuals are still looking for information, then this new information will be transmitted quickly through interpersonnal communication.

Observability, one of the factors mentionned by Rogers, also facilitates diffusion

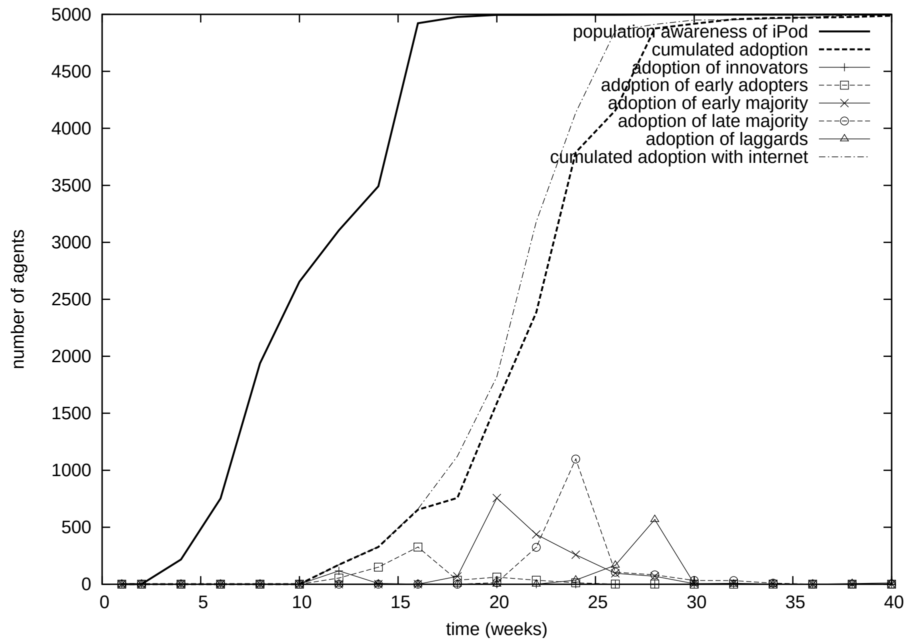


Fig. 5. Simulation of iPod$^{\mathrm{TM}}$diffusion in a population of 5000 agents

in this model. In the case of iPod$^{\mathrm{TM}}$, the white hearspeakers are easily identifiable, and are related to iPod$^{\mathrm{TM}}$based on advertisement campaigns. So potential adopters are aware of others adoption, leading them to follow this indirect recommandation. The importance of usage value, as in reality, is confirmed, because individuals who use the products are highly credible and can provoque adoption; it is of prime importance that they be satisfied by the product.

Social representations of the innovation appear in the model. In the beginning of diffusion, we can observe different social representations in the population (figure 6): individuals who have already adopted possess a large amount of information provided by experience, while others only have a representation created from advertisement.

Individuals who had no knowledge about mp3 players discover through word-of-mouth what are the criteria for evaluating the innovation. While late majority is initialized with no knowledge about mp3 players, all individuals end with general considerations about autonomy or storage capacity. We also observe examples of incomprehension: an individual who has no knowledge about storage capacity is unable to understand what "10Gb" means, but he will learn it through interpersonal messages or well-designed advertisement (with the slogan "1000 songs in your pocket").

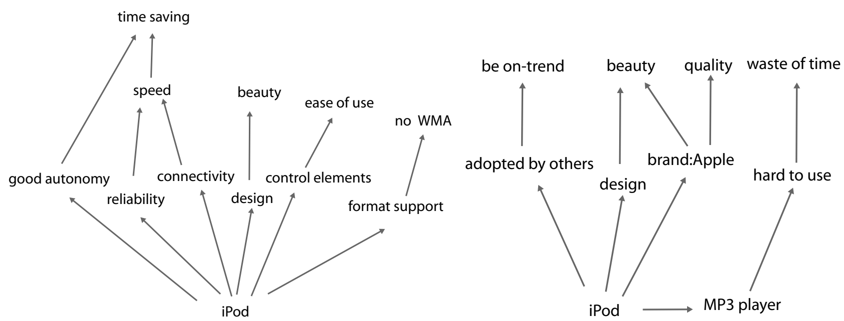


Fig. 6. Example of social representations. The left representation $SR^{earlyadopters,12}_{iPod}$ is shared by first adopters who have already an experience with iPod^{TM}. The right representation $SR^{latemajority,12}_{iPod}$ is held by several agents who just have information from advertisement, and know that iPod^{TM}is already widely adopted.

4 Discussion

Representing knowledge as associative networks permits to create models which can be tested against real datathat fit real data, and to represent both messages and individual knowledge in a computationnally tractable way. This representation is highly representative and manipulable, even for non-experts. Implicitly it allows to model misunderstanding of information, word-of-mouth or launch of related innovation in a more plausible way - in fact models that were expected by Rogers. Hence we could build models that represent the whole adoption process, from awareness to decision.

When a diffusion model is built to be used as a decision-support system, this approach is obviously more instructive. Through simulation the modeller is able to study the true parameters of diffusion of innovation (those mentionned by Rogers and used by marketers): what is the perception of products ? In what way are consumers aware of a product ? What is the background knowledge of individuals, and will they be able to understand information ? Why does an innovation provoke word-of-mouth ? Used before innovation launch, the model can be parameterized from interviews - for subjective perception of the innovation - and general information- for background knowledge -, giving one an efficient methodology to test possible diffusion cases.

The main limitation of such a model is the state of knowledge about human behavior and social phenomena: no sufficient information is available about the structure of real social networks, what provokes word-of-mouth, etc. Our future work will be focused on the validation of models based on associative networks, including interview protocols and statistical methods.

References

1. Valente, T.: Network Models of the Diffusion of Innovations. Cresskill NJ: Hampton Press (1995)
2. Rogers, E.M.: Diffusion of Innovations. 5th edn. New York: Free Press (2003)
3. Ellison, G., Fudenberg, D.: Word-of-mouth communication and social learning. The Quarterly Journal of Economics **110** (1995) 93–125
4. Banerjee, A., Fudenberg, D.: Word-of-mouth learning. Games and Economic Behavior **46** (2004) 122
5. Deffuant, G., Huet, S., Amblard, F.: An individual-based model of innovation diffusion mixing social value and individual benefit. American Journal of Sociology **110** (2005) 1041–1069
6. Roberts, J.H., Lattin, J.M.: Disaggregate-level Diffusion Models. In: New-Product Diffusion Models. Springer (2000) 207–236
7. Janssen, M.A., Jager, W.: Simulating market dynamics: Interactions between consumer psychology and social networks. Artifical Life **9** (2003) 343–356
8. Bass, F.M.: A new product growth for model consumer durables. Management Science **15** (1969) 215–227
9. Reynolds, T.J., Gutman, J.: Laddering theory, method, analysis, and interpretation. Journal of Advertising Research (1988) 11–31
10. Reynolds, T.J., Gengler, C.E., Howard, D.J.: A means-end analysis of brand persuasion through advertising. International Journal of Research in Marketing **12** (1995) 257–266
11. Audenaert, A., Steenkamp, J.: Means-end chain theory and laddering in agricultural marketing research. In: Agricultural marketing and consumer behavior in a changing world. Boston: Kluwer (1997) 217–230
12. Holbrook, M.B.: Introduction to consumer value. In: Consumer Value : A framework for analysis and research. Routledge, New York (1999) 1–28
13. Page, A.L., Cox, D.N., Russell, C.G., Leppard, P.I.: Assessing the predictive value of means-end-chain theory: an application to meat product choice by australian middle-aged women. Appetite **44** (2005) 151–162
14. Moscovici, S.: Psychologie Sociale. 7th edn. Presses Universitaires de France (1998)
15. Smets, P.: Probability, Possibility, Belief: Which and Where ? In: Handbook of Defeasible Reasoning and Uncertainty Management Systems. Volume 1. Kluwer, Doordrecht (1998) 1–24
16. Valente, T.W.: Network models of the diffusion of information. Hampton Press Inc., Cresskill: NJ (1995)
17. Ajzen, I.: The theory of planned behavior. Organizational behavior and Human Decision Processes **50** (1991) 179–211
18. Reppel, A.E., Szmigin, I., Gruber, T.: The ipod phenomenon: identifying a market leader's secrets through qualitative marketing research. Journal of Product and Brand Management **15** (2006) 239–249
19. Newman, M.E.J., Watts, D.J.: Renormalization group analysis of the small-world network model. Physics Letters A **263** (1999) 341–346
20. Carl, W.J.: What's all the buzz about? everyday communication and the relational basis of word-of-mouth and buzz marketing practices. Management Communication Quarterly **19** (2006) 601–634

Getting away from numbers: Using qualitative observation for agent-based modelling

Lu Yang and Nigel Gilbert

Centre for Research in Social Simulation, Department of Sociology, University of Surrey, Guildford, Surrey, GU2 7XH, United Kingdom
{l.yang, n.gilbert}@surrey.ac.uk

Abstract. Although in many social sciences there is a radical division between studies based on quantitative (e.g. statistical) and qualitative (e.g. ethnographic) methodologies and their associated epistemological commitments, agent-based simulation fits into neither camp, and should be capable of modelling both quantitative and qualitative data. Nevertheless, most agent-based models (ABM) are founded on quantitative data. This paper explores some of the methodological and practical problems involved in basing an ABM on participant observation and proposes some advice for modellers.

Keywords: agent-based models, qualitative data, ethnography

The right question isn¨t, does the number mean anything. The right question is, does the number correspond to a difference that makes a difference in the kind of world being modelled.[1]

1. Introduction

Most agent-based models (ABM) are intended to simulate some real-world phenomena and are therefore designed and validated by referring to data collected from the social world. Most often, such data are in numerical form. For example, the data may record the number of agents over time, obtained from organisational or government records; agents¨ opinions about some issue, derived from attitude surveys; or the number of transactions and their prices obtained from market records. There are well developed statistical techniques[1] for summarising and comparing such data with the outputs of simulation models.

Unlike most other modelling approaches, there is nothing inherently quantitative about agent-based modelling. It should be as easy to develop and validate ABMs using qualitative data as it is with quantitative data. Nevertheless, this is far less often done (a noteworthy exception is Agar [1, 2]). In the next section, a distinction is made between two types of qualitative data, one that can be used in almost the same

[1] Although the best known techniques make assumptions about the statistical distribution of the data that may not be appropriate for many agent-based models, e.g. those that generate power-law distributions.

way as quantitative data and one that for methodological reasons must be treated quite differently. Section 3 briefly describes an example of building a model based on qualitative data and section 4 considers the extent to which ethnographic data matches the ABM methodology. Section 5 concludes with a few tentative guidelines for using such data for agent-based modelling.

2. Qualitative data and ethnographic research

At its simplest, qualitative data cannot be converted into numeric values without distorting the information they contain. There are two types of such data and they are fundamentally different in nature. First, there are data that are collected using social measurement techniques such as surveys. The data may be gathered using nominal (e.g. black, white, green) or ordinal (e.g. agree, neutral, disagree) scales, rather than interval or ratio (i.e. numerical) scales[3], but, like conventional quantitative data, the measurements are made on variables with one value for each respondent and can be stored and analysed as a variable/actor matrix. Such qualitative data can be collected, managed and analysed using methods similar to those developed for quantitative data, with the exception of statistical techniques that require arithmetic operations on the values[4, 5]. For instance, with a categorical scale such as ¨political party voted for in the last election¨, it would be inappropriate to calculate the mean vote, although it is still possible to find the modal (most frequent) party voted for. Because categorical and ordinal data can be used in the construction and validation of agent-based models in much the same way as quantitative data, this type of data will not be further considered here.

The idea of qualitative data also encompasses an entirely different type of data (and the common use of the term ¨qualitative¨ to cover both often leads to confusion). These are data that are generally textual in form, such as fieldnotes obtained from observation, transcripts of interviews, and published documents, although they can sometimes consist of images, video or audio. Such data are not readily converted into a variable/actor matrix without losing information or doing an injustice to the data. The important feature of such data is that they are about the meanings that participants ascribe to events and actors. To distinguish such data from the categorical data mentioned above, they will be called ¨ethnographic¨ data[6].

Ethnographic data are often collected and analysed on the basis of epistemological and ontological assumptions that are significantly different from those typically adopted by quantitative researchers. At first sight, these assumptions may seem to get in the way of their application to the design and validation of agent-based models, but in fact the difficulties are not as great as they appear. The basis for most ethnographic data collection[2] includes:

1. Epistemological orientation: ethnographic researchers often espouse naturalism ¨ seeking to understand the particularities of a social phenomenon in its own

[2] Inevitably, not all ethnographic, quantitative and agent-based studies exactly match the descriptions suggested in this list, but the majority do.

terms in rich detail ¨ while quantitative researchers often, although not always, adopt a covering law model of explanation seeking deductive laws that are of universal or near-universal application. Agent-based modellers tend to adopt a middle way, expecting their models to have some degree of generality, but rarely proposing that the models could have universal application (although some game theorists come close to that ambition).

2. Causation: While quantitative research is generally concerned with identifying correlations between variables, ethnographic studies are more interested in causal mechanisms, that is, ¨by what intermediate steps, a certain outcome follows from a set of initial conditions¨[7].
3. Ontological perspective: Ethnographic studies are frequently based on constructivism[8], which considers that social actors continually (re-)construct their social world (rather than it being objectively available to them *a priori*). This contrasts with the positivist or, more frequently, realist orientation of most quantitative and agent-based modellers[9]. Critical realism[10] takes the approach that there are some ¨real¨ social processes or mechanisms existing that generate the observations that actors and researchers perceive.
4. Methodological orientation: ethnographic researchers generally adopt an inductive approach, in which theoretical categories are generated from consideration of the data (e.g. the procedure called ¨grounded theory¨[11]). In contrast, quantitative research usually involves testing hypotheses and, at least in their writings, modellers tend to adopt a deductive approach, in which a simulation, designed on the basis of some prior theory, is tested to see whether it conforms to observed data.
5. Importance of generalisation: Many ethnographic researchers adopt an idiographic stance, that is, aiming to understand the particular and contingent character of the case they are studying, and tend to be sceptical about the possibility and the value of generalisation to universals. Agent-based modellers often assume that their models express some general patterns and structures, although they will be instantiated with particular parameters to represent some specific case[3].
6. Emic versus etic orientation: an emic description is one that is formulated in terms that are meaningful to the actor[13]. Typically, ethnographic data are emic, in contrast to etic data which are formulated using concepts meaningful to the researcher, but not necessarily to the actor. Both quantitative theories and agent-based models often use concepts (for example, power law, evolution, utility) that are unlikely to be meaningful to those whose behaviour is being modelled[4].
7. Process and context: there is a strong emphasis in most ethnographic research on process, that is, on how and why things change over time. Mainly because it is often difficult or expensive to collect time-varying quantitative data, there is less emphasis on process in most quantitative studies. In this respect, agent-based modelling, in which the passage of time is represented by simulated time-

[3] For example, a ¨history-friendly¨ model[12].

[4] But this is not always true: see the Companion Modelling movement[14] and participatory modelling more generally.

steps, is closer to ethnographic than to quantitative research. Ethnography also emphasises context: that the meaning of actions depends on the context in which they are carried out. This is also a usual element in ABM, where agents are given rules of behaviour that are context specific (i.e. whether and how a rule is activated depends on the current state of the agent and its environment).

These similarities and differences between agent-based modelling and ethnographic research are summarised in Table 1 and illustrated in the next section, which describes research on the socialisation of members into service work teams.

	Quantitative	Agent-based	Qualitative
Epistemological orientation	Deductive-statistical laws		Inductive laws
Causation	Correlational	Mechanism	Process
Ontology	Positivist	Realist	Constructivist
Methodology	Hypothesis testing	Model validation	Grounded theory
Generalisation	Seeks general laws	Seeks middle range models[15]	Idiographic
Emic/etic	Etic	Etic	Emic
Temporal and context emphasis	Cross-sectional and context-free	Time important	Context-sensitive

Table 1. Characterisation of three types of social research

3. An example

Lu Yang has been conducting a study of the socialisation of newcomers into existing organisational groups, and examining how the entry of a new member reshapes the members¨ interactions and the structure of the group[16]. Data were gathered by participant observation of two service workgroups: one in a library where the group task was to record book loans and returns, stack returned books and perform routine administration; and the other in a nursing home, where the group task was to carry out personal care for a number of bed-ridden patients. Each observation was conducted for a period of six months by participation in the work group as a new team member. The study generated copious and detailed field notes recording the researcher¨s observations and providing direct evidence of the activities of members and changes in group behaviour.

The ¨translation¨ of these fieldnotes into a rule-based computer program involved two stages. First, a narrative account of the socialisation process was constructed. This began with a general overall picture of the setting, the culture and the members¨ everyday procedures. For example, the Nursing Home was situated in the countryside of southeast England. It is privately owned, medium in size, and the residents who live there are old and with various clinical needs. The functioning of the Nursing Home, including details such as how many nurses and carers were employed, what kind of work they engaged in and the procedures they used, were described.

Then, a number of themes were identified and illustrated with stories about specific events and incidents that were observed. The observation data were used to discuss, for instance, the high employee turnover rate, variations in the quality of service, the boundaries of friendship groups, the hostility of existing staff towards newcomers, etc. Explanations of these features of life in the Nursing Home were formulated in terms of themes, for example, that the care workers established their identities partly by reference to their ethnicity, that co-workers often got angry with others due to task conflicts, and that new carers learnt practical skills not from formal training but from working with more experienced carers.

Much ethnographic research stops at this point, with a plausible narrative that aims to explain the observations by reference to descriptions of social processes. However, such descriptions tend to be vague and difficult to validate. We wished to take the analysis a stage further by constructing an agent-based model that represented abstract versions of these processes.

The second stage of the analysis therefore consisted of the design and implementation of an agent-based model that encapsulated some of the essential features of socialisation observed in the two settings. As recommended by Dey[17], qualitative data analysis ¨involves a process of abstracting from the immense detail and complexity of our data those features which are most salient for our purpose¨. The ABM holds the promise of helping to identify which features are most salient, ensuring that the extracted features are consistent and coherent, and providing a way of validating our conclusions by running the simulation. A recurring problem, however, was how to model aspects of the settings without making them more precise or concrete than the observational data would warrant (compare Whitehead¨s ¨Fallacy of Misplaced Concreteness¨ [18]).

The model, implemented in NetLogo[19], was designed to have the following features:

1. Agents work on tasks. Agents and tasks are each represented as objects (that is, as ¨turtles¨ in NetLogo). A decision had to be made about how many of each type of object to create. For example, this involved ¨abstracting¨ from the number of carers that were employed in the Nursing Home (the actual number varied from one week to the next and by shift) to a constant number of agents (that could varied from 1 to 9 between runs) and categorising the activities of the carers into precisely 10 tasks.
2. Agents have attributes. It was observed that people¨s demographic characteristics (e.g. age, gender, race and ¨aggressiveness¨) were salient in their relationships with other team members. In the model, each characteristic is represented by the value of an agent attribute. The attributes are used to calculate the initial attraction that each agent has to the others. There are many social psychological studies that support the proposition that demographic attributes affect social integration (e.g. [20-22]). However, neither these studies nor our observations are able to quantify the degree of attraction given by these attributes, nor do they say anything about the attributes¨ relative importance or

the functional form of the relationship. Consequently, an arbitrary choice was made (a function based on the Pearson product-moment correlation coefficient) and a sensitivity analysis conducted to see whether variations made a substantial difference to the outcomes (it did not).

3. Tasks have a priority level and a completion time. Since there are a maximum of 10 tasks at any time, the priority level ranges from 0 to 9. When applied to the Library service team, priority level 9 represents frontline service; priority 8 means sorting book trolleys, and so on. The rank ordering of priority was based on what had been learned from observation of library work. The time taken to complete each task and the increase in speed at which a task is performed when more than one person works at it were also estimated from the observations. However, while in the model all tasks were assumed to take the same amount of effort, in reality different tasks took different times and the rates at which they were carried out depended on many factors beyond the number of workers assigned to them.
4. Agents that leave are immediately replaced by new, unsocialised agents. Observation confirmed that a new employee will be recruited to replace a worker who leaves. For simplicity, the model was designed so that one worker left and was replaced after a fixed number of time steps; in the settings, the timing of resignations was irregular.
5. Agents have ¨emotions¨. The field notes showed that emotion played a key role in the socialisation of new team members in both settings. A typical instance is shown in Figure 1, an extract from the observer¨s notes.

Day 10:
Markus started to arrange the working group.

'Gayle group 1, Flora group 2, Cora group 3, Shana group 4, Abigail group 5 and Lu group 6.' Gayle left the room.

'YOU CAN NOT DO THIS TO ME!' Abigail sat straight up, stared at Markus, and shouted at him. (I know, of course, it's me! Group 5 and 6 work in the wing of the home, which is the part that has the most resident care needs. This part is considered to be the hardest part of the job. It's normally the last part to finish every morning. The morning washing and dressing job should be finished by 10:30am, maximum 11am. Sometimes, if other groups finish earlier, they normally come and help the wing groups; however, they are not obliged to do so.) Markus didn't say a word, but looked pained. I can see he's trying to think how to deal with this. (Well, I don't think he has any choice though.) Then you heard different voices rose up, cross-netted (mainly Cora, Abigail and Flora), I had no clue what exactly they were talking about, the only thing I knew was that it concerned the groups! (I have no right to say, or, maybe more accurate, better not to say anything), I kept silent, but looked at them shouting to each other. (I felt sorry for them, but well, I am new, I am new! And that's the fact; I cannot change it in one night! I need to learn and learning takes time, simple as that. I'm already trying my best.) Finally, there ended up with a solution, Flora turned to me,

'Lu, there is a change about the groups, now, you are with Gayle upstairs and I'm with Abigail on the wing. OK?'

'Ok.' What else I can say. Abigail's still looking angry at Markus. (poor him! Sorry.)

Fig. 1. Extract from field notes recorded on the tenth day of participant observation at the Nursing Home. (All participants¨ names have been changed except for the observer¨s (Lu))

This was represented in the model by formulating rules of interaction such that if any one agent displays certain emotions, other agents check their attraction value towards this agent. If an agent decides to respond to the one who displays emotion, the attraction matrix is updated. Unfortunately, however, it is not clear how to decide on how much the attraction value should be increased. Nor is it easy to settle on a value for the emotion display threshold. There is no theory or empirical basis for these and other numerical manipulations.

Despite our efforts to reduce the number of unsupported numerical constants and arithmetical operations within the model, it did not prove possible to eliminate many of them. One could ¨cheat¨, removing numbers by tricks of coding, or increase the apparent quantification by using numerical instead of symbolic values for all constants. While the presence of unjustifiable, ¨magic¨ constants in the code were the clearest examples, they were not the only instances where the model had to be made more definite than the ethnographic record could justify. For example, in the model, agents that leave are immediately replaced (i.e. in the next time step). It looks as though there is no quantification here, but there is an implicit zero number of time steps between one agent leaving and the next arriving. In practice, in both the Library and the Nursing Home, there was almost always some time gap before a new member of staff could be recruited, trained and formally inducted.

To resolve this problem, it was necessary to engage in a further process of abstraction. As noted above, both ABM modellers and ethnographic researchers are interested in mechanisms, that is, the processes or sequences of events that causally link initial conditions and outcomes[7, 23]. A mechanism is a theoretical construction described in abstract terms (e.g. the mechanism of an epidemic involves a virus infecting a susceptible victim, who in turn infects another victim ¨ here the concepts, ¨virus¨ and ¨victim¨ are theoretical terms). The ethnographer observes an instantiation of the mechanism (Ann infects Bob, or in the Nursing Home, Charles has an argument with David) and infers the mechanism from her observations, through a process of abstraction. In doing so, many quantities, even if they are observed at all, are discarded (for example, the account of the epidemic may ignore the time taken for the virus to make the victim infectious; in the Library, the ethnographer may record but then disregard the number of arguments between Charles and David and simply note the existence of conflict). The modeller has to go the other way: given an abstract social process or mechanism, he or she has to invent some numbers in order to fill out the theory sufficiently to create a working implementation. The important issue in inventing these numbers is not their specific values, but that the implemented model must be an instantiation of the desired mechanism.

4. The relationship between ABMs and ethnographic data

The purpose of building an ABM to represent the Library and Nursing Home case studies was to formalise and verify the set of mechanisms that led to the main features of the settings that were observed. This involved creating a specification of the model and the behavioural rules to be followed by the agents, implementing and running the model, and comparing the outcomes with the ethnographic record. In terms of the issues listed in table 1 of section 2, this requires:

1. A realist orientation, in which observable social action is considered to be generated by the modelled social processes. Although critical realism is usually associated with an objectivist stance, it is possible to couple realism with constructivism if one allows for agents to construct the meanings and concepts they use as a response to the social environment in which they act. However, given the fairly primitive models of cognition and language use in almost all current ABMs, modelling of such aspects is unusual.
2. Either a deductive approach, in which the model is formulated and then tested against the ethnographic data, or an abductive one, in which a theory and thus a model is induced from the data and then the model is tested against the data and discrepancies used to improve the model.
3. A concern with generalisation. ABMs normally represent a general theory about how some social phenomenon works. Ethnographic data about one particular setting can be used as a case study or a ¨critical case¨ that can be used to help to validate a model ¨ more precisely, to falsify the model if the model is incapable of generating a representation of the behaviour observed in the setting.
4. An etic approach, that is, one that uses the analyst¨s concepts, but where actors¨ concepts are, as far as possible, allowed to emerge from the operation of the model.
5. The usual interest in process and context that can be found in most ABMs.

Working within these constraints, it seems possible and useful to construct an ABM that represents the processes or mechanisms that have been discovered from ethnographic research. However, the model cannot be expected to confine itself to representing the ethnographic data and no more, for an ABM will inevitably require many ¨numbers¨, or more accurately, modelling assumptions, that are not to be found in the data. This is not a difficulty provided that the model does represent at some abstract level, the same mechanisms as identified in the ethnographic analysis.

5. Conclusions

It has been suggested that it is possible to use ethnographic data, such as fieldnotes from participant observation in the design and validation of ABMs, but that there are number of areas where a modeller will find themselves having to move away from the conventional assumptions and methods of ethnographic research. While the agent-based modelling community needs to gain more experience of using ethnographies in their work, some advice can already be formulated for those contemplating using ethnographic data:

- Ethnographic data are usually very rich and detailed, while simulation models tend to be somewhat abstract. In order to bridge the gap, it will be necessary to develop analyses of the ethnographic data, possibly using computer aided qualitative analysis (CAQDAS) tools[5] to help manage, code and search the textual data.

- Although typically an ethnographic study describes just one particular setting, it can be used as a case study of a more general class of settings. It may be useful either to conduct more than one ethnographic study, so that settings can be compared, or to locate descriptions of other settings that can substitute for first-hand data collection.
- Ethnographic data are often good sources of information about relationships, among actors and between actors and objects. A model of the data will need to represent such relationships; that is, the model is likely to be based on networks of agents.
- Ethnographies usually contain descriptions of social processes that are equivalent to the kinds of mechanisms that are typically modelled by ABM. It helps to lay out these processes explicitly in the ethnographic analysis before attempting to model them.
- The modeller will have to add ¨concreteness¨ to the model beyond the evidence in the data (for example, choosing numeric values for thresholds and multipliers; selecting distributions for random number generators; assigning attributes to agents either at random or in some arbitrary way). This is an inevitable part of the process of using ethnographic data as the foundation of an ABM. As far as possible, such choices should be justified by conducting a sensitivity analysis and showing that they do not have a significant effect on the outcomes of the simulation, while recognising that it is impracticable to test the sensitivity of every aspect of a model.
- While much ethnographic data concerns the observation of individual action, some structural and group data can also be gathered (e.g. who talks to whom and when; what collective events occur; who is involved in rituals). These ¨macro¨ features of the setting are valuable for the modeller. The model may generate outcomes and patterns that correspond to these features, and if it does, this helps to validate the model.
- The modeller should consider the implications of a stance that proposes that the way that people structure their worlds is not based on an external reality, but a culturally generated and fluid categorisation that is self-organising and emergent. While this constructivist perspective is hard to accommodate, given the current state-of-the-art in agent-based modelling, one can at least try to represent some of its implications.

Ethnographic data, we contend, is no less relevant to building ABMs than quantitative data. However, while there is already much experience in using quantitative data to design and validate ABMs, there are relative few examples of ABM based on ethnographies and little advice to guide the modeller. We hope that this paper will inspire more consideration of how ethnographic data can best be used by modellers.

---

[5] See, for example, the reviews and resources available at http://caqdas.soc.surrey.ac.uk

Acknowledgements. We acknowledge partial support from the project, ¨EMIL: EMergence In the Loop: simulating the two way dynamics of norm innovation¨ funded by the European Commission Sixth Framework Programme - Information Society and Technologies - Citizens and Governance in the Knowledge Based Society.

References

1. Agar, M.: My kingdom for a function: modeling misadventures of the innumerate. Journal of Artificial Societies and Social Simulation **6** (2003) 8
2. Agar, M.: Agents in living color: towards emic agent-based models. Journal of Artificial Societies and Social Simulation **8** (2005) 4
3. Sirkin, R.M.: Statistics for the Social Sciences. Sage, Thousand Oaks, CA (1995)
4. Babbie, E.: The Practice of Social Research. Wadsworth (2004)
5. Stevens, S.S.: On the theory of scales of measurement. Science **103** (1946) 677-680
6. Agar, M.: The Professional Stranger: An Informal Introduction to Ethnography. Academic Press (1996)
7. Mayntz, R.: Mechanisms in the Analysis of Macro-social Phenomena. Philosophy of the Social Sciences **34** (2004) 237-259
8. Berger, P.L., Luckmann, T.: The Social Construction of Reality : A Treatise in the Sociology of Knowledge. Anchor (1967)
9. Gilbert, N., Chattoe, E.: Hunting the unicorn: an exploration of the simulation of small group leadership. In: Saam, N.J., Schmidt, B. (eds.): Cooperative Agents: applications in the social sciences. Kluwer, Dordrecht (2001) 109 - 124
10. Archer, M.: Realist social theory: the morphogenetic approach. Cambridge University Press, Cambridge (1995)
11. Glaser, B.: Basics of grounded theory analysis. Sociology Press, Mill Valley, CA (1992)
12. Malerba, F., Nelson, R., Orsenigo, L., Winter, S.: History-Friendly models: An overview of the case of the Computer Industry. Journal of Artificial Societies and Social Simulation **4** (2001) 6
13. Headland, T., Pike, K., Harris, M.: Emics and Etics: The Insider/outsider Debate. Sage Publications (1990)
14. Barreteau, O.: Our Companion Modelling Approach. Journal of Artificial Societies and Social Simulation **6** (2003)
15. Merton, R.K.: Social theory and social structure. Free Press, New York (1968)
16. Yang, L., Gilbert, N.: Case-based Model of Emotional Expression Influence on Work Group Socialization and Performance. 1st World Congress on Social Simulation, Kyoto, Japan (2006)
17. Dey, I.: Qualitative Data Analysis : A User-Friendly Guide for Social Scientists. Routledge, London (1993)
18. Whitehead, A.N.: Science and the Modern World. Free Press (Simon & Schuster) (1925)
19. Wilensky, U.: NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL (1999)
20. Hoffman, E.: The effect of race-ratio composition on the frequency of organizational communication. Social Psychology Quarterly **48** (1985) 17-26
21. McPherson, M., Smith-Lovin, L., Cook, J., M.: Birds of a feather: Homophily in social networks. Annual Review of Sociology **27** (2001) 415-444
22. Ward, R.A., La Gory, M., Sherman, S.R.: Neighborhood and network age concentration: Does age homogeneity matter? Social Psychology Quarterly **48** (1985) 138-149

23. Hedström, P., Swedberg, R. (eds.): Social Mechanisms: An Analytical Approach to Social Theory. Cambridge University Press, Cambridge (1998)

Parallel Sessions

Policy Session

A model to explore multi-dimensional change in an unsustainable farming system

Georg Holtz[1]

[1] Institute of Environmental Systems Research, Barbarastrasse 12, 49069 Osnabrueck, Germany
Georg.Holtz@usf.uos.de

Abstract. In the Upper Guadiana River basin in Spain, groundwater over-exploitation through farmers endangers ecologically valuable wetlands. A model to explore leverage points shifting the farming system towards more sustainable practices is presented. This model defines and relates three dimensions: 1) farmers' priorities, 2) potential practices and 3) rules. Farmers' choice among potential practices is modelled as weighted adding strategy. This model allows exploration of integrated solutions not restricted to regulatory approaches and is currently used to identify ways to achieve a permanent shift to more sustainable farming practices.

Keywords: Agent-based modeling, Farming-system, Upper Guadiana river basin, Pattern Oriented Modeling

1 Introduction

In the Upper Guadiana river basin in Spain, irrigation of farm land accounts for approximately 90% of total groundwater use. During the last decades, the amount of farming has increased and farming practices have changed towards water-intensive crops; leading to an over-exploitation of groundwater resources endangering wetlands of high ecological value ([1]). Policy interventions aiming at reducing groundwater use turned out to be either not very fruitful since restriction of water use could not be enforced to an effectual degree and many farmers took water illegally; or policy measures depended on lasting compensatory payments which could not be sustained over long periods of time. This paper presents a model aiming at enhancing the understanding of how the farming system in the Upper Guadiana could get "un-locked" from its current unsustainable path.

2 Modeling approach

This model builds on the sociological concept of context-dependence of actors ([2]), explaining behavior as depending on an actor but also on a context in which the actor acts. The context has a dual function: on the one hand it enables actions by providing options, possible behaviors. On the other hand it constrains actors by advising certain behavior and penalizing others. The model defines three dimensions to capture this: "priorities", representing an actor's evaluation scheme; "options" representing potential practices and "rules", representing constraints and incentives.

Using this concept it follows that a lasting shift in practices must be considered as emergent outcome of (a combination of) change in options available, in rules advising and penalizing options or in priorities governing the choice among options considering rules. The model is used to analyze which of the three dimensions (priorities, options, rules) is (not) sensitive to evoke lasting change of the farming system. The innovative contribution of this model is its openness for analyzing not only changes in rules but likewise in options and priorities which often are assumed to be fixed. This allows exploring solution strategies not restricted to mere regulatory approaches.

3 Implementation

This model is implemented as an agent-based model using a framework called Famoja ([3]). The agent-based modeling approach facilitates representation of heterogeneity of farmers and individual farm histories. This model is organized around four "topics" found to be relevant for farmers' decision on farming practices: 1) income, 2) acting legally, 3) saving water and 4) keeping the tradition.

Actors (farmers) are heterogeneous regarding their priorities which are represented as lists of weights attached to these four topics. Options score different regarding these topics. Options representing current farming practices comprise a combination of crop and irrigation technology; radical alternatives can easily be represented by specifying their scores. Rules, including the changing EU Common Agricultural Policy and national pumping restrictions, define which options allow staying legal and further influence income through subsidies and penalties. Hence rules modify options' scores.

Farmers' decisions among (modified) options are computed using a weighted adding strategy; that is for each farmer the scores of modified options are weighted by this farmer and summed up to a total score. Farmers only change behavior if the entailed increase in total score exceeds a threshold representing investments that are not yet written off, learning efforts and risk aversion.

4 Calibration and validation

The models' empirically unspecified aspects (e.g. distribution of farmers' priorities) are calibrated using data from the time period 1985-2000. As model input, policy measures incorporating EU and national policies are combined with scenarios of development in options (during that period, pumping technologies became more efficient and much cheaper). The 2000-2005 period is used to validate the gained model. The model is calibrated and validated using a pattern oriented modeling (POM) approach: multiple emergent pattern like water use and crop pattern are reproduced simultaneously ([4]). POM reduces uncertainty related to model structure and parameters.

5 Expected results

This model is designed to identify and understand leverage points of the farming system. It is currently used to identify ways to achieve a permanent shift in farming practices. Several scenarios representing possible future developments in the three dimensions options, rules and priorities are designed and resulting model behavior is explored.

Past attempts of solving the problem focused on rules. Restriction of ground-water use failed since acceptance of such rules among farmers is low and enforcement against the will of farmers is too costly (financially and politically). A key topic here might be achieving a higher level of farmers' awareness of water related problems like degradation of wetlands and depletion of aquifers since this could entail a shift in priorities. Then a social norm might establish which could prove to be more efficient than (only) a legal law.

Obviously inventing options meeting farmers' priorities without over-exploiting groundwater resources would be another solution for the problem. However, according to the model design, such options would not only need to provide adequate income but must also meet farmers need for keeping their identity as being farmers. Farm holidays or organic farming could be such options.

Acknowledgements

The author would like to thank the insightful comments of two anonymous reviewers. Many thanks also to Marcela Brugnach and Hendrik Nieße who closely cooperated in the development of several aspects of the model described here.

References

1. Llamas, R. and Martínez-Santos, P.: NeWater Report: Baseline Condition Report Upper Guadiana Basin, 2005, www.newater.info.
2. Schimank, U.: *Handeln und Strukturen*, Juventa Verlag, Weinheim und München, 2000, 366.
3. Famoja: Framework for Agent based Modelling with Java, http://www.famoja.net
4. Grimm, V., et al.: Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology, *Science* 310, (2005).

Modelling Rules, Norms, and Institutional Change Using the Grammar of Institutions

Eva Ebenhöh

Institute of Environmental Systems Research
University of Osnabrück
49069 Osnabrück
Eva.Ebenhoeh@usf.uos.de

Abstract This paper reports on a formal theory to analyse institutional statements - rules and norms - and how this theory can be applied in agent-based models to simulate and investigate institutional change. The theory is the grammar of institutions by Crawford and Ostrom [1,2]. This paper describes a possible application of this theory including an agent decision making mechanism, in which rules and norms that have been adopted by an agent serve as first filters to reduce the number of possible actions. The paper also reports briefly on two examples of models in which the grammar of institutions has been applied. This way of modelling institutional statements may help to model institutions and to investigate institutional change by means of social simulation.

1 Introduction

When agent-based social simulation is used for investigating institutional change, institutions, their effectiveness, and their emergence need to be endogenous in the models. This paper reports on a way of conceptualising institutions in order to transfer them from real-world problems into computer code. This is a step to facilitate modelling of institutions and institutional change, as well as of comparing institutions and their impacts across models.

One view on institutions regards them as constraints on decision makers' behaviour. As such they can be integrated in agent-based models. However, institutions vary greatly in both the way they are described and in their impact. The Grammar of Institutions [1,2] is a way of describing institutional statements using five distinct grammatical building blocks. This stringent view of analysing institutions reduces ambiguity while transferring rules and norms from the real world into computer code [3].

This paper presents the theory (Section 2) and its application to agent-based models, including an exemplary decision making process in which institutional statements constrain agent behaviour (Section 3). It includes a brief description of exemplary models in which this approach has been used (Section 4) and concludes with a short discussion on the usefulness of this approach.

2 Grammar of Institutions

The grammar of institutions [1,2] is a logical syntax, in which institutional statements can be represented as well as compared and contrasted. Institutional statements "describe opportunities and constraints that create expectations about other actors' behavior" [2, 137]. This syntax is useful for implementing institutions as constraints for agent behaviour in social simulation models.

This grammar can represent both formal and informal institutions [4]. Whether they are efficient or not, does not depend on their formality, but is determined by whether they are followed and enforced. Rule following and enforcement can happen in different ways in the real world, which can be represented in this theory. It allows for investigating consequences arising from not following a rule. Rule compliance can be modelled on this basis, including extrinsic (sanctions) and intrinsic motivation.

2.1 Five Building Blocks

According to this grammar, a *rule* is an institutional statement consisting of five components (ADICO): Attributes, Deontic, aIm, Conditions, and an "Or else" statement.

Attributes define any individual attributes needed to distinguish actors from each other, in order to define to whom the institutional statement applies.
Deontic A deontic statement is either permission, obligation, or forbiddance, that is "may," "must," or "must not", respectively. Note, that it is possible to define each deontic by the other two.
Aim defines what it is that may, must, or must not be done.
Conditions define states of the world in which the rule applies.
Or Else defines consequences for not following the rule.

Some examples throughout this paper illustrate the building blocks and some are later used as examples in a modelling context. A sign in a train could read: *If you don't have a valid ticket, you are charged 60 Euro.* Attributes in this case are implicitly referring to all people travelling in a train. However, the law probably specifies some age restriction, say age seven and older. The deontic is "obligatory." The aim is to *have a valid ticket.* Deontic and aim can be exchanged: *It is forbidden to travel in a train without a valid ticket.* Conditions are *in trains of a specified railway company.* The "Or Else" part is that you are charged 60 Euro. However, if you do not get caught, you do not have to pay. So, there is a chance that the sanction is not excerted. Re-phrasing the statement according to the grammar leads to: *All persons from age seven on - must - have a valid ticket - when travelling in trains of a specified railway company - or else suffer a fine of 60 Euro.*

2.2 Rules, Norms, and Shared Strategies

Three subsets of institutional statements can be distinguished according to which of the five elements are defined.

Shared Strategies - AIC Shared strategies consist of attributes, aim, and conditions. Consider for example the shared strategy: *Customers tip in a restaurant.* The attributes *customers* states that the shared strategy applies to all who eat a meal. The aim *to tip* refers to what is to be done. The condition *in a restaurant* states, where tipping is usual, for example this shared strategy says nothing about tips in a drive thru. Consider also: *Players of repeated prisoner's dilemma games - cooperate - in first moves* [2, 157].

Norms - ADIC Norms consist of attributes, deontic, aim, and conditions. Consider for example: *If you use the microwave, you must clean up your own mess* [2, 139]. The attributes refer to microwave users. The aim is to clean up. Cleaning up is obligatory, as the deontic "must" states. Conditions are if they caused a mess. The statement can be rephrased: *Microwave users - must - clean up - if they caused a mess* in order to bring the elements in the corresponding order of attributes - deontic - aim - conditions. Consider also this norm: *Players of repeated prisoner's dilemma games - must not - cheat - unless being cheated before.*

Rules - ADICO Rules consist of all five elements. Rules are distinguished from norms by adding explicit consequences for violating a rule. The train-riding example given above is an example for a rule. Another example would be the grim trigger: *Players of repeated prisoner's dilemma games - must not - cheat - unless being cheated before - or else suffer cheating in all following rounds.*

3 Applications in Agent-Based Models

3.1 Agent Decision Making

The agent decision making mechanism proposed here uses rules, norms, and shared strategies as first filters in a two-step filter process as displayed in Figure 1. From a list of possible actions, only those are considered, that are permitted by the institutional statements applicable in the current situation and individually considered to be important. From the remaining possible actions, which I call "considered actions", other mechanisms can be used to choose a single remaining action. These other mechanisms can include evaluative criteria or simple decision heuristics. By evaluative criteria, expected outcomes are judged, for example, highest expected return or equality of outcomes. Examples for simple decision heuristics include closeness to a previously made decision or choosing a focal point. The difference between evaluative criteria and decision heuristics is that the former refers to expected effects of an action, while the latter refers to the action itself. In our models, they are used in a similar way, but note, that the outcome of evaluative criteria depends on the way in which expectations are formed. Expectation formation may also be guided by institutional statements as suggested in [2] and modelled, for instance, in [5].

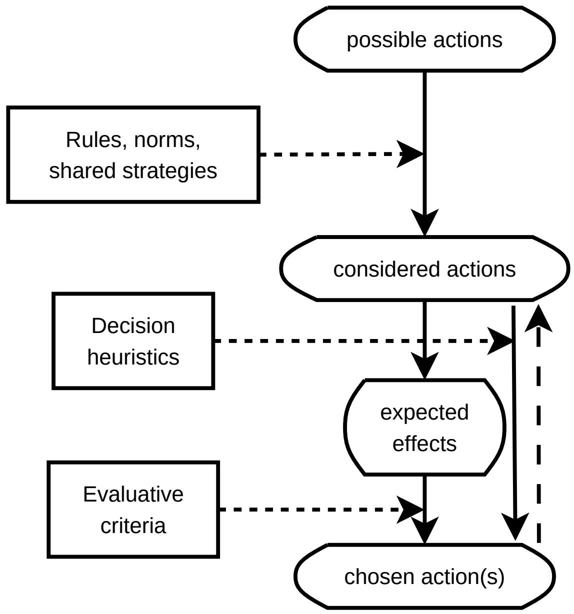


Figure 1. The decision making process suggested in this paper is a two-step process, which first filters out possible actions according to rules, norms, and shared strategies, and secondly determines the best option according to applied evaluative criteria or decision heuristics.

If a single action needs to be chosen, the second step of the process is repeated using a list of evaluative criteria and decision heuristics with decreasing priority to successively reduce the list of considered actions. This process stops, when there is only one choice left. In order to define an end-point to this process, in one model, I included a random choice decision heuristic with least priority to select a single action out of a list of actions, which could not be distinguished by an agents' evaluative criteria.

If more than one action can be chosen at the same time, the process gets more complicated. For now, in our models, only single decision processes have been modelled. Multi-decision processes have been modelled as sequences of single decision processes.

3.2 Reasons for Following Rules and Norms

As in the real world, the existence of a rule, does not automatically lead to it being followed. Different reasons exist for humans to follow rules. In order to use social simulation to investigate institutional change, these different reasons have to be taken into account in the respective models.

In the introduction I argued that this syntax reduces ambiguity. It can not, however, erase ambiguity completely. The prisoner's dilemma example for a rule, the grim trigger, can alternatively be stated as three shared strategies [2, 157]:

1. All players - cooperate - first round.
2. All players - cooperate - all rounds in which all players cooperated in the previous round.
3. All players - defect - all rounds after a defection.

Consequences of violating rules and norms, therefore, do not only include explicitly mentioned sanctions (which may or may not be excerted), but also behavioural changes in reaction to an agents' behaviour. Which of the two sets of institutional statements is more effective can not be seen, by merely looking at the corresponding formalisation. Rather, factors determining effectiveness are how many players follow the third shared strategy and how often the explicit sanction is actually exerted.

Sanctions One way of enforcing rules is by defining sanctions, monitoring rule compliance, and excerting defined sanctions after a rule violation has been perceived. Saam and Harrer model the Finder-Keeper norm as an institutionalised rule with sanctioning of observed violations [6, 4.15]. All three processes are of importance. Turning a norm into a rule by explicit "or else" statements, does not suffice if monitoring is not ensured or if there is no agent with the possibility and willingness to excert the sanction. The probability for being detected can be taken into account by agents, calculating expected returns of rule violations. In the train example, if tickets are rarely checked, the expected return is different than if tickets are checked on almost every ride. Nevertheless, pressure to follow norms can be modelled without explicitly modelling monitoring, for instance as "peer pressure" in [7].

Sanctions are not always pecuniary. They can be excerted in specific choices for actions in the future, as the grim trigger in the prisoner's dilemma example.

I do not propose, however, that agents should calculate subjective expected utility on the basis of monitoring and sanctioning probabilities. Instead, it seems sufficient that rules are regarded as effective or not, and are followed (or not) accordingly. Thinking about the train example this approach seems more realistic than calculating expected returns.

Reputation In repeated interactions, trust building or more generally expectation formation of others' behaviour, is one way to enforce norms or shared strategies, which lack an explicit sanction, as for instance through normative reputation in [8]. In repeated prisoner's dilemma games, the threat of defecting does not have to be stated as an explicit sanction, since it is in the interests of players to cooperate on the long run. Once cooperation has been established, trust is increased. Trust is modelled as the expectation that the other follows the shared strategy of cooperation [9]. Trust loss through one defection, on the other hand, can be hard to build up again. Future interactions in which another agents' behaviour is bound to be based on the reputation of a decision making agent, may be incentive enough to follow group norms or shared strategies.

Conformity As humans are social beings and socially embedded [10], there seems to be an incentive for some to follow a norm which is followed by most. Agents can be programmed to do the same using the "majority rule" [11]. In our models, this agent trait is called conformity [9,12]. The higher their conformity, the lower the percentage of other agents following a norm needs to be, in order for them to follow it as well.

Norm adoption, however, may also be induced by mirroring peer behaviour and not general behaviour, see for instance [13] on imitation.

Emotions Due to internalisation of norms, emotions such as guilt, shame, and anger can be made responsible for norm following behaviour [14]. These would be modelled as intrinsic motivations to follow a given norm.

Habit To reduce the amount of decisions or thinking a human or software agent has to do, many norms and shared strategies are followed out of habit. In that case, stopping to follow a norm, needs to be triggered. We model this, by giving an institutional statement an individual strength. This strength may be reduced or increased by observations of others or outcomes of previous applications. Thresholds exist which would trigger a re-evaluation of an institutional statement.

3.3 Transferring Institutional Statements into Computer Code

In order to be applicable for social simulations, the building blocks need to be transferred to building blocks as used in a simulation model. The examples discussed in Section 4 draw on a modelling template [15] based on the Institutional Analysis and Development (IAD) framework by Elinor Ostrom [16,17]. However, the following findings are rather generic and maybe even obvious, but still it is worth noting that all meaningful elements of institutional statements need also be meaningful elements in the model, if their impact on actors' behaviour is to be investigated.

Attributes For attributes to provide a meaningful distinction between agents, they must be based on those elements, that constitute agent heterogeneity. If, for example, heterogeneity of an agent population is solely based on one attribute, say cooperativeness, values for cooperativeness could be distinguishing characteristics of institutional statements: *Players with a cooperativeness of greater than 0.5 - cooperate - in first moves.* In most models, however, even if agents are very much alike, they differ in more than one respect, for instance they can be located in specific points in space and time, have different experiences, and different connections in a social network and the like. A shared strategy in a prisoner's dilemma could be: *Players directly connected to a peer cooperator - cooperate - in first moves.*

Conditions Conditions refer to the state of the world. What has been said about attributes, also refers to conditions. They have to be meaningful in the model and, furthermore, agents need to be able to perceive them. If an agent has no way of distinguishing a cooperator from a defector, the shared strategy *All players - defect - when encountering a defector* is meaningless. If there is no way of knowing, the way in which expectations are formed by an agent is of high importance. *All players - defect - when expecting to encounter a defector* could be the same *All players - defect - when the other defected in the last game* if expectation is based on the latest experience.

The difference between conditions and attributes is that the former refers to the world outside the agent and the latter to the world within the agent. Experiences and expectation formation are somewhat in between, but refer to the outside world, even if they are processes or states inside an agent.

The distinction is clearer, for a tag directly perceivable by an agent. Consider the Finder-Keeper strategy [8,6] *Normative agents - must not - attack - agents who eat their own food.* It seems obvious that the food needs to be tagged so that the agents can see to whom it belongs. To make agents perceive if other agents are also normative was explicitly modelled as a model extension with communication of normative reputation. Only then, the norm turned into *Normative agents - must not - attack - agents who are normative and eat their own food.*

Deontic and Aims Aims need to be actions that are actual options for a decision making agent, or else the institutional statement is meaningless in a given situation.

In the decision making mechanism suggested above, institutional statements serve as filters to reduce the number of possible actions. Since each deontic statement can be defined by the other two, we could choose "must not" and "may" (see below) as the only implemented deontics. In prisoner's dilemma games, the norm *All players - must - cooperate* is equal to *All players - must not - defect.* More generally if an agent is obliged to do something, all other actions are forbidden. This, however, assumes that the institutional statement refers to the complete list of possible actions. This is not in all models the case. If the set of possible actions evolves and, for instance, walking away from a prisoner's dilemma situation is introduced [18], *All players - must - cooperate* does not allow walking away, while *All players - must not - defect* does. If the list of possible actions is not closed, "must" and "must not" need to be explicitly modelled.

A permission, on the other hand, does either not change the list of possible actions, or it is an extension to this list. In the latter case, however, either the environment or another institutional statement can contradict this extension. The environment is assumed to define physical impossibilities and therefore rules out a permission. Another institutional statement would have to be more important to the agent in order to rule out a permission.

Or Else The "or else" statement needs to be something that the agent can calculate in its decision making. Also, expectations of the probability of getting caught are important in this respect. Possible sanctions are not always carried out, either because the rule violation has not been monitored or because it appears to be of minor importance to or is costly for those who could excert a sanction.

4 Examples

This kind of decision making has been used in two different models. One is a model of a series of economic experiments with varying institutional settings including the possibility to communicate and adopt self-designed shared strategies as well as agree on sanctioning mechanisms. The second is a model of an upstream-downstream problem of flood protection, in which decentralised, centralised, and collective action regimes are contrasted using different institutional statements to create behavioural constraints for decision makers.

4.1 Institutional Statements in an Economic Experiment

In a series of economic experiments [16] eight participants have to make anonymous choices of how many tokens, out of their 25 per round, they want to invest in a common pool resource. The return from the resource depends on total investment in a negatively quadratic way. In the parameterisation used in [16] group optimum is at 36 tokens total, 4.5 tokens per participant. In baseline experiments, no communication or sanctioning is allowed. Experiments have also been conducted with communication phases in which participants could decide on shared strategies, with sanctioning mechanisms, and with both communication and sanctioning, including the possibility to decide on adopting sanctioning mechanisms.

In the model of the baseline experiment, the only norm implemented is *All players - must not - cheat - always*. Cheating is defined as investing more than the symmetric Nash strategy, which is 9 tokens, in the parameterisation in [16]. Roughly half the agents start out following this norm, according to their individual conformity value.

The model allows suggesting and voicing support of shared strategies during communication rounds. Shared strategies are implemented as *All players - invest 4 (5, 6, 7 . . .) - as long as all others do so*. Individual complicance to these shared strategies depend on two things: The first is a combination of the individual conformity value and the perceived expressions of compliance to a shared strategy. The second is the expected returns of following the shared strategy and thus keep it going. Thus, if following a shared strategy pays, even agents with a low conformity can be drawn into mutual cooperation.

Shared strategies can turn into rules, even without explicit sanctioning mechanisms, for example as: *All players - must - invest 4 - as long as all others do - or else in the next round all invest 10.*

4.2 Institutional Statements in Transboundary Flood Protection

Three different institutional settings are contrasted in a model of flood protection as an upstream-downstream problem. A three bucket river is managed by three different administrative bodies.

The decentralised setting defines no institutional statement for taking the downstream effects into account, when deciding on flood protection plans.

The centralised setting defines a rule with varying "or else" parts: *All decision makers - must - take the whole basin into account - always - or else be relieved of office*; *or else pay a fine*; *or else do not receive money out of a joint fund for flood protection*. Different reasons for following these rules, are investigated in this setting, including but not limited to the sanctioning threat.

The third setting, a collective action setting, allows norms of other regard to emerge during communication phases, in which information is exchanged and trust is built. The norm says the same as the rule in the centralised setting without an "or else" part. The model is used to investigate circumstances under which a collective action situation can outperform a centralised regime.

5 Discussion

This paper describes the Grammar of Institutions, a theory which can be used to transfer institutional statements from the real world into computer code. The theory helps to analyse and compare institutions, by defining building blocks, by which institutions can be compared and their impacts analysed. This decomposition into building blocks helps to implement institutions and their impacts on agent behaviour in agent-based models. It further aids modelling institutional change by making explicit what elements need to be captured in the model.

Based on this conceptualisation, emergence of new rules as well as emergence of rule following may be modelled. With this kind of representation of institutional statements, agent-based models can be used to investigate institutional change.

However, as we have seen, institutions are not unambiguously represented in this theory. For instance, the grim trigger in prisoner's dilemma games can be modelled as a set of shared strategies or as an explicit sanctioning mechanism.

Also, the proposed decision making process does not pre-define all necessary processes. For instance, the adoption of norms and rules is so far modelled as a cue-based or learning direction mechanism. Even in the two examples, agent reactions on institutional change, have not been investigated thoroughly enough, to be modelled with confidence. Rather, the models mentioned in Section 4 were used to develop this technique. For social simulation models based on this template, careful data collection and qualitative confirmation with expert knowledge is necessary [19]. For the upstream-downstream model, this is currently done in close interaction with a case study.

References

1. Crawford, S.E.S., Ostrom, E.: A grammar of institutions. American Political Science Review **89** (1995) 582–600
2. Crawford, S., Ostrom, E.: A grammar of institutions. [17] chapter 5 137–174
3. Smajgl, A.: Advancing the capacity to simulate institutional change. [20] 189–196
4. Helmke, G., Levitsky, S.: Informal institutions and comparative politics: A research agenda. Perspectives on Politics **2** (2004) 725–740
5. Bhavnani, R.: Adaptive agents, political institutions and civic traditions in modern italy. Journal of Artificial Societies and Social Simulation **6** (2003) <http://jasss.soc.surrey.ac.uk/6/4/1.html> (March 22, 2007).
6. Saam, N.J., Harrer, A.: Simulating norms, social inequality, and functional change in artificial societies. Journal of Artificial Societies and Social Simulation **2** (1999) <http://jasss.soc.surrey.ac.uk/2/1/2.html> (March 22, 2007).
7. Thebaud, O., Locatelli, B.: Modelling the emergence of resource-sharing conventions: an agent-based approach. Journal of Artificial Societies and Social Simulation **4** (2001) <http://jasss.soc.surrey.ac.uk/4/2/3.html> (March 22, 2007).
8. Castelfranchi, C., Conte, R., Paolucci, M.: Normative reputation and the costs of compliance. Journal of Artificial Societies and Social Simulation **1** (1998) <http://jasss.soc.surrey.ac.uk/1/3/3.html> (March 22, 2007).

9. Ebenhöh, E., Pahl-Wostl, C.: Agent-based modelling with boundedly rational agents. In Rennard, J.P., ed.: Handbook of Research on Nature Inspired Computing for Economy and Management. Idea Group (2006)
10. Edmonds, B.: Capturing social embeddedness: A constructivist approach. Adaptive Behaviour **7** (1999) 323–348
11. Jaffe, K., Cipriani, R.: Culture outsmarts nature in the evolution of cooperation. Journal of Artificial Societies and Social Simulation **10** (2007) <http://jasss.soc.surrey.ac.uk/10/1/7.html> (March 22, 2007).
12. Ebenhöh, E.: A topology of agent attributes for modelling microbehaviour in economic experiments. [20]
13. Conte, R., Paolucci, M.: Intelligent social learning. Journal of Artificial Societies and Social Simulation **4** (2001) <http://jasss.soc.surrey.ac.uk/4/1/3.html> (March 22, 2007).
14. Staller, A., Petta, P.: Introducing emotions into the computational study of social norms: A first evaluation. Journal of Artificial Societies and Social Simulation **4** (2001) <http://jasss.soc.surrey.ac.uk/4/1/2.html> (June 10, 2006).
15. Ebenhöh, E.: Testing hypotheses on adaptive management using an agent-based model based on the institutional analysis and development framework. Deliverable 1.2.3 in the NeWater Project (2007) <www.usf.uos.de/∼eebenhoe/newater/Mileston1.2.3.pdf>.
16. Ostrom, E., Gardner, R., Walker, J., eds.: Rules, games, and common-pool resources. Ann Arbor, Michigan: The University of Michigan Press (1994)
17. Ostrom, E.: Understanding Institutional Diversity. Princeton University Press (2005)
18. Joyce, D., Kennison, J., Densmore, O., Guerin, S., Barr, S., Charles, E., Thompson, N.S.: My way or the highway: a more naturalistic model of altruism tested in an iterative prisoners' dilemma. Journal of Artificial Societies and Social Simulation **9** (2006) <http://jasss.soc.surrey.ac.uk/9/2/4.html> (March 22, 2007).
19. Moss, S., Edmonds, B.: Sociology and simulation: Statistical and qualitative cross-validation. American Journal of Sociology **110** (2005) 1095–1131
20. Terano, T., Takahashi, S., Sallach, D., Rouchier, J., eds.: Proceedings of the First World Congress on Social Simulation. Kyoto University (2006)

A Two-Layer Participatory Simulation to Support a Flexible Participation of a Consultative Council

Vinicius Sebba Patto[1], Paul Guyot[2], Jean-Pierre Briot[1,3], and Marta Irving[4]

[1] Laboratoire d'Informatique de Paris 6, Université Paris 6, Paris, France.
[2] National Institute of Informatics,Tokyo, Japan.
[3] CS Dept., PUC-Rio, Rio de Janeiro, Brazil.
[4] Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil.
{Vinicius.Sebba-Patto, Jean-Pierre.Briot}@lip6.fr, paul@nii.ac.jp & Marta.Irving@mls.com.br

Abstract. In this paper, we present a project at methodological and computer-based support for participatory management of protected areas, in order to promote biodiversity conservation, social inclusion, and conflict resolution. We discuss a new proposition to deal with temporal lapse during sessions of simulations and to improve the communication between participants through the decision making process. For our first case study, the Tijuca National Park, in Rio de Janeiro (Brazil), other approaches do not resolve our problem. To do it, we propose a two-layer simulation to grab information of participants in order to help the exchange of information in each layer and between both layers. We want to extend this method to other parks for conservation of natural resources and other fields which have to deal with social inclusion, participatory management, conflict resolution, and temporal lapse through the simulation.

Keywords: Role-playing games, multi-agent based simulation, protected areas, social inclusion, participatory management.

1 Introduction

Participatory decision making is often used to deal with biodiversity conservation, social inclusion, and economic development. Most social simulation methods are based on bottom-up participatory approaches to try to find solutions using the knowledge of local actors and communities [1], [2]. A limitation of these methods is that, in most cases, participants should be gathered in the same place at the same time. In this paper we propose to use a two-layer participatory simulation to deal with this temporal lapse and to help stakeholders make decisions.

2 Park of Tijuca and Participatory Management

In the urban National Park of Tijuca, different sorts of people and communities exert great pressure, especially through the human growth and illegal occupation [4]. To establish and to implement the predetermined actions to reach the goals of the park, a consultative council was composed of representatives of the civil society and the government. The council has a consultative role and is a department of the administrative structure of the park. One of its target is to promote the participatory management of the park through the inclusion of all involved sectors in order to improve the quality of the life for the communities around the park. In this system, we distinguish two groups: stakeholders and members of the council.

3 Two-Layer Participatory Simulation

Most social simulation methods have been developed and adapted to different circumstances in order to benefit from agent-based simulations, role playing games and participatory decision making [1], [2], [3], [5]. Some of them simulate systems before or during role playing game sessions [1], [3], other simulate systems within role playing game sessions [2], [5]. They help stakeholders make decisions and they have an educative role for stakeholders, considering that they improve the exchange of information and they help stakeholders understand each other. These methods group organizers and players in a single layer simulation, where everybody participate at the same time.

The context of the Park of Tijuca reveals several shortcomings of traditional methods: (a) in single layer simulations, the council would be represented by several players – the council must participate as a unique entity, not as a group of independent players; (b) it is not always possible to assemble members of the council and stakeholders in simulations; and (c) it is also not recommended because the council should be represented by all its members and all these members could not be always available to participate of all simulations.

We propose to split the simulation in two layers to help the council play its role advising stakeholders, to help stakeholders exchange information, and to help with the temporal lapse between sessions of simulations. On the first layer, stakeholders play their roles in a simulation. Assistant agents help stakeholders showing them the opinion of the council about actions. Stakeholders can explain their decisions to assistant agents and to other participants. This simulation grabs information about stakeholders, and help them exchange information and reveal their intentions. On the second layer, members of the council discuss the results of the first layer simulation, propose ideas, debate, and define the council position on questions which arose in the first layer. In each layer, at the end of simulations' turns, there is a debriefing session to discuss players' choices. After the simulation in each layer, assistant agents are updated in order to represent either the council (in the first layer) or the stakeholders (in the second layer). These processes are independent and can be iterated many times, in order to improve the exchange of information.

4 Conclusion

Assembling actors in two different groups, (a) does not require the whole council to attend every simulation designated for stakeholders and (b) favors, through the participatory simulation, the social learning and the exchange of information, and affects the stakeholders who take part in the decision process. In order to simplify the connection between the two layers we propose to use of artificial agents to help the council know about emergents questions which arise from the stakeholders' simulations (artificial agents are present in both layers) and to allow the council, as a unique entity, play its role in the first layer, through the artificial assistant agents.

Acknowledgments This research was funded by the ARCUS program (French Ministry of Foreign Affairs), the National Institute of Informatics, the Japanese Society for the Promotion of Science and the European AlBan program.

References

1. Barreteau, O.: Our Companion Modelling Approach. Journal of Artificial Societies and Social Simulation. Vol. 6. No 2. (2003)
2. Guyot, P., Drogoul, A.: Designing Multi-Agent Based Participatory Simulations. In: Proceedings of 5th Workshop on Agent Based Simulations. SCS Publishing House, Erlangen, San Diego (2004) 32–37
3. Barreteau, O.: The Join Use of Role-Playing Games and Models Regarding Negotiation Process: Characterization of associations. Journal of Artificial Societies and Social Simulation. Vol. 6. No 2. (2003)
4. Briot, J-P, Guyot, P., Irving, M.: Participatory Simulation for Collective Management of Protected Areas for Biodiversity Conservation and Social Inclusion. International Modeling and Multiconference 2007, Buenos Aires, Argentina. (2007) 183–188
5. Guyot, P., Honiden, S.: Agent-Based Participatory Simulation: Merging multi-agent systems and role-playing games. Journal of Artificial Societies and Social Simulation. Vol. 9. No 4. (2006)

Using Virtual Players in GMABS Methodology: A Case Study in Natural Resources Management

Diana F. Adamatti[1*], Jaime S. Sichman[1†], and Helder Coelho[2]

[1] Laboratório de Técnicas Inteligentes
Escola Politécnica da Universidade de São Paulo - Brasil
{diana.adamatti, jaime.sichman}@poli.usp.br
[2] Laboratório de Modelação de Agentes
Departamento de Informática da Universidade de Lisboa - Portugal
hcoelho@di.fc.ul.pt

Abstract. This paper proposes a new software architecture, called ViP-GMABS, which enables virtual players to be inserted in the GMABS methodology. This methodology combines RPG (Role-Playing Games) and MABS (Multi-Agent-Based Simulation) techniques in an integrated way. A prototype developed according to this architecture, called ViP-JogoMan, is presented. ViP-JogoMan was designed to be used as a Group Decision Support System (GDSS) in the natural resources management domain. This domain was chosen because it encompasses extremely important and complex negotiation processes, and GMABS methodology can be used to handle or mitigate conflict resolution.

1 Introduction

Role-Playing Games (RPGs) are a type of game where the players perform characters. These characters are created within a particular scene (an environment) which organizes the players actions, by determining the limits of what can or cannot be done [3]. In this way, RPGs are games where each player plays a role and takes decisions to reach its objectives. In fact, players use RPG like a "social laboratory", because they can try several possibilities, without real consequences [5].

Multi-Agent Systems (MAS) study the behavior of sets of independent agents with different characteristics, which evolve in a common environment. These agents interact with each other, and try to execute their tasks in a cooperative way by sharing information, preventing conflicts and coordinating the execution of their own activities [2]. Additionally, the use of simulation as an auxiliary tool for human decision-making is very efficient, because it allows the verification of specific details with great precision. The combination of both, multi-agent systems and simulation, generates a new research area called Multi-Agent-Based Simulation (MABS), that deals with problems that involve multiple domains [10]. An example of a MABS application domain is natural resources management, because it involves several knowledge areas, as sociology, hydrology and biology.

The use of MABS and RPG (isolated or in an integrated way) has been used in several works [8, 11] and it shows interesting results, such as to join the dynamic

*Supported by CNPq and CAPES, Brazil, grant number 141851/2004-0.

†Partially supported by CNPq, Brazil, grant numbers 304605/2004-2, 482019/2004-2 and 506881/2004-0.

capacity of MABS with the discussion and learning capacity of RPG techniques. We call the integration of RPG and MABS as GMABS (**Games and Multi-Agent-Based Simulation**) methodology [1].

Whenever a RPG is played, it needs a certain number of real players to be executed. However, many times the game cannot be played because it does not have the minimum number of players, even if in some cases players are in different places (if the game is played through the Web, for example) and/or at different time schedules (if the game is played in asynchronous way). In this way, the existence of **Virtual Players** would be very useful, because they could substitute the real players without damaging the game. By damaging the game we mean a situation where real players could easily identify the virtual players, because the decision-making results of the latter are not realistic, i.e., their actions are very different from the ones that real players are expected to perform.

The goal of this paper is (i) to present an architecture that enables to insert virtual players in the GMABS methodology, called ViP-GMABS and (ii) to present a prototype, based on this architecture, that was developed for the natural resources management domain. We have organized this paper in 7 sections. In Sections 2 and 3 we briefly present a short review about RPG and MABS, respectively. The GMABS methodology is presented in Section 4 . Section 5 details the ViP-GMABS architecture. In Section 6, the ViP-JogoMan prototype is described, as well as some preliminary test results. Finally, the conclusions and further work are presented in Section 7.

2 Role-Playing Games (RPG)

According to Barreteau, "*in-between games and theater, Role-Playing Games (RPGs) are group settings that determine the roles or behavioral patterns of players as well as an imaginary context. A RPG is the performance of a roughly defined situation that involves people with given roles*" [4]. Players genuinely use RPG as a "social laboratory". It is a way for them to experiment playing several roles and/or several strategies in a group with presumably few consequences in the real world .

Psychologists have analyzed RPG and its implications on players's behavior, since there are significative relationships between RPG and a kind of psychotherapy called *psychodrama*. In this kind of therapy, the patient lives a controversial or psychologically problematic situation according to other points of view, different from his own, to better understand them and live a cathartic experience [3].

RPG can be used for three different purposes: training, research or policy making. The first use is more predominant, since training tools are more and more used by professionals in market's training session. In this case, roles are strictly defined by the instructor's knowledge. In some cases, it is possible that more experts players use their own experience to get a better situation in the game, as in real practice.

3 Multi-Agent-Based Simulation (MABS)

The field of Multi-Agent Systems (MAS) is a well-established research domain in Artificial Intelligence (AI), which has emphasis in the resolution of problems by a society of agents. The distribution of the problem solving by several agents is necessary because these problems can be too complex or too large to be solved by a single process, or still, they may need knowledge from several different domains. One goal of these systems is to release the researchers from low-level technical-operational issues, allowing the researcher to concentrate his/her efforts on the relevant domain application level.

The computer simulation of social phenomena is a promising field of research at the intersection between the social, mathematical and computer sciences. The first developments of computer simulations in the social sciences coincided with the first use of computers in academic research, in the early 1960s. By this time, computer simulation was essentially used as a powerful implementation of mathematical modeling [10].

Multi-Agent-Based Simulation (MABS) is the union of MAS and Simulation and it is specially valuable to conciliate different interdisciplinary perspectives. Typically, it involves researchers from various scientific areas, such as social psychology, computer science, social biology, sociology and economics. The interdisciplinary character of MABS is an important challenge faced by all researchers, while demanding a difficult interlacement of different theories, methodologies, terminologies and points of view [10]. The MABS field is increasingly characterized by the study, design and implementation of computational platforms to simulate societies of artificial agents.

4 GMABS Methodology

In our work, we call the integration of RPG and MABS within a same methodology as GMABS (Games and Multi-Agent-Based Simulation) [1].

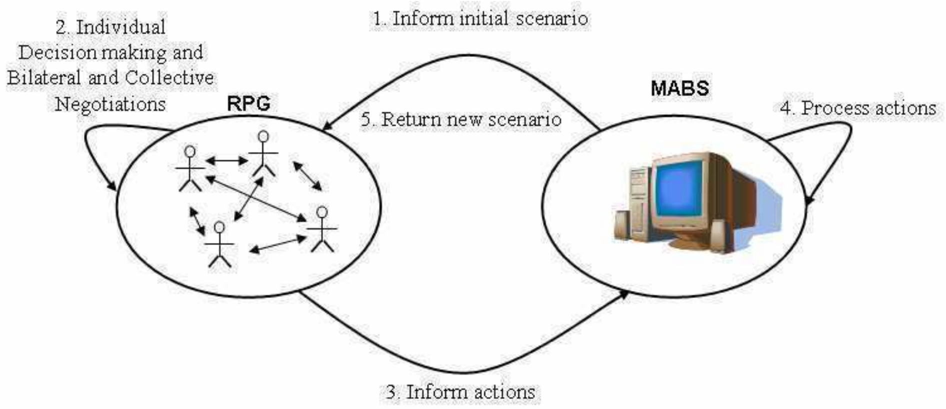


Fig. 1. GMABS Methodology.

The methodology steps, shown in Figure

1. Players receive all the information about the game: the roles they can play, the actions and rules available to these roles, the common environment, the topological constraints. When the game starts, each player defines the role he is going to play. Each player knows then what actions he can execute, and the benefits and/or damages its action can cause to the common environment. The initial scenario defines also where the players are physically located in the common environment and what are their initial possessions (money, land, etc.);
2. There are three different activities in this step:

(a) Players may reason and decide about individual actions that depend exclusively on themselves. As an example in the natural resources domain, land owners change their land use;
(b) Players have all the information necessary to initiate bilateral negotiations with other players. In order to do this, they may change information and make their decisions, according to the rules initially determined for their chosen roles. As an example in the natural resources domain, land owners can sell their lands. Normally, the duration of these two previous activities, which occur simultaneously, is defined in the beginning of the game;
(c) After deciding about their individual actions and concluding the bilateral negotiations, the players can then negotiate about collective strategies for the next rounds. These collective strategies should benefit all players or just a subgroup of them. As an example in the natural resources domain, players can demand improvements on infrastructures, more jobs, tax values, etc. This negotiation process of collective strategies is just a "predisposition" to define future actions: players are not really obliged to keep their word and really use these strategies in the further rounds. This process is hence very important for each player to better understand the objectives and strategies of the other players;

3. Players inform to the MABS tool which individual actions were chosen and which bilateral negotiations were concluded;
4. Data is computed by the MABS tool: the latter actions modify the initial scenario. The properties of the environment are then modified, which implies in the modification of each player data;
5. The MABS tool returns the new scenario. If the time deadline of the game is not reached or the maximum number of rounds has not been achieved, return to step 2.

More information about GMABS methodology can be found in [1].

5 ViP-GMABS Architecture

One goal of this paper is to present an architecture that enables to insert virtual players in the GMABS methodology. Differently of other works, as [11], where software agents were inserted in GMABS methodology as assistant agents to human players, we want to insert high-level cognitive agents (virtual players) to substitute people (real players) in the game. This architecture is called ViP-GMABS (*Virtual Players in GMABS*).

The main elements in GMABS methodology are the MABS tool and the players that play the RPG. In order to insert virtual players, we have divided the RPG element in two sub-elements: real players and virtual players. Since we have inserted virtual players in RPG, the MABS element should be modified, because the data stream must be addressed both to human and virtual players. We have therefore included an intermediate communication layer between the MABS and RPG elements. From the "MABS point of view", this communication layer treats the information exchange between the MABS and the players in the same way. This layer handles the communication between all players (real or virtual), during the game (see Figure 2).

The architecture presented in Figure 2 was defined to be both tool and domain independent. We really expect this architecture to be *generic*, since the GMABS methodology can be implemented for any knowledge area and for any MABS tool. We have

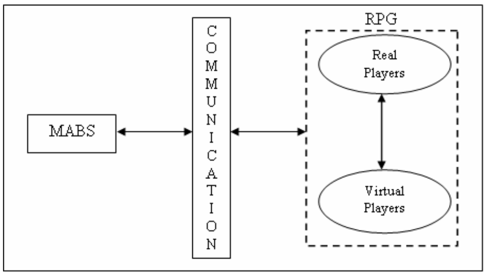


Fig. 2. Semi-Autonomous RPG architecture.

hence defined some requirements to handle the communication between the MABS and the RPG elements:

- The MABS element must supply, in some format (for example ASCII), the system data, such as the initial scenario and current players situations;
- The communication layer must receive the information from the MABS tool and forward it to the real and virtual players in an integral and uniform way, without losing any information. On the other hand, it must to return to the MABS tool the players's chosen actions, in order to process them in the simulator;
- The virtual players must receive information from the communication layer and must be able to manipulate them. Moreover, if necessary, they should communicate with other players (real and/or virtual) in order to get new information about the scene and/or players, to make their decisions.

6 ViP-JogoMan: A Case Study in Natural Resources Management

We have chosen to implement a prototype of the ViP-GMABS architecture in the natural resources management, specifically concerning water resources management. The most important aspect in the natural resources management is the negotiation process between the actors, because their objectives and strategies are different and many conflicts are generated. For example, a farmer and an industrial have different objectives and they do not necessarily reach a consensus very easily.

In fact, we have developed two prototypes to this domain: the first one, called JogoMan [1], was based on GMABS methodology but did not use virtual players. The second one, called ViP-JogoMan, used the ViP-GMABS architecture and virtual players. JogoMan has helped us to understand the chosen domain as well as the GMABS methodology. ViP-JogoMan, that is one of the focus of this paper, has as an objective to analyze the real use of ViP-GMABS architecture.

6.1 Game Description

The roles of the players and the rules of the game are the same in the two prototypes. They involve water, land use and urbanization pressure management of 3 different cities. In these games, there are four roles, each one having different goals:

- **Land Owner**: a land owner has some land portions, each one with a land use type, such as forest or agriculture. For each different land use type, there are different values related to their maintenance and financial return. Owners can sell or buy their portions or they can change their land use. Land owners should ask mayors to build infrastructure in their cities.
- **Mayors**: The game has different cities, each one having its mayor. The mayor goals are closely related to the city main activity (urban, agricultural, etc.) For example, if city "C" is a preservation area, then the player in the role of "Mayor C" should preserve this city. Mayors can invest on public infrastructure, such as portable water or sanitation networks or decide to build schools, hospitals or polices headquarters.
- **AguaPura Company Administrator**: This role can invest on public infrastructure to improve water quality, as portable water and sanitation net.
- **Migrant Representative**: This role has a special role in the game, since he/she must allocate a number of new homeless families. These families arrive to the cities (urbanization pressure), and they can be allocated in settlements or in slums. The quality and/or quantity of water of the region is modified depending on where these families are placed.

Each player chooses his/her actions individually, but he/she should know that these actions have consequences to other players, because the quality and quantity of water depends on the overall land use and infrastructure. For example, if a mayor decides to decrease the land taxes for land owners that preserve the forests, various land owners can consequently decide to maintain their areas with forest or even decide to plant forest (reforestation). This action influences every player, because the water quality probably will improve. In anther example, if a land owner decides to build an industry, the industry profit may be larger, but the water pollution will grow too.

6.2 Selected Tools

We have selected some tools to implement ViP-JogoMan, aiming to attend the previous requirements defined by the architecture proposed in section 5 (see Figure 3):

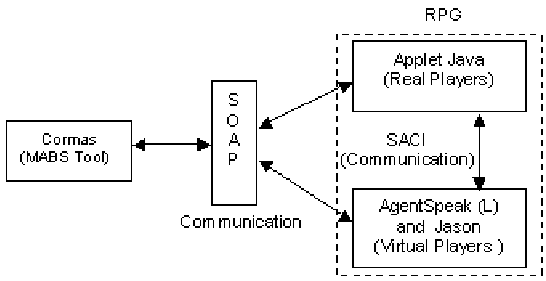


Fig. 3. Selected Tools to Semi-autonomous RPG.

- *MABS Tool*: *Cormas* is used as simulator [7], because it was used to implement the JogoMan prototype. *Cormas* has specific functions to extract system data in different formats (ASCII, Excel and several database formats like Oracle, MSAccess, MySQL or PostGre);
- *Virtual Players*: we have chosen the BDI architecture as the cognitive architecture for the virtual players, because we could use both a logic defined in *AgentSpeak(L)* language [16] and some implementation tools already developed for it, such as the *Jason* interpreter [6]. This interpreter allows that each step in BDI logic can be visualized and analyzed individually. It would also allow communication between virtual players, as well as between virtual players and the environment.
- *Real Players*: each type of player (Mayor, Land Owner, AguaPura Administrator or Migrant Representative) has a corresponding Java applet. In these applets, players can choose their actions and exchange information with other players (real and/or virtual);
- *Communication Layer*: we have chosen the SACI (*Simple Agent Communication Infrastructure*) tool [12] as the communication layer between real and virtual players. This tool provides agent communication infrastructure, using the KQML (*Knowledge Query and Manipulation Language*) language, and is used by *Jason*. The communication layer between MABS and RPG elements was implemented by using the SOAP (*Simple Object Access Protocol*) protocol [17], because the MABS tool (Cormas) and Jason were implemented in different programming languages (SmallTalk and Java, respectively), and the SOAP technology provides interoperability between both languages through the use of the XML language.

Nevertheless, the technology used was chosen to develop a web-based semi-autonomous RPG, meaning that it executes in a Web Server and it is accessed simultaneously by several players through web-browsers.

6.3 Virtual Players

In the ViP-JogoMan prototype, the development of virtual players is one of the most important aspects. As mentioned before, we have chosen the BDI architecture to model and implement the virtual players. The knowledge data base of this architecture is defined in terms of beliefs, desires and intentions of the players. In order to fill this knowledge base, we have mapped all the real players actions during the JogoMan prototype session tests. Our goal was to discover their objectives and strategies, aiming to build the virtual players knowledge data base. Having these information in hand, we have analyzed and discovered a sequence of actions that each real player executed. This experiment was very interesting, because we could see that real players defined autonomous strategies during the game, even if we did not explain them a priori, before starting the JogoMan prototype tests, how they should choose their actions.

For each role, different objectives were discovered, and we have defined different behavioral profiles for the virtual players:

- Land Owners:
  - *Economic* profile: must save and earn money;

    - *Ecologic* profile: must improve the ecological situation of region, having a big concern about reservoir pollution;
- AguaPura Administrator:
    - *Rational* profile: must improve water and sanitation networks with a rational use of money;
    - *Ecologic* profile: must improve water and sanitation networks;
- Migrant Representative:
    - *Economic* profile: must allocate families without worrying about the social conditions of these families;
    - *Social* profile: must allocate families in good places, with infra structures and near to urban areas;
- Mayors:
    - *Social* profile: must improve the life quality of its citys people;
    - *Economic* profile: must improve the life quality of the people if the city has enough money for it;
    - *Ecologic* profile: must improve the ecological situation in its city.

For example, the behavioral profile *Economic* Land Owner has as objective to save and earn money, and the strategies we have found during JogoMan results analysis were:

1. If the player has plots near to urban areas, he changes their use to Settlement to be able to sell to Migrant Representatives, since Migrant Representatives always want plots near to the urban areas;
2. If the player has plots where the land use is not Forest, he changes to Agriculture or Irrigated Agriculture, because these land uses have a low investment and a fast profit of money, comparing of the other land uses, as Industry;
3. If the player has plots where the land use is Forest, he changes to Plantation, to receive the suppression profit to cut the trees.

We have implemented each defined profile in *AgentSpeak(L)* language using the Jason interpreter. Figure 4 shows how the strategy number 1 of Land Owners with *Economic* behavioral profile was implemented in *AgentSpeak(L)* using Jason.

These profiles were analyzed and evaluated by specialists of natural resources to verify if the possible strategies and actions were similar to real player activities. In order to do that, some specific variables of the game were analyzed, to measure if the proposed objective of each profile was attained. For example, when evaluating the behavioral profile Economic Land Owners, we have analyzed the quantity of money in its *cash box* variable.

6.4 Negotiation and Protocols

In order to model and implement the communication between players (real and/or virtual), we defined a negotiation protocol, based on ACL-FIPA [9] and Raiffa's Theory of Negotiation [15]. Figure 5 presents the negotiation protocol associated to the action "buy a plot". In this protocol, there are four type of messages:

1. *Propose*: the buyer proposes to buy a plot;

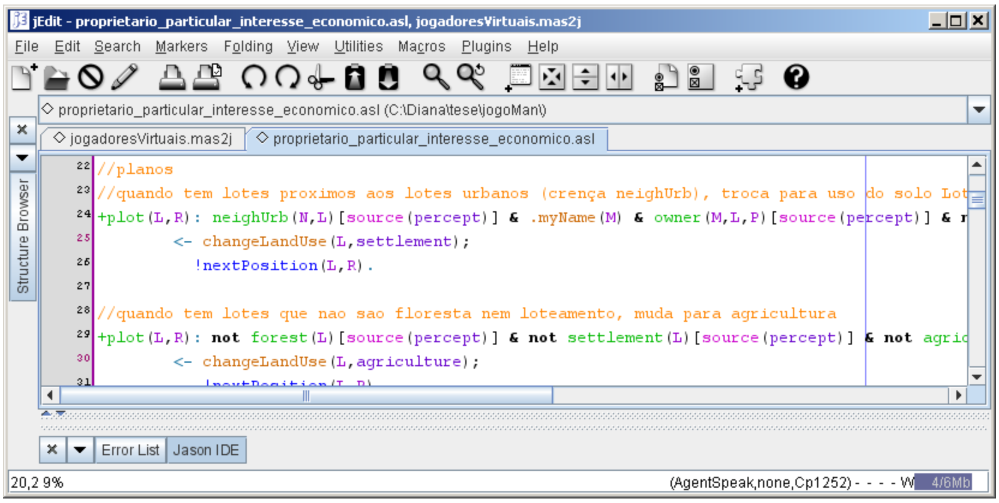


Fig. 4. Example of knowledge data base implementation in *AgentSpeak(L)* using Jason interpreter.

2. *Request*: the seller requests a new value to sell the plot;
3. *Accept_proposal*: the seller accepts the propose of the buyer;
4. *Reject_proposal*: the seller rejects the propose of the buyer;
5. *Inform*: the seller informs the environment (in our case, represented by the MABS tool) if the plot was effectively sold.

6.5 Tests and Preliminary Results

We have conducted three test sessions with ViP-JogoMan prototype, involving both people and virtual players. These tests had two objectives:

1. Verifying the influence of technology (computer) in GMABS methodology (using ViP-GMABS architecture) and the interaction between the players;
2. Analyzing the use of behavioral profiles of virtual players and their decision-making process.

We have defined three forms to evaluate our prototype: (i) pre and post questionnaires to be filled by real players; (ii) analysis of the behavioral profiles variables; and (iii) analysis of the message exchange flow between all players (virtual and/or reals) during the negotiation process.

By the answers given by the real players in the pre and post questionnaires, we have concluded that the ViP-JogoMan prototype gave the players interaction, entertainment and learning feelings. Moreover, real players did not discovery the virtual players easily in the game. In fact, real players just discovery some virtual players because these latter chose their actions very quickly and real players needed more time to think and to choose theirs actions, during negotiations.

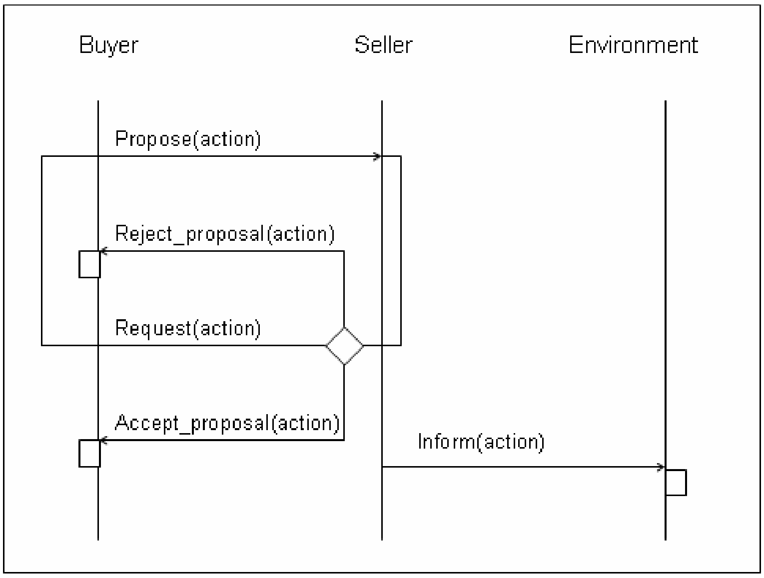


Fig. 5. Negotiation protocol associated to "buy plot" action based in ACL-FIPA.

By the analysis of the behavioral profiles variables, we have concluded that the defined strategies for each type of profile attended the proposed objectives. For example, all virtual players with *Economic* behavioral profiles finished the game with high cash box values, comparing with other players.

Finally, by the analysis of the message exchange flow between players during the negotiation process, we have concluded that all players interacted a lot with each other, because the number of exchanged messages was significant. According to Peppet [14], people feel more comfortable to express their opinions via Internet, because they do not have problems with shyness or prejudice. Another aspect important in ViP-JogoMan prototype is the storage of all negotiations between players (concluded or not). McKersie e Fonstad have stated that in Internet negotiations every data and action is stored, and in this way it is possible to analyze data with more attention and to better understand the negotiation process [13]. As an example, Table 1 presents some negotiations between players in the first round of a game session, either concluded or not. The concluded negotiations between players are represented in bold. An example of a concluded negotiation is number 2, where AguaPura *proposed* to buy plot 34 from Land Owner 5 by $1.500,00 and Land Owner 5 *accepted* the proposal, concluding the negotiation. An example of an unconcluded negotiation is number 1, where Administrator of AguaPura *proposed* to buy plot number 22 from Land Owner 2 by $1.500,00, but Land Owner 2 did not accept the offer, and has sent a request message for a proposal with a higher value. However, AguaPura did not propose new values and the negotiation was not concluded.

	Buyer	Seller	Type of Message	Plot	Offer
1	AguaPura	LandOwner2	propose	22	1.500
	AguaPura	LandOwner2	request	22	1.500
2	AguaPura	LandOwner5	propose	34	1.500
	AguaPura	LandOwner5	**accept_proposal**	34	1.500
3	AguaPura	LandOwner8	propose	47	2.500
	AguaPura	LandOwner8	**accept_proposal**	47	2.500
4	PrefeitoA	LandOwner8	propose	47	1.350
	PrefeitoA	LandOwner8	request	47	1.350
5	LandOwner9	LandOwner2	propose	22	4.000
	LandOwner9	LandOwner2	**accept_proposal**	22	4.000
6	LandOwner1	LandOwner2	propose	22	4.000
7	LandOwner9	LandOwner2	propose	23	4.000
8	LandOwner6	LandOwner2	propose	22	1.000
	LandOwner6	LandOwner2	request	22	1.000
	LandOwner6	LandOwner2	propose	22	1.500
	LandOwner6	LandOwner2	**reject_proposal**	22	1.500

Table 1. Negotiation to buy plots between players.

7 Conclusions and Further Work

One of our objectives to run the ViP-JogoMan prototype tests was to verify which was the influence of the technology in GMABS methodology and in the interaction between the players. We have concluded by these tests that the use of GMABS methodology by *Web* is efficient and practical, because the prototype becomes available in remote places and for a greater number of people. We have also concluded that the computer and the *Web* helped us to mapped the complete negotiation process of the players, as presented in Table 1. The second objective to run the ViP-JogoMan prototype tests was to judge if the use of the behavioral profiles and the BDI architecture in virtual players was adequate. This type of player has shown a non trivial decision-making process with respect to real players, and the use of these technologies to implement virtual players can be a good solution to bring realism and to help real players to learn about the domain.

An immediate improvement of the ViP-JogoMan prototype could be the implementation of a dynamic knowledge data base of virtual players. We implemented the virtual players in a static way, but we want to insert new beliefs and plans into the profiles according to the past actions of the players during previous rounds. This will probably improve the set of actions of each profile and the game will become more realistic.

Acknowledgments

The authors would like to thank Dr. Raphaèle Ducrot and the Negowat Project (Financed by European Community: grant number ICA4-CT-2002-10061), for the help in ViP-JogoMan definitions and tests; and Anarosa Brandão for the help in structuring and organizing this paper.

References

1. D. Adamatti, J. Sichman, P. Bommel, R. Ducrot, C. Rabak, and M. Camargo. JogoMan: A prototype using multi-agent-based simulation and role-playing games in water management. In N. Ferrand, editor, *Join Conference on Multi-Agent Modeling for Environmental Management. CABM-HEMA-SMAGET*, Bourg-Saint-Maurice, Les Arcs, France, 2005. IRD.
2. L. O. C. Alvares and J. S. Sichman. Introdução aos sistemas multiagentes. In *Jornada De Atualização Em Informática*, pages 1–37, Brasília - UnB, 1997. Sociedade Brasileira de Computacão.
3. S. Bandini, S. Manzoni, and G. Vizzari. RPG-Profiler: a MAS for role playing games based tests in employee assessment. WOA 2002, 2002.
4. O. Barreteau. The joint use of role-playing games and models regarding negotiation processes: characterization of associations. *JASSS*, 6(2), March 2003. http://jasss.soc.surrey.ac.uk/6/2/3.html.
5. O. Barreteau, C. Le Page, and P. D'Aquino. Role-playing games, models and negotiation. *JASSS*, 6(2), March 2003. http://jasss.soc.surrey.ac.uk/6/2/10.html.
6. R. Bordini and J. Hubner. JASON: A Java-based Agentspeak interpreter used with Saci for multi-agent distribution over the Net. http://jason.sourceforge.net/, 2004.
7. Cormas. Natural resources and multi-agent simulations, 2006. http://cormas.cirad.fr.
8. P. D'Aquino, C. Le Page, F. Bousquet, and A. Bah. Using self-designed role-playing games and a multi-agent systems to empower a local decision-making process for land use management: The selfcormas experiment in Senegal. *JASSS*, 6(3), June 2003. http://jasss.soc.surrey.ac.uk/6/3/5.html.
9. FIPA. Communicative act library specification. http://www.fipa.org, 2002.
10. N. Gilbert and K. G. Troitzsch. *Simulation for the Social Scientist.* Buckingham and Philadelphi: Open University Press, London, 1999.
11. P. Guyot and S. Honiden. Agent-based participatory simulations: Merging multi-agent systems and role-playing games. *Journal of Artificial Societies and Social Simulation*, 9(4), 2006. http://jasss.soc.surrey.ac.uk/9/4/8.html.
12. J. F. Hubner and J. S. Sichman. SACI: Uma ferramenta para implementação e monitoração da comunicação entre agentes. In M. Monard and J. S. Sichman, editors, *IBERAMIA/SBIA 2000, Open Discussion Track*, pages 47–56, Atibaia - São Paulo - Brasil, 2000. Springer Verlag.
13. R. B. McKersie and N. O. Fonstad. Teaching negotiation theory and skills over the internet. *Negotiation Journal*, 13(4):363–368, October 1997.
14. S. R. Peppet. Teaching negotiation using web-based straming video. *Negotiation Journal*, 18(3):271–283, July 2002.
15. H. Raiffa. *The art & science of negotiation - How to resolve conflicts and get the best out of bargaining.* Harvard University Press, MA, 1982.
16. A. S. Rao. AgentSpeak (L): BDI agents speak out in a logical computable language. In W. V. de Velde and J. Perram, editors, *Seventh Workshop on Modelling Autonomous Agents in a Multi-Agent World (MAAMAW'96)*, pages 42–55, London, January 1996. Eindhoven - The Netherlands, Lecture Notes in Articial Inelligence - Springer-Verlag.
17. W3C. W3C - World Wide Web Consortium - SOAP - Simple Object Access Protocol - Specifications. http://www.w3.org/TR/soap/, 2005.

Using Collective Rewards and Social Interactions to Control Agricultural Pollution: Explorations with FEARLUS-W

Nicholas M. Gotts and J. Gary Polhill

Macaulay Institute, Craigiebuckler, Aberdeen. AB15 8QH. United Kingdom
{n.gotts, g.polhill}@macaulay.ac.uk

Abstract. A major difficulty in reducing diffuse agricultural pollution is the cost of monitoring and enforcement. Spatially-explicit agent-based modelling work undertaken with the model FEARLUS-W in relation to the EU Water Framework Directive suggests a possible way forward, using social interactions between farmers in combination with a collective financial incentive.

1 Introduction

The paper reports exploratory work with FEARLUS-W[1], an extension of the spatially explicit agent-based model of land use change FEARLUS [1]. FEARLUS-W was designed to study interactions between environmental policy, rural social networks and pollution, originally in relation to diffuse agricultural phosphate pollution and the EU Water Framework Directive [2]. Controlling diffuse pollution of any kind raises monitoring and enforcement problems; we intend to investigate the feasibility of mutual monitoring among groups of farmers, offered a collective reward if measured levels of the pollution they produce between them are satisfactory. For phosphate or other water pollution, this would need to involve the farmers in a river catchment or sub-catchment, pollution levels being measured just downstream of the furthest downstream farms. However, the model is sufficiently general to be applied to airborne pollutants or greenhouse gases, if levels of these can be measured only, or at lower cost, over an area containing multiple farms.

2 Method

FEARLUS-W differs from original FEARLUS in several ways. Each *Land Use*[2] has an associated *Pollution Intensity*, and Land Managers select Land Uses in a different way, described below. In original FEARLUS Land Managers decide Land Uses for each *Land Parcel* they manage; their *Income* is calculated given this, the Parcel's

---

[1] This term refers to several versions of FEARLUS. FEARLUS-0.8.2.1 was used here.
[2] Terms for FEARLUS entities are given initial capitals, and italicised on first use.

Biophysical Properties, and the *Year*'s *Climate* and *Economy*; and those in debt sell Parcels to regain solvency. FEARLUS-W adds a phase in which total *Pollution* is calculated, and a *Government Agent* can issue a *Reward* to all Land Managers if it is below a specified threshold; and another in which Land Managers apply a *Social Approval Function* to each Neighbour, and receive information about their Neighbours' *Approval* or *Disapproval* of them.

The design of FEARLUS-W agents is motivated by features of human decision-making (and particularly the decision-making of farmers) which distinguish it from the *Homo economicus* picture of a computationally unlimited profit- or utility-maximiser.

First, FEARLUS-W Land Managers have preferences which cannot be reduced to a single measure of "utility". The Land Manager agents in the simulations in which the Government Agent pays a Reward if total Pollution is below a threshold face a form of social dilemma, if they are considered as profit maximisers: each individually will generally be better off if they adopt the Land Uses with high Pollution Intensity and mean profitability; but if all (or enough) do this, the Reward will not be paid. There is abundant empirical evidence [3,4] that people often act more cooperatively than if they were rational profit maximisers in such situations; and that this is more likely to occur if there is or has been social contact between those involved. In the context of farming, [5] describes how farmers are influenced by their desire to be seen as a "good farmer" by their peers. Hence our Land Managers "care" about their Neighbours' Approval or Disapproval, as well as about their Income. However, the relative importance given to the two factors varies from Year to Year, changes being prompted by specific *Events* – such as a bad harvest, or a specified degree of unpopularity in the preceding Year. While we cannot cite specific empirical support for such a "salience" effect, it is intuitively plausible that in weighing very different types of benefits and costs (concerning income and peer approval), farmers are influenced by recent events drawing their attention to one factor or the other.

Second, in making their decisions, farmers have to cope with uncertainty about the future, and limits to their own knowledge. Previous work with FEARLUS has concentrated on two ways of doing this: the use of *Aspiration Thresholds* for Income – levels of Income from particular Land Parcels that will lead to the same Land Use being employed again; and the imitation of Neighbours. In FEARLUS-W, we explore an alternative to imitation: case-based reasoning (CBR).

CBR is an approach to problem-solving based on work in cognitive science [6,7,8], and extensively used in artificial intelligence based decision-support systems [9,10]. The central idea is that problems can often be solved (and often are solved by human beings) by recalling and if appropriate reusing or adapting remembered solutions to similar problems encountered in the past. In the current context, that means recalling situations in which the relevant circumstances were similar to those now encountered, and the outcomes that followed the choice of specific Land Uses on those occasions.

FEARLUS-W Land Managers have Aspiration Thresholds for both Income (considered on a Parcel-by-Parcel basis) and Approval. If the latter is met, the same Land Use that was used last year is retained on every Land Parcel on which the Threshold Income was obtained, while a *Decision Process* is set up for every other Parcel. If the Approval Threshold is not met, such a Process is set up for every Parcel managed.

The Decision Process for a Land Parcel begins with CBR-based estimation of the levels of Income and Approval that will result from adopting each of the possible Land Uses in the considered Land Parcel. This estimation process will produce a map like that sketched in Figure 1.

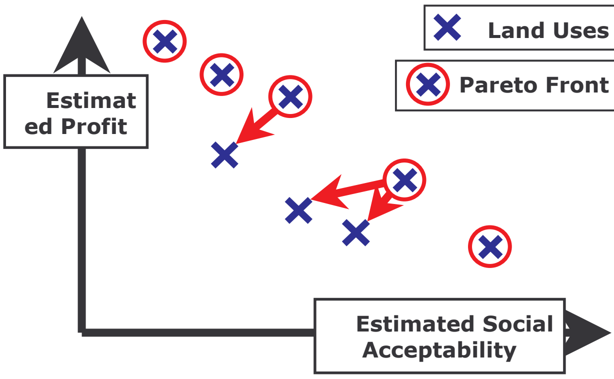


Fig. 1. Estimation of the social acceptability and the profit that different Land Uses provide in a certain Land Parcel.

In the general case there will not be a Land Use that outperforms all others in both dimensions of utility. Figure 1 shows a situation in which three Land Uses can be eliminated immediately, because at least one other Land Use is expected to outperform them on both dimensions. The remaining five form the "Pareto front": if any member of this is selected, it is impossible to find another which is expected to be at least as good on *both* dimensions. The choice within the Pareto front is made by calculating an overall index of worth or utility called Combined Utility. The Combined Utility of each Land Use *lu* is calculated using a linear function which guarantees that only Pareto front members are selected (as long as weights are positive).

$$Combined\_Utility[lu] = Income[lu] * income\_weight + Approval[lu] * approval\_weight \quad (1)$$

The weights for the two dimensions of utility may differ between Land Managers, and also vary in time for each particular Land Manager depending on the occurrence of certain Events, as mentioned above. Specifically, each Land Manager has a minimum value for each of the two weights, and a certain amount of units of "mobile worry" that they can assign to either dimension of utility to form the actual weights. These three values are specified at the Subpopulation level, using a probability distribution, and they satisfy the following constraint for each Land Manager:

$$\begin{aligned} incomeMinSalienceDist + approvalMinSalienceDist + salienceMarginDist = \\ = CONSTANT = \\ = income\_weight + approval\_weight \end{aligned} \quad (2)$$

The "mobile worry" is initially divided between the two dimensions in the ratio *incomeMinSalienceDist* / *approvalMinSalienceDist.* From then onwards, particular Events such as a *Bad Harvest* (low Income from the Land Use itself, leaving aside any Reward from the Government Agent), or a Neighbour's Disapproval, will cause a shift of a certain number of units of mobile worry, provided there are any to be shifted.

The main component of the CBR selection algorithm used to estimate Income and Approval outcomes is the Land Manager's Case Base, or Episodic Memory. This Episodic Memory is a set of ordered lists of cases. Each case represents the experience of the Land Manager in a certain Year *Y* on a certain Land Parcel. Thus, a case comprises:

- The Year *Y* when the experience occurred.
- The perceived state of the world in Year *Y*, which consists of some of the conditions that interacted with the decision (i.e. the particular Land Use applied) in determining the outcome of the decision.
- The decision: the Land Use applied to the Land Parcel at Year *Y*.
- The outcome of applying the selected Land Use on the Land Parcel under consideration. This consists of the Income and the net Social Approval (i.e. the overall balance between Social Approval and Disapproval from Neighbours) that the Land Manager got in Year *Y*.

The Episodic Memory is organized in different compartments, each of which stores all the cases pertaining to a particular Land Use. Cases within each compartment are stored in an ordered list, with the first cases in the list representing the experiences that are most fresh in the Land Manager's memory. Thus, at the end of a Year, the cases derived from experience of the current Year are stored at the head of the appropriate lists in random order. In each compartment of the Episodic Memory, immediately behind the current Year's cases, there are the other cases corresponding to previous experiences in an order such that the most recently used cases appear first. The size of the Land Managers' Episodic Memory is unlimited.

The estimates for each Land Use are taken directly from the most appropriate case for the specific Land Use in the considered Land Parcel given the current perceived state of the world. The two factors that affect which case is most appropriate are recency of use and similarity between the Land Manager's perception of the current state of the world and that at the time the case occurred. This is intended to (crudely) mimic some features of how human cognition works – that we are most likely to recall cases we thought about recently, and accept or reject them as input to current decisions on the basis of similarity.

The Land Manager's perception of the relevant aspects of the current state of the world consists of the following three descriptors, all of which consist of bitstrings:

- The Biophysical Characteristics of the Land Parcel.
- The most recent Climate (corresponding to the previous Year).
- The most recent Economy (corresponding to the previous Year).

Let *A*, *B*, and *C* be three states of the world.

- *A* is MORE similar to *C* than *B* is if and only if at least one descriptor in *A* is more similar to the corresponding descriptor in *C* than *B*'s corresponding descriptor is (i.e. A and C match at all the bitstring places B and C do, plus at least one m ore), and the rest of the bitstring descriptors are equally similar to *C* as *B*'s (match in exactly the same places).

– *A* is LESS similar to *C* than *B* is if and only if *B* is MORE similar to *C* than *A* is.

– *A* is EQUAL to *B* relative to *C* if and only if *A* is exactly equal to *B* (i.e. the three descriptors are the same bitstrings).

– *A* is INCOMPARABLE with *B* relative to *C* in the rest of the cases. This occurs when *A* is MORE similar to *C* than *B* in some descriptors, and LESS so in others.

Proceeding down the list of cases for a specific Land Use, the Land Manager will select the first case for which it is true that there is no other case in the list associated with a state of the world MORE similar to the current state of the world.

In the exploratory work reported here, four runs were undertaken for each of thirty-four FEARLUS-W parameter sets. All used a 20x20 grid of Land Parcels and five Land Uses (Pollution Intensity and mean profitability being inversely correlated), and ran for 100 Years. (This duration may seem unduly long, given the technological and social changes that could be expected over such a period. It was motivated by the time Pollution levels generally take to stabilise, and the importance of then running the model long enough to get some idea how robust this stability is.) Each parameter set used a different combination of the Social Approval Function and salience-changing Events used, and conditions on the Reward given. (The point of varying the Social Approval Functions and salience-changing Events is simply to provide an initial check on the sensitivity of the model to such changes.) The Government Agent either gave no Reward; or demanded total Pollution below 2000, 1750 or 1500 units (2000 is the expected level if all Land Managers choose among the five possible Land Uses at random and with equal probability), and gave a Reward of 50 or 25 monetary units per Parcel managed (breaking even required an Income of 55 units per Parcel). Land Manager Agents with a Social Approval Function ("SAF" agents; "NoSAF" Agents have none) differed in whether they disapproved of high absolute per-Parcel Pollution levels ("Abs"), approved of those with lower mean per-Parcel Pollution levels than their own and disapproved to an equal extent of those with higher mean levels ("Rel"), or disapproved of the latter to a greater degree than they approved of the former ("Dis"). All Land Managers increased the salience of profitability after a bad harvest, but some increased the salience of Social Approval when a Neighbour disapproved of them ("D"), others when no Reward was given ("N"). All Agents in a given run were of the same type. Note that it makes no sense to combine a no Reward condition with type N agents.

3 Results and Discussion

Table 1 indicates that both the Reward conditions and the agent type used affected total Pollution. Because periods of unusually high pollution may cause disproportionate damage, the mean (over the four runs) for the worst Year in a run is given, as well as the grand mean over all years in all four runs. Table 2, added for clarity, extracts the ordinal information implicit in Table 1: of the cells representing mean Pollution levels, that representing the parameter set producing the lowest mean

Pollution level contains “1”, the next lowest a “2”, and so on; cells representing mean worst Years are numbered similarly.

There are apparent differences both between Reward conditions and between Land Manager types, and these show similar patterns for mean and maximum Pollution levels. However, it should be reiterated that this is exploratory work: the simulation runs were carried out without prior decisions concerning the statistics to be gathered, or formulation of null and experimental hypotheses concerning possible differences between conditions – so the question of whether the apparent differences are statistically significant does not even arise. That said, it may be noted that among Reward conditions, a Pollution threshold of 2000 and Reward of 50 generally did best, and reducing either Reward or threshold increased Pollution, with the no Reward condition being clearly worst. Across Land Manager types, the differences are smaller, except with respect to the contrast between the NoSAF agents (those which do not care what their neighbours do, for which Pollution levels are highest) and the rest. Under most reward conditions the lowest Pollution levels are produced by the Abs-N agents, who approve or disapprove of their neighbours according to their absolute mean Pollution levels, and increase the salience of neighbours’ opinions when the Reward is not given. Differences between Land Manager types diminish as Reward and Pollution threshold fall.

The result that higher Rewards for not polluting decrease Pollution levels is entirely unsurprising, and the better results for higher Pollution thresholds could be expected as part of a larger pattern: clearly, in the real world, if a pollution threshold is so low as to be unachievable, or so high as to be impossible not to achieve, it will have no effect in lowering pollution: if any threshold-conditioned reward scheme is to be useful, usefulness might be expected first to rise, then to fall as the level of the threshold rises. For all that, these provisional findings are of some interest: if the simulation results conflicted with these expectations, it would be clear the model should be discarded. The finding that overall means and means of maximum Pollution levels show very similar patterns is less predictable, as is the fact that, over the parameter sets tested, mean Pollution levels are almost always above the Threshold, but that the existence of the Threshold nonetheless reduces this mean level. The general pattern of a simulation run in which the Reward had a marked effect was an initial rise in Pollution as the Land Managers learned which Land Uses were profitable, followed by a fall as the dislike of Social Disapproval came into play, then possibly more rises and falls until a fall took the Pollution level below the threshold and enabled the Land Managers to begin earning, and learning about, the Reward – but even after Years below the threshold, high Pollution levels could return, presumably at times when the low-Pollution Land Uses were giving very poor Income other than the Reward, under which conditions the CBR mechanism may produce a switch to high-Pollution Land Uses (recall that the salience of Income is increased by a Bad Harvest even if the Reward is keeping the Land Mangers solvent).

Table 2 suggests that there may in fact be some differences in the way overall mean and mean of the worst Year are affected by changes in the parameters used. Specifically, when combined with the less effective Reward conditions (the last three columns), Land Managers of the Dis-D class (disapproving strongly of Neighbours with higher mean annual Pollution levels than their own, and increasing the salience of Social Approval when a Neighbour disapproves of them) appear markedly more effective at limiting Pollution in the worst Years than mean Pollution over all Years. The high level of disapproval may act to restrain Pollution levels even though these are too high for the Land Managers to have the opportunity to learn the economic advantages of restraint by being rewarded.

Table 1. Annual total Pollution levels: grand means (upper row of each pair) and means of maximal values (lower row of each pair)

Agent Type	REWARD:	2000;50	2000;25	1750;50	1750;25	1500;50	None
Abs-D	Mean	2062.8	2083.1	2189.2	2233.6	2277.4	2355.6
	Mean Max	*3026.0*	*3216.0*	*3188.0*	*3292.5*	*3277.0*	*3323.5*
Abs-N	Mean	1987.7	2122.1	2034.1	2131.3	2156.9	N/A
	Mean Max	*2877.5*	*3174.0*	*3178.0*	*3190.0*	*3262.0*	*N/A*
Rel-D	Mean	2045.5	2191.6	2184.8	2329.1	2278.8	2379.5
	Mean Max	*2959.0*	*3195.0*	*3304.0*	*3299.0*	*3293.5*	*3342.0*
Rel-N	Mean	2043.0	2226.6	2199.4	2287.5	2223.0	N/A
	Mean Max	*3061.5*	*3239.0*	*3268.0*	*3269.5*	*3252.0*	*N/A*
Dis-D	Mean	1984.6	2143.1	2168.5	2271.2	2271.9	2324.0
	Mean Max	*2924.0*	*3196.5*	*3224.5*	*3227.0*	*3219.5*	*3265.5*
NoSAF	Mean	2209.9	2484.1	2509.0	2578.9	2634.3	2634.3
	Mean Max	*3262.5*	*3360.0*	*3393.5*	*3398.0*	*3437.5*	*3437.5*

Table 2. Ordinal information concerning annual total Pollution levels, implicit in Table 1: lower numbers indicate lower grand means (upper row of each pair) and means of maximal values (lower row of each pair)

Agent Type	REWARD:	2000;50	2000;25	1750;50	1750;25	1500;50	None
Abs-D	Mean	6	7	14	20	23	28
	Mean Max	*4*	*12*	*8*	*24*	*23*	*28*
Abs-N	Mean	2	8	3	9	11	N/A
	Mean Max	*1*	*6*	*7*	*9*	*18*	*N/A*
Rel-D	Mean	5	15	13	27	24	29
	Mean Max	*3*	*10*	*27*	*26*	*25*	*29*
Rel-N	Mean	4	19	16	25	18	N/A
	Mean Max	*5*	*16*	*21*	*22*	*17*	*N/A*
Dis-D	Mean	1	10	12	21	22	26
	Mean Max	*2*	*11*	*14*	*15*	*13*	*20*
NoSAF	Mean	17	30	31	32	33=	33=
	Mean Max	*19*	*30*	*31*	*32*	*33=*	*33=*

Future work will focus on testing these tentative conclusions using methods similar to those described in [1]; on exploring a wider range of Social Approval Functions (e.g. allowing Agents to disapprove of others with higher mean Pollution only if this is above an absolute threshold, or only if the Reward is not given – currently the latter can only influence salience); and on increasing realism by basing Pollution Intensities and profitability conditions on specific real-world land uses, using a new version of FEARLUS. The approach described in [1] consists of formulating specific hypotheses about how the model will behave in specific circumstances, and testing these hypotheses in experiments using non-parametric statistics. For example, we will test the apparent superiority of a Pollution threshold at the level expected if random Land Use chocies are made, over lower thresholds; and whether pushing this threshold higher gives better or worse results. Whether the optimal Pollution threshold itself changes as the size of the Reward varies will also be investigated. We will investigate whether Land Managers using completely different approaches to selecting Land Uses (e.g. imitation of Neighbours rather than CBR) give similar results: that is, how sensitive the model is to changing assumptions about how Land Managers make their decisions. Comparisons with work such as [11], which used a more detailed physical model but did not explore relations of social (dis)approval between farmers, will also be pursued.

Acknowledgements

This work was funded by the Scottish Executive Environment and Rural Affairs Department under the "Environment - Land Use and Rural Stewardship" Programme, and by the European Commission under the Sixth Framework Programme CAVES project, no. 012816. Thanks to Luis Izquierdo and Ben Davies for assistance, and to two anonymous reviewers, one of whom made very helpful comments.

References

1. Polhill, J. G., Gotts, N. M. and Law, A. N. R..: Imitative Versus Non-Imitative Strategies in a Land Use Simulation. Cybernetics and Systems 32 (2001) 285–307
2. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. Official Journal L 327 (2001) 0001-0073
3. Ostrom, E., Gardner, R., and Walker, J. M., Rules, games, and common-pool resource*s*. Ann Arbor: University of Michigan Press, 1994
4. Gotts, N. M., Polhill, J. G., and Law, A. N. R.: Agent-based simulation in the study of social dilemmas. Artificial Intelligence Review, 19 (2001) 3-92
5. Burton, R. J. F.: Seeing through the 'good farmer's' eyes: towards developing an understanding of the social symbolic value of 'productivist' behaviour. Sociologia Ruralis 44:2 (2004) 195-215
6. Schank, R.C.: Dynamic Memory: A Theory of Reminding and Learning in Computers and People. New York, NY: Cambridge University Press.
7. Aamodt, A., Plaza, E.: Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. AI Communications 7 (1994) 39–59

8. Lopez de Mantaras, R., McSherry, D., Bridge, D., Leake, D., Smyth, B., Craw, S., Faltings, B., Maher, M.L., Cox, M.T., Forbus, K., Keane, M., Aamodt, A. and Watson, I.: Retrieval, reuse, revision and retention in case-based reasoning. The Knowledge Engineering Review, 20:3 (2006) 215–240
9. Leake, D.B.: CBR in context: the present and future. In Leake, D.B. (ed.) Case-Based Reasoning: Experiences, Lessons & Future Directions. Cambridge, MA: MIT Press (1996) pp.3-30
10. Cheetham, W. and Watson, I.: Fielded applications of case-based reasoning. The Knowledge Engineering Review, 20:3 (2006) 321–323
11. Janssen, M.A.: An Exploratory Integrated Model to Assess Management of Lake Eutrophication. Ecological Modelling 140 (2001) 111-124

Agent-Based Land Markets: Heterogeneous Agents, Land Prices and Urban Land Use Change

Tatiana Filatova[1], Dawn C. Parker[2], Anne van der Veen [1,3]

[1] University of Twente, Drienerlolaan 5, 7522 NB, Enschede, The Netherlands
[2] Center for Social Complexity, George Mason University, 4400 University Drive MS6B2, Fairfax, VA 22030, USA
[3] International Institute for Geo-Information Science and Earth Observation, Hengelosestraat 99, P.O. Box 6, 7500 AA Enschede, The Netherlands

T.Filatova@ctw.utwente.nl

Abstract. We construct a spatially explicit agent-based model of a bilateral land market. Heterogeneous agents form their bid and ask prices for land based on the utility that they obtain from a certain location (house/land) and based on the state of the market (an excess of demand or supply). We underline the distinction between bid /ask price and individual willingness to pay/to accept and show that variations between them that reflect market conditions can influence land prices. Agents sort among locations with respect to distance from the city center and environmental spatial externalities. Aggregated outcomes such as land patterns and land prices are produced by the model. The basic model of buyers and sellers trading land in the urban area produces results identical to the monocentric urban model. However, more complex dynamics appears when environmental amenities and market-adjustment variable influence the formation of land prices.

Keywords: urban land market, heterogeneous agents, division of gains from trade, open-space amenities.

1 Introduction

The benefits of applying agent-based modeling (ABM) to study economic and ecological-economic systems are widely discussed in the literature [1-4]. Parker et al. [4] suggest that ABMs are appropriate when agent heterogeneity and interactions and cross-scale dynamics drive land-use outcomes. Further, ABMs may be useful for modeling the path of land-use change, whether or not that path may lead to equilibrium. Starting from the classical books [5, 6] the society of economists who adopt ABM tools to model economic problems is growing. ABM has been successfully used to model economic markets from bottom-up since mid 1990s [7-9]. Recent progress in designing markets with ABM is discussed at length by R. Marks and B. LeBaron [10, 11]. However, there is relatively little research done on simulating land markets with ABM. This can be explained by the fact that land markets are a very specific kind of markets. In comparison to other types of markets designed with ABM, a land market has not only (e.g., fundamentalists and chartists in financial markets) but also (each parcel of land or house has very different characteristics, which determine agents' willingness to pay and consequently market prices).

Several models study effects of hypothetical land markets, but with primary emphasis on the demand side. The SOME and SLUCE models allow agent to choose the parcel that maximizes their utility without competition from other sellers and assuming that the locating agent will outbid the current use [12]. Other models of agricultural land markets model both the demand and supply decisions [13-16]. While these models are becoming increasingly more sophisticated, they do not model differences between the buyer's willingness to pay –WTP – (underlying utility or payoffs for the land) and her bid or offer price for the land; nor do they model differences between the seller's willingness to accept – WTA – (opportunity cost of the sale) and his ask price.

This paper presents an agent-based bilateral residential land market, with a particular focus on the formation of WTP/WTA versus bid/ask price for land as well as the division of gains from trade. Moreover, we explore how heterogeneous preferences for green amenities influence the emergence of land prices via an endogenous land market. We model both the demand and supply sides and their negotiation over the final transaction price explicitly. This negotiation might depend on previous trades in the neighborhood and on some aggregate market information (relative excess of buyers or sellers). Our model also differs from previous ABM land market models through its focus on urban markets, rather than rural markets and farmer decision making. A WTP for land of a farmer can be directly expressed in monetary terms (derived as potential agricultural gains net of costs). In contrast, it is less obvious how to derive a WTP/WTA of a household in monetary units, since it is influenced by both market and non-market values. A household obtains a certain level of utility when she buys a spatial good (land lot/house). Her preferences, such as preferences for green amenities, are expressed in non-monetary units, which must be monetized to express her WTP and bid price.

The paper proceeds as follows. First, we discuss the economic approach to modeling residential land market and the advantages of an ABM for this problem. The main section of the paper presents the bilateral agent-based model of residential land market. Here we discuss interactions between traders and differences between willingness to pay/accept for a spatial good and the actual bid/ask price. Next, the results of several simulation experiments are presented. We conclude with a discussion of future elaborations of the model and a potential application to the case study in the Netherlands.

2 The traditional Economic Approach to Modeling Urban Land Use and Value Added of ABM

We underline the importance of using existing theoretical and empirical work done in the fields of spatial, urban and environmental economics in constructing ABM of land use with an endogenous market mechanism. ABM should be viewed as a way to supplement traditional scientific methods and expand the boundaries of science to test hypothesis and undertake experiments, rather then as substitute for traditional methods [4]. Many traditional models of urban land markets find their roots in the model of W.Alonso [17]. According to his bid-rent theory, households choose locations a certain distance from the central business district (CBD) based on the utility they get from the joint consumption of a spatial good (land lot/house) and a composite good (all other goods) under their budget constraint. Both demand for two types of goods (spatial and composite) and prices households are ready to pay are derived as a solution of a constrained maximization problem. Microeconomic studies usually use standard mathematical tools for that, such as a Lagrange function [18]. However, in order to be able to obtain analytical solutions for all variables, economics makes several assumptions, which may not hold for land markets with interacting buyers and sellers and explicit modeling of land price negotiation. There are several reasons why analytical solutions only cannot be used and ABM needs to be applied:

1. A representative agent approach is usually used in spatial economics models. This means that a demand curve of one agent can be extrapolated to represent the demand for the particular good (e.g., land) in the whole economy. However, the preferences for location differ among different economic agents, such as firms or households, and within these groups. The limitations of a representative agent approach are widely discussed by A. Kirman [19].
2. In a land market context, the utility of an agent is often dependent on the behavior of others, especially when actions of immediate neighbors change the local environment. However, traditional analytical models do not account for interactions among agents. The decision to convert open space into developed use or provision of a public good such as good quality schools influences the utility of households living in the neighborhood. These spatial externalities/neighborhood effects are difficult or impossible to model analytically in a spatially explicit context.
3. In addition to agent heterogeneity and interactions, the spatial environment is also heterogeneous. Spatial goods differ in their availability of environmental and social (schools, theatres, etc.) amenities, soil quality, administrative affiliation (and consequently taxes), and zoning and spatial planning regulations. Analytical methods used in economic theories dealing with space usually treat these characteristics of spatial goods as being distributed , denying heterogeneity. However, different characteristics of land produce different land use patterns and land prices, which are usually not uniformly distributed.
4. In an analytical setting a market clears only at the equilibrium state. While finding analytical solution we have to assume that system comes into equilibrium in one shot, as if all agents make optimal decisions at one moment. In reality, decisions of different agents to buy/sell spatial goods are separated in time, meaning that in time step situation on the market and spatial neighborhood will be different than in time step and agent in time step t will be willing to pay different price for the same land lot/house. Economic markets, including land markets, are dynamic adaptive systems rather than static models [20], and the temporal interdependency of agent decisions may produce path dependence [12]. Often real-world economic systems are out-of equilibrium

[21], and systems may move towards, but not reach, equilibrium (for instance, in a case where cyclical adjustments of changing neighborhoods, changing patterns of land use, and land price adaptations occur). The dynamic path towards equilibrium can be elucidated through agent-based market modeling.

Keeping some main axioms of economic theories dealing with space (such as the fact that people want to minimize travel cost and value green amenities) as a basis, ABM makes a step forward by allowing agents and spatial environment to be heterogeneous, modeling agents' interactions, and modeling dynamics in a 2D spatially explicit setup. At the same time, by modeling spatial and market interactions between buyers and sellers explicitly, ABM allows modelers to obtain patterns of land prices and total urban territory occupied, as economic theory does.

3 An Artificial Land Market

The ABM of an Artificial Land MArket (ALMA), which model explicitly interactions between buyers and sellers of spatial goods and feedbacks of the market transactions, is based on the following conceptual scheme (see Fig. 1).

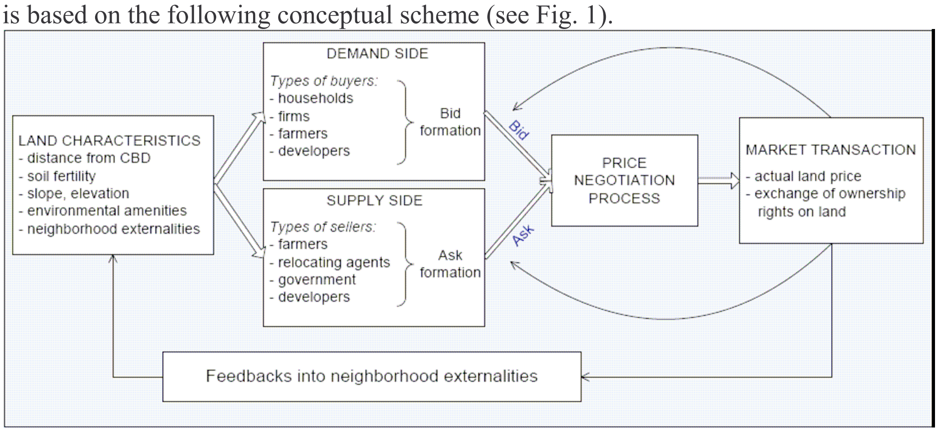


Fig. 1. Conceptual scheme of the land market

Depending on the characteristics of the spatial good, buyers and sellers form their "bid" and "ask" prices for land. Both supply and demand sides are represented by several types of agents. Buyers and sellers participating in land market may differ in nature, in their motivations to buy or to sell a land parcel or a house, and in their preferences for location. Each type of land market actor might appraise the same land parcel or house differently. When two trading partners are able to agree upon transaction price of a spatial good (decide how gains from trade will be divided), then the land is transferred to a new user or may be converted into another land use (e.g., developers buying land to convert it into the developed state). Thus, prices for spatial goods are formed endogenously via interactions of heterogeneous economic agents. If a spatial good changes its owner, then the structure of the neighborhood changes as well and feeds back into spatial externalities. At the same time, actual transaction prices affect the formation of ask and bid prices in the future. The ALMA model produces urban land patterns and land prices (rent gradients) as a result of market allocation of land between competitive users.

3.1 Types and Characteristics of Economic Agents

The main agents in ALMA model are land users operating in an urban area. The main good they exchange via market mechanisms is a spatial good, which can be viewed as a land lot or a house. The landscape is represented by a grid of equal cells. Each can be owned by one economic agent. The ALMA model was developed using the NetLogo 3.1.4 environment.
In this first implementation, households and developers are the main economic agents. Households are assumed to be buyers and sellers, and developers may convert open-space parcels. We do not aim to model all potential land market actors as presented in Figure 1 and leave it for future research.

The land market in the ALMA model is presented as a two-side matching market. Sellers represent households who want to relocate. Buyers are households searching to buy a house/land lot. At the moment of model initialization, the number of buyers and sellers is defined, each seller holds one land lot, and each buyer has a certain budget. The location choices of agents are based on their preferences expressed in the form of utility function (Formula 1).

$$= \alpha \cdot \ln \quad + \beta \cdot \ln \frac{\quad}{+1} \cdot \tag{1}$$

Here: α is agent's preference for green amenities, which can be heterogeneous (uniformly distributed in the range from 0 to 1) or homogeneous among agents;
β is an agent's preference for commuting; at this point βis equal to one for all agents;
k is a constant.
The utility of a spatial good (house/land) decreases with the distance ( ) from the CBD and increases if environmental amenities ( ) are available. Green areas, which could be considered as parks, forest areas or open space, are defined at the initialization stage. One can choose whether to have green areas only at the city fringe or to spread them over the whole city. Both the type of location of green areas and their amount can be defined by a user.
The choice of location is constrained by the budget (Formula 2), where is an agent's disposable budget for housing (it is assumed to be equal to some constant Y since we do not assume that market clears in one shot and, thus, cannot determine housing budget share via constrained maximization solution), equals transport costs, and is price for land the person can afford.

$$= \quad (\quad) + \quad . \tag{2}$$

Agents are assumed to maximize utility by choosing the optimal location under the budget constraint. However, there are two important distinctions from neoclassical utility-maximization problem. We assume that economic agents are boundedly rational (agents are not able to predict how the neighborhood will develop in the future) and do not possess perfect information about their environment (agents look for a local maximum among the sample of spatial goods they have randomly chosen). The algorithm, which artificial traders follow, is presented in Figure 2.

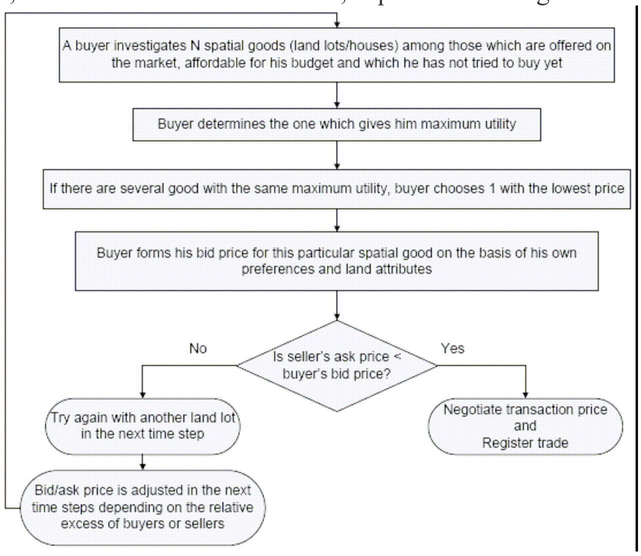


Fig. 2. Algorithm of trade

3.2 Ask and Bid Price Formation

An important question to be asked while modeling agent-based markets is the question of price formation. Economic theory suggests that buyers form their willingness to pay for land (WTP) based on their preferences for different land attributes under their budget constraint, and sellers form their willingness to accept (WTA) based on their profit expectations. WTP/WTA can be derived by monetizing utility (Formula 1). The difference between WTP and WTA defines the gains from trade—the economic surplus that can be captured from the market transaction. For homogeneous goods, the market-clearing price is assumed to be the price at which each trade will occur, and it is thus used to define the way in which gains from trade will be divided. However, residential land is generally sold through bilateral bidding and negotiation. A clear distinction should be made between WTP and a bid price, and WTA and an ask price. WTP/WTA is that threshold level (reservation price), above/below which a trader is not able to buy/sell. A bid/ask price is the price that the agent announces in the market when it comes on the stage of negotiation (see Figure 1). Since economic agents try to maximize the proportion of surplus that they capture from the transaction, they tend to form their final bid/ask prices for land lower/higher than WTP/WTA, as in Formula 3, where the WTP is adjusted by ε for buyers and ε for sellers.

$$= \quad -\varepsilon \ , \qquad = \quad +\varepsilon \ . \tag{3}$$

However, if a bid price will be too low or an ask price will be too high, the trade will not occur. Thus, the variable ε should not be random number but should reflect the situation on the market (for example, excess of demand or supply [10], or rate of land prices change). In perceived "seller's" markets (when sellers have an advantage), realized transaction prices are often higher than ask prices. In "buyer's" markets, prices are often lowered over time. In the current version of the ALMA model we define ε as in Formula 4.

$$\varepsilon = (\quad - \quad)/(\quad + \quad) \ . \tag{4}$$

Here: NB – number of buyers, NS – number of sellers. For sellers variable also accounts for the average price in the Moore neighborhood of the cell, which seller owns.

3.3 Price Negotiation and Market Transactions

Location choice and the price for the desirable site are determined jointly. However, the market transaction may not take place if two sides are not able to come to a price agreement (e.g. if the sellers ask price is higher than the bid price of a buyer). In the opposite case when a bid price of a buyer is higher than a seller's ask price, there are several possible ways to determine the actual transaction price for the spatial good. Price negotiation mechanisms in the existing ABM of markets vary from simple arithmetic/geometric average of bid and ask prices [14] to sophisticated algorithms, such as auctions [16, 22]. For simplicity at this stage of the ALMA model the price negotiation procedure is implemented as a calculation of the arithmetic average of the seller's price and a bid of a single buyer who finds the offered land lot attractive and can afford to buy it.

After the matching buyer and seller have agreed upon the actual price of the land, the market transaction takes place. At the moment when trade is registered in the ALMA model both buyer and seller update their status (the seller will not sell in the next period and the buyer will not search for a land lot to buy), the ownership rights on the spatial good are transferred from seller to buyer, and the transaction price is registered as the actual price for this specific land lot. The numbers of buyers and sellers left on the market after the transaction will influence the determination of bid and ask prices via the variable ε (see Formulae 4 and 3). The actual transaction price influences future ask price in the neighborhood. The state of the sold land (urban use or green area) feeds back in the spatial characteristics of the neighborhood and influence WTP/WTA for the surrounding cells in the future trades.

4 Simulation Experiments

We run the model with different assumptions and different numbers of agents.

: In the first set of experiment we tried to reproduce the traditional monocentric model proposed by Alonso. The model assumes that the city will grow until the moment when urban rent will be equal to the agricultural rent. We do not model the agricultural profit calculation directly, but as in the model by Alonso assume that agricultural rent is fixed and equal to some constant R_a=24). This potential agricultural profit serves as an ask price for sellers. It is assumed that buyers bid exactly their WTP, so that variable ε (see Formula 3) is equal to zero. In other words, the market situation is not taken into account, only the pure preferences of the economic agents. We run the model on a small cell grid (11x11 cells) and a number of traders equal to 265 (120 sellers and 145 buyers). Buyers' disposable budget Y is equal to 100. The result is that we always get the typical rent gradient for the monocentric city as in Alonso type of models (see Figure 3).

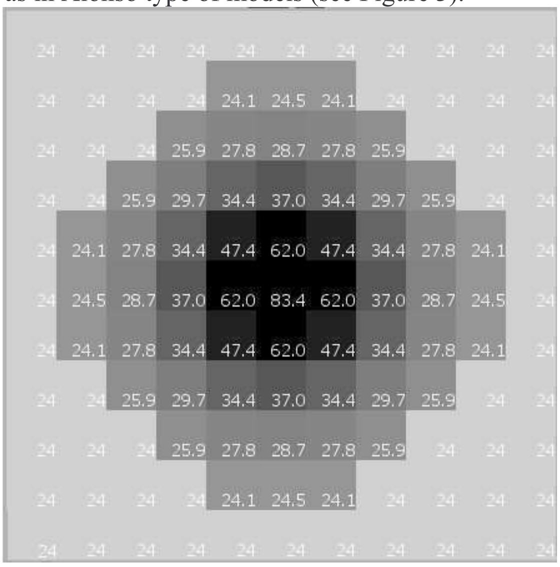


Fig. 3. Rent gradient (bid price is equal to WTP; no green amenities)

In Figure 3, the number in each cell shows the actual market price on which buyer and seller agreed. The intensity of grey color symbolizes the price of the land: the darker the color, the higher the land price. The rent is equal for all the cells that are situated an equal distance from the CBD (as in the Alonso model). The expansion stops expand when the bid price of a buyer is less than the agricultural rent R_a. The lightest-grey area shows the beginning of agricultural (non-sold area) and symbolizes the city border.

Thus, the model of a bilateral land market, with buyers and sellers negotiating over land price is able to produce the traditional spatial economics model. However,

.

Let us assume that buyers form their bid prices not only on the basis of utility but also taking into account market situation (variable ε Formula 3). All other conditions stay the same as the previous exercise. The new assumption implies that we make a distinction between buyer's WTP and her bid price. If bid price is higher or lower than in the Experiment 1-1 for the same property, then the final transaction price will be different as well. Whether land prices will be higher or lower depends on whether there is a "seller's" or a "buyer's" market. In our settings (where the amount of buyers is higher than the amount of sellers) land prices will grow. The result of the rent gradient and actual values of the rent is shown in the Figure 4 below.

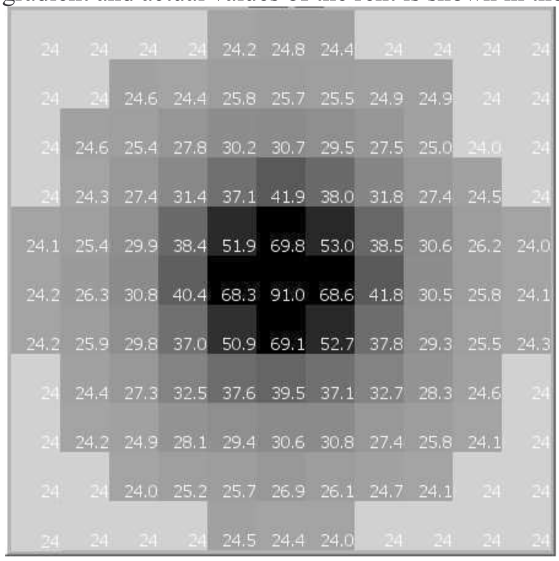


Fig. 4. Rent gradient (bid is a sum of WTP and ε, no green amenities)

As one can see, first of all the boundary of the city expands, as expected. It is caused by the fact that the market structure favors for sellers. So, buyers incrementally raise their bid prices and are able to buy the land further from the CBD even if they originally would have bid less than R_a. The rent gradient still follows the Alonso-von Thunen predictions, i.e. the rent decreases with the distance. However, the structure of the land prices is different than in the previous example: the

, since buyers bid it at different time steps within different market structure (different value of ε). The bigger the gap between the number of buyers and sellers on the market, the larger the deviation of land prices from Figure 3 above (where bid price is equal to WTP and is not biased by the agent's perception of the market situation).

Land prices in general are higher in Figure 4 in comparison with Figure 3. However, sellers' WTA did not grow: it is still the same fixed R_a. Buyers bid higher prices and capture fewer gains from trade. Thus, ε

.

Here we concentrate on investigating environmental externalities and their influence on land prices. Let us bring in 15 green areas spread over the city. The agents have uniformly distributed preferences for green amenities (variable αin

Formula 1). Both buyers and sellers take market information into account while forming their bid and ask prices. Figure 5 shows the rent gradients in the city with green amenities. The cells with letter “G” show where the green areas are. The white cells are the ones which nobody either can afford with the current disposable budget (defined by user) or wants to buy. Here we do not assume the existence of the agricultural land rent equal for the whole landscape as a reservation price of the seller. Each seller has individual ask price (a sum of her WTA, which is determined as a monetized utility, and market variable ε).

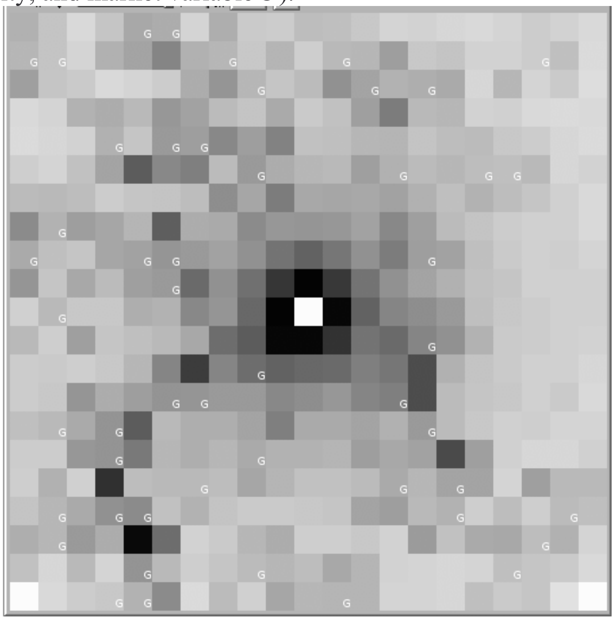


Fig. 5. Rent gradient (with green amenities and ε included in determination of bid and ask)

Prices are no longer strictly decreasing with distance from CBD. We can see some “islands” of higher land prices, clustered around the green areas in the city. Theoretical and empirical literature gives the evidence that people value green amenities and that this value is reflected in the prices for land or for houses. Our results show that

. With the help of the ALMA ABM we can explore how prices evolve over time and how systems moves towards equilibrium.

5 Case Study and Future Work

This study was motivated by the case study of land use in the costal zone in the Netherlands. The coastal zone is a very densely populated area, and traditionally

Dutch government guarantees a certain safety level. The main two policy objectives in the coastal zone area are to provide safety from flooding and efficient spatial allocation. However, since 2002 parts of some coastal towns appeared to be in the area beyond the dikes and are considered legally unprotected [23]. According to the Commission of Poelmann decision (declared in 2005) future developments in these areas are at the risk of individuals. Nevertheless, some studies showed [24] that the perception of the risk of coastal flooding is very low. In this situation individual location choices bounded by low flood risk perception might produce outcomes that contradict policy objectives. Neither land prices nor patterns of land use will reflect the existing risk of flooding if risk perception bias is present at micro level. A better understanding of individual behaviour and aggregated outcomes of this behaviour is crucial to achieve policy objectives. Thus, a proposed ABM with endogenous land price and land pattern formation might serve to simulate emergence of land use patterns in coastal zone area caused by individual location preferences. The utility function of an agent can be easily extended to include the notion of risk of flooding (for example in a form of preferences for safety). Safety can be considered as a public good.

At the current moment we have included another attribute of the spatial good – so-called social amenity (S). By social amenities we mean the availability of public services (such as schools of a good quality, or higher standards of safety from flooding) in the neighborhood. Social amenities are implemented by local governments and are financed by the taxes the local government gathers from its citizens. The ABM land market model with social amenities presented elsewhere [25] is inspired by the logic of the Tiebout model [26]. The Tiebout model assumes that citizens choose where to locate according to their preferences for public services provided by local governments, which compete for citizens by offering different quality of public services. These settings very well represent the situation of our case-study in the Netherlands.

Acknowledgments. Funding from the US National Science Foundation grant 0414060 is gratefully acknowledged. The authors are also grateful to Prof. Robert Axtell, Maciej Latek, brown bag seminar participations, and students from the Spatial Agent-based Modeling of Human-Environment Interactions course from the Center for Social Complexity at George Mason University for discussions and valuable advice.

References

1. Bousquet, F. and C. Le Page, Ecological Modelling, 2004(176): p. 313–332.
2. Grimm, V. and S.F. Railsback, . 2005: Princeton University Press.
3. Judd, K.L. and L. Tesfatsion, 2006: Elsevier B.V.
4. Parker, D.C., T. Berger, and S.M. Manson. . in . 2002. USA: University of California.

5. Arthur, W.B., S.N. Durlauf, and D. Lane, . 1997: Addison-Wesley.
6. Epstein, J.M. and R. Axtell, . 1996, Washington, DC, Brookings and Cambridge, MA: The MIT Press.
7. Arthur, W.B., et al., , in , W.B. Arthur, S.N. Durlauf, and D.A. Lane, Editors. 1997, Westview Press. p. 15-44.
8. Gode, D. and S. Sunder, The Journal of Polictical Economy, 1993. **101**(1): p. 119-137.
9. Lux, T., Journal of Economic Behavior & Organization, 1998. **33**: p. 143-165.
10. LeBaron, B., , in K.L. Judd and L. Tesfatsion, Editors. 2006, Elsevier B.V. p. 1187-1233.
11. Marks, R., , in K.L. Judd and L. Tesfatsion, Editors. 2006, Elsevier B.V. p. 1339-1380.
12. Brown, D.G., et al., Geoforum, 2007.
13. Balmann, A. and K. Happe. . in . 2000. Corvallis, Oregon USA.
14. Berger, T., Agricultural Economics, 2001. **25**(2-3): p. 245-260.
15. Happe, K., K. Kellermann, and A. Balmann, Ecology and Society, 2006. **11**(1): p. 49.
16. Polhill, J.G., D.C. Parker, and N.M. Gotts, , in , C. Hernandez, K. Troitzsch, and B. Edmonds, Editors. Forthcoming.
17. Alonso, W., . 1964, Cambridge, MA: Harvard University Press.
18. Varian, H., 1992, USA: W.W. Norton & Company, Inc.
19. Kirman, A.P., Journal of Economic Perspectives, 1992. **6**(2): p. 117-136.
20. Tesfatsion, L., , in K.L. Judd and L. Tesfatsion, Editors. 2006, Elsevier B.V. p. 831-880.
21. Arthur, B., , in K.L. Judd and L. Tesfatsion, Editors. 2006, Elsevier B.V. p. 1551-1564.
22. Miyake, M., Regional Science and Urban Economics, 2003. **33**: p. 721-743.
23. Ministerie van Verkeer en Waterstaat Rijkswaterstaat, . 2002, The Hague: Direct Dutch Publications BV.

24. Kaiser, G. and D. Witzki,
Geographie der Meere und Küsten. Coastline Reports, 2004(1): p. 101-108.
25. Filatova, T.,
, in
. 2007, George Mason University: Fairfax, VA, USA. p. 1-26.
26. Tiebout, C.M., The Journal of Political Economy, 1956. **64**(5): p. 416-424.

Modelling Social and Economic Influences on the Decision Making of Farmers in the Odra Region

Friedrich Krebs, Michael Elbers & Andreas Ernst

Center for Environmental Systems Research (CESR)
University of Kassel, Kurt-Wolters-Str. 3, D-34109 Kassel, Germany
{krebs, elbers, ernst}@usf.uni-kassel.de
www.cesr.de

Abstract. Farmers in the Odra region of Poland are caught in a social dilemma: While, in principle, the existing land reclamation system (LRS) of ditches and canals can absorb the negative effects of extreme weather conditions, its proper functioning requires collective action as regards maintenance. However, such a collective effort is undermined by the asymmetrical nature of the dependency of farmers resulting from the different locations of their land parcels along the LRS. In this paper, we report on an agent-based approach which models the farmers' decisions, whether to maintain their local section of the LRS or not, in terms of their economic success and the social support they receive from acquaintances. Simulation results indicate that the frequency of LRS strategy changes - an indicator for overall volatility - is reduced with the introduction of social influence and that further social pressure leads to a positive social lock-in that even prevents free-riding behaviour.

Keywords: social networks, collective action, volatility, water use, informal institutions

The Context of the Model: Land Reclamation in the Odra River Region

The CAVES (Complexity, Agents, Volatility, Evidence, and Scale)[1] project aims at describing the emergence, the characteristics and long-term behaviour of social networks of people who use natural resources such as land or water. It includes case studies in Great Britain, Poland, and South Africa to acquire data on real world evidence of social networks.

The Polish case study (with input provided by Wroclaw University[2] and the Wroclaw University of Technology[3]) is concerned with issues of land use in the Odra

---

[1] see http://cfpm.org/caves/. The authors wish to thank the European Commission for funding under the FP 6 NEST programme.

[2] The authors are much indebted to Karolina Królikowska for conducting and evaluating interviews with regional stakeholders and for her invaluable collaboration regarding the compilation of abstracted decision rules.

river region. More specifically, it focuses on those parts of the region that are prone to regular flooding due to a lack of maintenance of an old land reclamation system (LRS). Maintaining or re-establishing this land reclamation system which consists mainly of channels and ditches used for drainage and irrigation purposes requires social mobilisation of the farmers concerned. Thus it is important that the acquaintance and/or friendship links that exist amongst them are utilised appropriately. Moreover, it is suspected that land reclamation possesses the structure of a social dilemma [3], and that it therefore represents a collective action [10], [11]. If so, existing theoretical insights can be used to investigate this issue further.

The overall simulation model that is currently being developed and which is to reproduce to some extent the phenomena found in the Odra region consists of two main components: a biophysical model of the environmental conditions and an agent-based model of the decision making and activities of the actors in the area. The biophysical model developed by the Wroclaw Institute of Technology provides a first basic insight into the costs and benefits of farming and land reclamation under certain climatic and water regime conditions. Complementing this, the agent-based model called SoNARe (Social Networks of Agents' Reclamation of land), that is the focus of this paper, seeks to capture key aspects of the reasoning of the actors involved and their interactions with their biophysical and social environment. SoNARe is based on an explicit representation of social influence on the one hand and the individual agents' perceptions of economic success on the other hand. The agents exert and perceive social influence in social networks while at the same time deriving their perception of economic success from feedback of the coupled biophysical model. These two central dimensions – the social and the economic - drive and determine the agents' decision making as regards farming and LRS maintenance as well as social behaviour.

The following sections provide a description of the SoNARe model including some details of the case study, the abstract structure of the situation found there, the technical realisation of the model, and how it interfaces with the biophysical model. Subsequently, some results of initial simulation runs for three basic scenarios are presented and discussed. Finally, an outlook is given on how the model is to be further enhanced in the future.

The SoNARe Model

The SoNARe agent-based model is part and parcel of the attempt to provide a useful and plausible abstraction of key features of the CAVES Odra case study for the purpose of simulation. First and foremost, these features are:

- *Environmental shocks/extreme weather conditions.* Flooding or drought leads to a substantial loss of crop yield.

---

[3] The biophysical model referred to in this paper – *S*imple *H*ydro-*A*gricultural *M*odel (version 0.2.3) – was implemented by Grzegorz Holdys. The authors greatly appreciate his contribution creating a lean yet sophisticated environmental model.

- *Water stress: Local coping strategies.* The farmers in the Odra region have two means at their disposal for coping with water deficiency and water excess stress, namely maintaining their local section of the land reclamation system (LRS) and/or operating the sluice gates in the ditches. The LRS consists of channels and ditches, which drain the soil directly, or through a drainage pipes system and thus protects a field against flooding. The LRS maintenance process mainly involves the periodic cleaning of channels and ditches, e.g. by removing vegetation and sediments from the channels' and ditches' beds. In addition, an LRS may be equipped with sluice gates to facilitate irrigation in dry years. However, this passive irrigation can only be used for pastures and meadows. As this papers focuses on arable land with a (hypothetical) crop, we only consider LRS maintenance and excess water stress. Therefore, farmers can decide either to
  - participate in the LRS, i.e. maintain the LRS locally on their respective land parcel and thus increase the level of protection against environmental shocks, or
  - neglect the LRS, leading to degradation and subsequently to a decreased level of protection against environmental shocks.

  In principle, if viewed independently from other channel sections, the maintenance of the local section of the land reclamation system (LRS) serves to mitigate or even eliminate the negative effects of extreme weather conditions, especially excess water stress in the case of flooding, whereas neglecting it will only increase these effects.
- *Water stress: Global asymmetrical dependency.* LRS maintenance must, however, be regarded as a collective task that requires social mobilisation of the participants, i.e. the farmers whose land parcels are located along a ditch or a communicating ditch system. This is because the difficulties concerning land and water use in the Odra case study region result mainly from the fact that the conditions encountered on individual land parcels depend highly on the amount of LRS maintenance (and/or the sluice gate operations) performed on other (adjacent) land parcels. In wet periods, for example, LRS neglect leads to a loss of yield on neighbouring land parcels upstream since the runoff of excess water is blocked, whereas LRS maintenance has the opposite, beneficial effect since the runoff is facilitated. The latter effect arises even if the upstream neighbours do not themselves maintain their section of the LRS (free riding). Downstream effects can be observed mainly in dry periods, if the LRS on a particular land parcel is not maintained (or the sluice gate on that parcel is closed). This causes a decrease in the amount of water flowing down to adjacent parcels resulting in a loss of yield on those parcels. For the reasons stated above, sluice gate operation will not be considered in the presented version of the model.

Maintenance of the land reclamation system thus enables to overcome flooding and drought with reduced or even without loss of harvest, but it requires a collective effort. Moreover, the asymmetrical dependency entails a social dilemma structure, in turn providing incentives for such problematic types of behaviour as free riding. It is expected that it is this social dilemma structure that hinders and in some cases prohibits the installation of a functioning LRS.

Against the background of these problem features it is the particular purpose of SoNARe to reflect the reasoning of the actors involved on an adequate level of abstraction and to seamlessly interface with an underlying biophysical model from which agents perceive the state of the environment and which they act upon. In order to do so, the implementation is based on results of interviews with actual farmers and other stakeholders in the Odra region. More importantly, however, a set of abstract decision rules for different types of actors was compiled which forms the basis for the implementation of the agent decision rules in SoNARe.

Technically, the SoNARe model uses production rules implemented in JESS (the Java Expert System Shell; http://herzberg.ca.sandia.gov/jess/), and JESS's reasoning engine to simulate the cognitive control structure and decision making of farmers and other relevant actors, while the actors themselves and their interaction with the social and natural environment is treated in a straightforward way in Repast (Recursive Porous Agent Simulation Toolkit, [8]). This is consistent with a long-standing rule-based representation of problem solving and cognition [1]. It opens the possibility to provide Repast agents with full cognitively plausible capabilities. Since JESS is written entirely in Java and allows for calling Java methods from rules, it integrates well with any Java software. JESS consists of a rule interpreter which can apply both forward and backward chaining, using an improved version of the fast but memory-intensive RETE algorithm to match facts from the fact base to rules in the rule base. Declaring facts and rules is done via a script language with a LISP-like syntax. Actions are buffered and later executed by the Repast part of the model. The details of the model will be described next, followed by some scenario results and their discussion.

Model Setup

Environment (the Biophysical Model)

The biophysical model simulates the effects of different weather conditions, LRS maintenance and LRS neglect as well as sluice gate operations on the water levels and thus on the crop yield of individual land parcels along a channel. It offers a number of parameters to be varied across simulation runs, most importantly for our purposes:

- *Weather sequence.* This is a repeated sequence of normal, wet and/or dry years, which differ in the mean water levels in the channel per month.
- *Land parcels.* The number of land parcels per channel and the type of land use of each land parcel can be set. Currently, the biophysical model allows for two types of land use, namely arable and fish-pond, of which only the former is used in the simulation runs described in the next section. Moreover, the initial condition of the LRS section on each individual parcel can be specified.

The model is run at monthly steps with crops for arable land parcels being planted in month 5 and harvested in month 10. The condition of the local LRS slowly degrades when it is not maintained and (at present) it fully recovers within a month of an agent first maintaining it. By instantiating the biophysical model multiple times it is possible to simulate any number of channels in parallel without any interrelations among them.

Agents

As a first step of agent modelling, generalised farmer types were determined based on elicited knowledge, i.e. the storylines derived from the transcripts of the interviews. These farmer types differ along dimensions like land ownership, activity related to LRS maintenance, social network integration or agricultural knowledge. The current version of the model considers two of those actor types: the farmer, i.e. a general prototypical farmer type, and the water partnership initiator (WPI). Farmer agents are embedded in a dependency network according to the location of their land parcel along their respective channel and in an overall acquaintance network which is randomly superimposed on the dependency networks thus spanning all channels. The acquaintance network not only includes all farmer agents, but also a WPI agent, which is linked to all farmer agents in a star-like manner. Water partnerships are formal institutions with the aim of maintaining the common LRS. Presently, the WPI is not a farmer itself (e.g. the village mayor).

In order to capture the decision dynamics of the farmers a rule-based approach was chosen. Some of the compiled decision rules could be abstracted from the storylines while others were deduced from domain expertise. Examining the decision rules of farmers revealed that the internal dimension of past economic success and the external dimension of perceived social influence appear to play an essential role. Therefore, we decided to explore the effects of contrasting social influences and economic success in an isolated way. Accordingly, in the model version presented in this paper, agents' decisions have been reduced to a binary decision of participating in the water partnership and hence locally maintaining the LRS or not.

In order to explicitly contrast social and economic influences on the decision making of farmers with regard to LRS maintenance two dimensions of agent perception are introduced that drive a farmer agent's decision making: *economic success* and *social support*.

- The perception of economic success is shaped by several factors: First of all, every year a farmer agent appraises its current yield as "good" or "bad" with respect to a fixed yield perception threshold. It then stores either a positive value ("good") or a negative value ("bad") – values are currently symmetrical – in its yield memory which has a fixed capacity, i.e. the agent forgets yields after some time. The extent to which agents attach importance to economic factors can be influenced by increasing or decreasing the so-called *economic sensitivity*, i.e. the weight given to the values for "good" and "bad" yields. The *economic success* is simply the sum of all yield appraisals stored in memory.

- The perception of social support is a function of the agreement/disagreement between farmer and acquaintance concerning LRS maintenance, i.e. whether it receives support or negative pressure, weighted by the individual level of social influence of each acquaintance.

On the basis of these two factors farmer agents decide on their LRS maintenance strategy, changing it if their perceived economic success and their perceived social support fall below a certain level (see pseudo-code below). Thus, in addition to the impact of the actual yields the model also reflects general opinion dynamics amongst farmers [9], [6]. This exertion of social influence is strictly symmetrical in the sense that each farmer agent supports each farmer agent in its acquaintance network that uses the same strategy by the same amount as it imposes pressure on each farmer agent that uses the opposite strategy.

```
Farmer:

if(socialSupport + economicSuccess < 0)
switch_LRS_strategy();

for(allAcquaintedFarmers) {
          if(otherLRSStrategy == myLRSStrategy) then
supportOther(AMOUNT);
          else putPressureOnOther(AMOUNT);
}
```

The opinion dynamics are extended to account for the social influence of WPIs on the pervasiveness of the LRS maintenance strategy and thus on the formation of a working water partnership (WP). A WPI agent exerts its social influence in favour of LRS maintenance once it perceives at least three farmers who have big losses; it does not exert any influence otherwise. The WPI's level of social influence is independent from that of the farmers. The institution of the water partnership is active as long as at least three farmers maintain their respective LRS; it is inactive otherwise. Farmers are automatically members of the WP as long as they maintain their LRS (see below). However, at present, WP membership does not have any effect and there are no costs modelled for LRS maintenance.

```
Farmer:

if(exists(WP) AND (myLRSStrategy == MAINTAIN)) then be-
MemberOfWP();
else doNotBeMemberOfWP();
```

Model Execution Cycle, Sequence of Events

For each year the model executes the following sequence of events: In month five the biophysical model simulates the planting of crops for each land parcel and in month ten it simulates the harvesting of these crops. Finally, at the end of every year the agent-based model is run performing the following sequence of steps: all agents perceive and memorise their individual yields, exert their social influence, perceive the social influence exerted on them, revaluate their economic success and then make their decisions for the next year. It is important to note that agents are synchronised at every step. Above all, this means that they take their decisions simultaneously.

Simulation Results

This section provides an overview of the simulation results reported here that were produced with the described model. We present three scenarios: The first scenario functions as a baseline by showing the behaviour of the biophysical model on its own, i.e. without any intervention on the part of the agents. The second scenario then looks into the effects of having agents decide on LRS maintenance annually and with the perceived economic success as the sole basis for that decision. Finally, in the third scenario both economic success and social support take effect as criteria in the agents' decision making.
All simulations are initialised with rows of ten land parcels that are located along one channel. It is assumed that each land parcel is owned and managed by exactly one farmer and that each farmer owns and manages exactly one land parcel. The model is scaled up to simulate 100 farmers in order to have a more realistic population size for the underlying acquaintance network. It is important to note that we scale up the model by increasing the number of channels instead of increasing the number of land parcels per channel, because the latter is not backed by observation, i.e. it is unrealistic to assume a channel size well in excess of ten parcels. All three scenarios start off with no LRS maintainers. The same weather sequence is used throughout: Two years with normal weather conditions are followed by one year of wet weather. This pattern is then repeated for the whole run.

In all performance diagrams the time axis shows months, since they are the smallest simulation step considered in the model. Data points in the diagrams below are displayed every twelfth month and thus represent one respective year.

Due to the model's level of abstraction, it has to be stated that the simulation results shown here do not claim to be exact predictions or forecasts of future developments. E.g. when results are discussed in terms of years until a certain process has finished, this should be interpreted as being in reference to an abstract time span of "model" years. Nevertheless, scenarios may be compared with respect to differences in temporal dynamics.

Scenario 1 – Baseline Scenario

The baseline scenario simulates one hundred land parcels along ten channels. This scenario excludes both effects of social influence between agents and agents' rating of their economic success. Thus, agents do not change their LRS maintenance strategy (they never maintain). Because of that, the results of all ten channels are equal; so only one channel is depicted here (see figure 2). The results shown here demonstrate the implicit spatial dependencies between the farmers as generated by the coupled biophysical model.

Figure 1 shows the farmers' yield statistics over the simulation period. Because all farmers keep to their passive LRS strategy, a repeated pattern of yield losses every third (i.e. wet) year can be observed. Furthermore, under wet weather conditions the mean deviation of the farmers' yields is much greater than in normal years. As can be seen in figure 2, in cases of flooding and all farmers neglecting their LRS, yields of farmers at the top of the channel are considerably worse than the yields of those further at the bottom. Thus, the biophysical model shows that under flooding conditions farmers located further downstream obtain a certain degree of implicit flood protection if upstream neighbours neglect their LRS and absorb most of the effects of flooding. Moreover, the farmers at the bottom experience only minor differences between normal years and wet years in this scenario.

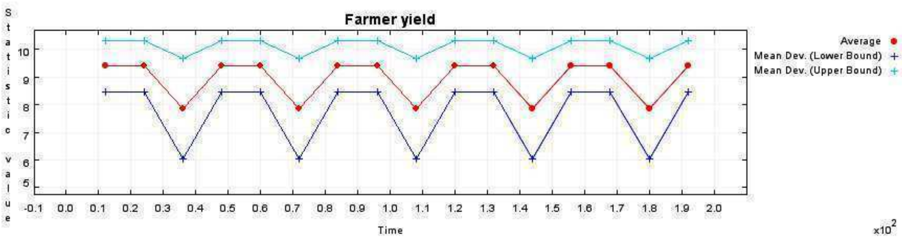


Fig. 1. Average yields of farmers bounded by the mean deviation. The time axis shows months*100 as ticks. Each dot thus marks the average yield for one year, i.e. sixteen years in total are shown.

Scenario 2 – The Influence of Economic Success

The second scenario differs from the baseline scenario only in that the farmer agents now appraise their economic success in the way described in the model setup. Again, social influence is excluded. Although farmers' decision making in this scenario is based solely on their internal and subjective perception of their respective economic farming success, farmers' decisions may well affect other farmers' economic success (due to the hydrological dependencies).

Figure 3 shows the development of LRS strategy adjustments over time, i.e. the proportion of farmers who change their opinion about LRS maintenance in either direction. This volatility indicator rises for about 8 to 9 years of simulation time and then falls back to zero. Figure 4 depicts the corresponding convergence of the number of LRS maintainers to a stable state of 80% after 20 years (240 on a months scale). Figure 5 shows the resulting increase in the average of farmers' crop yields.

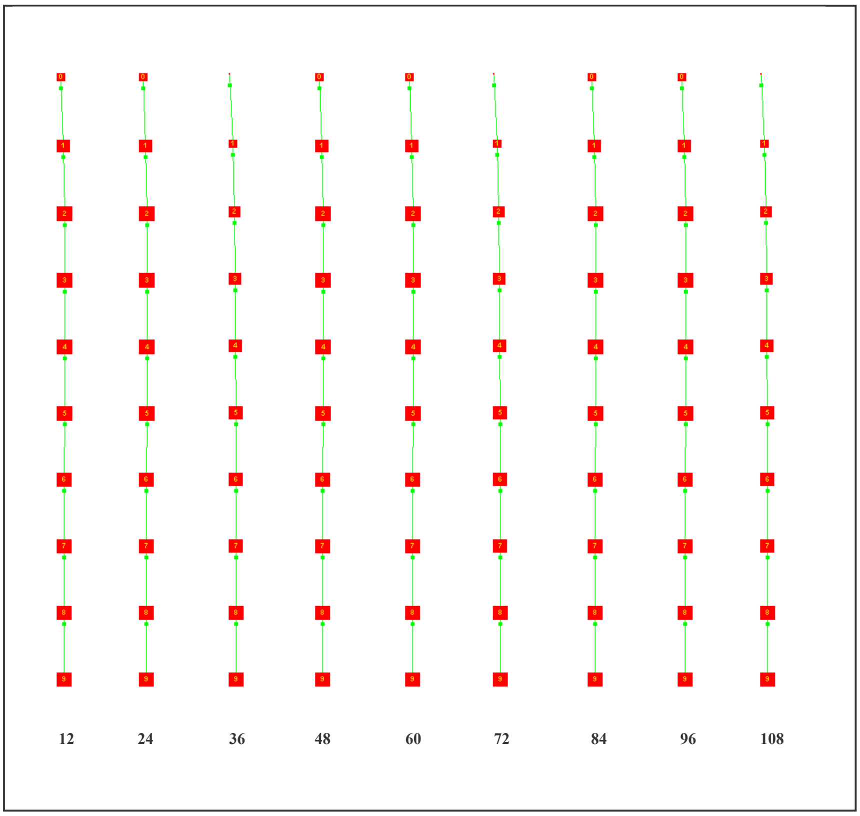


Fig. 2. Yields and LRS strategies on individual land parcels along a single channel over time (years 1 to 9, i.e. months 12 to 108). Land parcels are represented as squares connected by green lines indicating the dependency relation (flow direction from top to bottom). The size of each square is proportional to the yield for the respective land parcel. The red colour indicates LRS neglect.[4]

[4] The numbers inside the squares denote agent IDs (0-9 in scenario 1 and 2, 0-99 in scenario 3). They are, however, irrelevant for all intents and purposes.

Fig. 3. Strategy adjustments over time as an indicator for volatility. The time axis shows months*100 as ticks. Each dot marks the proportion of farmers who have switched their LRS strategy in one year (30 years in total).

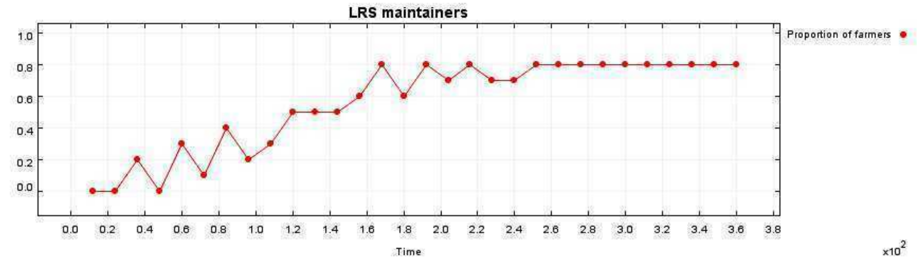


Fig. 4. Proportion of LRS maintaining farmers over time. The time axis shows months*100 as ticks.

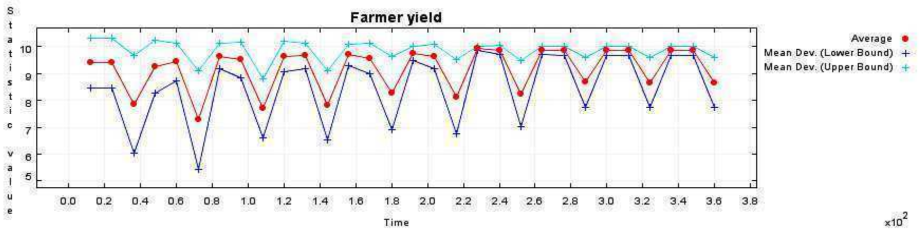


Fig. 5. Average yields of farmers bounded by the mean deviation. The time axis shows months*100 as ticks. Each dot marks the average yield for one year (30 years in total).

Figure 7 illustrates the spatial distribution of opinion shifts over time. In the course of the simulation, the topmost farmer first starts maintaining LRS after a wet year. This is probably due to the combination of the positional disadvantage even in normal years (see scenario 1) and the bad yield in a wet year. As may be seen in figure 6, the perceived economic success of LRS maintainers increases substantially and then settles on a slightly higher level than the corresponding values for non-maintainers.

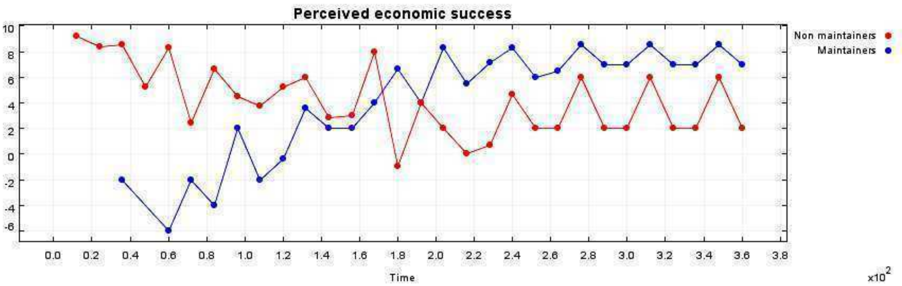


Fig. 6. Mean perceived economic success of maintainers and non-maintainers over time with the yield threshold set to 9.0, an economic sensitivity of 2.0 and a yield memory capacity of 5 years, i.e. the perceived economic success has a lower bound of -10 and an upper bound of 10. The time axis shows months*100 as ticks.

Scenario 3 – The Combined Influence of Economic Success and Social Support

In the third scenario agents use both their past economic success and the social influence of other agents as a basis for their decision making. As in the scenarios above, ten independent channels with ten land parcels each are considered which amounts to 100 farmers altogether. Since we now include effects of social influence in this scenario, farmers are now embedded in an acquaintance network that spans all ten channels. In addition, the WPI agent is linked to all farmers as described in the model setup.

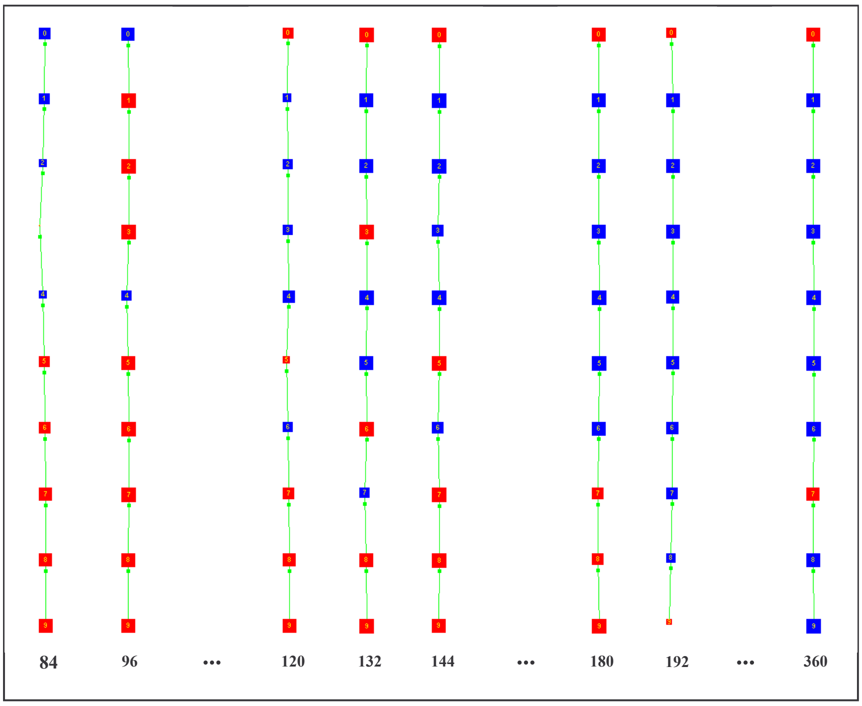


Fig. 7. Yields and LRS strategies on individual land parcels along a single channel over time (years 7-30; numbers refer to months). Land parcels are represented as squares connected by green lines indicating the dependency relation (flow direction from top to bottom). The size of each square is proportional to the yield for the respective land parcel. Red squares indicate LRS neglect, blue squares LRS maintenance.

While in earlier simulations (see [5], in press) acquaintances networks were generated as Small-World networks with the algorithm of [12] in the simulations shown here a scale-free network topology with an average node degree of 10 is used. The assumption of a scale-free topology is supported by Odra case study narrative storylines and by many other studies on social networks [2], [7]. The scale-free network used in the simulations is generated by an algorithm described by [4]. This algorithm allows generating a sufficient proxy of a scale-free network for 100 nodes and an average node degree of 10.

As before, Figure 8 shows the proportion of farmers that change their opinion about LRS maintenance as an indicator of the system's volatility. Since the WPI agent exerts its social influence this activity slowly pushes opinions towards LRS maintenance. Figure 10 shows the corresponding convergence of the number of LRS maintainers to a stable state of 100% after 34 years. Around year 20 (month 240 in the diagram) the proportion of LRS maintainers exceeds 50% which triggers an avalanche pro LRS which is reflected in figure 9 as a much higher opinion shift towards maintenance. Figure 11 again shows the resulting increase in the average of farmers' crop yields.

Figure 12 contrasts the development of economic success and social support over time (average values over the 100 agents are shown). The perceived economic success starts off with unrealistically high values because agents' yield memories are initialised with 5 "good" years. This value decreases as soon as agents have experienced the first years of the simulated weather sequence. When the shift in LRS strategies starts (see figure 12, years 3 and 4) the average social support indicator falls steeply from 4 to below 2. These low values of social support persist throughout the phase of high volatility. As more and more agents switch to LRS maintenance social support rises again until month 360 when the WPI becomes passive (see figure 12). Social support parallels the perceived economic success in that it, too, rises continuously.

Fig. 8. Strategy adjustments over time as an indicator for volatility. Each dot marks the proportion of farmers who have switched their LRS strategy in one year (40 years in total). The time axis shows months*100 as ticks.

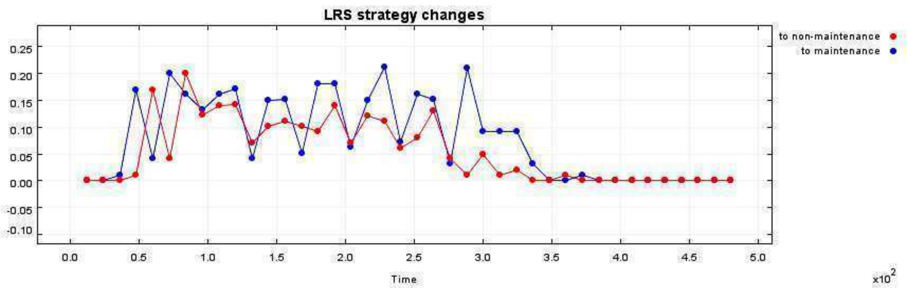


Fig. 9. Proportion of strategy changes to maintenance and to non-maintenance strategy over time. The time axis shows months*100 as ticks.

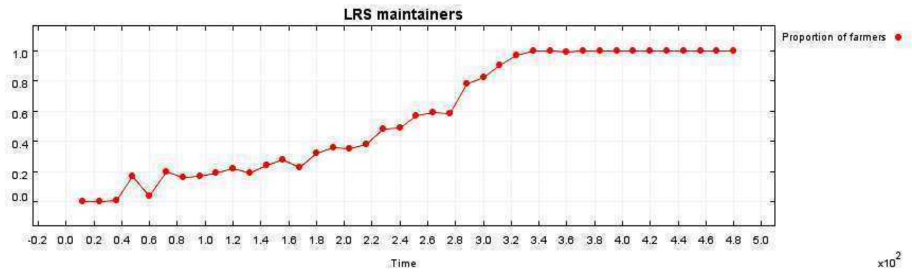


Fig. 10. Proportion of LRS maintaining farmers over time. The time axis shows months*100 as ticks.

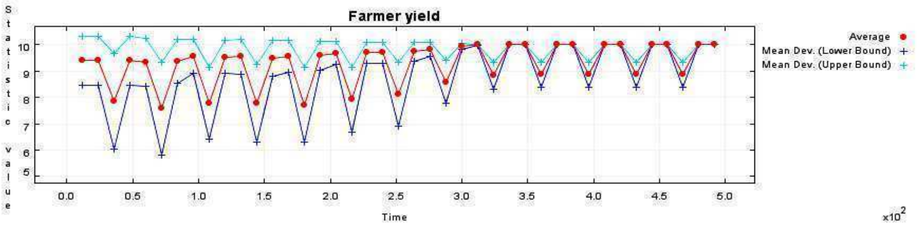


Fig. 11. Average yields of farmers bounded by the mean deviation. Each dot marks the average yield for one year (40 years in total). The time axis shows months*100 as ticks.

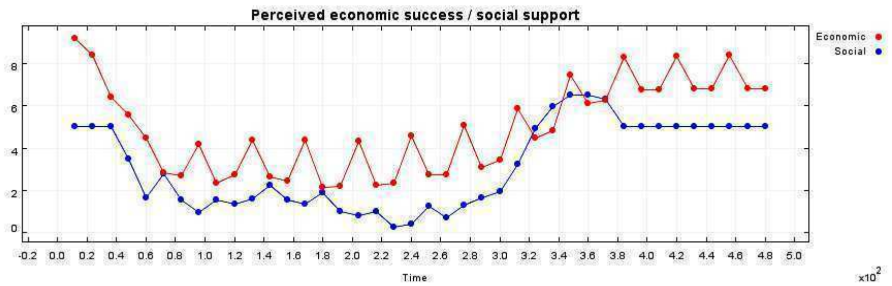


Fig. 12. Mean perceived economic success and mean perceived social support over time. The time axis shows months*100 as ticks.

Figure 13 shows the development of the perceived economic success of maintainers and non maintainers. LRS maintainers start with very little economic success: this is what caused them to shift to maintenance. After that and similar to figure 12 maintainers become more and more successful. Figure 14 contrasts the perceived social support of LRS maintainers and neglecters. As expected, during the phase of volatility both economic success and social support are on very similar levels for both groups of agents.

Finally, figure 15 displays the spatial distribution of opinion shifts along the ten channels over time. We show 4 snapshots at months 36, 48, 168, and 300. The general dynamics is that agents begin maintaining LRS from top to bottom. This reflects the perception that topmost agents are in general more severely affected by flooding.

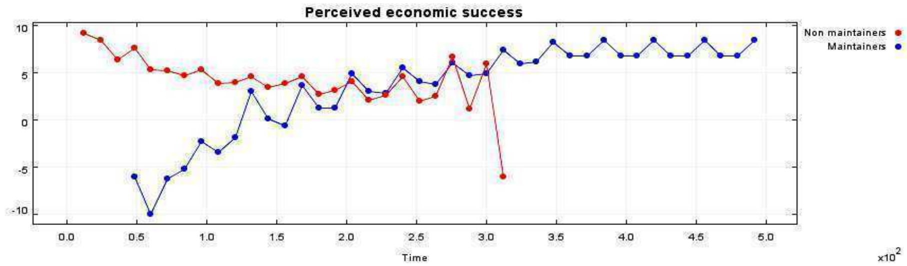


Fig. 13. Mean perceived economic success of maintainers and non-maintainers over time with an economic sensitivity of 2.0 and a yield memory capacity of 5 years, i.e. the perceived economic success has a lower bound of -10 and an upper bound of 10. The time axis shows months*100 as ticks.

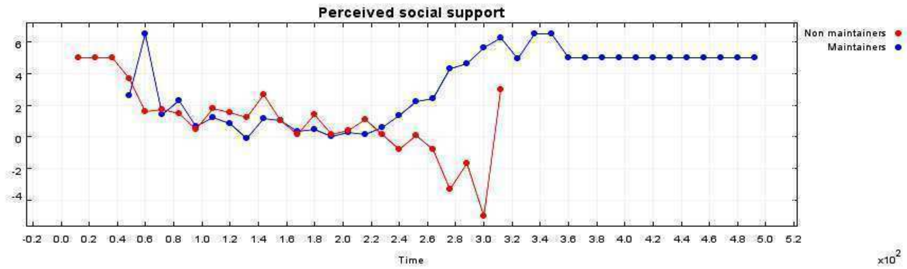


Fig. 14. Mean perceived social support of maintainers and non-maintainers over time with a social influence level of 0.5 for farmers and 1.5 for the WPI and an average acquaintance degree of 10. The time axis shows months*100 as ticks.

Discussion

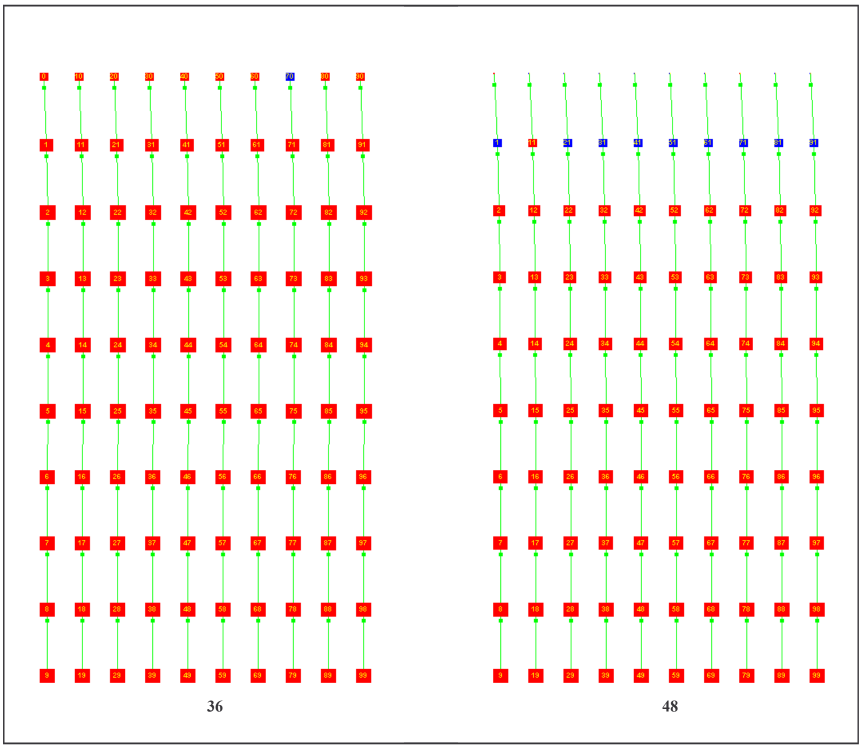
In the work reported here, it has been a guiding assumption that it is fruitful to build models on a medium level of abstraction, i.e. to keep just between too much detail and too much abstraction. On one hand, there has to be a certain amount of recognisable empirical characteristics of the domain being modelled. On the other hand one has to make sure the model is interpretable in terms of complexity indicators, network characteristics, and basic theoretic or phenomenological structures, like social dilemmas. The SoNARe model has made a step in this direction. It comprises a social dilemma structure derived from a geographical structure, a lean yet sufficiently powerful geo-physical model, stylised behavioural rules represented in actor types and well-founded psychological assumptions about social influence, memory span and social networks. While the model is being tested with only a smaller number of actors, it is easily scalable to several hundreds of actors without losing the basic environmental or social structure.

The SoNARe model produces, besides other behavioural indices, a measure of volatility (the amount of strategy changes by the actors). When comparing the development of the volatility indicators of scenarios 2 and 3 (see figures 3 and 8) it can be seen that for scenario 2 the indicator has peaks to above 0.5, whereas in scenario 3 the indicator is always well below 0.4. As a conclusion, this might indicate that under the given circumstances the presence of an active social network and of mechanisms of social influence dampens phases of high volatility in opinion dynamics and instead lead to a coherence effect. This effect will be investigated further.

From the point of view of the underlying social dilemma structure, it is most interesting to note the emergence of a positive social lock-in. Due to the WPI's activity and the subsequent economic success of LRS maintainers, a self-sustained WP can be installed in scenario 3. In spite of the situation's underlying dilemma structure prone to free riding, a social "activity seed" together with some social and economic pressure exerted on the participants is sufficient to trigger this process and to keep it alive. The intertwining of social and economic processes and their long-term effects will have however to be investigated further. Still, it is safe to attribute some effectiveness to the modelled institution of WP and the leading WPI. This result will be discussed with the local stakeholders of the case study.

Outlook

It is intended to further approximate the situation in the case study region and press ahead with the investigation of the covariance of network properties and collective behaviour as well as possible phases of volatility.


36
48

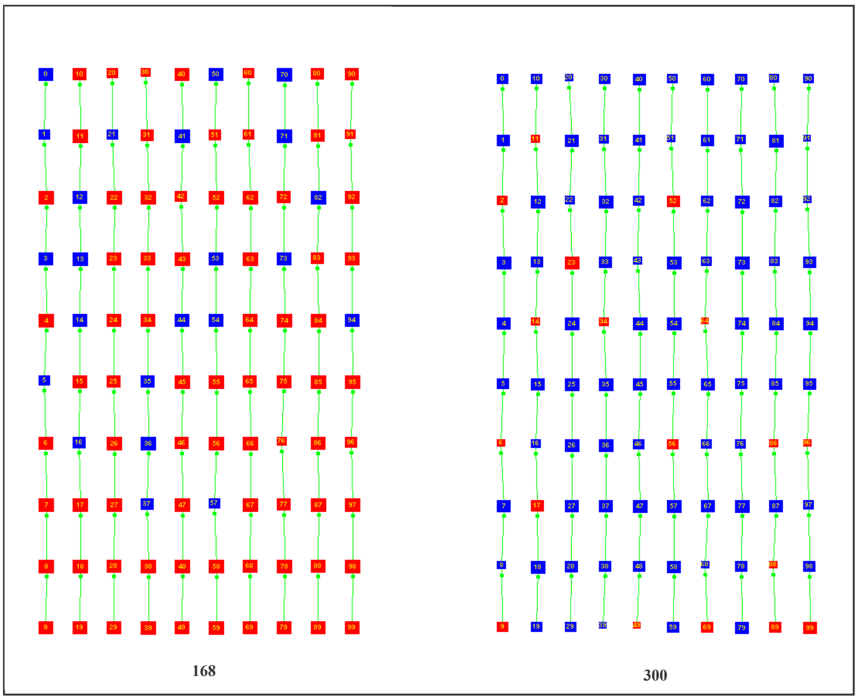

168
300

Fig. 15. Yields and LRS strategies on individual land parcels along ten channels (left to right) over time (months 36, 48, 168 and 300). Land parcels are represented as squares connected by green lines indicating the dependency relation (flow direction from top to bottom). The size of each square is proportional to the yield for the respective land parcel. Red squares indicate LRS neglect, blue squares LRS maintenance.

To these ends, we plan to enhance and study the model in a number of respects. LRS maintenance costs, allowances for LRS maintenance and compensation payments in case of yield loss will be introduced as economic factors. Moreover, additional land use types (fish ponds, in particular) are to be included and farmer types are to be modelled that differ in terms of their economic and social perception and the sets of decision rules they employ. In this context the effects of varying heterogeneous distributions of such farmer types will also be studied. Finally, it is planned to analyse possible sluice gate operation strategies of farmers on individual land parcels and their relation to LRS strategies and the general dynamics of the model. Considering the effects both of flooding and drought leads to a double social dilemma with bidirectional (up- and downstream) dependencies. Its social and economic effects will have to be investigated further.

References

1. Anderson, J. R. (1983). *The Architecture of Cognition*. Cambridge: Harvard University Press.
2. Barabási, A.-L. (2002). *Linked: The new science of networks*. Perseus Publishing, Cambridge.
3. Dawes, R.M. (1980). Social dilemmas. *Annual Review of Psychology, 31*, 169-193.
4. Ebel, H., Davidsen, J., & S. Bornholdt (2002). Dynamics of social networks. *Complexity 8*(2):24–27.
5. Ernst, A., Krebs, F. & Zehnpfund, C. (in press). Dynamics of task oriented agent behaviour in multiple layer social networks. In T. Terano, S. Takahashi, D. Sallach & J. Rouchier (eds.), *Advancing Social Simulation: The First World Congress*. Tokyo: Springer.
6. Friedkin, N. (1998). *A Structural Theory of Social Influence*. Cambridge University Press, Cambridge.
7. Newman, M.E.J. (2003). The structure and function of complex networks. *SIAM Review, 45*(2), 167–256.
8. North, M.J., N.T. Collier, and J.R. Vos (2006). *Experiences Creating Three Implementations of the Repast Agent Modeling Toolkit*, ACM Transactions on Modeling and Computer Simulation, Vol. 16, Issue 1, pp. 1-25, ACM, New York, New York, USA.
9. Latané, B. (1981). The Psychology of Social Impact. *American Psychologist 36*: 343–56.
10. Olson, M. (1965). *The logic of collective action*. Cambridge, MA: Harvard University Press.
11. Ostrom, E. 1990. *Governing the commons: The evolution of institutions for collective action*. New York, NY: Cambridge University Press.
12. Watts, D. & Strogatz, S. (1998). Collective dynamics of small-world networks, *Nature, 393*, 440 (1998).

Finance Session

Agent-Based Computational Finance: A Practical Application

Arvid O. I. Hoffmann * and Wander Jager #

University of Groningen
Faculty of Economics – Department of Marketing
P.O. Box 800
9700 AV Groningen
The Netherlands

* Contact Author
Email: a.o.i.hoffmann@rug.nl
Telephone: +31 (0) 50 363 28 02

Presenting Author
Email: w.jager@rug.nl
Telephone: +31 (0) 50 363 4007

Abstract.

This paper uses agent-based computational finance as a method to link micro level investor behaviour and macro level stock market dynamics. Empirical data from an online investment survey on individual investors' decision-making and social interaction is used to formalize the trading and interaction rules of the agents of the artificial stock market SimStockExchange. Multiple simulation runs are performed with this artificial stock market, which generated macro level results, like stock market prices and returns over time. These outcomes are subsequently compared to empirical macro level data from real stock markets, and qualitative as well as quantitative agreement between the simulated asset returns distributions and the real stock markets' asset returns distributions are found.

Introduction and Background

In recent years, agent-based computational finance has developed into a growing field in which researchers rely on computational methods to overcome the inherent limitations of analytic methods (LeBaron, 2000). Amongst the advantages that agent-based computational models offer are (1) the ease with which it is possible to limit agent rationality, (2) the facilitation of heterogeneity in the agent population, (3) the possibility of generating an entire dynamical history of the processes under study, and

(4) the ease with which it is possible to have agents interact in social networks (Axtell, 2000). Furthermore, as a specific instance of the broader field of social simulation, agent-based computational models are well adapted to developing and exploring theories concerned with social processes and easily represent dynamic aspects of change. Using these kind of models can help to increase our understanding of the relationship between micro level attributes and individuals' behaviour and macro level aggregate effects like stock market dynamics (Gilbert & Troitzsch, 1999). More specifically, the application of multi agent models for financial markets research has been promoted by a number of empirical puzzles or stylized facts (e.g., time series predictability, volatility persistence/clustering, and fat tails in the asset returns distribution) that are difficult to explain using traditional representative agent structures (LeBaron et al., 1999).

In this presentation, we will elaborate on the artificial stock market SimStockExchange™ [1]. This model, which is more extensively described in Hoffmann, Jager and Von Eije (2007), establishes a more direct link between micro level investor behaviour and macro level stock market dynamics than existing models and offers a number of contributions to the field.

The first contribution is that our model - like the model of Takahashi and Terano (2003) - is theory-driven. Yet, in contrast to these authors, we have adopted a *multi-disciplinary* approach, in which (behavioural) finance, social-psychological and consumer behaviour theories are combined. Amongst the theoretical concepts that are used in developing this model are the general notion of boundedly rational investors as well as more specific concepts like the Prospect Theory of Kahneman and Tversky (1974). Other theories that we use consider the different personal needs of people, investors' conformity behaviour, and the way decision-makers in general deal with uncertainty and risk. Moreover, social network theory is used in developing the model.

The second contribution resides in the fact that the model is not merely based on the before-mentioned theories, but we have used these theories to develop specific hypotheses with respect to individual investors' trading and interaction behaviour which were subsequently tested using empirical studies. In Hoffmann, Von Eije, and Jager (2006), we report on part of the results of these studies in which we investigated e.g., to what extent individual investors have needs that deviate from a risk/returns perspective. Moreover, we measured the differences in the amount of investment-related knowledge and experience of these investors and the effect of these differences - which result in different levels of confidence - on investors' conformity behaviour. It was found that individual investors do have other, more social needs apart from their financially oriented needs. In fact, investors that gave a higher importance to socially oriented needs and/or who had lower levels of investment-related knowledge and experience, displayed more informational and normative conformity behaviour. Subsequently, we used these results to develop empirically plausible agent trading and interactions rules and to parameterize the model in order

[1] Visit www.simstockexchange.com to download a demonstration version of the model.

to achieve an artificial stock market that is a closer match with reality than many existing models on investor behaviour.

The third and last contribution is that we not only investigate how the aggregation of micro level investor behaviour results in macro level stock market results, but we also estimate the *empirical plausibility* of the macro level stock market price and returns data that are generated by the model. Amongst other things, we determine the *occurrence* of stylized financial market facts (Cont, 2001) in the model's returns time series, and also the extent to which these stylized financial market facts agree with those of the overall Dutch stock market.

In this presentation, we will not only present the model formalization, but also a number of key results that we have achieved with SimStockExchange. Most importantly, we will compare the occurrence of several stylized financial market facts as like volatility clustering, fat tails and autocorrelation in the returns time series as well as the overall stock market volatility for markets in which agents are positioned in a regular torus network and a Barabasi and Albert (1999) scale free network, respectively.

Key References

Axtell, R. L. (2000). *Why Agents? On The Varied Motivations For Agent Computing In The Social Sciences.* (Rep. No. 17).

Barabasi, A.-L. & Albert, R. (1999). Emergence of scaling in random networks. *Science, 286,* 509-512.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance, 1,* 223-236.

Gilbert, N. & Troitzsch, K. G. (1999). *Simulation for the Social Scientist*. Buckingham: Open University Press.

Hoffmann, A. O. I., Jager, W., & Von Eije, J. H. (2007). Social Simulation of Stock Markets: Taking it to the Next Level. *Journal of artificial societies and social simulation, 10*.

Hoffmann, A. O. I., Von Eije, J. H., & Jager, W. (2006). *Individual Investors' Needs and Conformity Behavior: An Empirical Investigation.* (Rep. No. http://ssrn.com/abstract=835426). SSRN Working Paper Series.

LeBaron, B. (2000). Agent-based computational finance: suggested readings and early research. *Journal of Economic Dynamics and Control, 24,* 679-702.

LeBaron, B., Arthur, W. B., & Palmer, R. (1999). Time series properties of an artificial stock market. *Journal of Economic Dynamics and Control, 23,* 1487-1516.

Takahashi, H. & Terano, T. (2003). Agent-Based Approach to Investors' Behavior and Asset Price Fluctuation in Financial Markets. *Journal of artificial societies and social simulation, 6*.

Tversky, A. & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and Biases. *Science, 185*.

Market Selection of Competent Venture Capitalists

David Mas*

ERMES-CNRS, 12, place du Pantheon, F-75230 Paris Cedex 05
David.Mas@u-paris2.fr

Abstract. This paper presents an original model of venture capital as a market of heterogeneous interacting agents. Simulating this model, I investigate the market selection of venture capitalists and show how an efficient venture capital industry can emerge from an initial random population of heterogeneous venture capitalists. I also identify and study the role of a particular institution of venture capital, the limited partnership, in venture capitalists selection. I show that the size of the limited partnership has almost no influence on the final distribution of competence. And I study the optimal size of the limited partnership for an institutional investor.

1 Introduction

Venture capital is actually considered a very efficient mean for financing innovation. Kortum & Lerner (2000) estimate that a dollar invested in venture capital is three times more effective in stimulating patenting than a dollar invested in traditional R&D. Understanding venture capital is therefore a central matter for designing innovation policies. Venture capital has a very peculiar functioning. Venture capitalists finance young firms whose only activity is to develop radical innovations (start-ups). These new firms have no access to the banking system because they are too risky and have no collateral. But their future is also too uncertain to allow them to enter the financial market. Venture capitalists are assuming this uncertainty because they are looking for high-risk/high-reward investments. Venture capitalist are not wealthy individuals risking their own money (business angels), but fund managers. This means they provide institutional investors (pension funds, insurance companies, investment bank) with the possibility to invest in an asset, the venture capital fund, whose risk is manageable with traditional financial methods, like portfolio diversification. Venture capital is in fact turning the uncertainty of investing in radical innovations into a simple, though high, risk.

I propose a model that explicitly describe this essential feature of the venture capital market. This model is intended to understand what are the conditions for

* The author thanks CO3 (Common Complex Collective Phenomena in Statistical Mechanics, Society, Economics and Biology), European targeted project, for financial support.

the existence and for an efficient functioning of a venture capital market. Following Carlsson & Eliasson (2003), I consider that venture capitalists are competent actors whose function is to select and finance the most promising start-ups. This means that venture capitalists reduce uncertainty in choosing the right start-ups to invest in. To formalize this idea, I propose a model of venture capital as a market of heterogeneous interacting agents. On the one hand there are start-up projects of different qualities that need financing, and on the other hand there are venture capitalists of different levels of competence trying to detect and finance the best start-up. The evaluation of start-up quality by venture capitalist is imperfect, and its accuracy depends on venture capitalist's competence. Despite the simplicity of the model, the interaction of heterogeneous agents in a stochastic environment makes it tractable only with a simple distribution of agents' qualities and some restrictive hypotheses. The use of simulation allows to overcome this limitation.

In a previous paper Mas & Vignes (2006) show under which conditions the competence of venture capitalists allow for an efficient screening of start-ups. In this paper I investigate the market selection of venture capitalists and show how an efficient venture capital industry can emerge from an initial random population of heterogeneous venture capitalists. I also identify and study the role of a particular institution of venture capital, the limited partnership, in venture capitalists selection. I show i) that the accuracy of the selection as well a the risk taken by institutional investors increases with the size of the limited partnership, i.e. with the number of investments a venture capitalist can make before having to raise an other fund, ii) that the size of the limited partnership has almost no influence on the final distribution of competence, iii) that the optimal choice for the size of the limited partnership can be determined by the computation of selection costs taking type I and type II errors into account.

The remainder of this paper is structured as follows. Section 2 reports the related literature, Section 3 presents the model, Section 4 presents a short formal analysis of the selection process, Section 5 gives the results of the simulations and Section 6 concludes.

2 Related literature

Kaplan & Stromberg (2001) distinguish three main roles for venture capitalists, which are screening, contracting and monitoring. The contracting role has been extensively studied both from a theoretical point of view, based on the work of Aghion & Bolton (1992), Dewatripont & Tirole (1994) and Admati & Pfleiderer (1994), and from an empirical point of view. Kaplan & Stromberg (2000) for instance propose an enlightening comparison between the efficient contracts predicted by financial theory and the real contracts signed between venture capitalist and entrepreneurs. The optimal contract approach has successfully proposed rationales for some stylized facts of venture capital, like the control right allocation in venture capital contracts (Hellmann 1998) and the staging of venture capital investment (Gompers 1995). The monitoring role of venture capitalists is often considered as the specific added value provided by venture capitalists. The early empirical studies (Gorman & Sahlman 1989, Sapienza 1992) emphasize the time spent by venture capitalists interacting with the firms of their portfolio. The later one (Lerner 1995, Hellmann & Puri 2000, Hellmann & Puri 2002) show evidences of the active involvement of venture capitalists in the management of the start-ups and the positive impact of this involvement on the start-up success. But the screening role of venture capitalists has received much less attention. I consider though, like Carlsson & Eliasson (2003), that the screening of start-ups is the main role of venture capitalists. The model presented in this paper places this role at the core of the venture capital market.

My work can be related to two others model. Chan (1983) proposes a model of the venture capital market in which information about the quality of investment is costly. After showing that investors with null search costs are needed in order to allow for the market to exist, he claims that intermediaries like venture capitalists may be these null cost investors, since they make their client pay for their search. Here the cost of information is the only rational for the existence of venture capitalists. In this paper I propose an alternative hypothesis for the role of venture capitalist : instead of costly information, I consider that information about start-up quality is only (imperfectly) available to competent venture capitalists. Amit, Brander & Zott (1999) propose a model of venture capital that takes all agency risk (moral hazard and asymmetry of information) into account. They also consider the monitoring and screening role of venture capitalists. For the later they propose a screening competence, called 'due diligence', that would allow to predict the quality of a given start-up project (and thus to minimize the asymmetry of information) : but they have not developed the model based on this hypothesis. In this paper I adopt a very similar hypothesis for the screening competence of venture capitalists, I build a model based on that idea and I simulate its functioning.

3 The Model

The model contains two types of heterogeneous agents, the start-ups and the venture capitalists. The aim is to keep the model as simple as possible in order to track more easily the effect of each parameter. The model only take two major characteristics of the venture-capital into account : the high-risk/high return type of investment, and the uncertainty about the quality of start-up projects. I first present each agent then the dynamics of the model.

3.1 The Start-ups

Hypothesis 1. *The quality of start-up projects*

Start-up projects are heterogeneous in quality. The quality of each start-up project determines the probability of its success if undertaken. For each start-up project i, the probability of success p_i is given by:

$$p_i = q_i \, \bar{p} \tag{1}$$

With the mean quality of start-up projects $\bar{q} = 1$.

Among the factors explaining the success of a start-up, the main ones are the technological feasibility of the new product, the as-yet-unknown market demand for it, and the capability of the management team. For simplicity, we condense all these factors into one unique quality level. And we choose a very simple linear relationship between quality and the probability of success.

With such a linear relationship, if we assume that the mean start-up project quality is equal to 1, then $\bar{p}$ is exactly the mean probability of success for the population of start-up projects. Thus, the choice of the distribution for q parametrizes the dispersion of the quality of the start-up projects and $\bar{p}$ the mean probability of success. The simplest distribution for q would be a uniform distribution in $[0; 2]$, but we can choose any other distribution of mean 1.

Hypothesis 2. *The return on investment*

For an initial investment I *in a start-up* i *the expected profit is:*

$$E(\pi_i) = (p_i \, g - 1) \, I \tag{2}$$

Hypothesis 2 models the "hit or miss" characteristic of venture capital investment. A start-up is reduced to an investment that may succeed with probability p_i or fail. Financing a start-up means investing an amount I with a risky outcome. In the event of success, the start-up generates a gain of g times the investment, otherwise the investment is a pure loss.

3.2 Venture capitalists

Hypothesis 3. *Competence*

Competence is defined by the ability of a venture capitalist to evaluate the level of quality of a start-up project. For each venture capitalist of competence c_j*, the evaluation of the start-up project quality* q_i *is :*

$$\tilde{q}_{ij} = c_j \, q_i + (1 - c_j) \, u_{ij} \tag{3}$$

The mean-preserving mix we choose here is similar to the specification used in Voorneveld & Weibull (2004), except that we use a distribution for q instead of two discrete levels. Here, u is a noise with a distribution identical to q, and $(1-c_j)$ is the level of noise. Hence c_j is the proportion of actual information in the evaluation of quality. With c_j equal to 1, the venture capitalist is very competent and has an exact evaluation of the start-up quality. With c_j equal to 0, he has no specific competence and the quality he observes is totally uncorrelated with the true quality.

Hypothesis 4. *Screening*

Each venture capitalist invests in a unique start-up. Each start-up is financed by only one venture capitalist.

The screening is sequential. Let S be the set of start-up projects, and S_{j-1} the start-ups selected by the previous $j-1$ venture capitalists. Each venture capitalist finances the best start-up s_j according to its evaluation in the set $S \backslash S_{j-1}$:

$$s_j = argmax_{i \in S \backslash S_{j-1}} \left(\tilde{q}_{ij} = c_j \, q_i + (1 - c_j) \, u_{ij} \right) \tag{4}$$

3.3 The Market

From the four preceding equations, it follows that the market is defined by the following six parameters: the mean probability of success of the start-up projects $\bar{p}$, the gross return in case of success g, the number of start-up projects S, the number of venture capitalists V, the distribution of the quality of start-up projects q and the distribution of competence of the venture capitalists c.

Let us now consider $\bar{q}_s$, the average level of quality of the start-ups. From hypotheses 2 and 1, we obtain $\bar{E}(\pi) = (\bar{q}_s \bar{p} g - 1) \, I$. Thus the product $\bar{p}g$ allows us to determine the minimum level of average quality required to achieve positive profit:

$$\bar{E}(\pi) > 0 \Leftrightarrow \bar{q}_s > \frac{1}{\bar{p} \, g} \tag{5}$$

We now consider the ratio of the number of start-up projects to the number of venture capitalists $l = \frac{S}{V}$. Since each venture capitalist finances only one start-up, only one out of l projects is financed. Thus, l allow us to determine the intensity of screening, defined as the probability for a start-up project of being financed. The higher l, the lower this probability.

Hence, the market is only defined by $\bar{p}g$, l the screening intensity, q the distribution of the quality of start-up projects, and c the distribution of the competence of venture capitalists.

3.4 Dynamics

Hypothesis 5. *Demography*

Start-ups only live for one period. They are created as projects and screened by venture capitalists. If they are selected they may generate may generate profits in case of success.

Venture capitalists stay on the market as long as they have enough capital to invest. The failed venture capitalists are replaced by new ones.

The number of start-up projects S and of venture capitalists V is constant in time. This is warranted for venture capitalist by the replacement of exit. The number of start-ups projects created at each period is a function of the number of venture capitalists : $S = l * V$.

Hypothesis 5 illustrate the fundamental difference between start-ups and venture capitalists : the former are unique investment opportunities while the later are permanent actors of the venture capitals market. Thus start-ups have to be evaluated ex-ante when venture capitalists may be evaluated on past performance. Competence is required to evaluate start-ups but market selection is enough for venture capitalists.

The size of the venture capitalist population and the number of start-up projects are kept constant in order to avoid increasing the complexity of the model with more specific entry mechanisms. A constant population size is a strong assumption but also a good reference case. The new venture capitalists are created with the same rule as the initial population : their competence is randomly chosen with the same distribution and they start with a new fund to invest.

Hypothesis 6. *Venture capitalists fund (Limited Partnership)*

Each venture capitalist raises a fund LP (Limited Partnership) that he invests totally, always financing one start-up each period, and always investing the same amount I in each start-up. The eventual profits are accumulated by the limited partners. Once the fund is totally invested the limited partners may decide to give the venture capitalist a new fund to manage according to its performance. Let $K_{j,t}$ be the capital of the venture capitalist j a period t and $\Pi_{j,t}$ the accumulated profits :

$$K_{j,t} = \begin{cases} LP & \textit{if } (t \bmod LP) = t_0 \\ K_{j,t-1} - I & \textit{else} \end{cases} \tag{6}$$

$$\Pi_{j,t} = \begin{cases} 0 & \textit{if } t = t_0 \\ \Pi_{j,t-1} + \pi_{j,t} & \textit{else} \end{cases} \tag{7}$$

Limited partnerships are funds raised for a determined time. The limited partners provide the fund but abandon the control over it for the given time period. The generals partners manage the fund but has to liquidate all assets before the end of the limited partnership. Here the finite time horizon is modeled by the finite number of investments. The limited partnership provide incentive to the general partners both with a substantial percentage on the realized profit and with the open possibility of not giving them a new fund to manage. In this model the limited partnership is essentially considered as a tool for venture capitalists selection. The replacement procedure of the venture capitalists may be interpreted as a constant aggregate level of fund available for venture capital investments.

Hypothesis 7. *Time*

At each time period t, we repeat the following steps:

1. *start-up projects are created.*
2. *every venture capitalist invests the same amount I in one and only one new start-up. A start-up may only be financed by one venture capitalist.*

3. *the financed start-ups generate profits or losses. The other start-ups are discarded.*
4. *venture capitalists accumulate the profits or losses realized by their start-up.*
5. *venture capitalists whose capital falls to zero are replaced by new ones*

The time structure incorporates two selection processes. The screening of new start-ups occurs at each time period. The market selection of venture capitalists occurs at each end of the limited partnerships. I assume that the size of the limited partnership is always a multiple of i, the amount invested in each start-up. Then the size of the fund determines a number of start-ups to finance, and therefore the number of periods before the venture capitalist is evaluated. Venture capitalist whose performance is sufficient are given new funds to manage, the other are replaced by new ones.

4 Formal analysis

The aim of this section is to find the best estimator of the performance of venture capitalists and to study its influence on the selection process.

4.1 Best selection criterion

The competence of a venture capitalist determines in a non linear way the expected quality of the start-ups he finances, and hence their expected probability of success. This probability is also affected by the market conditions and structure (Mas & Vignes 2006). Thus I can define for each venture capitalist j the expected probability of success p_j of his start-ups s_j as a function of his competence c and of the market conditions M. With given market conditions M, p_j only depends on c_j, and $\frac{\partial p_j}{\partial c_j} > 0$.

$$p_j = E(q_{s_j})\,\bar{p} = f(c_j, M) \tag{8}$$

The outcome of an investment for the venture capitalist j is a Bernoulli trial of probability p_j. Thus the number of successful investments out of n follows a binomial law $B(n, p_j)$. From these considerations it follows that the average profits of all the investments made by a venture capitalists converge towards a value that depends linearly on p_j.

$$\hat{\pi_j} = \frac{1}{n} \sum_t \pi_{j,t} \xrightarrow[n\to\infty]{} (p_j\ g - 1)\ I \tag{9}$$

$$E(\hat{\pi_j}) = \frac{1}{n}(E(B(n, p_j))g - n)\ I = (p_j\ g - 1)\ I \tag{10}$$

$$var(\hat{\pi_j}) = \frac{1}{n^2}\ var(B(n, p_j))\ g^2\ I^2 = \frac{1}{n}\ p_j\ (1 - p_j)\ g^2\ I^2 \tag{11}$$

The expected profits of a venture capitalist depend through p_j on his competence c_j. After any number of investments made by a venture capitalist, the average of his realized profits is the best estimator of his performance. Therefore the best selection criterion for venture capitalists is $\hat{\pi_j} > 0$.

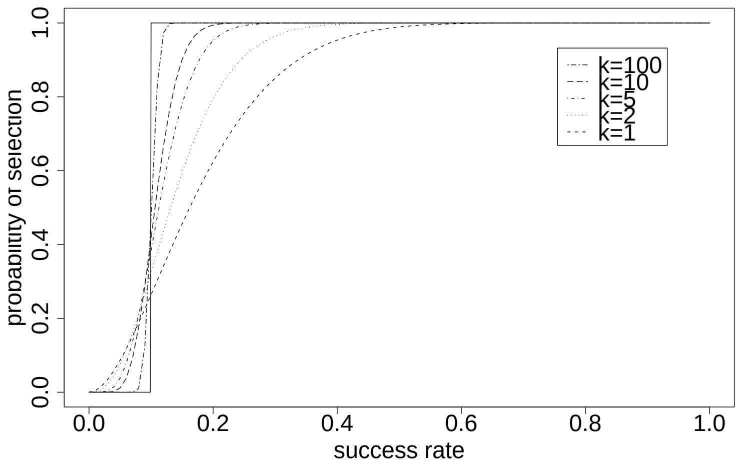


Fig. 1. Probability of selection of a venture capitalist at the end of his first limited partnership as a function of his expected success rate p_j for various number of investments per fund $LP = k\ g$ with $g = 10$.

4.2 Influence of the size of the limited partnership

For the institutional investor who gives the venture capitalists a fund to manage, the finite time horizon of the limited partnership provides both a powerful incentive to the venture capitalist to do his best and an opportunity to monitor his performance. The institutional investor can chose the size of the fund LP which correspond to a given number of investments after which he can choose to give an other fund to the venture capitalist or to give his chance to a new one. What is the impact of this choice on the selection of venture capitalists ?

Let assume for simplicity that the size of the limited partnership is a multiple of the return in case of success $LP = k\ g$. From the previous analysis I can compute the probability of meeting the selection criterion at the end of the first fund.

$$P(S|p_j, LP) = P\left(B(k\ g, p_j)\ g - k\ g)\ I > 0\right) \tag{12}$$
$$= P(B(k\ g, p_j) > k) \tag{13}$$
$$= 1 - F_{B(k\ g, p_j)}(k) \tag{14}$$

For a given size of the limited partnership $LP = k\ g$ the probability of selection can be expressed as the value at k of the complementary cumulative distribution function (ccdf) of a binomial law of parameters $k\ g$ and p_j. Figure 1 shows the evolution of the selection function for various size of funds. As k increases it converges towards an Heaviside's function $h(p_y - \frac{1}{g})$. The bigger the size of the limited partnership, the more accurate is the selection. But a bigger size of the fund corresponds also to a bigger risk for the institutional investor. Thus the choice of the size of the limited partnership is a balance between the accuracy of the selection and the risk taken by the institutional investor.

4.3 Global selection function

The outcome of the complete selection process is not tractable. In order to study it, I will simulate the model. After the end of the run it is possible to retrieve the global selection function using Bayes' rules. Let S be the event that the venture capitalist has been selected. Then the global selection function is :

$$P(S|c) = \frac{P(c|S)\;P(S)}{P(c)} \tag{15}$$

Where $P(c|S)$ is the final distribution of competence, $P(S)$ the ratio of the number of selected venture capitalists over the number of the venture capitalists that once entered the market and $P(c)$ is the initial distribution of competence.

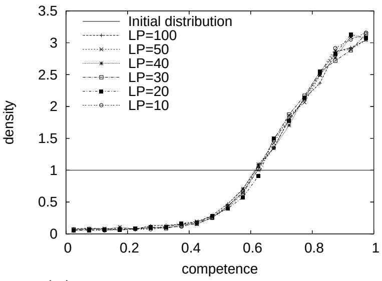
5 Simulations Results

I simulate the model with four different settings, with either a gaussian or an exponential distribution for the quality of start-ups and with a uniform or triangular distribution for the competence of venture capitalists. I then vary the size of the limited partnership. Following the formal analysis, I choose $LP = k\ g$, with $g = 10$. In each case I ran one hundred simulations of five thousand periods, with a population of one hundred venture capitalists that screen five hundred start-ups projects at each period.

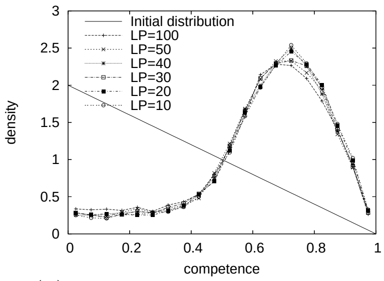
The parameters of the model have been chosen such that the minimum level of competence required to achieve positive profits is $c_G = 0.5$ for the gaussian distribution of quality, and $c_E = 0.6$ for the exponential distribution. The essential difference between the two distributions of quality is the spread of the function $p_j = f(c_j, M)$. With the exponential distribution the difference between more competent investors and lesser ones is bigger. It should leads to a more discriminative selection. The difference between the distributions of venture capitalists' competence, uniform and triangular, is the proportion of very competent venture capitalists. With the triangular distribution sufficiently competent venture capitalists are rarer, making the selection more challenging.

5.1 Final distribution of competence

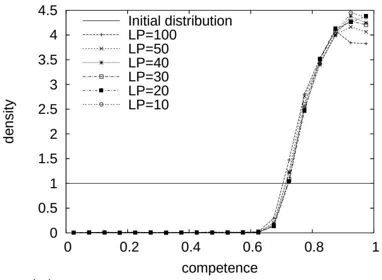
Figure 2 shows the resulting distribution of competence after five thousand time periods. These distributions are averaged across the hundred simulation run with each setting. The plain line represents the initial distribution of competence.



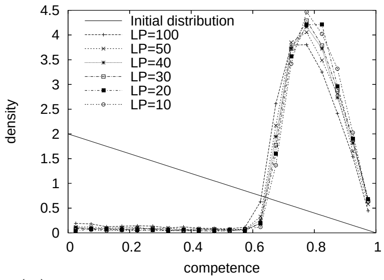
(a) Q: gaussian - C: uniform



(b) Q: gaussian - C: triangular



(c) Q: exponential - C: uniform



(d) Q: exponential - C: triangular

Fig. 2. Distributions of competence after 5000 periods

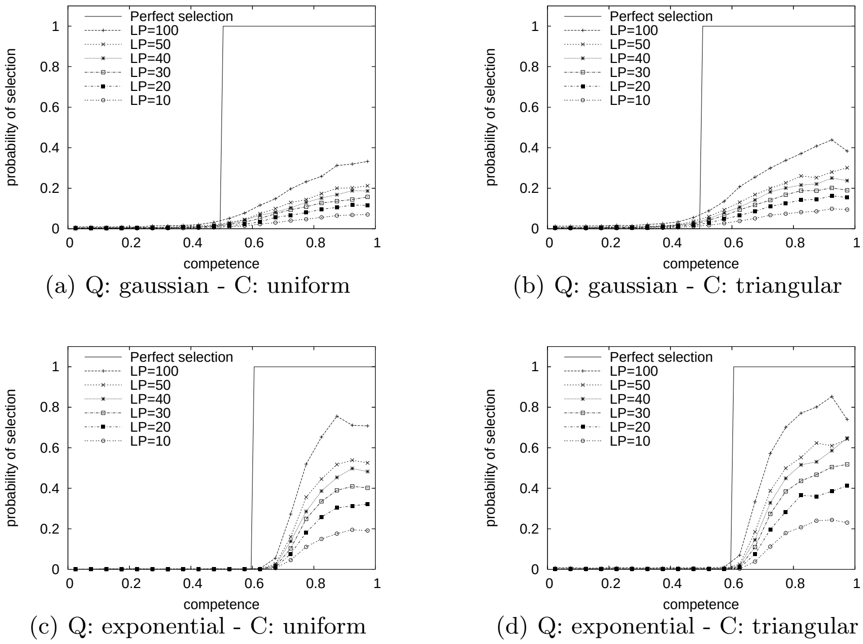
The first result is that the selection is efficient. For each setting, even with few very competent venture capitalists in the initial distribution, the final population in the market is essentially composed of sufficiently competent venture capitalists.

The second result is that the final distributions of competence are almost the same for all sizes of the limited partnership. This is surprising because the formal analysis shows that bigger sizes should lead to a more accurate selection. This means that even if each selection step is different, the result of the global selection process does not depends of the size of the limited partnership.

5.2 Market selection function

Figure 3 shows the global selection function for each process, computed using Equation (15). The plain line represents the perfect selection function. The selection functions are approaching the perfect selection function, getting closer to it as the size of the limited partnership increases.

The selection functions are clearly better with the exponential distribution of quality, which corresponds to a more discriminative spread of the performance of venture capitalists. The market has its own selection capability, determined by the market conditions which are essentially the distribution of agents' characteristics. The choice of the size of the limited partnership magnifies this selection capability.



(a) Q: gaussian - C: uniform (b) Q: gaussian - C: triangular

(c) Q: exponential - C: uniform (d) Q: exponential - C: triangular

Fig. 3. Selection functions after 5000 periods

This analysis confirms that even if the size of the limited partnership does not change the final distribution of competence, it affects the accuracy of the selection.

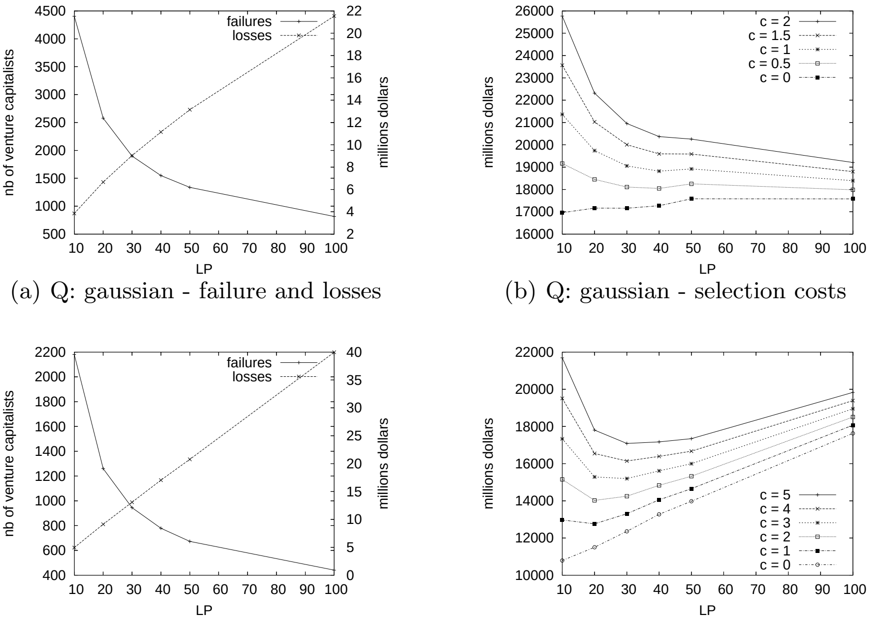
5.3 Optimal size of the limited partnership

The optimal size for the limited partnership LP corresponds to the optimal balance between selection accuracy and risk. In order to determine it, I compute the total selection costs paid to obtain the (identical) final distribution of competence with the following formula :

$$C_{selection} = E(losses) * failures + c_{search} * failures \tag{16}$$

The first term is the total losses incurred by the institutional investors before finding a sufficiently competent venture capitalists. It corresponds to the costs associated with type I errors (financing a bad venture capitalist). The second term is the total search costs incurred when the institutional investors have to look for a new venture capitalists to replace a failed one. It corresponds to the costs associated with type II errors (missing a good venture capitalists).

Figure 4 shows on the left the evolution of the number of failures and of the expected losses in case of failure with the size of the limited partnership. In



(a) Q: gaussian - failure and losses

(b) Q: gaussian - selection costs

(c) Q: exp. - failures and losses

(d) Q: exp. - selection costs

Fig. 4. Selections costs as a function of the size of the limited partnership $LP = k\ g$, $g = 10$

both cases, gaussian and exponential distribution of start-up qualities[1], expected losses increases linearly with LP, while the number of failures fails as a negative power of LP. As already stated, increasing LP increases both the risk and the accuracy of the selection.

The graphs on the right of Figure 4 allow to determine the optimal size for the limited partnership for different values of the unitary search cost c. In both cases the institutional investor will prefer the smallest size in the absence of search costs. If they don't pay for type II errors, it is rational for them to minimize the number of type I errors only. As the search costs increase, the optimal size for LP also increases. In the gaussian case, the selections costs with no search costs are very close. Thus the optimal size increases very rapidly with c. In the exponential case, on the contrary the initial difference is much bigger, and only slowly compensated by increasing search costs. With the same value of unitary search cost $c = 1$, the institutional investors will prefer the size $LP = 20$ in the exponential case when they already prefer the size $LP = 100$ in the gaussian case.

[1] For concision the results are only showed for the uniform distribution of competence. They do not change with the triangular distribution of competence.

6 Conclusion

The analysis of this agent based model provides a characterization of the selection process of competent venture capitalists with a noisy selection criterion. It shows that the market selection based on venture capitalists' past performance is efficient, it leads to a market with only enough competent venture capitalists. It also shows that the accuracy of the selection increases with the size of the limited partnership, i.e. with the number of observations used for the evaluation. The final distribution of venture capitalists' competence, however, is independent of the size of the limited partnership.

The choice of the size of the limited partnership for the institutional investor is a balance between the accuracy of the selection and the risk he incurs. The optimal size can be determined by the computation of selection costs. It depends on the relative costs of type I and type II errors.

References

Admati, A. & Pfleiderer, P. (1994), 'Robust financial contracting and the role of venture capitalist', *The Journal of Finance* **49**(2).

Aghion, P. & Bolton, P. (1992), 'An incomplete contracts approach to financial contracting', *The Review of Economic Studies* .

Amit, R., Brander, J. & Zott, C. (1999), 'Venture-capital financing of entrepreneurship: Theory, empirical evidence and a research agenda'.

Carlsson, B. & Eliasson, G. (2003), 'Industrial dynamics and endogeneous growth', *Industry and Innovation* **10**(4).

Chan, Y.-S. (1983), 'On the positive role of financial intermediation in allocation of venture capital in a market with inperfect information', *The Journal of Finance* **37**(5).

Dewatripont, M. & Tirole, J. (1994), 'A theory of debt and equity: Diversity of securities and manager-shareholder congruence', *The Quarterly Journal of Economics* **109**(4).

Gompers, P. A. (1995), 'Optimal investment, monitoring and the staging of venture-capital', *The Journal of Finance* **50**(5).

Gorman, M. & Sahlman, W. (1989), 'What do venture capitalists do ?', *Journal of Business Venturing* .

Hellmann, T. (1998), 'The allocation of control rights in venture capital contracts', *The Rand Journal of Economics* **29**(1).

Hellmann, T. & Puri, M. (2000), 'The interaction between product market and financing strategy : The role of venture capital', *The Review of Financial Studies* **13**(4).

Hellmann, T. & Puri, M. (2002), 'Venture capital and the professionalization of start-up firms: Empirical evidence', *The Journal of Finance* **57**(1).

Kaplan, S. N. & Stromberg, P. J. (2000), Financial contracting theory meets the real world: An empirical analysis of venture capital contracts, Working paper, University of Chicago, Graduate School of Business.

Kaplan, S. & Stromberg, P. (2001), 'Venture capitalists as principals: Contracting, screening, and monitoring', *The American Economic Review* **91**(2).

Kortum, S. & Lerner, J. (2000), 'Assessing the contribution of venture capital to innovation', *The Rand Journal of Economics* **31**(4).

Lerner, J. (1995), 'Venture capitalists and the oversight of private firms', *The Journal of Finance* **50**(1).

Mas, D. & Vignes, A. (2006), Why do we need venture capitalists ? the role of competent intermediaries in the creation process, Working paper, ERMES-CNRS.

Sapienza, H. (1992), 'When do venture capitalists add value ?', *Journal of Business Venturing* **7**.

Voorneveld, M. & Weibull, J. (2004), Prices and quality signals, Working Paper Series in Economics and Finance 551, Stockholm School of Economics.

Analysis of Random Agents for Improving Market Liquidity Using Artificial Stock Market

Shigeto Kobayashi[1] and Takashi Hashimoto[1]

[1] School of Knowledge Science, Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi, Ishikawa 923-1292 Japan
{s-kobaya, hash}@jaist.ac.jp

Abstract. Securing liquidity in a market is indispensable for the market mechanism to work efficiently. While there is a discussion that random agents have the effect to ensure the liquidity, the mechanism of this effect has not been clarified. We show that, using an artificial market system U-Mart, the random agents can improve the liquidity of a market by maintaining the balance of supply and demand. We also suggest that it is necessary for the random agents to manage their position in order to work as a market maker.

Keywords: Market Liquidity, Random Agent, Artificial Market, Institutional Design, Market Maker

1 Introduction

Extremely low liquidity in a stock market may cause severe problems of the market functions, since the market should serve as a place of fund collecting for every company and should be appropriately designed. We consider the market liquidity as the extent of contract without large price fluctuation and as a necessary condition for market stability. For example, in Osaka Securities Exchange that has only 1/40 volume of trading than Tokyo Stock Exchange, stock prices are sometimes not decided smoothly, and it may make the district economy sluggish. Shiozawa [1] calls such a market with low liquidity "thin market," and discusses that a market maker must play an important role in order to avert the problems caused by the thinness. Ueki et al. [2] point out that those who decide order prices randomly around the latest price can improve the market liquidity and that such random agents have an effect of the market maker. The mechanism to ensure the liquidity, however, has not been clarified. In this paper, we study how the random agents increase the liquidity using an artificial market simulation.

2 Simulation

We considered how the random agents affect a stock market using an agent-based artificial market simulator called "U-Mart" [3] (http://www.u-mart.org/). We conducted simulation experiments by changing the ratios of random agents to other agents, A_r, between 0% and 100%. The strategy of the random agents is to place orders at random around the latest price. We adopted the daily Nikkei Index that showed a downward trend as the spot price series given to the U-Mart system.

Firstly, we made sure that the contract rate increased with the ratios of the random agents (Fig. 1), as Ueki et al. [2] already showed. We also checked that the total trading volume also grew as the ratios of random agents became larger. Secondly, we measured the historical volatility. As indicated by the dashed line in Fig. 1, it decreased along with the ratios of the random agents. These evidences convinced us that the random agents affected the market to provide ample liquidity.

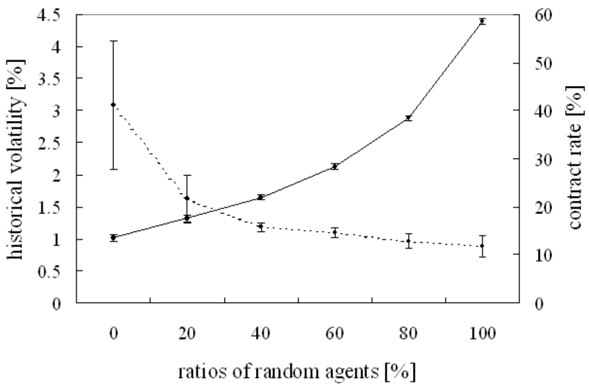


Fig. 1. The change of the contract rate (the solid line and the right vertical axis) and the historical volatility (the dashed line and the left vertical axis) related to the ratios of random agents. The error bars show the standard deviation of 100 runs of the simulation.

In order to clarify the mechanism to promote the liquidity by the random agents, we inquired into the way of ordering of both the random and the non-random agents (Fig .2). We found that, on the one hand, the non-random agents put the orders correlated with price changes. Namely, the buy orders are larger than the sell orders under positive price changes, and vice versa. Accordingly, the non-random agents are likely to skew supply and demand and to enhance trends. On the other hand, the random agents put the orders of both sell and buy irrespective to the price fluctuation. When the market is mainly occupied by the non-random agents, the left column of Fig. 2, the orders are not contracted, especially under the large price fluctuations. Although the random agents can potentially absorb such skewed orders, the volume is not enough. When the ratio of the random agents is large, the contract rate improves due to the balance between the order volumes by the non-random and the random agents, as shown in the right column of the Fig. 2.

In our simulations, the random agents often gained lower profit than the others and sometimes went bankrupt. In order for the random agents to efficiently work for providing liquidity, the agents with such function should exist in a market always. In order to do this, the agents should manage their position appropriately.

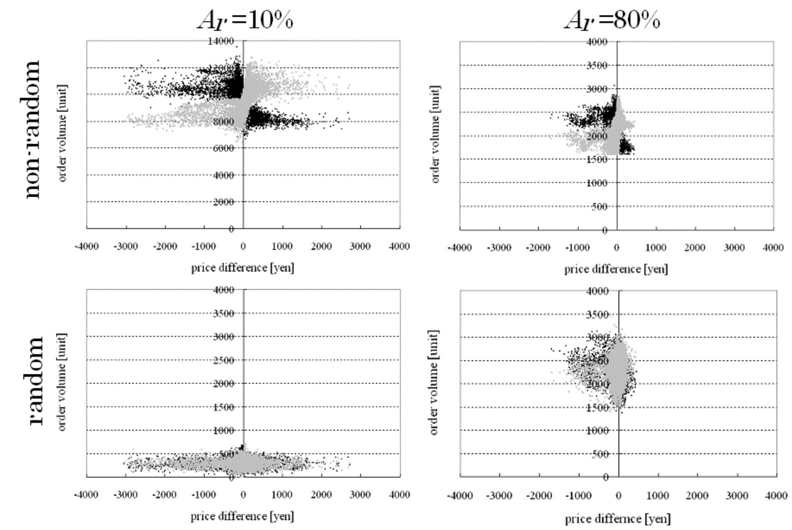


Fig. 2. The scatter diagram of order volume (the vertical axis) related to the price difference from the latest period (the horizontal axis). The left and the right columns are $A_{r'}$ =10% and 80%, respectively. The black and grey dots indicate the sell and buy orders, respectively. Note that the left upper graph has the larger scale of the vertical axis than the others.

3 Conclusion

We have showed that the random agents promote the liquidity in a stock market by improving the skewed state of the orders and maintaining the balance of supply and demand. We think institutional design is essential for market with enough stability. Accordingly, the characteristics of the random agent should be taken into consideration as the institutional design to provide liquidity for a market. We should further study the strategy for the random agents capable of managing their position.

Acknowledgments. This work was financially supported in part by foundation for the Fusion Of Science and Technology (FOST). We are grateful to FOST.

References

1. Y. Shiozawa: Making a Thin Market Intentionally: a Challenge of the U-Mart Project. 9th Workshop on Economic Heterogeneous Interacting Agents (2004)
2. J. Ueki et al: Simulation Study Using U-Mart (in Japanese). Papers of Annual Conference of the Japan Association For Evolutionary Economics. 6 (2002) 323-329
3. T. Terano et al: U-Mart: An Artificial Market to Bridge the Studies on Economics and Multi-agent Systems, Proceedings of Fourth Pacific Rim International Workshop on Multi-agents. (2001) 371-385

Firms Session

Transaction Cost Economics meets ABSS: a Different perspective on Asset Specificity in the IT-Outsourcing context

Bogdan Werth, Scott Moss

Manchester Metropolitan University Business School, Centre for Policy Modelling, Aytoun Street, Aytoun Building, M1 3GH Manchester, United Kingdom
{bogdan, scott}@cfpm.org

Abstract. Past research on IT outsourcing has mainly focused on the transaction itself with the use of the Transaction Cost Theory as a primary investigation framework. The notion of asset specificity was blindly accepted in its primary definition, which embodies a bone of contention for many academics due to its looseness of definition. This paper introduces an alternative metric of asset specificity in a qualitative way. It goes down the line of Agent-Based Social Simulation and represents actors of the outsourcing process as heterogeneous agents. This approach is contrary to a widely held perception of economic standard literature with its assumption of homogeneous actors.

Keywords: transaction cost theory, asset specificity, agent-based social simulation, outsourcing.

1 Introduction

The strategic importance of the IT in the banking sector is commonly accepted ([1], [2]), yet banks still continue to outsource parts of their information services assiduously and some even outsource them entirely [3]. A Recent report in the annual series of BCG benchmarking publications stipulates that *´Relatively few European banks have benefited from IT [Information Technology] outsourcing - obtaining IT services from external companies – to the extent they anticipated. Nonetheless, most intended to increase outsourcing activities in pursuit of reduced labour costs, specialized skills, process expertise, superior technical resources, and increased ability to focus on core business.´* [4]. This citation mirrors current tendencies in the financial industry precisely.

At first glance this trend seems somewhat counterintuitive. In compliance with the classical theory of the firm, organisations ought to have a constant aspiration to autonomy, thus trying to take as many essential business activities under their wing as possible in order to maintain relative independence [5]. On the other hand, the economic downturn of late 2000 reinforced re-engineering issues in many of the corporations in order to survive in the face of volatile competition. Therefore, the practice of outsourcing in modern corporations is proliferating and is likely to continue [6]. This trend forced both practitioners and academics to theorise and speculate on the underlying momentum towards outsourcing.

The topic of Information Technology/Systems (IT/S) outsourcing has been around in academic research for the last fifteen years [7]. A huge body of academic literature and research has emerged in that time and is continuing to evolve rapidly. The research was partly initiated by the disappointing effects of the first outsourcing attempts in the early nineties and continues to be stimulated in recent years by inexplicable successes and failures of firms.

Past research on IT outsourcing has mainly focused on the transaction itself, without investigating the strategic characteristics of the organization [8], not to mention a built-in social framework of the firm or social context it's located in. This trend captures the widely held perception that organisational members make sourcing decisions based upon an economic rationale and regard social factors as negligible in their influence on the overall picture of outsourcing. Therefore, a representative strand of research on IT/S sourcing has used the Transaction Cost Theory (TCT) [9] to investigate make-or-buy decisions.

However, it has been widely criticized that TCT contains some ambiguities with respect to the lack of precision in terms used. Dissatisfaction with different measures, dimensions or metrics used in TCT, or for its evaluation, is not new to outsourcing research ([10], [8]). One such bone of contention is the notion of *asset specificity*, which is central to the whole concept of Williamson's transaction based framework. It is rather crucial for TCT e.g. to understand what is meant by asset specificity to make a credible statement, whether the given theory accounts for empirical evidence or not.

This paper introduces a novel approach for investigation of the TCT. It goes down the line of Agent Based Social Simulation (ABSS) and tries to model actors of the outsourcing process as autonomous and heterogeneous agents that can act according to changes in the environment they are located in. Social structures emerge from the interaction and information exchange between individuals in the market. This approach is contrary to a widely held perception of economic standard literature with its assumption of homogeneous actors.

As proposed by Edmonds in [11] a methodological process of developing a simulation can be thought of as having different stages[1]: *abstraction, design, inference, analysis* and *conclusion.*

The presented model can be classified as being in-between the inference stage and the analysis stage. Abstraction of the system is partly done due to the usage of a theory (which is an abstraction itself already). Chosen theory is formalized by means of model development, thus initiating the stage of model execution. Evidence incorporated into the logic reasoning of the agents comes through constant interviewing and evaluation process with stakeholders from the industry.

[1] Most steps are not carried out in the consequent order and this is not implied by Edmonds. The order of steps is a subject to longstanding discussions on various social simulation mailing lists.

2 Transaction Cost Economics

Opponents of the TCT frequently tell the story of two hikers camping in the tiger country. One morning they wake up and find a tiger lurking close to their tent. While one of the hikers is puzzled as to what to do, the other one immediately reaches for his running shoes. On the reminder of his partner that he could not possibly outrun the tiger, he responds that he does not need to outrun the tiger at all. The only thing he has to do is to outrun the partner. This somewhat macabre joke is a powerful reminder on the assumption which is incorporated into Williamson's transaction cost framework – opportunistic behaviour. By reaching for his shoes, the hiker behaves in an opportunistic manner in deciding to abandon his partner. The reason for this behaviour is rather simple – in the world where everybody is opportunistic the hiker who decides to outrun the partner cannot be, *ex-ante*, sure that his partner will not behave opportunistically. Therefore, he doesn't spare any thought for the cooperative actions but tries to forestall his partner.

In TCT, as developed by Williamson, the central notion is given to transaction specificity of assets in combination with opportunism and bounded rationality. If there are assets which are specific to the particular transaction – i.e. have no or substantially limited use outside given transaction – this will cause dependence between transaction partners which yields transaction costs if there is a risk of opportunism[2]. In such a case, primarily, Williamson assumes symmetric dependence [12]. Thus if a producer uses assets that are specific to the transaction he will obtain a unique, or at least differentiated product, but then the discontinuity of the transaction will be a problem, not only for the producer but also for the customer[3].

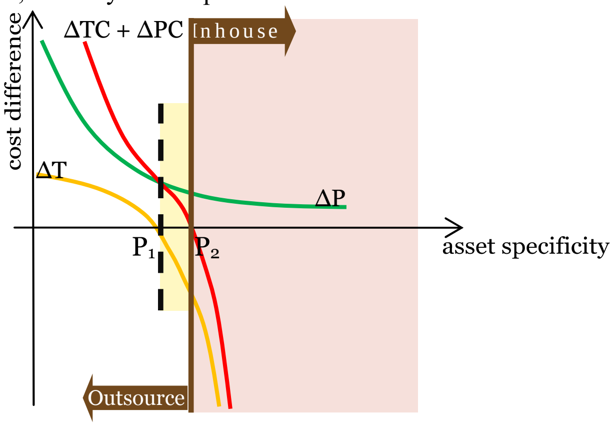


Fig. 1. Relationship between asset specificity, production and transaction costs. ΔTC – transaction cost, ΔPC – production cost. (Source: in the style of [13])

Due to the central role of the notion asset specificity in the model presented, it needs to be explained more detailed. In simple terms, we can assume that an asset is

[2] The situation is described under assumed conditions where rationality is bounded.

[3] It is assumed that the customer will not find an alternative supplier of an equivalent product immediately and therefore will suffer discontinuity and higher costs of production.

classified to be specific[4] if there is no alternative use or demand for a given asset outside of the relationship. Thus the asset is required for the particular transaction it is used in and is worthless in any other transaction. The claim of the present article is that in standard TCT there is too much looseness in the application of the concept of specificity, leading to misrepresentation of relations of dependence between buyer and supplier. The concept of transaction specificity of assets and its consequences seemed simple enough at the start but, on closer analysis, one finds himself asking questions about exact the meaning of this term.

Figure 1 shows a simplified relationship between asset specificity and a decision to produce in-house or outsource resulting from this relationship. From the point of origin to P_1 outsourcing is an advisable move; the Area from P_1 to P_2 represents the trade-off area between production cost advantages and transaction cost disadvantages, but it is still lucrative to outsource; beginning with P_2 the disadvantages of outsourcing outweigh the advantages and production has to be shifted in-house.

It has to be emphasised that it is not intended to claim that asset specificity is an insufficient condition for dependence. The presented model is still in its early stages for such statements to be made. The only claim made at this stage is that asset specificity does not have a commonly accepted metric due to its looseness of the definition. The model attempts to reproduce behaviour forecasted by TCT without using a numerisation of the asset specificity. Furthermore, it is hoped to be able to show that asset specificity is not central for the notion of dependence[5] and, more importantly, it is hoped to give an alternative metric of asset specificity, which is not based on the numerical value but has a qualitative character.

3 Methodology

A methodology of evidence-based modelling was adapted. The rules for the agents were derived partly from the relevant reports and partly from qualitative insight into the modelled target system. These insights were gained from semi-structured interviews with domain experts. Internal and published support documents were collected. It is intended to use stakeholders, not only as a primary source of qualitative data but also incorporate them into the model validation process. It should be remembered that the sample used for qualitative studies was opportunistic.

Subsequent sections describe the evidence-driven approach closer and introduce the modelling technique used.

3.1 Evidence driven

Common practice in the outsourcing projects is the dissemination of the best practices by consultants involved in the project allowing for perception of the coordination

[5] It is hoped to obtain this kind of insight at the later stage of the development when a validated model of TCT behavior will get incorporated with social components.

processes with respect to IT as increasingly commoditised [14]. This is exactly the type of stakeholders that were used for qualitative data. The anecdotal evidence incorporated into the model comes from interviews with stakeholders involved in outsourcing projects, either on the client side or on the vendor side.

The modelling and the filed work were carried out by the same party with data engineering involving stakeholder interactions. A constructive and modular approach to model design was adopted. The first research stage involved gathering of relevant data in collaboration with industry partners. Subsequently, a mock-up declarative agent-based coarse grained model, based on gathered, scarce data, was developed. Building a mock-up model first was intended in order to point out data requirements, and help to determine which data is important and which can be dismissed. In a second research stage additional interviews were conducted in order to fill gaps, which became evident through the analysis of the mock-up model.

Exploration and validation of the models' results will happen in a feedback-loop manner, together with domain experts. Thus through constant cross validation with stakeholders and domain experts, we claim to be on the right path for achieving "good social science" according to [15].

3.2 Declarative Modelling

The rules are implemented declaratively as this eases the handling of qualitative data in social simulation. The evidence obtained from domain expert interviews is best described qualitatively and the rules individuals give for their behaviour are of a qualitative nature as well. Therefore, as argued by Moss and Edmonds in [ibid], validation of software agents as adequate representations of real actors is made smoother by designing agents to perceive events specified by qualitative descriptions, maintain the qualitative terms during processing these qualitative perceptions and then act in a way that can be characterised qualitatively. A proposed way of maintaining the qualitative link between the language of the actors and the language of the agents is to use systems where action conditions and actions followed on these conditions can be defined in rules. These rules can then be used by some inference engine, with actions specified by consequences of these rules. One such rule for inference based systems used for the described model is an expert system allowing infer on a given set of rules.

Furthermore, expert systems allow considerable freedom in design of sophisticated agents as these systems perform well under a large amount of rules. They can also considerably mitigate difficulties in terms of information preparation and presentation, as the information stored in these systems is closer to natural description used by people. This fact makes it easier for non programmers and modellers to identify themselves with the model. On the back side, it has to be mentioned that a declarative approach implies a steep learning curve.

An Evidence driven modelling approach requires constant cross-validation with stakeholders and domain experts in both stages – development and validation. Having a tool which will enable better cross-validation will increase model credibility. This approach differs fundamentally from the usual approaches to statistical research and extends the usual approaches to qualitative research.

4 The Simulation Model

The model presented in this paper is a pure representation of the extreme selfishness notion incorporated in Williamson's theory aforementioned in the section 2. The main idea is to develop a model which reproduces the behaviour of individuals inspired by TCT and willingly abstract from any thoughts of social components that might matter in the target system. After such a model is validated, it is intended to feed it with social modules gradually while observing the clustered / volatile behaviour that might emerge as a consequence of actions. The resulting distribution of economic activity across different organisational forms emerges *bottom-up* from processes of interaction between agents, and their adaptation of future decisions according to agents' past experience.

There are two sets of rules in the model *general rules* – which describe local actions of individuals independent of their capacity to interact with each other, and *topological rules* – which describe the possibility of interactions through connections between individuals.

4.1 (Production) Space

Central to the Williamson's transaction cost framework is the notion of asset specificity. Specificity in its own turn leads to product differentiation and mutual/unilateral dependencies. One way to model it is to use the n-dimensional product characteristics space (PCS) [16] as was done, in the only agent based model on TCT found during literature review, by Klos in [17]. In Lancaster's proposed approach consumers and products are placed on an n-dimensional PCS, the location of a product represents a characteristics portfolio it offers and the location of the consumer represents his "ideal product". Klos locates products as well as consumers as points in a multidimensional PCS. Through the product differentiation and agent-specific variable Klos determines the returns for the producers.

This Lancastrian PCS approach was adopted by the author in such an altered way that instead of products and consumers, the space has been populated with agents representing vendors and consumers. Both product differentiation and asset specificity are defined upon this space. Vendors and clients are located in the production space. Location of the vendor agents represents the service he offers and the location of the client agent represents his in-house service he would like to outsource.

The evidence from stakeholder interviews suggested five most frequently outsourced IT functions: production, operations, application development, desktop services and local support. Therefore, a 5-dimensional product space was chosen.

Location on the grid is not exclusive. It is very possible for more than one agent (vendor/client) to be allocated at the same spot in the space.

Services are identified by an *n*-dimensional vector of salient characteristics [16]. Clients face a variety of vendors each offering a different set of characteristics. Once a client agent is allocated, his service characteristics are fixed. Behind this assumption is the belief that it is highly unlikely for a bank to change its in-house IT services. It is clear that by moving to multiple characteristics representation it becomes possible to incorporate many of the intrinsic qualities of vendor-client interaction.

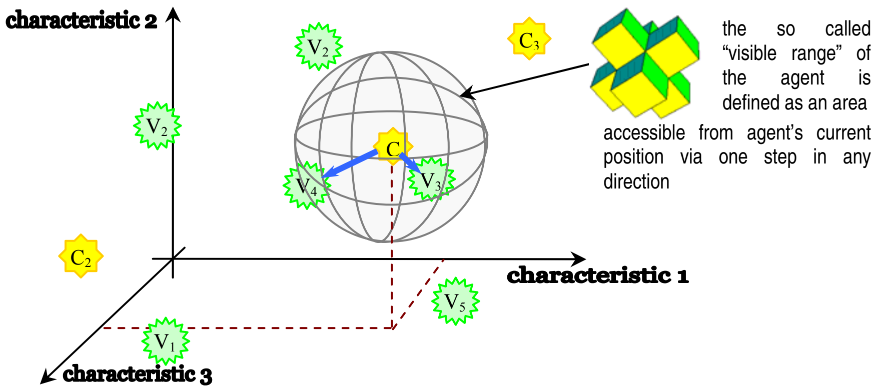


Fig. 2. Simplified 3D production space; octagon represents a client agent, 16-gon represents a vendor agent; the sphere around the C_1 agent represents his "visible" [6] space.

Figure 2 shows a sample client agent C_1 with corresponding acceptance range. Two vendors V_4's or V_3 are currently in client C_1's "visible" space. The arrows from client C_1 to vendors V_4 and V_3 stay for the distance vectors, which represent the cost of "transportation" C_1 has to incur in order to obtain either V_4's or V_3's service.

The acceptance range is also used for the representation of the bounded rationality. The essence of bounded rationality is that agents' computational and information processing capacities are limited [18]. Therefore, allowing agents to perceive the "outer world" though the limited "eyes" of acceptance range, the concept of bounded rationality is modelled.

4.2 Opportunism[7]

The client re-evaluates his visible range in every time step of the simulation. The more vendors become "visible" (get into his range) for the client, the more likely he is to terminate the ongoing relationship in case he is in one. The ulterior motive for this is the notion of standardisation. As there are more vendors who satisfy the outsourcing needs of a particular client, he starts thinking that his services become more standardised. This transforms the symmetric transaction dependence of the vendor and client into unilateral dependence on the part of the vendor – switching cost for the client decreases whereas vendor's switching cost in its turn increases.

A client expects the vendor to adjust his characteristic portfolio in order to experience economies of scale and experience. Therefore, the client will send out regular requests for characteristic's adjustment. By complying, the vendor starts to

---

[6] While talking about one agent's acceptance range the term *visible* will be used in order to reflect how many other agents would satisfy his demands with respect to characteristics.

[7] The sort of opportunism described in this section deals with opportunistic behaviors, which may occur during a transaction between two agents.

transform his assets in a way specific to the transaction he is in. Analogously, the vendors are assumed to be entirely opportunistic too. On that account the vendor would comply with the request as long as it does not put him at a disadvantage with the client in terms of losing potential customers, or getting too specialised on the needs of the contracted client. In other words, the vendor agent will gradually move towards the client agent until he starts losing other clients, which were "visible" in his range before. This means that with respect to his reasoning, the vendor will start thinking that his adjustment in the service characteristics became too specialized and are not covering the same broad range of clientele needs as before. Thus he is more willing to break up the transaction after repeated requests.

4.3 Preferences and Matching

The choice of partners is initiated by the clients. It is assumed that if vendors become "visible" in the clients acceptance range, this equals to the vendor advertising himself. The client would choose his contractual partner on the basis of the best match to his characteristics (i.e. vendor with shortest distance to the client). Thus, the client would send an arbitrary number[8] of requests to the best matched vendors.

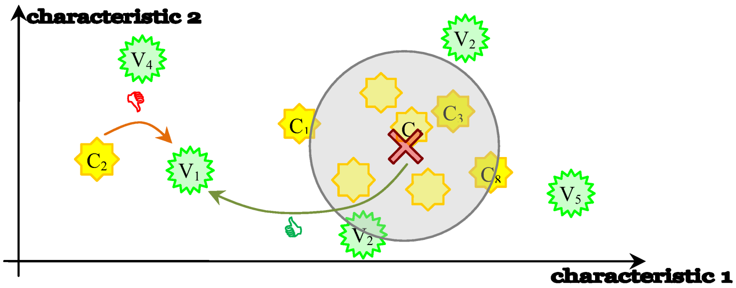


Fig. 3. Simplified 2D-production space; octagon represents a client agent, 16-gon represents a vendor agent; all clients in the figure are assumed to be in the "visible" range of the vendor V_1, darkened area in the circle around the cross is the preference area for the vendor V_1 where he is likely to accept a request from.

The vendor in his turn, due to his selfish attitude and behaviour, would evaluate all requests about possible contracts sent to him and go for one which comes from the client located in the biggest conglomeration of other clients. This case's hidden agenda is to portray the vendor's drive to remain attractive to as many clients as possible. As the vendor is expected to adjust his characteristics in compliance with his contractual partner by moving closer to him, the vendor aims for conglomerations of clients. Thus, by moving closer to the client in this setting, the vendor wins potential clients in case his current client will opt out of the transaction.

Figure 3 shows a simplified example of the matching process for the vendor agent V_1. For simplistic reasons it is assumed that all agents depicted are "visible" to each

[8] The maximum number of request a client agent can process is initialized for each single agent at the beginning of the simulation.

other. Despite the request from the client C_2, the vendor will turn down his request if there are any requests from clients allocated in the darkened area around the cross.

In accordance to the vendor's assumption that clients will not change their service characteristic portfolios, he will rather try to match his characteristics to the contractual partners based in such conglomerates as shown in the picture.

In case a client could not find any suitable vendor, even after altering his acceptance range, he will be matched to himself, thus preferring in-house production over outsourcing.

4.4 Adaptation

Some Vendors, who did not get any requests, will relocate randomly one step in any direction in the production space. This is an attempt to alter services' characteristics in order to look more appealing to the clients which might have not considered this vendor otherwise. Thus moving in any direction along the axis might bring the vendor in the acceptance range of one or more additional clients.

The Customers are assumed to be stationary and alter their acceptance range only. They will never relocate in the production space during the whole simulation, as it is unlikely that clients will adjust their IT services only to get a better match to the vendor(s). In case a client could not find an appropriate vendor match in the time step (be it due to the refusal of the vendor or lack of vendors in his acceptance range), he will change his preferences with respect to what is an acceptable service level agreement to him. Thus he may adjust the width of his acceptance range – it will get bigger. One should bear in mind the fact that the further the client goes in his search for an appropriate vendor, the less attractive the relationship appears to him and the more likely he is to break the relationship prematurely in spite of better alternatives.

In the case of the contractual relationship the client's acceptance range will either stay the same – e.g. if the client is not forcing the vendor to become more cost or production efficient – or decrease gradually – e.g. if the client expects vendor's compliance to the agreed customizations for the particular transaction.

5 Preliminary results

Two categories of results are presented – and excerpt out of random agent's life (micro perspective) and the rule activation's statistics of a single simulation run (macro perspective).

5.1 Micro perspective

Due to the space constraints only statistical analysis of the excerpt can be given. The data in the table 1 depicts the typical vendor-agent's actions during the whole simulation run of 300 time steps. The names of the rules are self-explanatory.

Table 1. The set of rules fired for the “vendor-7” during the simulation run of 300 time steps

Rules	Amount of activations
init-acceptance-range / init-service-location	1
adjust-service-due-2-client-demands	12
adjust-service-due-2-no-requests	27
reject-request-due-2-disadvantaged-location-of-client	74
accept-request	11
reject-request-due-2-no-capacity	14
refuse-adjustment-for-client-due-2-own-disadvantage	33
terminate-transaction-prematurely	7
terminate-transaction-naturally	2

The excerpt illustrates perfectly the selfish nature incorporated into the vendor-agent, when he repeatedly refuses complying with client’s demands to adjust his service levels and finally breaks the relationship. Vendor breaks the relationship because he thinks that further adjustment of his services for his contractual partner will be disadvantageous. According to vendor’s reasoning, next service adjustment action will lose him two possible customers “client-2” and “client-38”.

5.2 Micro perspective

Same simulation run as the excerpt in the section 5.1 is presented from the macro perspective. This evaluation of results should shed light on the overall activity and dynamics within the model.

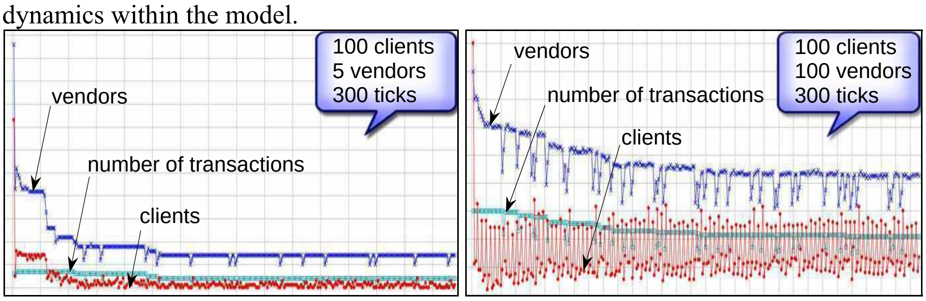


Fig. 4. Number of rule activations per tick for client and vendor agents

Generally speaking, the diagram of activations in the figure 4 exhibits that the model is in constant flux, thus agents are applying either topological or/and general reasoning rules.

At the start of the simulation, a large number of rule activations are observable. This is partly due to the setup rules which fire at the beginning of the simulation but also partly due to the high degree of interaction between vendors and customers at the start of the simulation as they are seeking for the contractual partners. The use of agents’ memory economises on re-evaluations at later stage of the simulation.

Later on, a rash decline of activity is to observe. The constellation of 100 vendors and 5 clients produces a long and stable periods of during contracting time as one can see on the vendor’s line. The constellation with similar number of clients and vendors, however, produces a more volatile overall picture. A graduate increase in activity due to frequent break-ups of the transactions – either because of the natural or premature cancellation of contracts – can be recorded. The critical mass of free vendors reaches the peak and due to the selfishness there is a mass termination of contracts by clients to observe.

Admittedly, both constellations are marked by slight but steady decrease in the overall number of transactions. This can be explained together with the decline in the relocation activities of vendors. After some time vendors located themselves in the areas surrounded of highest client congestion, leaving some other clients alone.

6 Discussion

The example of the TCT showed that ABSS can be utilised as an appropriate tool for theory examination. The ever proliferating research on IT/S outsourcing seems to find its state of equilibrium in a sort of standard deviation with the TCT as the mean. Therefore we need a more critical observation of theories borrowed from different disciplines before these can be established. The methodology presented in the paper does not aim to criticise concepts of TCT as such, it just aims to point out obscurities where the theory is not clear enough. As the process of formalisation is a discipline which needs to be precise ABSS claims to compensate for the lack of precision in the IT/S research so far. ABSS uses models to devise precise statements about these theoretical points, which are not clear in the theory a priori.

The alternative metric of asset specificity used for agent's reasoning succeeded in reproducing results forecasted by TCT and offered a qualitative description of the term. Organisational structures emerged from the interaction and information exchanges between individuals in the artificial world. It was shown that organisational structures depend on connectivity and information exchange between agents.

Nonlinear social behaviour needs to be included in the outsourcing research on the more elaborate basis rather than just excluding it from the outset. Nonlinear behaviour is excluded in most econometric observations as the concept of nonlinearity gives problems to the statistician and their forecasts. It needs to be emphasized that this paper does not proclaim the illegitimacy of the concept of asset specificity as such. On the contrary, the author is of the opinion that it is a useful concept although it lacks an exact definition and metric, which are needed for its proper instrumentalisation.

The next step in the elaboration of the presented model will be the gradual enrichment of the agents with rules for social interaction on top of already implemented rules stemming from the selfish roots of TCT.

References

1. BCG - Blumstengel, A., Leibrock, E., Minz, R., Möllenkamp, H.,van Laak, C., Wolter, A.: Renewing Core Banking IT Systems: Open-Heart Surgery for European Banks, BCG report May 9. (2006)
2. OECD - Organization for Economic Co-operation and Development: Banks Under Stress. (1992)
3. Ang, S., Detmar, S.: Costs, Transaction Specific Investments and Vendor Dominance of the Marketplace: The Economics of IS Outsourcing, in Information Systems Outsourcing: Enduring Themes, Emergent Patterns and Future Directions, Hirschheim, R., Heinzl, A. and Dibbern, J. (eds.), Springer, Verlag: Berlin, pp.47-76. (2002)
4. BCG - Blumstengel, A., Leibrock, E., Minz, R, Möllenkamp, H., Neufeld, E., van Laak, C., Wolter, A.: IT Outsourcing and Offshoring: Hype or Opportunity? IT Cost Benchmarking in the European Banking Industry, BCG report, June 28. (2005)
5. Gouldner, A.: Organizational Analysis, in Sociology Today, Merton, R., Broom, I. and Cottrell, I. (eds), Basic Books, New York. (1959)
6. Tettelbach, B.: CIO Update: How Measurement Can Help in Evaluations of IT Outsourcing, Gartner Group document #IGG05032000-04, 3 p. (2000)
7. Dibbern, J., Goles T., Hirschheim R., Jayatilaka B.: Information systems outsourcing: a survey and analysis of the literature, ACM SIGMIS Database, Volume 35 , Issue 4, Pages: 6 - 102, 2004, ISSN:0095-0033, ACM Press. (Fall 2004)
8. Aubert, B. A.; Croteau, A.: Information Technology Outsourcing from a Business Strategy Perspective, Cahier du GReSI # 05-01, February. (2005)
9. Williamson, O.E.: Markets and hierarchies, analysis and antitrust implications: A study in the economics of internal organization, New York: Macmillan. (1975)
10. Aubert, A.B., Weber, R.: Transaction Cost Theory, the Resource-Based View, and Information Technology Sourcing Decisions: A Re-Examination of Lacity et Al.'s Findings, GreSI publications, May. (2001)
11. Edmonds, B.: The Use of Models - making MABS more informative, in Moss, S. and Davidson, P. (eds.) Multi Agent Based Simulation 2000, Lecture Notes in Artificial Intelligence, 1979:15-32. (2001)
12. Nooteboom, B.: Research Note: An Analysis of Specificity in Transaction Cost Economics Nooteboom Organization Studies.(1993); 14: 443-451
13. Dibbern, J., Güttler, W., Heinzl, A.: Die Theorie der Unternehmung als Erklärungsansatz für das Outsourcing der Informationsverarbeitung, Entwickling eines theoretischen Bezugsrahmens, Universität Bayreuth, Lehrstuhl für Betriebswirtschaftslehre VII (Wirtschaftsinformatik), Arbeitspapier 5/1999 (Fassung vom 20.07.1999)
14. Poppo, L., Lacity, M.C.: The normative value of transaction cost economics: What managers have learned about TCE principles in the IT context, in “Information Systems Outsourcing: Enduring Themes, Emergent Patterns and Future Directions”, Hirschheim R., Heinzl A., Dibbern J. (eds), Springer, Dordrecht, Netherlands, 253-276. (2002)
15. Moss, S., Edmonds, B.: Towards Good Social Science, JASSS vol. 8, no. 4. (2005)
16. Lancaster, K.: A New Approach to Consumer Theory, Journal of Political Economy, 74(april), 132-157. (1966)
17. Klos, T., B., Nooteboom, B.: Agent-based computational transaction cost economics, Journal of Economic Dynamics and Control, Elsevier, vol. 25(3-4), pages 503-526, March. (2001)
18. Moss, S., Dixon, H. D., Wallis, S.: Evaluating Competitive Strategies. (1994)

Investment Strategies in Innovation Competition – a Simulation Analysis of the Pharmaceutical Industry

Tino Schütte[1]

[1] Chair of Managerial Economics, University of Technology Dresden
01062 Dresden, Germany

Situated in the research field of market structure and innovation a model of product innovation competition is developed to outline the effects of different R&D-investment strategies. Placed in a multi-firm, multi-product setting the decisions of firms to invest in innovation, improvement or imitation with regard to their probability of success is investigated. Specific features of the drug market are taken into consideration. A simulation approach is used to explain how the benefits, cost of R&D, interaction between competitors and the influence of market characteristics determine individual dynamic R&D spending over time. The model claims to be history friendly by adjusting inputs to innovation races observed and taking place in the pharmaceutical industry.

Keywords: agent-based modelling, innovation races, pharmaceutical industry

1 Introduction

The pharmaceutical industry is a science-based industry confronted with the peculiarities of the drug market. Research, preclinical and clinical development processes of new drugs are full of uncertainties, expensive and of long duration. Nevertheless there are opportunities for high profits because of protection by patents.
Based on the game theoretic modelling of patent races, later enlarged to multi-stage games e.g. Fudenberg et al. (1983), Reinganum (1985), or Harris/Vickers (1987)) a dynamic asymmetric product innovation competition model is built, which includes technological as well as market uncertainties. The model tries to answer questions concerning the influence of rivalry not just by its quantitative amount but also by the firms' strategies and their history of R&D-activities.
In line with simulation analysis of industry evolution in general and the development of the pharmaceutical industry in special (e.g. Malerba/Orsenigo (2002)) empirical findings are taken as inputs for the parameters in use (e.g. data from the EFPI).

2 The model

Placed in a multi-firm, multi-product setting the decisions of firms to invest in innovation, improvement or imitation, with regard to their probability of success, are investigated. Actors have to decide whether to take part in innovation races by allocating R&D-investments in an appropriate time and amount. A firm's probability to be successful with an innovation and the probability of being the first depends directly on its knowledge stock. Incorporating prior participation in innovation races, an individual learning process is considered to display firms' experiences for current decisions. Integrating uncertainty (when and who will be awarded a patent) the value of the turnover is simulated by evolving the demand depending on the market volume, marketing efforts and the offered product portfolio in the industry. Investments in marketing activities are used to increase the market penetration rate of a drug as well as for improving the image of the company. Taking "close to reality"-scenarios, investment strategies are defined to analyse the profit policy of firms and the market structure of the whole industry in the short and longer run.

The supply side of the market is categorized by three basic types of pharmaceutical companies found in the industry, firms with a focus on the development of new molecular entities, firms with a focus on "me-too" drugs and producers of generics. Firms have to allocate each period their available financial resources to research, development and marketing activities. Production cost are neglected because of there minor importance especially in research-based pharmaceutical companies.
The model claims to be history friendly by adjusting to innovation races observed and taking place in the Pharmaceutical Industry. Specific features of the drug market, e.g. regulated prices, inelastic demand curves and strong competition in drugs without patent protection are taken into consideration.

As modelling follows an evolutionary approach with bounded rational subjects an agent-based simulation environment, NetLogo, is used. This done with the aim to sketch some stylized facts of the pharmaceutical industry evolution.
First results show that there is some sensitivity to parameter variations with respect to the assumed probability distribution of making an invention, competitors' investment behaviour, and differences in the available budget or barriers to market entry. It seems that firms which relate their investment decisions to successful behaviour in the past (reinforcement learning) can defend or improve their market position whereas firms following time-invariant decision rules are more likely to face market share decline.
In a model extension a simple algorithm for mergers and acquisitions of firms, based on a company's available budget, is implemented. By integrating recent developments especially with regard to the growing importance of biotechnological firms a discussion about the importance of economies of scale and scope and predictions about possible future developments are tried. Most importantly recommendations for investment decisions are to be given for selected scenarios.

Agent-based Modeling of Human Organizations

Alexei Sharpanskykh

Vrije Universiteit Amsterdam, Department of Artificial Intelligence,
De Boelelaan 1081a, 1081 HV Amsterdam, the Netherlands
sharp@few.vu.nl

Abstract. At present the agent paradigm is often used for computational modeling of human behavior in an organizational setting. However, in many developed models only a limited number of (unrelated) organizational aspects are represented. Furthermore, some of these models make little use of a rich theoretical basis developed in social science. This may undermine the practical feasibility of such models. This paper proposes a formal approach for modeling of characteristics and behavior of agents in organizations, diverse aspects of which are represented using an expressive formal framework. The approach is based on the theoretical findings from social science and enables analysis of how different organizational and environmental factors influence the behavior and performance of agents. The approach is illustrated by a simulation case.

1 Introduction

The agent paradigm has been extensively used for modeling and analysis of both human and artificial organizations. In particular, in the area of Multi-Agent Systems (MAS) the representation of a system as an organization consisting of roles and groups can help to handle high complexity and poor predictability of the system dynamics [11]. Although organizational models of MASs can be computationally effective, nevertheless most of them have a limited ontological expressivity required for modeling of human organizations. Furthermore, such models only rarely make use of an extensive theoretical basis developed in social science.

Modeling of individuals in a social setting using the agent has gained popularity in the area of computational social science [3]. In contrast to the traditional methods (e.g., based on system dynamics [8]) that abstract from individual events and entities and take an aggregate view on the social dynamics, the agent-based approaches take into account the local perspective of a possibly large number of separate agents and their specific behaviors in (formal) organizational structures. Agent-based social simulation has been used for investigating organizational structures and dynamics at macro- (e.g., market fluctuations [1]), meso- (e.g., interactions between organizations [5]) and micro- (e.g., personal traits and organizational performance [25]) levels. In many approaches that identify and exploit relations between different levels much attention has been devoted to analyzing, predicting and improving the effectiveness and efficiency of the allocation and the execution of organizational tasks to/by different types of agents. In particular, the frameworks TAEMS [6] and VDT [14] provide the elaborated models for (collaborative) task environments and the computational means to analyze the performance of agents and of a whole organization with respect to the task execution. The agents in these and other similar frameworks are represented as autonomous entities with such characteristics as skills, competences, experience, and, sometimes, goals. In task-oriented agent-based modeling it is often assumed that agents comply with organizational goals and will perform tasks in such a way that a high level of organizational performance is ensured. However, in some cases such an assumption may not be valid. In particular, for feasible modeling of human organizations various (sometimes conflicting) interests of different organizational actors should be explicitly considered, as they often (significantly) influence the organizational performance. In general, to stimulate productive work of employees, an organization should reconcile (or align) its goals with the (key) goals of its employees. Furthermore, the organization should arrange work and provide incentives to its employees in such a way that they are constantly motivated to adopt the behavior that ensures the satisfaction of the essential organizational goals. The topic of work motivation has received much attention in Organization Theory [10, 13, 15, 18, 19]. Also, different computational motivation models and the mechanisms for manipulating them have been proposed [4]. However, only a little research has been done on the computational modeling of motivation and intentional attitudes of agents situated in the organizational context. Organizational factors that influence the behavior of agents are diverse: e.g., norms and regulations related to the tasks execution, to communication, a power (authority) system, a reward/punishment system etc. Furthermore, many of these factors are interrelated (e.g., a power structure influences the execution of tasks). However, often models that are used in social simulations consider only a limited number of organizational aspects and do not reveal (inter-) dependencies that exist between these aspects. This results into limited evaluation possibilities of effects of different organization processes and may undermine the practical feasibility of such models.

In this paper, a formal agent-based approach for modeling of characteristics and behavior of individuals in the organizational context is proposed. The approach makes use of a rich theoretical basis developed in Organization Theory. In particular, the motivation modeling of agents is based on the expectancy theory (the version of Vroom) [26] that has received good empirical support. The formal motivation modeling has an advantage that automated tools can be developed using which (human resource (HR)) managers can make estimations of how different organizational factors influence the motivation and performance of different types of employees (agents). Agents are situated in a formal organization modeled using the general organization modeling and analysis framework proposed in [12]. This framework comprises several interrelated views: the performance-oriented view [21] describes organizational and individual goal and performance indicators structures; the process-oriented view [20] describes task and resource structures and dynamic flows of control; within the organization-oriented view [12, 24] organizational roles, their power and communication relations are defined. Concepts and relations within every view are formally described using dedicated languages based on an order sorted predicate logic [16]. Temporal relations within and across the views are formalized using the Temporal Trace Language (TTL) [23], which is an extension of an order sorted predicate logic that allows reasoning about dynamic properties of systems.

Both the order sorted predicate logic and TTL are also used for specifying the structural and temporal aspects of agent-based models correspondingly.

The paper is organized as follows. Section 2 introduces the proposed modeling approach. The application of the approach is illustrated by a simulation case study in Section 3. Section 4 concludes the paper.

2 An Agent-based Modeling Approach

Using the general modeling framework an organizational model that comprises concepts and relations from different views is specified. The elements of the model are related as follows: Organizational goals are structured into a hierarchy using the refinement relations. Goals are satisfied by execution of certain tasks. Different sets of organizational tasks are associated with roles. Interaction (e.g., communication) and authority structures are defined on organizational roles with respect to tasks. To enable effective and efficient execution of tasks, agents with appropriate characteristics should be allocated to roles. In this Section, a description of professional, psychological, and intentional agent characteristics is provided (Section 2.1), followed by the introduction of a motivation model of an agent (Section 2.2).

2.1 Characteristics of agents and allocation to roles

For each role a set of requirements on agent *capabilities* (i.e., knowledge and skills) and *personal traits* is defined. Requirements related to knowledge define facts and procedures with respect to organizational tasks, confident understanding of which is required from an agent. Skills describe developed abilities of agents to use effectively and readily their knowledge for tasks performance. In the literature [18] four types of skills relevant in the organizational context are distinguished: technical (related to the specific content of a task), interpersonal (e.g., communication, cooperation), problem-solving/decision-making and managerial skills (e.g., budgeting, scheduling, hiring). More specific requirements may be defined on skills reflecting their level of development, experience, the context in which these skills were attained. To enable testing (or estimation) of skills and knowledge, every particular skill and knowledge is associated with a performance indicator(s) (PI) (e.g., the skill ‘typing’ is associated with the PI “the number of characters per minute”). Notice that some indicators may be soft (not directly measurable) (such as the level of flexibility); the value of such indicators may be established by indirect evidences (e.g., from the agent’s history and achievements). Moreover, a skill may be associated with a compound PI built as a weighed expression on simple PIs.

Personal traits may also influence the successfulness of the execution of tasks. The traits are divided into five broad categories discovered in psychology [13]: openness to experience, conscientiousness, extroversion, agreeableness, and neuroticism. In some cases agent personal traits may be evaluated through psychological tests and by consultations with agents’ referees. Some agent’s traits may mediate the attainment of agent’s skills. For example, extroversion and agreeableness play an important role in building interpersonal skills.

Agent capabilities and traits can have different levels of importance. Whereas required for a role capabilities and traits are compulsory for taking the role, desired capabilities and traits considered as an advantage. In some cases an organization may tolerate the deficiency in (or insufficient level of development of) some skills if a feasible guarantee is provided that this gap will be filled during a certain time period.

Most of the approaches on personnel management used currently are based on the HR-models [19]. In contrast to the traditional scientific management models [17], the HR-based approaches pay a special attention to the needs, desires and goals of employees and to the alignment of the individual goals with the organizational ones. Therefore, during the evaluation of agents-candidates for a role, also the goals of the agents should be taken into consideration (to a possible extent) to identify similarities and conflicts with the organizational goals.

In modern social science behavior of individuals is considered as goal-driven. A goal is defined as an objective to be satisfied describing a desired state or development of the individual. It is recognized that high level goals of individuals are largely dependant on their needs. These needs are to a great extent determined by the individual behavioral and biological history (i.e., biological and social background). Currently the following division of needs is identified in social science: (1) *extrinsic needs* associated with biological comfort and material rewards; (2) *social interaction* needs that refer to the desire for social approval, affiliation and companionship; (3) *intrinsic needs* that concern the desires for self-development, self-actualization, mastery and challenge. Such a categorization has some similarities with the hierarchy of needs proposed by Maslow [10]. However, a number of empirical studies showed that the Maslow's key hypothesis that the high-order (intrinsic) needs cannot motivate behavior of an individual until the lower-order (extrinsic) needs are satisfied does not hold for all individuals. Empirical evidences confirmed that the importance (or the priority) of different types of needs (and the associated goals) often changes over time in different life phases of an individual. The characteristics of an agent can be formalized using the sorted first-order predicate logic as it will be shown in Section 3.

In general, the efficiency of allocation of an agent to a role is dependant on how well the agent's characteristics (i.e., capabilities and traits) and goals fit with the role specification and the requirements. However, modern organizations implement very diverse allocation principles (e.g., based on equality, seniority or stimulation of novices) [10]. Such principles can be formalized as allocation policies comprising executable (temporal) rules. An example of such a policy is given in Section 3.

When an individual is allocated to a role, the identification of his/her specific lower level goals is performed in cooperation with a managerial representative of the organization. During this process, the high level goals, based on the agent's needs are refined into more specific goals aligned with organizational goals using AND- and OR- relations as it is shown in [21]. Many authors argue that the lower level goals should as detailed and specific as possible [9, 19]; furthermore, such goals should be attainable by agents. Often two types of such goals are distinguished: development (or learning) and performance goals. Development goals reflect wishes of agents to gain certain knowledge or some skills that are also useful for the organization. For example, the attainment of the skills required to perform task(s) interrelated with the task(s) already assigned to the agent may enable the allocation of the agent to a more global and essential (composite) task. Individuals vary in the abilities and desires to learn; therefore, this type of goals is particularly dependent on the individuals' traits and goals. Performance goals usually concern the effectiveness and efficiency of the execution of the tasks already allocated to the agent. Both development and performance goals may change over time.

Within the performance-oriented view of the modeling framework [21] the formal specification of a goal is based on a mathematical expression over a PI(s). The characteristics of a goal include, among others: *priority*; *horizon* – for which time point/interval should the goal be satisfied; *hardness* – hard (satisfaction can be established) or soft (satisfaction cannot be clearly established, instead degrees of *satisficing* are defined); *negotiability*. For example, the hard performance goal "it is required to maintain the time for the generation of a plan < 24 hours" is based on the PI "the time for the generation of a plan". Another example is the development goal "it is desired to achieve the state in which the framework JADE is mastered". In the latter example the goal is desirable, which points at its low priority.

The satisfaction of goals in the organizational context is associated directly or indirectly with the performance of tasks. In particular, goals associated with intrinsic needs are often satisfied by intrinsic rewards that are a natural consequence of the agent behavior related to the execution of a task. While externally provided rewards (e.g., salary, bonuses, group acceptance) serve to the satisfaction of goals related to extrinsic and social interaction needs. At any time point the (level of) satisfaction of a lower level goal may be established by the evaluation of the PI expression, on which the goal is based. Further, using the rules defined in [21] information about the satisfaction of lower-level goals is propagated to determine the satisfaction of high-level goals.

Many organizations have reward/sanction systems contingent on the satisfaction of goals. Furthermore, besides general organizational policies also particular individual policies (e.g., concerning promotions, bonuses etc.) can be defined. Such policies can be also formalized by sets of executable rules. Many studies showed that making explicit rules based on which rewards and sanctions are provided increases the motivation of an agent to perform certain actions (tasks) [18]. The motivation of agents to perform certain tasks is important to ensure the satisfaction of both individual and organizational goals related (directly or indirectly) to these tasks. Therefore, the motivational aspect of the agent behavior should be explicitly represented in the models of organizational agents.

2.2 Modeling the motivation of an agent

The topic of motivation in work organizations has received much attention in social science. In [26] the motivation is defined as a *process governing choice made by persons among alternative forms of voluntary activity*. There exist many different theories of motivation [15, 18, 19]. In this paper we adopt the Vroom's version of the expectancy theory [26] that has received a good empirical support.

According to this theory, when an individual evaluates alternative possibilities to act, s/he explicitly or implicitly makes estimations for the following factors: *expectancy*, *instrumentality*, and *valence* (see Fig.1).

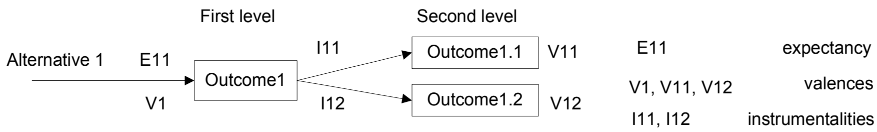


Fig. 1. An example of the motivation model by Vroom [26]

Expectancy refers to the individual's belief about the likelihood that a particular act will be followed by a particular outcome (called a first-level outcome). In the organizational context expectancy of an agent related to successful task execution is determined by the characteristics of the task and the agent, and by the organizational and environmental conditions. Tasks can be characterized along multiple dimensions: (a) complexity and predictability; (b) specialization; (c) formalization; (d) interrelation with other tasks; (e) collaboration required from agents. Usually agents that possess knowledge and the skills required for some task have a high level of expectancy of the successful task execution. Furthermore, agents with highly developed skills tend to assign a high value to expectancy associated with complex and not completely predictable tasks. On the opposite, inexperienced agents decrease their expectancy when dealing with complex tasks and especially with tasks with low predictability. For such agents the formalization of a task (e.g., by detailed procedure descriptions and guidelines) will increase their expectancy level. If a task requires from an agent a contribution from or collaboration with other agents, then the agent's belief about reliability and trustworthiness of these agents will play an important role in his/her expectancy estimation. Furthermore, other organizational factors, such as internal policies, rules and constraints (e.g., temporal, authority-related constraints) may influence expectancy of the task execution. Many modern organizations actively interact with the environment, which is often highly dynamic and unpredictable. The less certainty and knowledge about the environment an agent has (e.g., market fluctuations, resource availability), the less his/her expectancy level. As expectancy is defined as a subjective perception (or a belief) of an agent, the agent's personal traits also have influence on his/her expectancy.

Instrumentality is a belief concerning the likelihood of a first level outcome resulting into a particular second level outcome; its value varies between -1 and +1. A second level outcome represents a desired (or avoided) by an agent state of affairs that is reflected in an agent's goal(s) (e.g., bonus receipt, imposition of a sanction). Although the notion of instrumentality can be perceived as probability, in contrast to the latter instrumentality may take negative values, in case a second-level outcome does not follow a particular first-level outcome. If an organizational reward system is defined explicitly, instrumentality between a performance level and the corresponding material reward/sanction is perceived as high (>0.5) by agents.

Note that the agent's experience gained by the execution of tasks influences the values of expectancies and instrumentalities associated with these tasks. For example, if despite high performance the agent did not get the promised/expected (amount of) rewards, then his/her instrumentality between the level of efforts and the previously identified reward will decrease. Similarly, the agent adjusts the expectancy value associated with a task based on his/her actual amount of efforts put into the task execution.

Valence refers to the strength of the individual's desire for an outcome or state of affairs. While second level outcomes are directly related to the agent's goals, the valence values associated with these outcomes refer to priorities of these goals. Thus, similarly to goal priorities, the values of valence change over time (e.g., depending on the satisfaction of goals).

While in most cases the correspondences between actions of agents and rewards provided externally can be specified in a straightforward way, the prerequisites for obtaining intrinsic rewards are less obvious. One of the conditions for intrinsic rewards identified in literature [9] is that a task assigned to an agent should represent a reasonably complete peace of work, to the outcomes of which the agent could attribute his/her efforts. Some agents receive intrinsic rewards from the very process of task execution irrespectively of the execution results. While intrinsic rewards for other agents are contingent upon the execution outcomes. In the latter case if the actual task result equates to or exceeds the agent's expectation, then the agent receives an intrinsic reward. Furthermore, as follows from [9] the amount of intrinsic reward is dependent on the task complexity.

In the Vroom model *the force on an individual to perform an act is a monotonically increasing function of the algebraic sum of the products of the valences of all outcomes and the strength of his expectancies that the act will be followed by the attainment of these outcomes* [26]. Hence, the motivational force to perform act i can be calculated as:

$$F_i = f\left(\sum_{j=1}^{n} E_{ij} \times V_j\right), \quad V_j = \sum_{k=1}^{m} V_{jk} \times I_{jk} \qquad (1)$$

Here E_{ij} is the strength of the expectancy that act i will be followed by outcome j; V_j is valence of first-level outcome j; V_{jk} is valence of second-level outcome k that follows first-level outcome j; I_{jk} is perceived instrumentality of outcome j for the attainment of outcome k.

3 A Simulation Case Study

In this Section we shall investigate the behavior of the employees of a small firm that develops web graphics by request from external clients. Such an organization manages all its activities using a cohesive team structure. Teams have a flat power structure, which allows achieving high responsiveness to the environmental dynamics. Although the role of a leader (or manager) is identified, all important decisions are made with the assistance of all team members. The manager is responsible mostly for organizing tasks: e.g., searching for clients, distribution of orders, monitoring of the order execution. The firm consists of highly motivated members and has a very informal and open working style. The risky, environment-dependant nature of the firms of such type may cause financial insecurity and deficiency for their members. In the following the model used for the simulation is introduced. Due to the space limitation the model introduction will be mostly informal, providing the formalization only for the most essential parts. Subsequently, the simulation results are presented.

Modeling tasks and the environment

Tasks received by the firm are characterized by: (1) *name*; (2) *type*; (3) *required level(s) of development of skill(s)*; (4) *average / maximum duration*; (5) *extra time delay per unit of each skill development*; (6) *material reward*; (7) *intrinsic reward*; (8) *development level increment per skill*. For this case study the generalized PI "the development level" for each skill is used, which is an aggregated quantity (a real number in the range 0-5) reflecting the skill-related knowledge, experience, task execution context etc. The task average duration is the time that an agent that possesses the skills satisfying the task requirements will spend on the task execution. Agents with insufficient development levels of skills will require additional time for the execution. This is specified by the extra time delay characteristic per deficient unit of each required skill. The maximum task duration specifies the maximal time allowed for the task execution. For the successful performance of tasks agents are granted with material rewards; also the development level(s) of their skill(s) is (are) increased by the experience increment amount(s). Note that for simplicity the intrinsic rewards associated with the tasks in this case study are made independent of the specific characteristics of the agents who execute these tasks.

The task types used in the simulation are specified in Table 1. When detailed data about the task execution are available, more precise dependencies between task durations, extra delays and the skill development levels and traits can be established.

Table 1. The characteristics of the task types A1/A2 (create a simple/complex web illustration) and B1/B2 (create a simple/complex Flash animation) used in the simulation

Type	A1	A2	B1	B2
Required skill(s)	S1: 2	S1: 4	S2: 1	S2: 4
Average (max) duration (hours)	14 (18)	30 (38)	12 (15)	50 (60)
Extra time delay per skill (hours)	S1: 2	S1:4	S2: 3	S2: 8
Material reward	10	20	7	25
Intrinsic reward	1	3	1	4
Development increment	S1:0.1	S1:0.2	S2: 0.08	S2: 0.2

In the simulation we suppose that tasks arrive in accordance with a nonhomogeneous Poisson process $\{N(t), t \geq 0\}$ with a bounded intensity function $\lambda(t)$. Here $N(t)$ denotes the number of events that occur by time t and the quantity $\lambda(t)$ indicates how likely it is that an event will occur around the time t. We use the thining or random sampling approach [22], which assumes that $\lambda(t) \leq \lambda$ for all $t \leq T$, where T is the simulation time (2000 working hours (1 year) for this case study). Furthermore, for $T \leq 1000$: $\lambda_{A1}=\lambda_{A2}=\lambda_{B1}=\lambda_{B2}=0.05$ and for $T > 1000$: $\lambda_{A1}=\lambda_{A2}=2 * 10^{-5}$; $\lambda_{B1}=\lambda_{B2}=0.05$.

Organization modeling

The firm has two high level long-term goals with the same priority: "it is required to maintain a high profit level" and "it is required to maintain a high level of satisfaction of the employees". These goals are imposed on the organizational structure that comprises the role Manager and the generalized role Task Performer. The latter is instantiated into specific roles-instances associated with the tasks received by the firm. An instantiated role is assigned to one of the agents representing the employees using the following policy: Agents that can be potentially allocated to a role should be *qualified* for this role. An agent is qualified for a role under two conditions: (1) the agent is not involved into the execution of any other tasks; (2) agent possesses the skills required for the task associated with the role; and the level of development of these skills will allow to the agent to finish the task before the task deadline (i.e., maximum duration). To formalize these conditions, for each task and agent characteristic a predicate is introduced. Some of these predicates are given in Table 2. To express the temporal aspects of the agent qualification rule the language TTL is used [23]. TTL specifies the dynamics of a system by a trace, i.e. a temporally ordered sequence of states. Each state corresponds to a particular time point and is characterized by a set of state properties that hold in this state. State properties are defined as formulae in a sorted predicate logic using state ontologies. A state ontology defines a set of sorts or types (e.g., TASK, AGENT), sorted constants, functions and predicates (see Table 2). States are related to state properties via the satisfaction relation |=: state(γ, t) |= p, which denotes that state property p holds in trace γ at time t. Dynamic properties are specified in TTL by relations between state properties.

Table 2. Predicates for the formalization of agent-based models

Predicate	Description
task_arrived, task_started, task_finished: TASK	Specifies the arrival, start and finish of a task
role_for_task: ROLE x TASK	Identifies a role for a task
agent_allocated: AGENT x ROLE	Specifies an agent allocated to a role
agent_qualified_for: AGENT x ROLE	Specifies an agent qualified for a role
agent_requested: AGENT x ROLE	Identifies an agent that requested a role

The agent qualification rule is formally expressed in TTL as follows:

∀γ ∀t:TIME ∀a1:TASK ∀ag:AGENT ∀r1:ROLE ∀tp1: TASK_TYPE
state(γ, t)|= [task_arrived(a1) ∧ role_for_task(r1, a1) ∧ task_type(a1, tp1) ∧
¬∃r2:ROLE r2≠r1 ∧ agent_allocated(ag, r2) ∧ sum([sk:SKILL], ∃VALUE:n, m, k case(state(γ, t)|= task_requires_skill(a1, sk, n) ∧ agent_possesses_skill(ag, sk, m) ∧ m≥0.5∧ task_extra_delay(tp1, sk, k), k • (n-m), 0)) < (task_max_duration(tp1) - task_average_duration(tp1))
⇒ ∀t1:TIME t1>t state(γ, t1)|= agent_qualified_for(ag, r1)

Here in sum([summation_variables], case(logical_formula, value1, 0)) logical_formula is evaluated for every combination of values from the domains of each from the summation_variables; and for every evaluation when logical_formula is true, value1 is added to the resulting value of the sum function.

Further, since the firm recognizes the importance of wishes of its employees, a role can be only allocated, when a qualified agent has voluntarily requested the role. Furthermore, the firm established the rule that in case several qualified agents requested a role, then the agent with the most distant (i.e., the earliest) previous allocation time among these agents will be allocated to the role. This rule is also formalized using TTL:

∀γ ∀t, t1:TIME ∀ag:AGENT ∀r1:ROLE ∀a1:TASK
state(γ, t) |= [agent_qualified_for(ag, r1) ∧ agent_requested(ag, r1) ∧ role_for_task(r1, a1) ∧ latest_allocation(ag, t1) ∧ ∀ag1:AGENT ∀t2:TIME ag1≠ag ∧ agent_requested(ag1, r1) ∧ latest_allocation(ag1, t2) ∧ t1 < t2
⇒ agent_allocated(ag, r1) ∧ task_started(a1)]

Here latest_allocation(ag1, t1) is a short notation for:
∃t1:TIME ∃a2: TASK ∃r2: ROLE state(γ, t1)|= task_finished(a2) ∧ role_for_task(r2, a2) ∧ agent_allocated(ag1, r2) ∧ ∀t2:TIME t2> t1 ∀r3:ROLE state(γ, t2)|= ¬agent_allocated(ag1, r3)

For the successful execution of tasks the agents are provided with material rewards on the following basis: 50% of the reward is given to the agent who performed the task and the rest is divided equally among all other employees.

Modeling agents

The firm consists of three members and the manager modeled as agents. As in the most firms of such type, the employees are intrinsically motivated by their work, and pursuit high performance standards. For each agent two high level long-term hard goals are defined that also comply with the organizational goals: g1: it is required to maintain the level of income not less than 50; g2: it is required to maintain the level of intrinsic satisfaction not less than 5. It is assumed that the goal g1 when unsatisfied has higher priority than the goal g2. When g1 is satisfied, g2 becomes more important.

Two agents ag1 and ag2 possess the skill S1 to perform purely graphical work: agent_possesses_skill(ag1, S1, 4) and agent_possesses_skill(ag2, S1, 3). Here the third argument denotes the level of the skill development. The agent ag3 has the skill S2 to make Flash animations: agent_possesses_skill(ag3, S2, 4). Furthermore, ag1 has the general knowledge related to S2 (agent_possesses_skill(ag1, S2, 0.1)), which however is insufficient for the performance of tasks that require S2. By mutual consent of the firm and ag1 the development goal for ag1 without a strict deadline has been set: it is desired to achieve the level of development of S2 ≥ 0.5. When ag1 decides to gain the minimum level of the skill S2 development that is necessary for the task execution (0.5), s/he will be given one week for the training, during which no other tasks will be assigned to him/her. The motivation of the agents to attain their goals is represented by the motivation models, two examples of which for ag1 are given in Fig. 2.

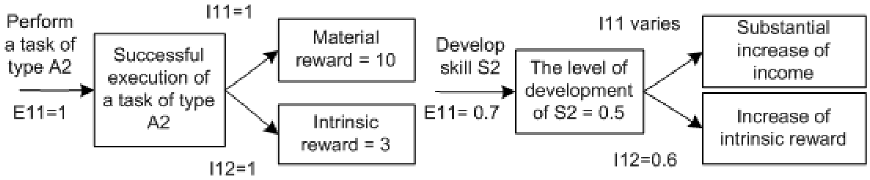


Fig. 2. The examples of two motivation models for the agent ag1 used in the case study

The parameters of the motivation models are defined as follows: Expectancy of an agent ag for the successful execution of a task tk is defined as a weighed average of the quotients $pos(sk_i)/req(sk_i)$ for each skill sk_i required for tk; here $pos(sk_i)$ is the development level of the skill sk_i possessed by ag and $req(sk_i)$ is the level required by tk. Instrumentality for each second level outcome associated with the successful execution of a task is equal to 1 for every agent qualified for this task. This is because the reward system is defined explicitly and the qualified agents have a clear estimation of the intrinsic reward associated with the task. The instrumentality value of ag1 for the skill S2 development is reevaluated in the end of each month and is equal to 1, when $n/m > 50$, and is equal to $n/(m*50)$ otherwise; here n is the amount of the material rewards provided by the tasks of types B1 and B2 received by the firm up to the current time point, and m is the amount of months of the simulation time (the initial instrumentality value is 0.35). The valence values of second level outcomes change over time. In particular, when the goal g1 of an agent ag is not satisfied, then the valence values of ag for all outcomes related to material rewards will become 1, and the valence values of outcomes related to intrinsic rewards will become 0.5. When g1 is satisfied, then the valence values for material outcomes will decrease to 0.5, and for intrinsic outcomes will increase to 1. An agent generates a request to perform an action specified in a motivation model (e.g., request for a role), when the motivational force associated with this action calculated using the formula (1) is greater than 0.5.

The initial income value is 20 for all agents, and the initial intrinsic satisfaction level is 3. Each agent consumes 0.05 units of the received material rewards per day and the amount of the received intrinsic rewards decreases by 0.03 each day.

The simulation is performed using the dedicated tool [2]. Fig. 3a shows how the motivational force of ag1 to attain the skill S2 changes over time. After the time point 1000, when the amount of tasks of type A diminishes significantly, the force transgresses the threshold 0.5, and ag1 begins the attainment of S2. After some time ag1 possesses the skills required to perform the tasks of both types A and B and both his/her goals g1 and g2 become satisfied (see Fig. 3b).

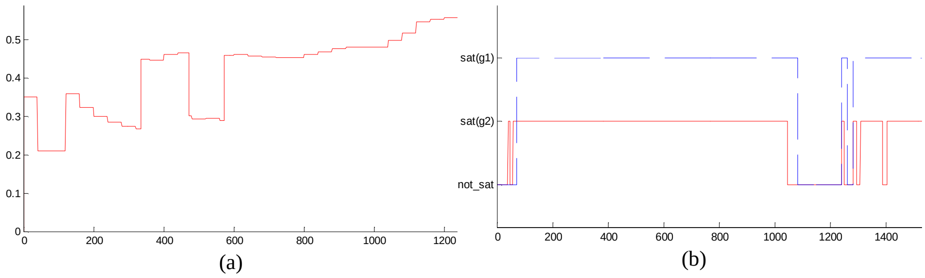


Fig. 3. (a) The change of the motivation force (the vertical axis) of agent ag1 for the attainment of skill S2 over time. (b) The change of the satisfaction of the goals of agent ag1 over time.

4 Conclusion

The paper proposes a formal approach for modeling of agents situated in an (formal) organization that accentuates the intentional and motivational aspects of agent behavior. The proposed quantitative motivation model of an agent based on the expectancy theory allows estimating the agent's motivational force to attain certain (organizational or individual) goals. Since the goal expressions are based on performance measurements, using the proposed approach it is possible to analyze how different organizational factors that affect the parameters of the motivation model influence the organizational or agent performance. An example of such analysis is demonstrated by a simulation case study in this paper.

Based on a large corpus of empirical social studies a great number of dependencies between organizational and environmental factors and the agent's motivation have been identified. In general, to create a feasible and valid model for a complex organization, a large number of variables and functions representing these factors and dependencies should be specified. This causes such undesirable properties of a model as a high complexity and the loss of tractability [7]. Therefore, it is recommended that an organization analyst depending on the organizational type and the purpose of analysis should choose only the most relevant organizational and environmental factors that have a direct impact on the agent behavior in the considered organizational setting. Such a choice may be based on the results of empirical studies for organizations of the considered type.

In the future research the behavior of various types of agents situated in organizations of different types (e.g., mechanistic, organic [17]) will be investigated.

References

1. Bertels, K., Boman, M.: Agent-Based Social Simulation in Markets. Electronic Commerce, 1(1-2) (2001) 149 – 158
2. Bosse, T., Jonker, C.M., Meij, L. van der, Treur, J.: A Language and Environment for Analysis of Dynamics by SimulaTiOn. International Journal of Artificial Intelligence Tools (to appear, 2007)
3. Carley, K.M.: A comparison of artificial and human organizations. Journal of Economic Behavior & Organization, 31(2) (1996) 175-191
4. Coddington, M. and Luck, M.: A Motivation Based Planning and Execution Framework, International Journal on Artificial Intelligence Tools, 13(1) (2004) 5-25
5. Davidsson, P., Henesey, L., Ramstedt, L., Tornquist, J., Wernstedt, F.: An Analysis of Agent-Based Approaches to Transport Logistics. Transportation Research Part C: Emerging Technologies, Vol. 13(4) (2005) 255-271
6. Decker, K.: TAEMS: A Framework for Environment Centered Analysis & Design of Coordination Mechanisms. Foundations of Distributed Artificial Intelligence, Chapter 16, O'Hare, G. and Jennings, N. (eds.), Wiley Inter-Science (1996) 429-448
7. Dooley, K.: Simulation research methods. In: Baum, J. (ed.): Companion to Organizations. Blackwell, London (2002) 829-848
8. Forrester, J. W.: Industrial dynamics, Waltham, MA: Pegasus Communications (1961)
9. Galbraith, J.R.: Designing organizations. Jossey-Bass, San Francisco California (1995)
10. Hackman, J.R.: Work redesign and motivation. Professional Psychology, 11 (1980) 445-455.
11. Horling, B. and Lesser, V. A Survey of multi-agent organizational paradigms. The Knowledge Engineering Review, Vol. 19(4) (2005) 281-316
12. Jonker C.M., Sharpanskykh, A., Treur, J. and Yolum, P., A Framework for Formal Modeling and Analysis of Organizations, Applied Intelligence (in press).
13. Katz, D and Kahn, R.: The social psychology of organizations. Wiley, New York (1966)
14. Kunz, J.C., Levitt, R.E., and Jin Y. The Virtual Design Team: A computational simulation model of project organizations. Communications of the Association for Computing Machinery 41(11) (1999) 84-92.
15. Lawler, E.E.: Motivation in Work Organisations, Cole Publishing (1973)
16. Manzano, M.: Extensions of First Order Logic. Cambridge University Press (1996)
17. March, J.G. and Simon, H.A.: Organizations. Wiley, New York (1958)
18. Pinder, C. C. (1998). Work motivation in organizational behavior. Upper Saddle River, NJ: Prentice-Hall.
19. Porter, L.W., Bigley, G.A., Steers, R.M. (eds.): Motivation and Work Behavior, 7th edition. New York: McGraw-Hill (2003)
20. Popova, V., Sharpanskykh, A.: Formal Analysis of Executions of Organizational Scenarios Based on Process-Oriented Models. To be published in Proceedings of ECMS'07 (2007)
21. Popova, V., Sharpanskykh, A.: Formal Modeling of Goals in Agent Organizations. In Proceedings of AOMS workshop at IJCAI'07 (2007) 74-80
22. Ross, S.: Simulation. 2nd edn. Harcourt Academic Press, London Boston New York (1998)
23. Sharpanskykh, A., Treur, J.: Verifying Interlevel Relations within Multi-Agent Systems. In Proceedings of the 17^{th} European Conference on Artificial Intelligence, IOS Press (2006).
24. Sharpanskykh, A.: Authority and its Implementation in Enterprise Information Systems. In Proceeding of the 1st Int. Workshop on Management of Enterprise Information Systems, MEIS 2007, INSTICC Press (2007)
25. Tyler G.P., Newcombe P.A.: Relationship Between Work Performance and Personality Traits in Hong Kong Organizational Settings. International Journal of Selection and Assessment 14 (1) (2006) 37–50
26. Vroom, V.H.: Work and motivation. Wiley, New York (1964)

Consumers' Behaviour Session

Recycling or product-life extension? An evolutionary modelling

Eric Brouillat
Groupe de Recherche en Économie Théorique et Appliquée (GREThA) UMR CNRS 5113, Montesquieu - Bordeaux IV University, Avenue Léon Duguit, 33608 Pessac Cedex, France
eric.brouillat@u-bordeaux4.fr

Abstract. This paper presents a model-based analysis of the development and diffusion of so-called 'green' products, i.e. those which are easily recyclable and which have a long lifetime. We address this problem by developing an simple evolutionary simulation model to describe the behaviour of business firms as interacting with that of consumers and recyclers. The simulation results show that investing both in product recyclability and product-life extension can be effective. Adopting such a strategy means that the firm may then market green products, which may provide it with a competitive advantage. From an environmental point of view, the diffusion of green products will reduce both waste flows in the economy and pressure on virgin resources. However, evolution of their recyclability and lifetime can mean slowing down the flow of recycled materials and may lead to even greater quantities of unrecycled waste. Our results suggest introducing regulation policies aiming at encouraging firms to invest in developing green products and getting consumers to buy them.

Key words: Recycling – Product-life extension – Environmental innovation - evolutionary modelling

JEL Classification: O33, D21, Q53

Introduction

When trying to deal with the continuous increase in waste volumes, one solution is to close the "material loop". A *linear industrial economy* is an economy where no loop exists (figure 1 – based on Stahel, Reday, 1976). Such a configuration leads to an accumulation of end of life products and an exhaustion of resources. One idea, therefore, is to build one or several loops making it possible to reintroduce into the production–use process, those products which have arrived back in post-consumption phase. It is thus necessary to distinguish two types of loops: short loops and long loops (figure 2 – based on Stahel, 1983).
Short loops refer to product-lifetime extension by means of re-using products as well as by maintenance, repair, reconditioning and technological upgrading. This makes it possible to avoid the waste phase over several periods for any given product.

Extending the use period of products will thus reduce the flow of materials used during the production-use process (Stahel, 1994).

Long loops involve material recycling. We refer here to post-use or post-consumption recycling. This means collecting waste, sorting it to recover materials which may be recycled, and then transforming these materials (physically and/or chemically) to obtain new raw materials. Consequently, recycling does not avoid a waste stage. So recycling, contrary to lifespan extension, does not reduce material flows and energy flows through the economy (Bellmann, Khare, 1999).

Fig. 1. The linear structure of industrial economy

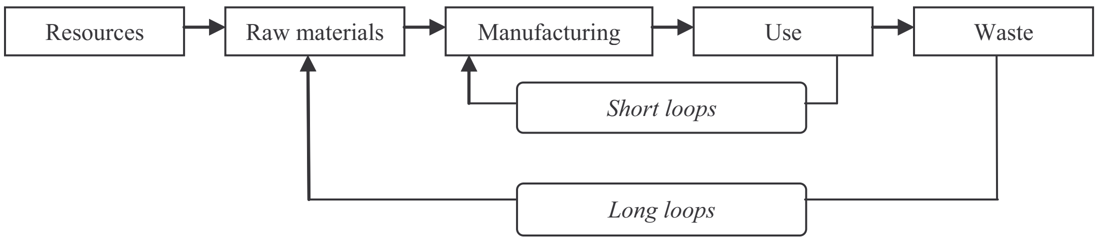


Fig. 2. Closing the material loop

From a strictly environmental point of view therefore, the problem of waste growth could be solved by implementing both short and long loops. But short loops are much more complex to implement than long ones. They imply significant changes in product design and the production process which requires new knowledge and competences. Inversely, long loops have the short-term advantage that they preserve the existing economic structures and are therefore easy to implement. Besides, product-lifetime extension can lead to slowing down product replacement speed by consumers which will mean lowering sales over that period. For these reasons, firms in general are usually reluctant to adopt a lifetime extension strategy. Today, the tendency is rather to see a fall in product-lifetime and, in order to attenuate the ecological impact of their end of life products, firms tend mainly to concentrate more on recyclability.

Furthermore, product life extension can equally slow down technical progress. In fact, while product-lifespan may be longer, new technology incorporated into these products will be distributed and used over longer periods and so will be replaced later.

Our study, however, will focus solely on the relations between, on the one hand, waste production and recovering, and, on the other hand, environmental product innovation. Consequently, we will not be taking into account the problem of any technology built into the product and this aspect will not, therefore, be explicitly defined in our analysis.

We will be assuming the relevance of long and short loops. Thus, our analysis is centred on an environmentalist perspective. We consider that there are two strategies

that already stand out and which need to be combined i.e. recycling and product-life extension. The question is then to study firms' strategic choices regarding the characteristics of their products while these firms are interacting with consumers and recyclers. We are investigating firm's economic incentives in extending product life and marketing recyclable products. We aim to identify the conditions favouring diffusion of products with low environmental impact, i.e. easily recyclable products with a long lifetime. We call these products *green products*. Furthermore, we will be studying the impact of such diffusion on waste management and ecological variables. To this end, we present in this paper an evolutionary simulation model that studies the issue of emergence and diffusion of green products in order to analyse co-evolution of industrial and environmental dynamics in the field of waste management. Our model describes the behaviour of firms as well as that of consumers and recyclers.
The paper is presented as follows. In Section 2 a description of the simulation model is presented. In Section 3 results of model experiments are discussed. Section 4 concludes.

The model:

Before describing the model, a warning note is required. We do not aim at building a realistic model, but one that can provide us with insights. Real world markets have so many peculiar aspects that, even if we were able to build a good approximation of one of them, we would face the same problems of generalizations and understanding that we have looking at real data. The model presented here is a very basic model and many aspects of reality are intentionally neglected.
We take into account three categories of actors: firms (*i*) who all market a single finished product, end consumers buying those products, and recyclers recovering and recycling those end of life products used by consumers. To simplify, we are supposing that there is one single recycler in the economy. This agent will represent all the downstream actors in the supply chain. These three categories of agents will be interacting at the various stages of the model.

Product's characteristics

The product *i* is characterized by its environmental performances. The product's environmental performance will develop and change depending on the firms R&D strategies (Saint Jean, 2005). Consequently, the combination of the product's performance will reflect whichever strategy is being adopted. Performance will be divided into recyclability (*R*) and life time (*LT*) (Boons, 2002).
R is the recyclability degree of the product. This reflects the firm's ability to design a product that can be easily recycled. We are presuming that there is an *Rmax1* recyclability threshold characterizing the design change needed to increase product recyclability. *Rmax1* represents the maximum recyclability reachable through incremental innovations. Crossing this threshold will require completely reviewing product design in order to take into account its end of life from the design phase.

LT is the lifespan of the product. It reflects any firm's ability to extend the lifetime of its product, in particular by marketing a reusable product. We are also presuming that there is an *LTmax1* lifetime threshold. This threshold characterizes the adoption of a new product design needed to extend the product's lifetime. The implementation of product life extension equally supposes the design of a radically new product: a more robust and more positively developing product. A product will need to be designed to be used over long time periods. This new product design must be backed up with new services: return, maintenance and repair of products and technical and technological update. Adopting policy for a product life extension strategy requires then radical innovations both in product design and in production structure. *LT* and *R* are limited by fixed upper limits *LTmax2* and *Rmax2*.

Producers' R&D investment

Each firm invests for each period in R&D a fixed proportion (μ) of its profits (Π) for the previous period:

$$RD_{i,t} = \mu_i . \Pi_{i,t-1} \tag{1}$$

with $0 \leq \mu_i \leq 1$.

μ is a firm specific parameter. The part of the profits not invested in R&D is supposed to be used to remunerate shareholders (this will not be investigated in this model).

R&D investment seeks to improve the environmental performance of any product. *RD* is divided into expenditure aiming at increasing the product's life time (RD^{LT}) and its recyclability (RD^{R}):

$$RD_{i,t}^{LT} = \delta_i . RD_{i,t} \tag{2}$$

$$RD_{i,t}^{R} = (1-\delta_i) . RD_{i,t} \tag{3}$$

with $0 \leq \delta_i \leq 1$.

In this model firms are seen as only investing in environmental R&D. δ is a firm specific parameter reflecting its distribution choice of R&D expenditure. Consequently, this parameter reflects the firm's innovation strategy regarding its product's environmental characteristics.

The innovation process

R&D expenditure determines the probabilities of access to new values within the range of product characteristics. Access probabilities to a new recyclability performances ($Prob^{R}$) are logistic functions of their R&D investment allocated to *R*. In a similar way, access probabilities to new lifetime performances ($Prob^{LT}$) are logistic functions of their R&D investment allocated to *LT*. These two functions are as follows:

$$\mathrm{Pr}ob_{i,t}^{R} = \frac{\pi 1}{\pi 2 + \pi 3 \exp\left(-\pi 4 RD_{i,t}^{R}\right)} \tag{4}$$

$$\mathrm{Pr}\ ob_{i,t}^{LT} = \frac{\pi_1}{\pi_2 + \pi_3 \exp^{\left(-\pi_4 RD_{i,t}^{LT}\right)}} \tag{5}$$

with π_1, π_2 and π_3 parameters limiting the logistic function and π_4 a parameter determining the speed to which the upper limit of access probabilities may be reached. To simplify, we are supposing that these parameters are the same for both kinds of access probabilities. The logistic function implies increasing and then decreasing R&D outputs (Silverberg, Verspagen, 1995).

The environmental potential of the product is a two dimensional space defined in the strictly positive R^2 mathematical space. Parameters *Rmax1*, *Rmax2*, *LTmax1* and *Ltmax2* represent the borders of the different potentials on the *R* and *LT* dimensions corresponding to the different product designs. The environmental innovation process means increasing the value of at least one of these two characteristics according to functions specified hereafter.

If the innovation draw for *R* is a success, the improvement of the value of this characteristic is given by:

$$\Delta R_{i,t} = \alpha_R . \left(RD_{i,t}^{R}\right)^{\gamma_1} \left(\mathrm{Rmax} - R_{i,t-1}\right)^{\gamma_2} \tag{6}$$

with $Rmax = Rmax1$ if $R_{i,t-1} \leq Rmax1$ and $Rmax = Rmax2$ if $R_{i,t-1} > Rmax1$.

The equation (6) implies that the value of the increase in *R* depends on R&D expenditure allocated to *R* and the distance of the achieved design to the frontier (Malerba & al, 1999). The parameter α_R is a scale parameter. Parameters γ_1 and γ_2 are selected so that their sum is equal to 1. These two parameters respectively reflect the impact of the R&D expenditure and the impact of the exhausting of innovation opportunities on the extent of improvement to the considered characteristic.

In the same way, if the innovation draw for *LT* is a success, the improvement of the value of this characteristic is given by:

$$\Delta LT_{i,t} = \alpha_{LT} . \left(RD_{i,t}^{LT}\right)^{\gamma_1} \left(\mathrm{LTmax} - LT_{i,t-1}\right)^{\gamma_2} \tag{7}$$

with $LTmax = LTmax1$ if $LT_{i,t-1} \leq LTmax1$ and $LTmax = LTmax2$ if $LT_{i,t-1} > LTmax1$.

α_{LT} is a scale parameter and we assume by simplification that the values of parameters γ_1 and γ_2 are the same as those used for *R*.

On the other hand, if the innovation draw is a failure, performances remain unchanged.

Supply - demand interactions

In this very basic model, consumer's behaviour is very simplified. In fact, no interaction between consumers is assumed and consequently, the social aspect of consumer's decision making is intentionally neglected.

Each consumer uses one single product at the same time and renews its purchase only when this product is at the end of its lifetime, i.e. when its lifespan *LT* expires. The consumer will then choose a new product according to products' visibility (*V*):

$$V_{i,t} = \left(LT_{i,t}\right)^{\beta_1}.\left(R_{i,t}\right)^{\beta_2}.\left(\frac{1}{p_{i,t}}\right)^{\beta_3}.\left(S_{i,t-1}\right)^{\beta_4} \tag{8}$$

The visibility of the product is a specification of the total performance for the product. β_1, β_2, β_3 and β_4 are parameters whose values are selected so that the sum of the four is equal to 1. This function implies that visibility increases with the environmental performance of the product and decreases with its selling price. Parameters β_1, β_2 and β_3 represent sensitivity of the visibility to lifespan, recyclability and price. *S* is the market share of the firm and the parameter β_4 reflects the bandwagon effect. The parameters β_1, β_2, β_3 and β_4 represent the consumer's preferences with respect to the product's characteristics. We suppose that two types of consumers exist: green and non-green consumers. We assume that products' recyclability and lifetime are believed to be considered by non-green consumers as secondary purchasing criteria ($\beta_3 \geq \beta_1 + \beta_2$). Inversely, they are the main criteria for green consumers ($\beta_3 < \beta_1 + \beta_2$). We assume that the consumer's behaviour is a bounded rational one. The consumer cannot take an optimal decision consisting in choosing the product with the best visibility. The rule to choose a product is random with probabilities proportional to products' visibility.

Once the consumer has bought the desired product, they become a customer of the selected firm. From the firm's point of view, if the product of one of its customers reaches the end of its lifetime, a loss is recorded, and so its stock of customers will drop by one unit. On the other hand, each time its product is bought, the firm records an additional sale and its stock of customers will increase by one unit. At the end of the purchase cycle, each firm counts the number of sales (*Q*) and the number of lost users and consequently determines the current number of users of his product, i.e. its stock of customers (*U*). The market share of the firm (*S*) is given by this stock of customers:

$$S_{i,t} = \frac{U_{i,t}}{\sum_{i=1}^{n} U_{i,t}} \tag{9}$$

The manufacture of products requires a certain quantity of inputs. We suppose that there are two categories of perfectly substitutable inputs: recycled inputs and virgin inputs[1]. To produce the quantity *Q*, the firm needs a quantity $\omega.Q$ of recycled inputs

[1] We are actually supposing that virgin inputs can be used as a substitute for recycled inputs.

and a quantity *(1 - ω).Q* of virgin inputs. Parameter *ω* represents the share of recycled inputs constituting the product ($0 \leq \omega \leq 1$). Recycled inputs are provided by the recycler, the virgin inputs by suppliers external to the model. The recycler offers for each period a quantity Z^s of recycled inputs. If the recycler offers a sufficient quantity of recycled inputs to face the demand, each firm buys the quantities of recycled and virgin inputs desired. If the recycler does not offer a sufficient quantity of recycled inputs to satisfy the whole of the demand, he will sell off the whole of his stock, the distribution between firms being carried out according to their requirements in recycled inputs. Then, each firm facing a recycled inputs shortage will buy in an additional quantity of virgin materials. Consequently, we assume that there is no constraint in the quantity of virgin inputs.

Selling price

The selling price of the product is defined as a minimum price (*pmin*) to which the firm adds a fixed mark-up (λ):

$$p_{i,t} = p\min_{i,t}.(1+\lambda) \tag{10}$$

The minimum price is the product unit cost of production from the previous period. Two cases arise.
If $LT_{i,t} \leq LTmax1$, the firm markets a product with a short lifetime design. In this case *pmin* is given by:

$$p\min_{i,t} = CM_{i,t-1} = \frac{\left(pr.qr_{i,t-1}^{d} + pv.qv_{i,t-1}^{d}\right)}{Q_{i,t-1}} \tag{11}$$

CM is the average production cost of a product. *pr* is the price of recycled inputs and *pv* the price of virgin inputs. *pv* is exogenic, fixed by virgin inputs suppliers. By simplification, we assume that *pr* is also fixed.
If $LT_{i,t} > LTmax1$: the product is designed to have a long lifespan. In this case *pmin* is equals to:

$$p\min_{i,t} = CM_{i,t-1}.(1 + AddC_{i,t}) \tag{11'}$$

AddC is an extra cost. This reflects the fact that products with a long lifetime design are more expensive to produce (Janssen, Jager, 2002). The additional service offered by the firm (return, maintenance, repair of products, technical and technological update) is an additional cost. This service is incorporated into the selling price. However, by training and through accumulated experience, the firm will be able to lower *AddC*. In fact, once the firm crosses the *LTmax1* threshold and markets a product with a long lifetime design, its R&D investment allocated to *LT* (RD^{LT}) will not only contribute to increasing the lifespan of that product, but will also lower the extra cost. For each period, there will be two innovation draws for *LT*: the first to know if *LT* has increased, the second to know if *AddC* has decreased. The first draw has been presented previously (*see* The innovation process). Regarding the second draw, if it is a success, the fall in the value of the extra cost will be given by:

$$\Delta AddC_{i,t} = \alpha_{AddC} . AddC_{i,t-1} \tag{12}$$

with α_{AddC} a parameter whose value is lower or equal to zero. This reflects the speed at which *AddC* will decrease. If the draw is a failure then *AddC* will remain the same. We can then fix the profits (Π) for each firm:

$$\Pi_{i,t} = p_{i,t}.Q_{i,t} - p\min_{i,t}.Q_{i,t} = \lambda.p\min_{i,t}.Q_{i,t} \tag{13}$$

Profits play the role of financial constraint by determining the budget allocated to R&D.

The recycler

The recycler collects the complete range of end of life products which he recycles and sells to the producers as recycled inputs. By "end of life products" we mean products whose lifetime (*LT*) is over. The recycler does not recycle every part of the end of life product, but only that part which can be recycled which, in turn, will depend on the recyclability (*R*) of that product. We are assuming that the part which cannot be recycled is incinerated or stocked in a waste disposal site. Thus, starting from a unit of product, the recycler manufactures and sells *R* units of recycled inputs. Consequently, the sum of the recyclabilities (*R*) of the various end of life products determines the recycled inputs supply created by the recycler over the period. Then, this supply created over the current period adds to the exiting stock of recycled inputs to give the total recycled inputs supply.

The model defined in this manner allows, by the means of simulations, to study the evolution of firms' performances and changing trends in waste production and recovery, such changes being guided in particular by the division of R&D expenditure between RD^R and RD^{LT} and thus by the combination of the product's characteristics.

Simulation results

Experimental set-up

We use the LSD 5.5 simulation platform[2] to compute and run the model. In the simulation experiments we formalise 10 firms, 1000 consumers and 1 recycler. In this model neither entries nor exits of firms and consumers are assumed. A methodology close to the Monte-Carlo method is adopted. We run 10 000 simulations where the number of periods is randomly chosen between 1 and 200. The results of the last period of each simulation are studied. The simulations are initialized with a randomly drawn vector of values for the main parameters of the model. As a result, we obtain a

[2] http://www.business.aau.dk/~mv/Lsd/lsd.html

set of 10 000 observations covering quite a diversified subset of the parameter space. To analyse the simulations results we use box plots, Wilcoxon-Mann-Whitney tests and regression trees[3].
In the following sections, we will present the simulation results. In the first section we will discuss the firm level results of simulations. Then, we will complete this analysis by the study of the market level results.

Firm level results

As far as economic performances of firms are concerned, figure 4 shows the regression tree of market share. A regression tree (Venable, Ripley, 1999) establishes a hierarchy between independent variables using their contributions to the overall fit (R^2) of the regression. It splits the set of observations in sub-classes characterized by their values in terms of their contribution to the overall fit and their predictions for the dependent variables. This value is validated against a fraction of the sample that is not used during the estimation. The tree gives a hierarchical sequence of conditions on the variables of the model: the higher the role of a condition in the classification of the observed case, the higher its status on the tree. For each condition, the left branch gives the cases for which the condition is true and the right branch gives the cases compatible with the complementary condition.
The figure 3 shows that the main determinant of the market share of a firm is the firm's innovation strategy regarding its product's environmental characteristics which is reflected by *Delta* (δ). Market share is the lowest when *Delta* is very weak (< 0.1418) and the higher *Delta* is, the greater the market share of the firm will be. When *Delta* is high (≥ 0.2359), the part of the profits invested in R&D (*Mu* (μ)) plays a central role. The greater this part is, the higher the market share will be. Eventually, figure 4 shows that the highest market share is obtained when the firm invests a major share of its profits in R&D and when this investment is mainly used to increase the lifetime of its product. In this case, consumer's preferences with respect to the product's lifetime (*Av_Beta1*)[4] play a positive role: when those preferences are high (≥ 0.38), the expected market share is greater.

[3] We use Eviews 5 and R to analyse the simulations results. Interested readers may obtain a full copy of the initialisation of the parameters and variables and the results of the Wilcoxon-Mann-Whitney tests by writing to the author.
[4] *Av_Beta1* reflects the average preference of the population of consumers with respect to product-lifetime. The higher *Av_Beta1* is, the higher this preference will be.

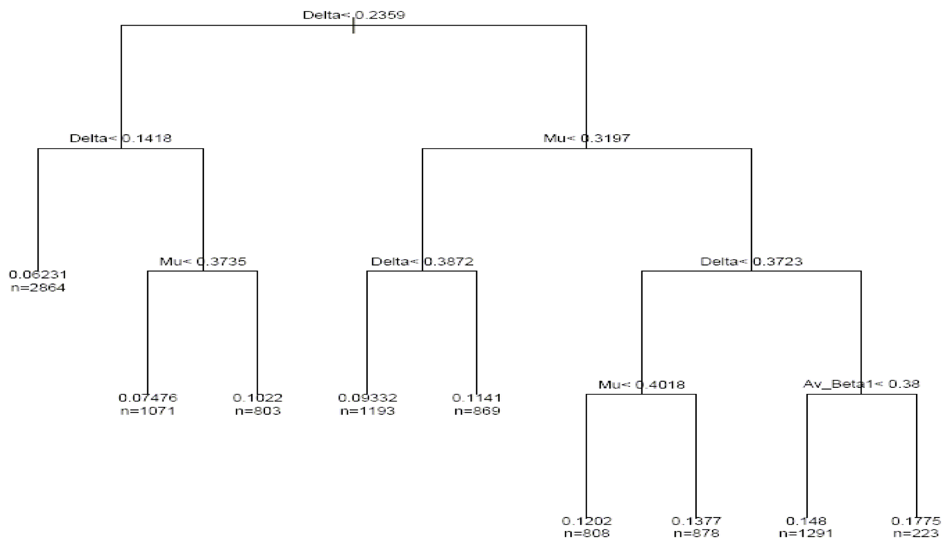


Fig. 3. Determinants of market share

Figure 4 shows that the main determinant of profit is *LTmax1* (*LifeTimeMax1*). A low value for *LTmax1* means that incremental improvements of the initial design can provide just a small rise in product-lifetime. Radical changes in the product design and the production process are needed to significantly extend the product's lifetime. Inversely, when *LTmax1* is high, product-life extension does not require such radical changes. Incremental innovations are enough to reach a high level in product-lifetime. Figure 4 shows that the highest profits are obtained when *LTmax1* is low (< 2.964). It is easy to explain this result by the fact that when a firm crosses this threshold, its price rises because of the cost of the additional service offered, and, consequently, its profits also rise[5]. In both main branches, *Delta* plays a crucial role. The greater *Delta* is, the higher profits will be.

We call "MIX firms" the firms whose *Delta* is higher than the second quartile of its distribution and "RECY firms" those firms whose *Delta* is lower.
As regards product designs adopted, figure 5 shows that environmental performance of MIX firms are better than those of RECY firms and consequently, MIX firms are the more inclined to market green products[6]. Besides, demand preferences do not play a crucial role in environmental performance of firms: the Wilcoxon-Mann-Whitney test shows no significant difference between green and non-green configurations.

[5] The fall in sales following the rise in price is always lower than this rise in price because $0 \leq \beta_3 \leq 1$. Consequently, the profits rise.

[6] Product-lifetime of MIX firms is much higher than RECY firms. Product recyclability of MIX firms is lower than RECY firms, but the performance of both categories of firm on this last dimension are close.

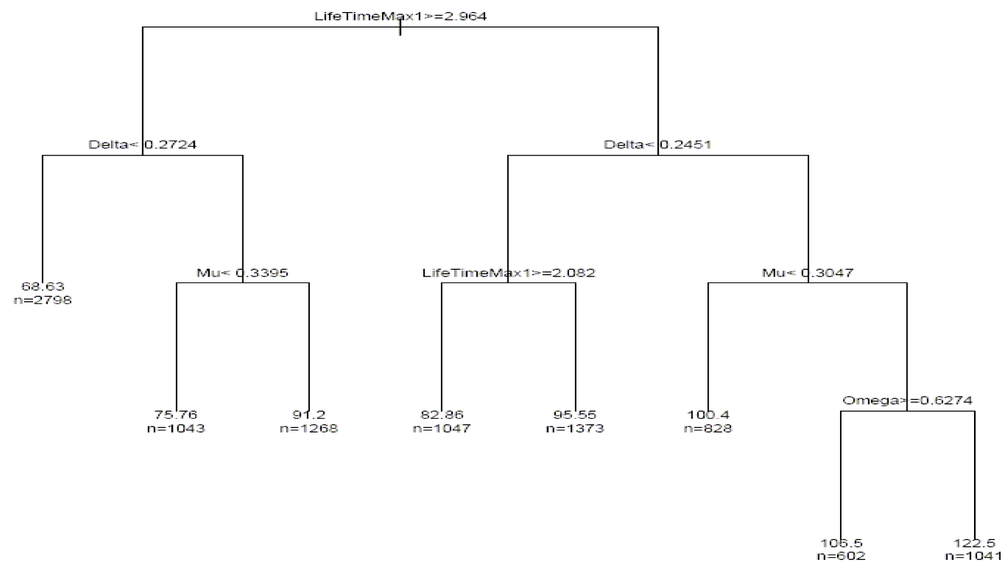


Fig. 4. Determinants of profits

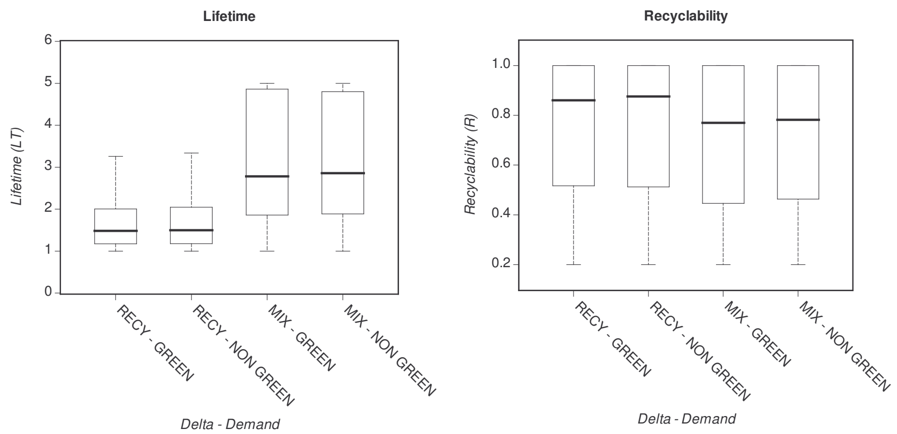


Fig. 5. Box plots of any product's environmental performances

Our analyse of firm level results shows that the innovation strategy of the firm regarding its product's environmental characteristics is crucial. Investing both in product life extension and in product recyclability (high value of *Delta*) means that the firm will come to be inclined to market green products and to obtain the best economic performances. This economic advantage will be made even stronger by a high investment of the firm's profits in R&D and the development of a green demand on the market.

Market level results

Four types of flows will be discussed here: the flows of products sold each period, the flows of recycled waste, the flows of unrecycled waste and the flows of virgin inputs.
Figure 6 shows that the main determinant of the total quantity of products sold over each period is *LTmax1* (*LifeTimeMax1*). The higher the threshold is, the lower the quantity will be. It is easy to explain this result by the fact that a high level for *LTmax1* involves a great average lifetime of products sold in the economy. From a strictly environmental point of view, this high value for *LTmax1* is a good point because it will lower flows of products, and consequently, it will bring about lower quantities of waste to manage. However, a high level for *LTmax1* tends to lower economic performances of firms, in particular their profits (*see* Firm level results).
The second main determinant of the flows of products sold each period is the innovation strategies of firms regarding their products' environmental characteristics. The tree shows that the greater the average value of the parameters *Delta* (*Av_Delta*) is, the lower the flows of products sold each period will be. This result is explained by the fact that a high level for *Av_Delta* involves a great average lifetime of products sold in the economy. Finally, the more firms diversify their R&D expenditures by investing both in product recyclability and product-life extension, the lower the total quantity of waste to be collected and recovered over the economy will be. Furthermore, the Wilcoxon-Mann-Whitney tests show that green demand leads to lower flows of products than non-green one.

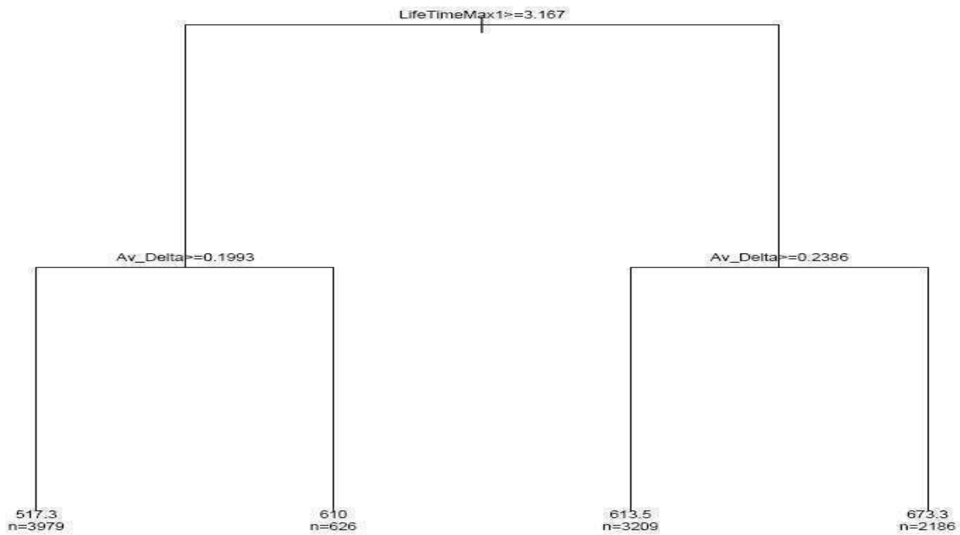


Fig. 6. Determinants of product quantities sold

As assumed previously, a part of the waste produced will be recycled. The other part will be incinerated or stocked in a waste disposal site. Two kinds of effect actually come into play here: a *volume effect* linked to the total quantity of products sold and a *substitution effect* linked to product recyclability. Regarding recycled waste flows, both volume and substitution effects are positive. In fact, the larger the quantity of products sold and the higher the recyclability of these products, the larger these flows will be. As regards unrecycled waste flows, volume effect is positive and the substitution effect negative: the larger the quantity of products sold over each period and the lower the recyclability of these products, the larger these flows will be.

The regression tree of the recycled waste flows (figure 7) shows that the main determinant of these flows is *Rmax1* (*RecyclabilityMax1*). A low value for *Rmax1* means that incremental improvements of the initial design can provide just a small rise in product recyclability. Radical changes in the product design (design for recycling) are needed to significantly extend any product's recyclability. Low value for this threshold tends to lower recycled waste flows. This observation is explained by the fact that when *Rmax1* is high, the average recyclability of the products sold is high because it is easier for firms to sharply increase their product recyclability (they just need to improve the initial design without radically changing it). A high average recyclability means large recycled waste flows (substitution effect) *ceteris paribus*. *LTmax1* (*LifeTimeMax1*) plays also a crucial role here. A high value for this threshold tends to lower recycled waste flows. It is easy to explain this result by the fact that a high level for *LTmax1* involves a low total quantity of products sold over each period, and consequently, it lowers the flows of waste to be treated (volume effect). The regression tree shows that high value of *Av_Delta* emphasizes this phenomenon[7].

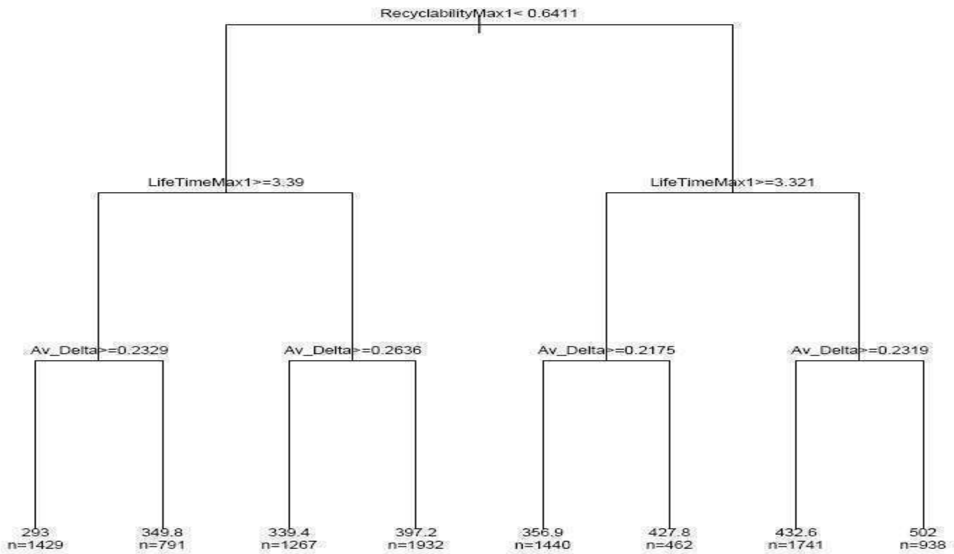


Fig. 7. Determinants of recycled waste flows

The regression tree of the unrecycled waste flows shows the same kind of configuration as the recycled waste flows tree. The main determinant of these flows is *Rmax1*, but here, low value for this threshold tends to increase unrecycled waste flows (substitution effect). The second main determinant is *LTmax1* too. When the level of this threshold is high, unrecycled waste flows tend to be low (volume effect). But this effect only comes into play when the substitution effect is low, i.e. when *Rmax1* is low (< 0.6824). Unlike the recycled waste flows case, *AvDelta* seems not to be crucial here. However, the Wilcoxon-Mann-Whitney tests show a significant difference between RECY and MIX supplies[8]. They show that the lowest flows of unrecycled waste are obtained with a RECY supply.

Finally, the *Rmax1* and *LTmax1* thresholds play a central role in waste recovery. A high *Rmax1* means a sharp substitution effect and a low *LTmax1*, a sharp volume effect. From an environmental point of view, a high value for *Rmax1* seems to be the best situation because it leads to low unrecycled and great recycled waste flows. In the same way, a high value for *LTmax1* seems to be the best situation because it leads to low unrecycled waste flows (in particular when *Rmax1* is low), but it could mean lower recycled waste flows. Better still, innovation strategies of firms impact waste recovering performances. By investing both in product recyclability and product-life

[7] A high level of *Av_Delta* involves a low total quantity of products sold over each period.

[8] We call "MIX supplies" the configurations whose *Av_Delta* is higher than the second quartile of its distribution and "RECY supplies" the cases whose *Av_Delta* is lower.

extension, firms are likely to sharply slow down recycled waste flows and increase flows of unrecycled waste because their recyclability performance tends to be lower than the one obtained with a bigger investment in RD^R.

The growth in recycled waste flow can lead to stockpiling. The production of recycled inputs by the recycler can exceed demand and the surplus will need to be stocked. This stock is greater with a RECY-type supply than with a MIX-type. This increasing stock will involve a social cost for the economy which can be easily borne given that at least a large part much of this stock will be used by other industries.
Innovation strategies of firms and consumer's preferences impact pressure on virgin resources. Figure 8 shows that this pressure is sharper when firms invest mainly in product recyclability (RECY supply). Rise in recyclability combined with an extension in the product's lifetime means using less virgin materials for each period. The lowest pressure is obtained when firms invest both in product recyclability and product-life extension and when demand is green[9].

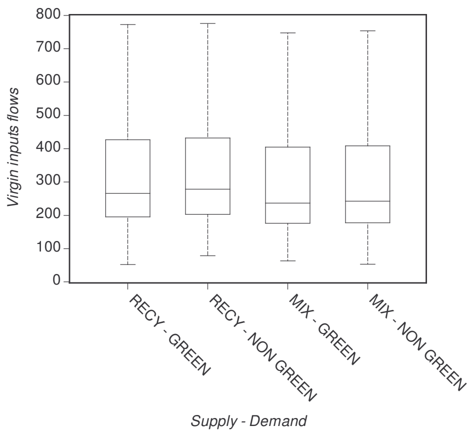


Fig. 8. Box plots of virgin inputs flows

Conclusions

The model presented in this paper provides a very simplified vision of the problem studied and, needless to say, the hypotheses being assumed here are fairly restricted. In particular, there are only two kinds of unadaptive R&D strategies being taken into account. In addition, in this study, consumers' preferences are assumed to be fixed. Clearly, in the real world, preferences may vary in time and firms' strategies are more complex and varied: they change to deal with fluctuations in the economic

[9] The Wilcoxon-Mann-Whitney test shows that the type of demand is not significant in the MIX supply case.

environment. Furthermore, in this model, firms do not face fixed adoption costs when they are adopting (Malerba & al, 1999) which means their profits are always positive. However, despite this simplification in the modelling, our simulations yield some interesting conclusions about the diffusion of green products and its consequences on the market and the environment.

First, the model's dynamics show that firms investing both in product life extension and in product recyclability will tend, more often than not, to market green products. This strategy is effective and actually means that the firm will have the opportunity to dominate the market. Better still, this economic advantage will be strengthened further by a major investment of the firm's profits in R&D and the development of a green demand on the market.

Green products mean reducing both waste flows in the economy and pressure on virgin resources. However, because recyclability and lifetime evolve together, this can slow down recycled input flows and can even cause more unrecycled waste quantities. In this experiment, performance thresholds *Rmax1* and *LTmax1* were seen to play an ambiguous role. When radical change is needed to increase environmental performance of products, high expected values of profits can be obtained, but at the same time, this can lower waste recycling.

Our results underline that local authorities should introduce regulations aiming at encouraging consumers to choose green products and getting firms to invest in development of this type of product.

This very simple model calls for more research into the modelling of the development and diffusion of green products. We now plan to improve the basic model by adding new elements in order to provide a more realistic vision of the behaviour of firms, consumers and recyclers. We look forward to being able in the future to insert environmental regulations, notably those based on extended producer responsibility (Lindhqvist, 2000), in order to be able to study their impact on the development and diffusion of green products.

References

Bellmann, K., Khare, A., 1999, Economic issues in recycling end-of-life vehicles, *Technovation*, 20: 677-690

Boons, F., 2002, Greening products: a framework for product chain management, *Journal of Cleaner Production*, 10: 495-505

Janssen, M., Jager, W., 2002, Stimulating diffusion of green products – Co-evolution between firms and consumers, *Journal of Evolutionary Economics*, 12: 283-306

Lindhqvist, T., 2000, *Extended Producer Responsibility in Cleaner Production : Policy Principle to Promote Environmental Improvements of Product Systems*, IIIEE, Lund University

Malerba, F., Nelson, R., Orsenigo, L., Winter, S., 1999, « History-friendly » models of industry evolution : the computer industry, *Industrial and Corporate Change*, 8(1): 3-40

Saint Jean, M., 2005, Polluting emissions standards and clean technology trajectories under competitive selection and supply chain pressure, *Working Papers of GRES*, 2005-16

Silverberg, G., Verspagen, B., 1995, "Evolutionary theorizing on economic growth", in K. Dopfer (ed), *The Evolutionary Principles of Economics*, Norwell, MA, Kluwer Academic Publishers

Smallwood, D.E., Conlisk, J., 1979, Product Quality in Markets where Consumers are Imperfectly Informed, *Quarterly Journal of Economics*, 93(1): 1-23

Stahel, W., 1983, *The Product Life Factor*, Orr, Ed. NARC, Texas

Stahel, W., 1994, "The utilization-focused service economy : ressource efficiency and product-life extension", in B.R. Allenby, D.J. Richards (eds), *Greening of industrial ecosystems*, National Academy of Engineering, Washington, DC

Stahel, W., Reday, G., 1976, Jobs *for tomorrow : the potential for substituting manpower for energy*, Report to the Commission of the EC, Vantage Press, New York

Venables, W., Ripley, B.D., 1999, *Modern Applied Statistics with S-PLUS*, Third edition, Springer, New York

Different Ways of Modelling Phone Adoption

Lynne Hamill

Centre for Research in Social Simulation, Department of Sociology, University of Surrey, Guildford, Surrey, GU2 7XH, United Kingdom
lynne.hamill@surrey.ac.uk

Abstract. Systems dynamics and agent-based models are used here to examine the spread of fixed line phones in Britain over 120 years. Both models approximately fit the data and it is shown that in this case the two approaches can be used to complement each other. The SD model is simpler and produces a better fit while its deterministic nature facilitates sensitivity analysis. The agent-based model provides greater explanatory power, which can in turn be used to fine tune the systems dynamics model. Together, they can be combined to tell a plausible story about the adoption of telephones.

Keywords. Systems Dynamics. Agent-based modelling. Technology adoption.

1. Introduction

This paper follows Scholl's recommendation [1] that system dynamics (SD) and agent-based (AB) models be compared on identical topics. The relative strengths of the two models are explored by looking at the adoption of fixed line phones in Britain from their introduction in 1880 to the end of the twentieth century. The aim of the paper is to explore methodological issues rather than to provide a robust explanation of the observed adoption pattern, which will be the subject of later work.

Section 2 provides the historical background to the adoption pattern that the models are to be tested against. The theory is set out in Section 3. Sections 4 and 5 describe the SD and AB models respectively. The models are compared in Section 6. Section 7 shows ways in which the two approaches can be used to complement each other and Section 8 concludes.

2. History

Britain's first public telephone exchange opened in London in August 1879 to serve eight subscribers. By the end of the year the number of subscribers in London totalled 200 and exchanges had been opened in seven other cities. The first phone directory was issued in 1880 and "contained details of over 250 subscribers" plus details of 16 provincial exchanges [2]. By 1882, there was one phone for every 3,000 people in London: by 1890, the ratio was up to one in about 800 but it did not reach one per 100 until 1905 [3]. However, these numbers included business as well as private subscribers and it is likely that the adoption rate for households was very much lower, even in London, and was lower still outside London. However, the data series for households only starts in 1964, by which time 21.6% of British households had fixed line phones. This percentage peaked at 95% in 1999, after which it started to drop as mobiles were substituted [4].

Thus phones took some 80 years to spread from virtually no households to around 20% and almost 120 years to reach 95%. Perry [3] suggests that this slow take-off was due to price and poor regulation of the nascent phone industry. Also, there was limited geographical coverage: London's first trunk line was not opened until 1884 and it was not possible for Londoners to call "the Midland and Northern Counties" until 1890 even though the first line to Paris opened in 1891 [2]. Rural areas were not well covered and even by 1913, one third of all telephones were in London [3]. Another complicating factor was that public phones were available from 1886 [2]: this means that it was not necessary to have a phone at home in order to make a call, but person-to-person calls using a public system were, of course, difficult to manage.

3. Theory

The models presented in this paper focus on two important factors that underlie the adoption of phones: the network effect and affordability.

The network effect. In his 1969 seminal paper Bass [5] argued that except for first adopters, take-up of new "generic classes of products" (as opposed to new models of older products) is related to the number of previous buyers. If the new product happens to be a link to a communication network, this effect is particularly important. Metcalfe's Law states that "the value of a communications network is proportional to the square of the number of its users" and "the law is said to be true of any type of communications network" [6]. Essentially "the idea is that a network is more valuable the more people you can call" [6]. Fischer [7] noted that when phones were introduced in the US they were used "to widen and deepen existing social patterns rather than to alter them". Valente [8] argued that there are two processes at work: one reliant on the "entire social system" and the other on "an individual's personal network". The first is a matter of following changes in society in general, such as opinion leaders who are not personally known, while the other implies that an important determinant of phone adoption is whether your family and friends, that is, those in your 'personal network', already have phones.

Affordability. Initially, only the better off could afford phones. In the US "the more affluent households were the earliest subscribers" [7]. It was the same in Britain. Phones were expensive: in 1901 "When you could employ a maid for £20 per year, having unlimited phone use for £17 per year did not seem to be a bargain" [2].

4. A Systems Dynamics Model

The SD model is based on Verhulst's logistic equation [9]:

$$dp/dt = rp(1 - p)$$

where p is the proportion of adopters, t is time and r is a parameter controlling the speed of adoption. This equation means that the rate of adoption is determined by the existing proportion of adopters. Thus this SD model can be said to model the network effect implicitly: the greater proportion of households that have phones, the more likely any given non-adopting household will adopt.

The adoption curve is therefore determined by two factors, a start point and the growth rate, r. As explained above, we only know that the telephone service started in 1879 and took some 80 years to reach about a fifth of households. Although by 1882 there was one phone per 3,000 people in London, this ratio is clearly too high for households because it includes business phones and takes no account of the very few phones outside London. In the absence of better data, it was therefore assumed that one household in 10,000 were adopters in 1880. On this basis, a Verhulst equation with r set at 10.3% provides the best fit (using OLS) for the whole period: but for the period from 1960 to the mid-1970s setting r at 10% gives a better fit, while for the later years 10.5% is better. Also, the SD model gives no reason for the take-off of adoption after 80 years other than a simple network effect i.e. the more people who have a phone the more attractive it is for others to have one too. Nor does it give any indication why the growth rate should be around 10%. (The results are shown in Fig. 1 later in the paper alongside the results of the AB model to facilitate comparison.)

5. An Agent-Based Model

In this AB model 10,000 agents are spread randomly across a toroidal grid of just over 99,000 cells. There are two types of agents: 'Blues', who represent the affluent early subscribers and are all located in one quadrant; and 'Greens', who represent the rest of the population and are spread randomly throughout the whole 'world'. These assumptions are designed to reflect both the affordability and the concentration of adopters both socially and geographically noted above.

To be consistent with the SD model, one of the Blues is designated the first adopter so that one in 10,000 had a phone. In each time period the adopting Blues 'persuade' another Blue within a fixed radius, representing their network of family and friends, to adopt. It was not presumed necessary for all the Blues to adopt before the Greens start adopting because, depending on the distribution of Blues, it may be that not all

Blues would adopt. Instead, it was assumed that once the adoption rate among the Blues stops rising, adoption spreads to the Greens on the same basis.

The percentage of Blues in the population will, by definition, be small and information on personal networks suggests that the number of people contacted frequently is measured in tens (see e.g. [10]). The model was therefore run with the percentage of Blues at 2.5%, 5%, 7.5%, 10% and 15% and the personal network radius increasing in increments of 1 from 5 to 10. (Note this is not the size of the personal network but the radius defining the network.) Each of these 30 combinations was run 10 times, giving a total of 300 runs, because the output of each run varies due to the different random distribution of agents across the 'world'. The average of each group of 10 runs was taken to facilitate comparisons. Table 1 indicates that only six combinations produced a curve that approximated the actual series by producing an adoption rate of 21% by 1960.

Table 1. Predicted adoption rate by 1960 using the AB model (%.)

Personal network radius	Percent Blues 2.5	5	7.5	10	15
5	14	9	7	6	6
6	**18**	10	9	6	10
7	**19**	10	7	8	46
8	**20**	11	8	45	77
9	**21**	8	60	82	97
10	**19**	**21**	95	100	100

Closer examination of the results from the six combinations shown in bold in Table 1, including taking OLS, revealed that only when it was assumed that 5% were Blues and the agents had a personal network radius of 10 did the curve approximately match the observed pattern. Typically, this radius implied a personal network of about 30 agents. To confirm this result a further 30 runs were undertaken using these same assumptions. Fig. 2 shows the results of all 40 runs (grey), with the average (a solid black line) and one standard deviation (dashed lines) alongside actual take-up (black squares).

6. Comparing the Models

Figs. 1 and 2 show the results of the models. So which method provides the best model for phone adoption in Britain? Parunak et al [11] argued that AB modelling "is most appropriate for domains characterized by a high degree of localization and distribution and dominated by discrete decisions" and that the choice between the two approaches should be made on a case-by-case basis. But on what basis should that choice be made?

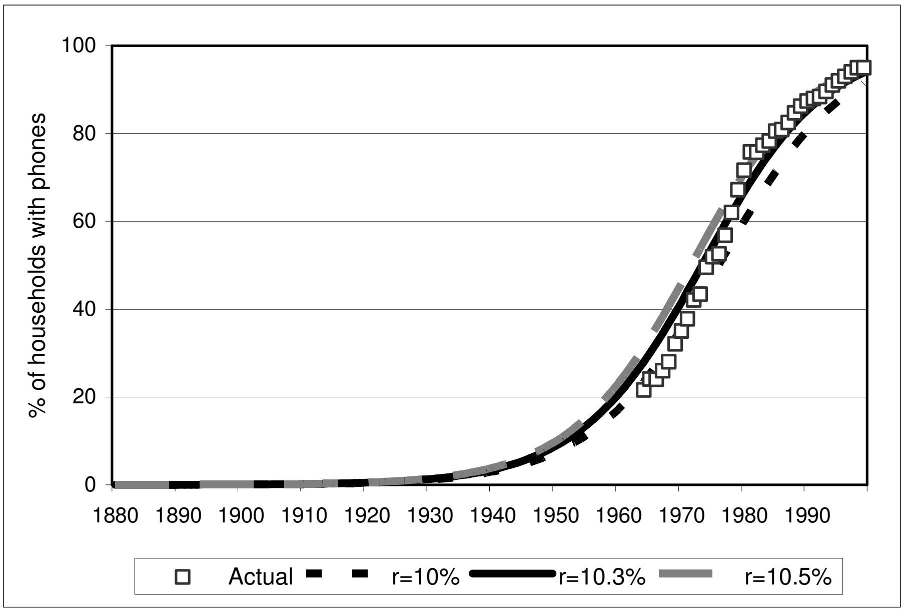


Fig. 1. Household adoption of phones: predictions of the SD model compared to actual.

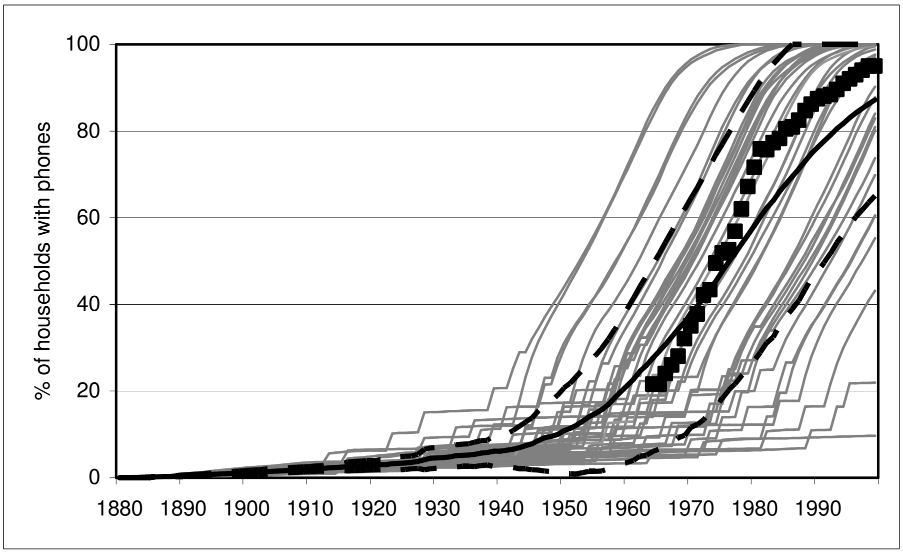


Fig. 2. Household adoption of phones: 40 predictions (grey), the average (solid black line) and one standard deviation (dashed lines) of the AB model assuming 5% Blues and a personal network radius of 10 compared to actual (black squares).

> "A standard modelling principle is that the level and complexity of a model should be chosen so that it answers the questions and embodies the theoretical elements we are interested in, but is otherwise as simple as possible" [12].

How can that principle be applied? I suggest three basic criteria can be used: goodness-of-fit, fitness-for-purpose and simplicity. These criteria are now applied to the two models described above: the SD model with *r* set at 10.3% and the AB model with Blues set at 5% and the personal network radius set at 10.

Goodness-of-fit. The SD model provides a unique result for each set of parameters. However, the results for the AB model vary between runs because the distribution of the agents across the 'world' varies and for this reason, the average is used to measure goodness-of-fit. Fig. 3 shows that in both models it takes some 80 years for adoption to reach about 20%. But the SD model replicates the rise in the following 40 years more accurately. Overall, the SD has a much lower OLS, as illustrated in Fig. 3. Thus the SD model provides the best fit although fine-tuning the parameter values with the AB model might improve its fit.

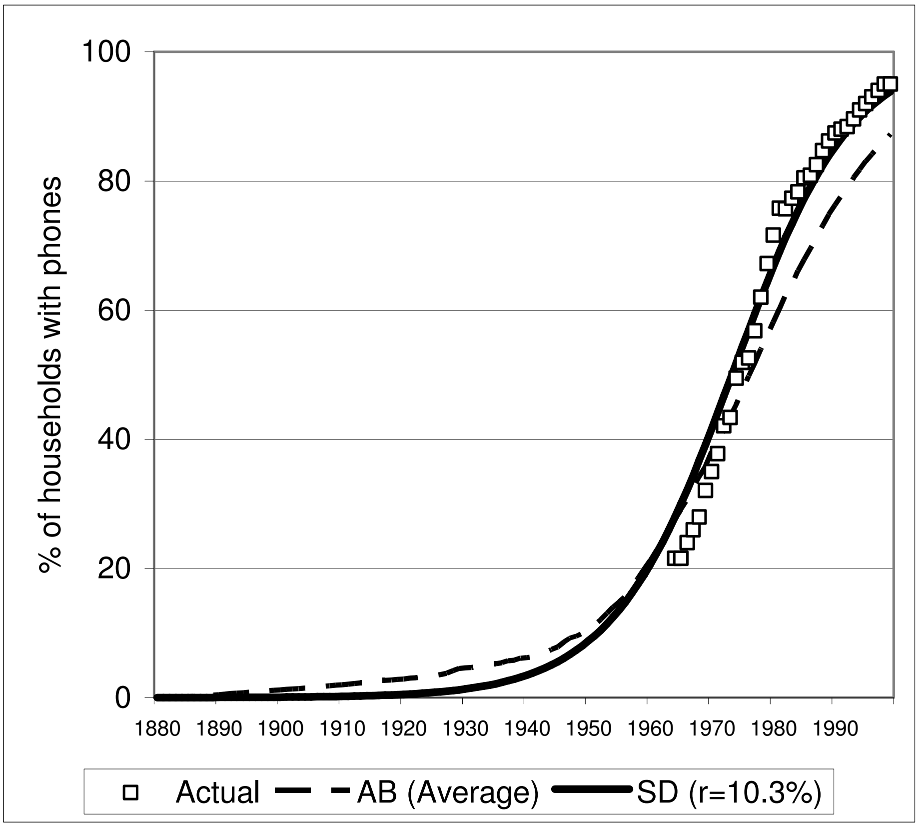


Fig. 3. Comparison the SD and AB models with the actual data.

Fitness-for-purpose. As Forrester pointed out long ago [quoted in 13] the validity of a model should be judged by its suitability for a particular purpose. Many models are built for forecasting or planning. However, a model that predicts well may not add much to our understanding. Quite accurate short term economic forecasts can be made by looking at the time trend of the variable in question and extrapolating it a little way ahead, with no understanding of what is underlying the forecast changes. Indeed, Coleman [14] observed that:

> "macroeconomic predictions based on leading indicators having known statistical association with subsequent system performance may give better predictions than will economic models based on interactions among parts of the system".

In this case the SD model simply reflects the network effect. It provides no insight to understanding the underlying processes whereas the AB model suggests an explanation, telling "a story" that is consistent with the literature.

Simplicity or Occam's razor: *entia non sunt multiplicanda praeter necessitatem* or "entities are not to be multiplied beyond necessity" [15]. The SD model is undoubtedly the simpler to describe and simpler to program. It is also runs faster.

To sum up: in this example, SD scores well on goodness-of-fit and simplicity but low on explanatory power. It is therefore not surprising to see that this is the approach chosen by UK phone supplier BT to model phone uptake (see [16]). The AB model scores lower on fit and simplicity but much higher on explanatory power: phone adoption spread through a geographically and socially close group of affluent early adopters before reaching a wider population. While the fit of the AB model may be improved by adding more parameters, too much fine-tuning of this kind could result in reduction of explanatory power by making interpretation difficult. (Some work was also done on cellular automata but it is not reported here as measured against these criteria, the method was no more than a 'poor man's' AB model: it was poorer fit, offered less explanatory value but was not significantly simpler.)

The models were implemented using NetLogo [17] and can be found at:
www.hamill.co.uk/misc/essa07.zip.

7. Combining the Models

The SD versus AB debate seems often to be presented as an either/or choice, top-down versus bottom-up, macro versus micro. Yet as Fishwick [18] noted, models of different types can be combined to answer different types of question about a given process. Möhring and Troitzsch [19] took an SD model and started to break it down, moving it towards an AB model. How this might work in the case of UK phone adoption?

The SD model assumed everyone was identical whereas in the AB model, there were two types of individuals, the affluent early adopters (Blues) and the rest (Greens), and each individual had their own unique personal network. While the SD model cannot readily deal with 10,000 households, it can deal with two types of people: early adopters and the rest. Now the AB model suggested 5% of the population were affluent early adopters and it so happens that assuming the growth rate for these early adopters was 9¼% and then 13% for the others, a better OLS fit to the data can be obtained using the SD model than simply using a constant rate throughout (see Fig. 4). Thus information from the AB model has been used to improve the SD model.

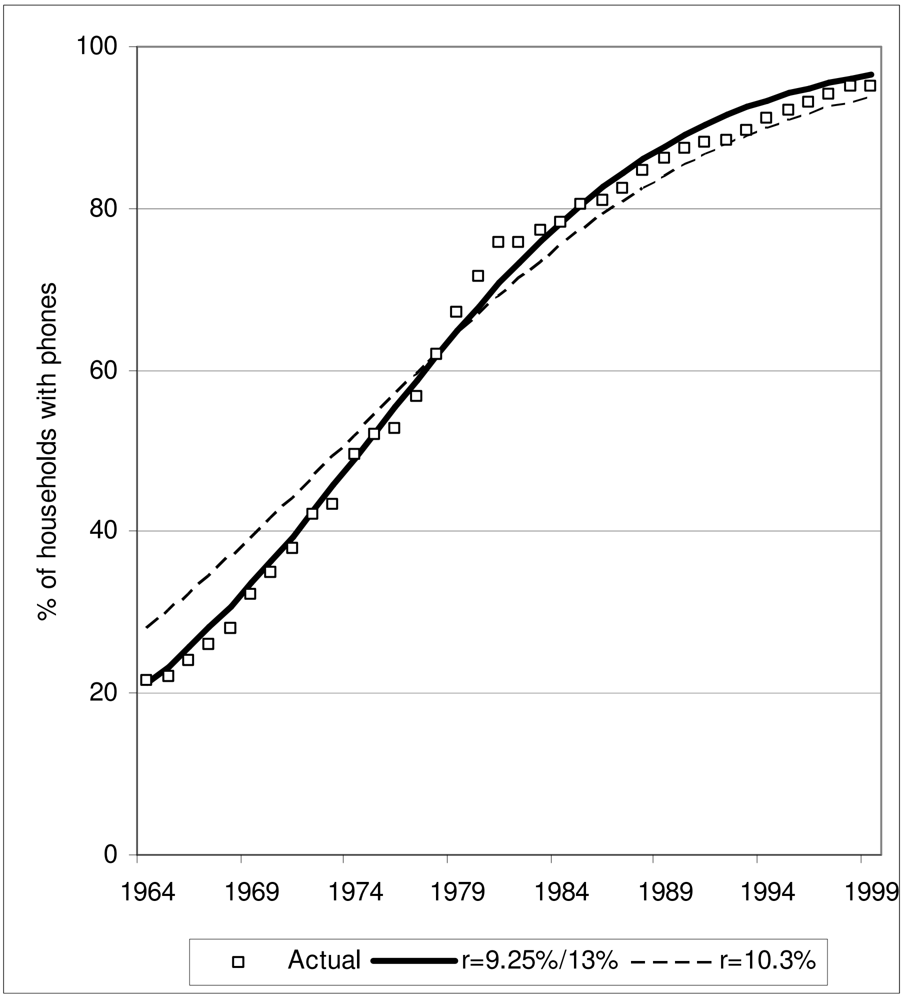


Fig. 4. Household adoption of phones: predictions of the improved SD model compared to the simple SD model and actual.

In turn, the SD model may be used to improve the AB model. The fact that this SD model is deterministic facilitates sensitivity analysis. For example, given the adoption rate of 9¼% for the early adopters and 13% for the rest, varying the percentage of affluent early adopters from 1% to 15% makes a noticeable difference to the adoption rate between 1950 and 1990 but little difference in the early years or as the take-up approaches saturation (see Fig.5). Put another way: the assumed percentage of early adopters, by definition a small group, did not matter much for 70 out of the 120 years being studied!

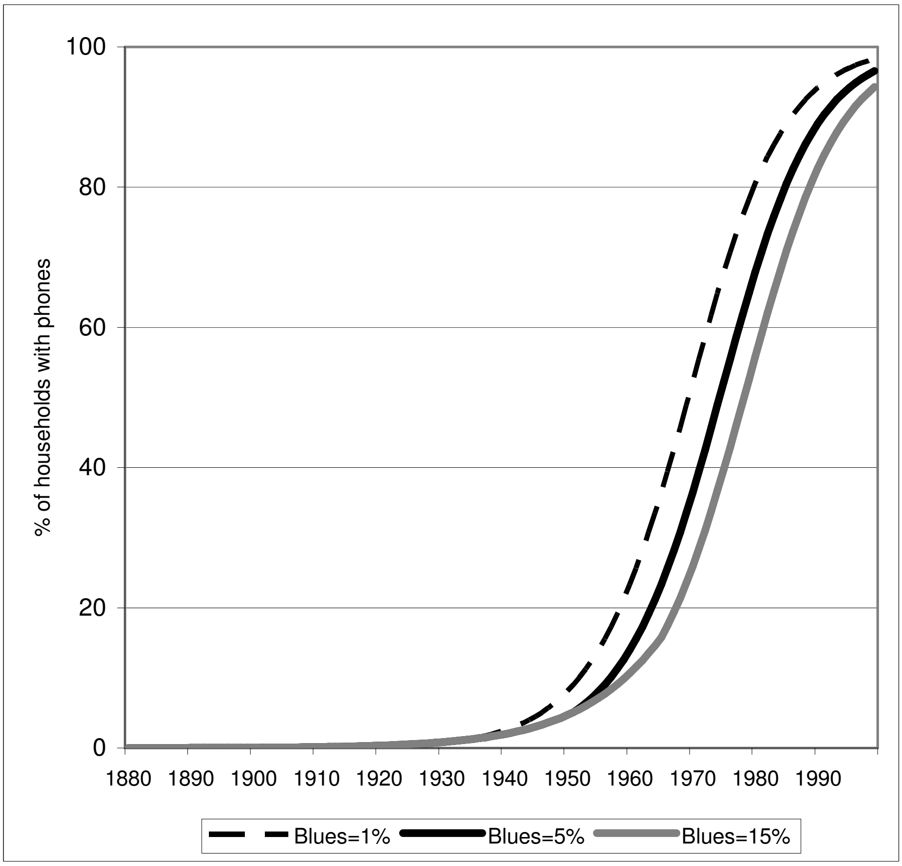


Fig. 5. Results of varying the percentage of early adopters using the SD model.

Because each run of the SD model takes seconds rather than minutes as for the AB model and because using the AB model more runs are needed due to the stochastic processes, it is possible to do more experimenting, more sensitivity analysis with the SD model. For example, the SD model can be used to demonstrate that if 5% of the population were affluent early adopters then unless their growth rate is at least 7% it is not possible for overall take-up to exceed 90% in 120 years (see Fig. 6). In other words, a slow initial take-up can have very long-term consequences! These 121 runs reported in Fig. 6 took under a minute with the SD model: doing the equivalent with the AB model, which would have required perhaps a thousand runs, would have taken many hours, even days.

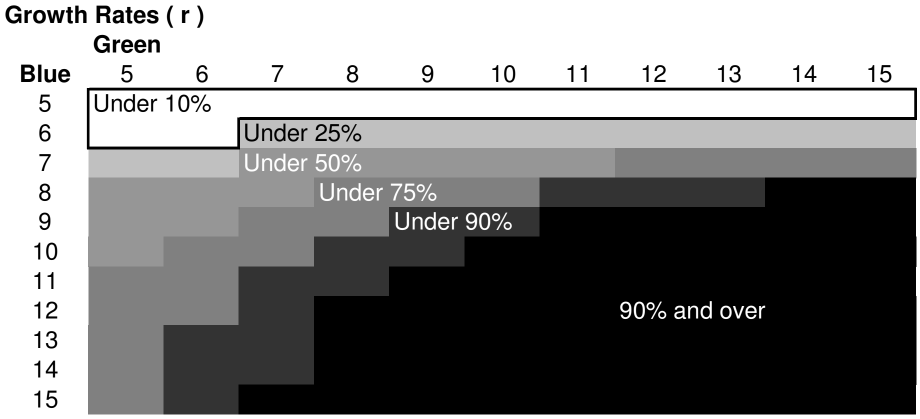


Fig. 6. Results of varying the two growth rates using the SD model.

I suggest that the greater insight provided by the AB model has been used to improve the SD model and that the SD model has been used to test the sensitivity of the assumptions in a way that would have been impractical with an AB model because of its heavy computational demand and its stochastic nature.

8. Conclusion

The main aim of the paper was not to explore in detail the adoption of fixed line phones in Britain. Nevertheless combining the history, theory and the results of the two models an interesting story emerges. The slow adoption of phones for the first 80 years followed by a fast rise to saturation can be explained as follows. About 1 in 20 geographically and socially close affluent households were early adopters. These households had average personal networks of about 30 and they persuaded one member of their network to adopt each year. Once adoption stopped spreading among this group, it started to spread in the same way to the rest of the population. This is consistent with a Verhulst growth rate of about 9¼%, which rises to 13% for the later adopters. With these growth rates, for the first 70 years the outcome is not sensitive to the percentage of early adopters assumed, given that by definition it is a small group. However, if the growth rate for the early adopter group is ‘too low’ – less than 7% on these assumptions – the product will never ‘take-off’. Further work is needed.

More importantly, this paper has shown that SD and AB models can, in some cases, produce similar results and that they appear to complement each other, each having its own strengths and weaknesses. So I suggest that rather than choose between SD and AB models, in some cases, both can usefully be used. Further work will explore the use of SD models to assist formulating and verifying AB models.

Acknowledgements
This work was supported by Microsoft Research through its European PhD Scholarship Programme. Thanks, too, to Nigel Gilbert for advice and encouragement.

References

1. Scholl, H.J.: Agent-based and System Dynamics Modeling: A Call for Cross Study and Joint Research. Pr. of 34th Hawaii International Conference on System Sciences. IEEE. (2001)
2. BT Archives: Events in Telecommunications History: (2007) http://www.btplc.com/Thegroup/BTsHistory/Eventsintelecommunicationshistory/Eventsintelecommunicationshistory.htm Accessed 6 June 2007.
3. Perry, C.R.: The British Experience In de Sola Pool (ed), The Social Impact of the Telephone. MIT Press, Cambridge, Mass.(1977)
4. 1964-93: CSO: Family Spending. HMSO London (1994)
   1994-97: ONS: Family Spending 1997-98. TSO. London (1998)
   1998-99: ONS Family Spending 2005-06. Palgrave Macmillan. Basingstoke (2007)
5. Bass, F.M.: A New Product Growth for Model Consumer Durables. Management Science, Vol 15. No 5, Theory Series. Pp 215-227 (1969)
6. Briscoe, B., Odlyzko, A. & Tilly, B.: Metcalfe's Law is Wrong. (2006) IEEE Spectrum (July 2006) 26-31
7. Fischer, C.S.: America Calling: A Social History of the Telephone to 1940. University of California Press. Berkeley. (1992) 261-3
8. Valente, T.W.: Social network thresholds in the diffusion of innovations. Social Networks, (18) (1996) 69-89
9. Strogatz, S.: Nonlinear Dynamics and Chaos. Westview, Cambridge MA (1994) 22-23
10. Boase, J., Horrigan, J.B., Wellman, B. & Rainie, L.: The Strength of Internet Ties. Pew Internet & American Life Project. Washington (2006)
11. Van Dyke Parunak, H. et al.: "Agent-Based Modeling vs. Equation-Based Modeling: A Case Study and Users' Guide" in "Multi-Agent Systems and Agent-Based Simulation": Book Series Lecture Notes in Computer Science, Springer, Vol. 1534, 1998, 10-25.
12. Doran, J. & Palmer, M.: The EOS Project. In Gilbert, N. & Conte, R. (eds) Artificial Societies. UCL Press. London. (1995) 103-125
13. Qudrat-Ullah, H.: Structural Validation of Systems Dynamics and Agent0Based Simulation Models. Proceedings 19th European Conference on Modelling and Simulation (2005)
14. Coleman, J.S.: Foundations of Social Theory. Harvard University Press, Cambridge, Mass. (1994) 4
15. Oxford Dictionary of Philosophy. OUP. (1994)
16. Lyons, M.H., Burton, F., Egan, B., Lynch, T. & Skelton, S.: Dynamic modeling of present and future service demand. Proceedings of the IEEE, 85(10), (1997) 1544-1555
17. Wilensky, U.: NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL. (1999)
18. Fishwick, P. A.: Simulation model design and execution: building digital worlds. Prentice Hall Englewood Cliffs, N.J (1995) 285-332
19. Möhring, M. & Troitzsch K.G.: Lake Anderson Revisited by Agents. Journal of Artificial Societies and Social Simulation 4.3. http://www.soc.surrey.ac.uk/JASSS/4/3/1.html (2001)

Modeling Essential Micro Interactions for Analyzing Emergent Phenomena in Market

Kotaro Ohori[1] and Shingo Takahashi[2]

[1,2]Department of Industrial Management and Systems Engineering, Waseda University
[1]oohori@fuji.waseda.jp, [2]shingo@waseda.jp

Abstract. This paper introduces a new framework for market analysis, and discusses essential micro interactions among economic entities described in the framework. Finally we observe market emergent phenomena in simulations using a model supported by the framework, and emphasize the importance of describing micro interactions. Since various descriptions of interactions on agent-based modeling have directly adopted existing simulation technologies based on computer science and artificial intelligence etc., it is difficult for modelers to describe the essential interactions on a model that correspond to intentions of actions in real markets. So we need to propose novel simulation technologies.

Keywords: Essential micro interaction; Emergent phenomenon; Genetic operation; Market analysis

1 Introduction

Some conventional studies of market analysis have focused on market characteristics concerning product diffusion and standard competitions[7].The Innovator's Dilemma[1] and User-innovation[2] that form novel trends of market studies have shown various influence of innovation on a market. Since these studies are ex post case analysis, they seem unable to analyze future market conditions that will change rapidly through globalization and technological innovation.

On the other hand, recent studies try to analyze market phenomena with virtual experiments of computer simulation. Agent-based modeling has been applied to various fields such as society, economy, or organization. The progress of social simulation has been able to simulate some phenomena in a real world.

Some studies with agent-based modeling have focused attention on market dynamics. Deguchi[3] has formalized a lock-in model that different two populations interact each other in various markets with technological innovation, and analyzed a learning process, which is called social learning dynamics. Struben[4] has focused on decisions of consumers and firms, and proposed a transitions model of technologies. Most of models for analyzing market dynamics, however, have supposed the premise of economics, and considered only population level learning. These models were not able to describe various emergent phenomena observed in a real world.

In this paper, we introduce a new framework for market analysis that can explain various phenomena in real markets. The framework represents essential micro interactions among economic entities and evolutionary learning processes. Finally we discuss the possibilities of these representations and the observation of emergent phenomena.

2 Framework

In this chapter we describe the new market framework using a model of coevolutionary processes of firms' technologies and consumers' preferences[8]. The framework has the following features: 1) economic entities as firms or consumers are regarded as autonomous agents, 2) essential interactions among these agents in a market are described, 3) consumers' preferences and firms' technologies co-affect their evolutionary behavior.

This framework considers the bounded rationality and heterogeneity of economic entities. Each consumer recognizes a product space as an environment, and then he/she chooses and purchases a product for maximizing his/her utilities. Each firm develops their technologies for gaining higher market share. Consumers and firms affect each other through the launching or choice of products. We call this framework CAMCaT (Coevolutionary Agent-based Model for Consumers and Technologies) (see Fig.1). In the subsequent sections, we describe internal model (2.1), activity model (2.2), and evolutionary learning process (2.3).

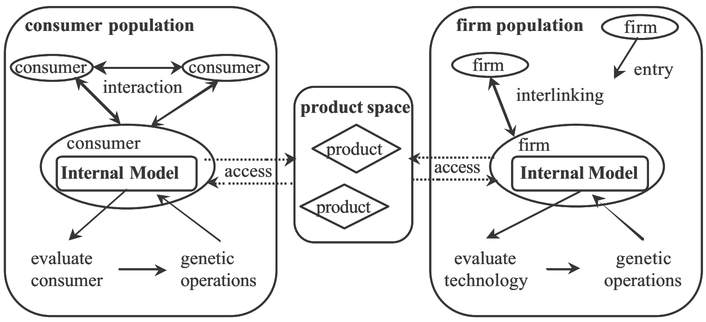


Fig. 1. CAMCaT framework.

2.1 Internal Model

An agent has an internal model consisting of the chromosomes that describe market situations and the agent's characteristics in his/her mind. Each economic entity, based on its internal model, makes decisions of economic actions that are the choice or launch of products in a market that has high uncertainty.

Each consumer has some chromosomes that are rules for selecting products. The chromosomes of a consumer are consisted of Preference, Consumption Propensity, Possessed Technology, and Market Situation (see Fig.2).

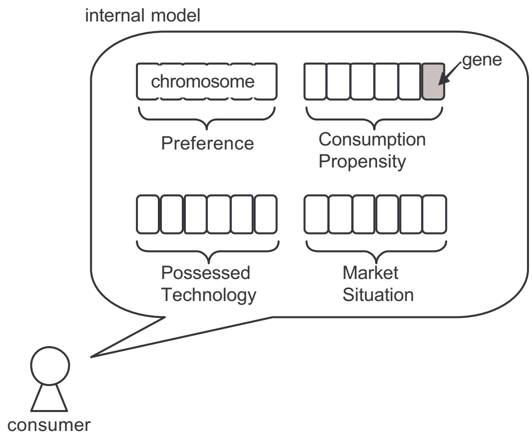


Fig. 2. Internal model of a consumer.

Each firm also has some chromosomes that are rules for developing a new product, and gaining higher market share. The chromosomes of a firm are consisted of Management Strategy, Technology Strategy, Possessed Technology, and Market Situation.

We selected these chromosomes by considering various decisions in firms or consumers in real markets. Modelers do not have to adopt all chromosomes, and can set parameters of partial chromosomes depending on the problem situations or target of modeling in detail. This framework, therefore, has no restriction on building the internal model.

2.2 Activity Model

Consumers and firms, based on their internal models, make their decisions in a market. Each consumer evaluates products and chooses a product in the product space (see Fig.3). Each firm decides whether to develop and launch products. These decision makings are considered as micro level activities in a market.

2.3 Evolutionary Learning

After micro level activities, each consumer evaluates his/her own choice and internal model by using a fitness function. Similarly, each firm evaluates its own development, launching, and internal model by using a fitness function. The fitness function defines how desirable each economic entity is in a market. This is formulated depending on problem situations. As for examples, the consumer's fitness function consists of preference values and sensibilities to trend etc. The firm's one consists of technology values and market share etc.

Based on the self-evaluation results, various genetic operations realize a series of learning of economic entities. Recent studies have considered the necessity of learning mechanism, and adopted genetic algorithm (GA) or reinforcement learning etc. We will uphold the GA, since the reinforcement learning dose not guarantee that population level learning converge.

To take an example of consumers, bandwagon-effect, information exchange and gathering information correspond with the selection, crossover, and mutation in GA [5], respectively. We call a series of operations "evolutionary learning process." The long term result by evolutionary learning processes achieves the coevolution of consumers and firms through the choice and development of products.

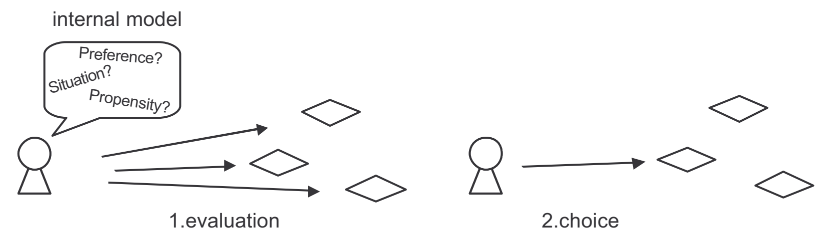


Fig.3. Activities of a consumer.

3 Micro Interaction and Micro-macro Link

The primary feature of CAMCaT framework is to realize the essential micro interactions among economic entities. The micro interactions represent neither a two player game nor a change of decisions depending on macro information like interactions described in conventional studies. We represent the micro interactions as essential individual economic actions that are described as the revision of internal model as rules for an agent's action (see Fig.4). So the micro interactions do not change the agent's action itself directly.

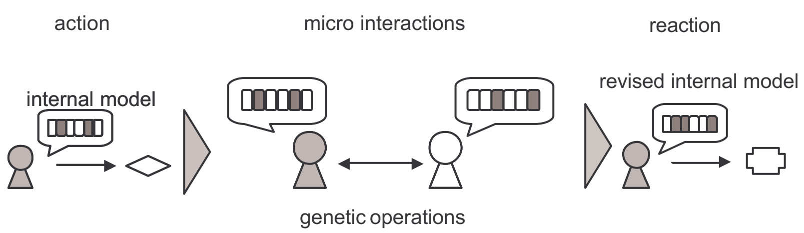


Fig. 4. Micro interactions in CAMCaT framework.

Some internal models in a population that are revised by the micro interactions change some agent's actions. This change results in a macro level behavior, then the macro level behavior affects the micro interactions and decisions. Therefore, this process represents a micro-macro link that corresponds to the change of economic actions and behaviors in real markets.

3.1 Micro Interactions

In this section we discuss evolutionary learning of consumers' preferences as micro interactions. Our framework applies the analogy of evolution not to the population level learning but to the individual level one. The population level learning can be achieved as a result of the individual level one. We have originally reinterpreted and proposed genetic operations as technologies for the learning in the framework[8][9]. The economic actions represented by the operations can be essentially effective in market analysis since we were successful in explaining the micro interactions and macro phenomena on the analogy of evolution in our previous studies. In following subsections we explain intention-driven genetic operations in our previous studies (3.1.1), and discuss the genetic operations on evolutionary learning process (3.1.2) and necessity of novel genetic operations (3.1.3).

3.1.1 Correspondence between Intentions and Genetic Operations

The micro interactions are composed of three phases using genetic operations. First, a consumer population initially set out in the model selects the consumers' preferences that realize higher evaluation values. This implies a kind of bandwagon effect. Second, a pair of consumers is combined at random in the consumer population, then they uniformly crossover each partial gene in their chromosome. This implies the information exchanging between two consumers. The crossover process has an effect to spread various preferences over the consumer population. Finally, each consumer mutates partial genes in his/her chromosome. This implies gathering information through advertising media (see Fig.5).

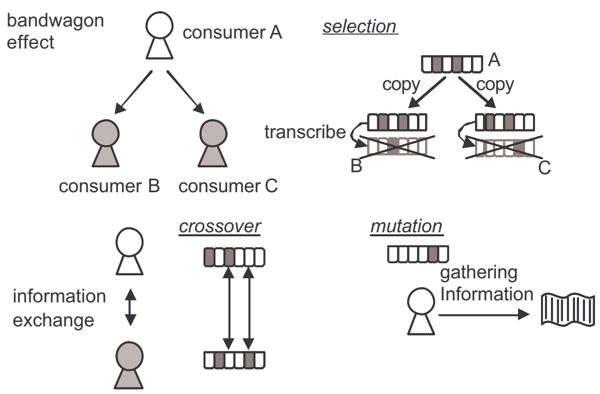


Fig. 5. Implications of genetic operations.

3.1.2 Evolutionary Process by Genetic Operation

We here point how the evolution of a consumer population has achieved in our previous studies with CAMCaT framework. The evolutionary learning was performed with roulette selection, uniform crossover of a random pair in GA, and mutation. The uniform crossover could temporarily decrease genes values in the chromosome of preferences (see Fig.6). Supposing that the genes values show the preferences values, the decrease seems to prevent the evolution. It nevertheless did not fail from the viewpoint of the evolution by combining the selection and the uniform crossover. Though genes values were temporarily decreased by the uniform crossover, varieties of the consumer population increased. And consumers who had a higher evaluation value were selected with the roulette method. Finally, evolution was achieved in the consumer population (see Fig.7).

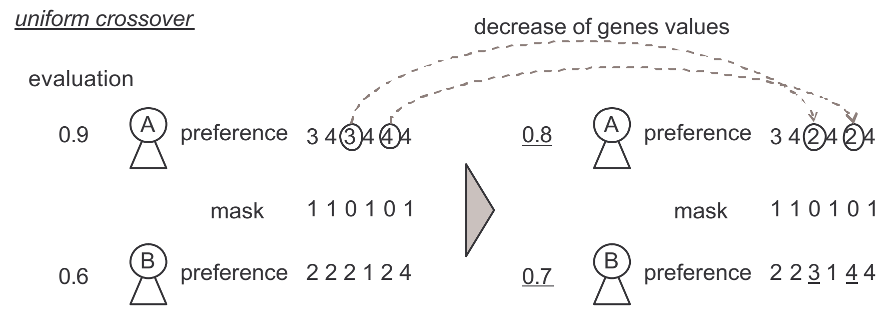


Fig. 6. Decrease of genes values caused by uniform crossover.

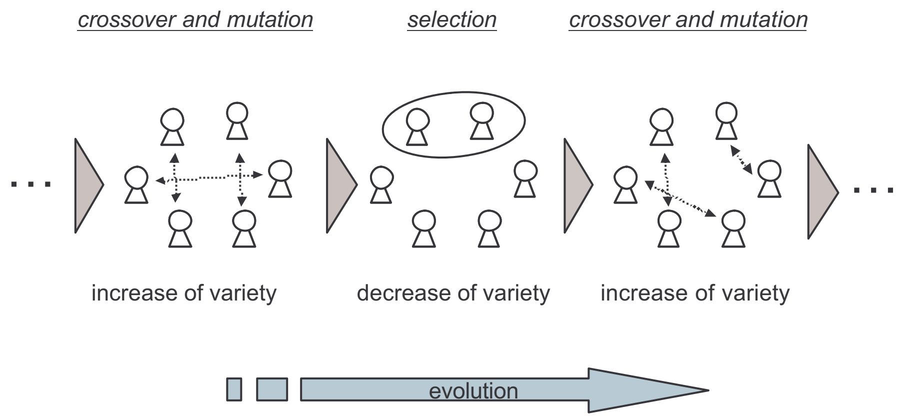


Fig. 7. Change of variety in evolutionary learning process.

3.1.3 Necessity of novel intention-driven simulation technologies

As has been noted the genetic operations are capable of corresponding to intention of economic actions. Since our previous studies[8][9] has described some market characteristics observed in real markets, the analogy of evolution using genetic operations achieved essential individual level learning of economic entities. However we here should notice that the existing genetic operations based on classical GA cannot be promised to succeed in future analysis by using a model built in different types of problem situations.

We will assume that a market model has a network structure. A recent study with an agent-oriented model has adopted the viewpoint of social network that a market structure determines the activities of economic entities[10]. A model with a network structure describes links among economic entities, which define feasible interactions. In this case selection based on classical GA cannot imply economic actions directly, since it operates two consumers who are not linked(see Fig.8). Copy(1) expresses that a consumer affects linked another one, and can imply spread of information by bandwagon effect. However copy(2) cannot imply in real markets, since a consumer is affected by not linked another one directly. We therefore have to develop a genetic operation improved for achievement of economic entities' learning.

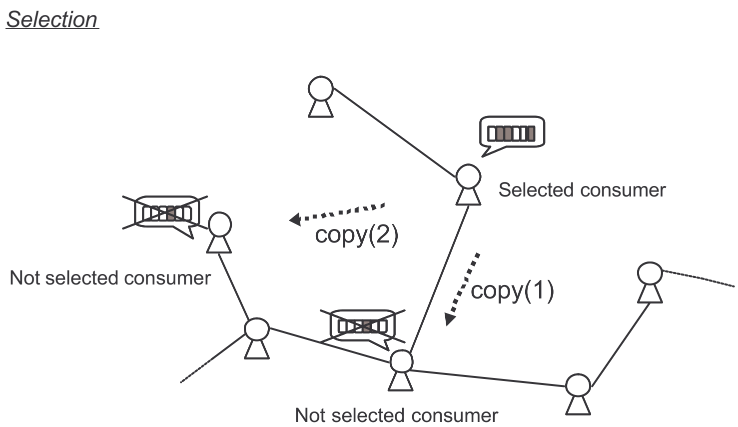


Fig. 8. Non-correspondence between bandwagon effect in network structure and selection based on GA.

Let us consider an example of a novel genetic operation that is helpful to learning in a market network structure. The operation is a fusion of selection and uniform crossover, and can imply the bandwagon effect in real markets (see Fig.9). Though the operation is based on the uniform crossover, it does not perform exchange of the genes between linked two consumers. The genes values of consumerB who has lower evaluation values copy from the genes values of consumerA who has higher ones. The operation, without adopting selection, will achieve evolution in population by selecting the high evaluation chromosomes and gain high variety. Therefore the operation will achieve learning of consumers in real markets.

In the preceding argument we have insisted that agent-based modelers have to develop novel simulation technologies when they would like to build the model that has restraints on agent activities like the network structure. In particular we expect genetic operations as the simulation technologies to incorporate learning mechanism. The operations will enable us to build an intention-driven model. It is not, however, enough to consider only intention when we build a new model. As Allen[6] has discussed modeling, we also need to consider other related theoretical studies which are economics and marketing science etc.

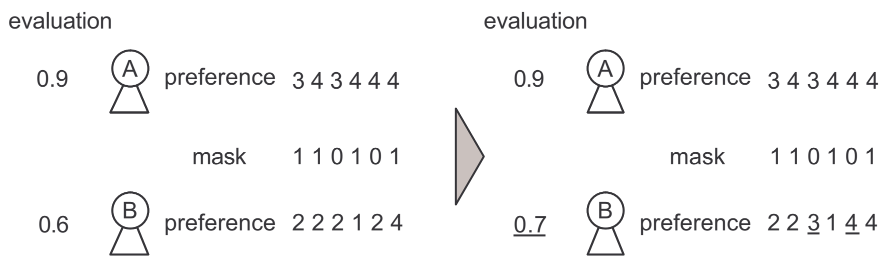


Fig. 9. An example of a novel genetic operation

3.2 Micro-macro Link

Micro interactions described as evolutionary learning revise agents' internal models in populations. The revision should imply economic actions in real markets (see Section 3.1), and change agent's decisions as choice and launch. The change in populations affects market products, market share, and a trend product in a product space. The whole market results in different behavior with which are compared a few generations ago. This macro market behavior affects economic actions in the next generation. So micro-macro link is described.

4 Modeling and Emergent Phenomena in Simulations

In this chapter, we provide a market model by using CAMCaT framework and show the macro phenomena emerged on a market simulation. The target of the provided model is competitions among some standards.

4.1 Model

According to the CAMCaT framework, the market model consists of a consumer population, a firm population and a product space. The size of the consumer and firm population is 100 and 20, respectively. The number of product standards is 5.

4.1.1 Model of Product

In the product space there are a lot of products that firms launched. We describe the attributes of the product that are consisted of specs a_{ik} and standard s_i defined by $A=(a_{ik},s_i)$ where $a_{ik}\in\{1,2,\ldots,100\}$, $s_i\in\{1,2,\ldots,5\}$ $k=1,2,\cdots,5$, i identifies an individual product, k is the spec number of a product.

4.1.2 Model of Consumer

1. Internal Model

The internal model of a consumer i consists of the sensibilities to other consumers d_i, a favorite standard s_i, cutoff value c_{ik}, and purchasing weight for product attributes w_{ik}, defined by $IM=(d_i,s_i,c_{ik},w_{ik})$ where $d_i\in\{x\mid 0\le x\le 1\}$, $s_i\in\{1,2,\ldots,5\}$, $c_{ik}\in\{1,2,\ldots,100\}$, $\sum_k w_{ik}=1$, $i=1,2,\cdots,100$, $k=1,2,\cdots,5$, i is a consumer index, k is an attribute index.

2. Activity Model

Each consumer evaluates products by the evaluation rule of products and chooses one having the maximum utility bigger than the cutoff values. The evaluation rule is defined by the utility function u_{ij} of consumer i for product (1).

$$u_{ij}=\Big(\sum_k b_k * a_{ijk} * d_i + \sum_k w_{ik} * a_{ijk} * (1-d_i)\Big) * c_j * s_j \tag{1}$$

$$where\ c_j=\begin{cases}0 & if\ consumer\ i\ cutoff\ product\ j\\ 1 & otherwise\end{cases},\quad s_j=\begin{cases}0 & if\ s_i\neq s_j\\ 1 & otherwise\end{cases},$$

b_k represents the attributes of the trend product.

3. Evolutionary Learning Process

After choice of a product, each consumer i evaluates his/her own decision and internal model by using the fitness function fc_i (2).

$$fc_i = w_a*(1-ncut)+w_b*sumcut+w_c*(1/maxcut)+w_d*(network) \tag{2}$$

ncut is the number of non cutoff products, *sumcut* is the sum of cutoff values, *maxcut* is the maximum cutoff value, *network* is the number of consumers who adopt the same standard with consumer i. The weight parameters are set to $w_a=0.50, w_b=0.40, w_c=0.05, w_d=0.05$. The fitness function represents that a consumer has higher evaluation can reduce recognition effort of products, and can gain advantage by network effect.

We adopt GA (roulette selection, uniform crossover, and mutation) in this model. According to the evaluation of consumers, internal models of consumers are selected with roulette selection. And the revision of the internal model is performed with uniform crossover and mutation. Crossover and mutation rate is 0.6 and 0.05, respectively.

4.1.3 Model of Firm

1. Internal Model

The internal model of a firm i consists of the standard to adopt s_i, technological concept c_{ik}, and possessed technology t_{ik} defined by $IM = (s_i, c_{ik}, t_{ik})$ where $s_i \in \{1,2,\ldots,5\}$, $\sum_k c_{ik} = 1$, $t_{ik} \in \{1,2,\ldots,100\}$, $i = 1,2,\cdots,20$, $k = 1,2,\cdots,5$, i is a firm index, k is an attribute index.

2. Activity Model

Each firm i launches a new product, according to the launching rate that is 0.05. The attribute A of launch product j is set by calculating from the possessed technology T of the firm.

3. Evolutionary Learning Process

After launching the product, each firm i evaluates its own decision using the fitness function ff_i (3).

$$ff_i = w_a * share + w_b * othervalue + w_c * selfvalue + w_d * network \tag{3}$$

$$where\ othervalue = \sum_k c_{ik} * t_{jk},\ selfvalue = \sum_k c_{ik} * t_{ik}$$

j is a firm index, which the firm interlinked with firm i.

share is the share of the products firm i launches, *othervalue* and *selfvalue*(3) shows the fitness of other firms' and own technology with its own technological concept, and *network* is the number of firms that adopt the same standard with firm i. The weight parameters are set to $w_a = 0.25, w_b = 0.05, w_c = 0.65, w_d = 0.05$.

The internal models of firms are selected with Baker's linear ranking selection after the evaluation. This implies licensing or M&A. The revision of internal model is performed with crossover and mutation. This implies cross license and R&D. Crossover and mutation rate is 0.6 and 0.05, respectively.

4.2 Simulation results

In this section, we simulate competitions between five standards in a market using the model (see Section 4.1) and show emergent phenomena in the market. Fig.10 shows two market share transitions of the standards in simulations using the same parameter.

As the simulation result in Fig.10-(a), the market share of standard5 reaches about 100% at about 50 generation, and keeps high share. This state is called "lock-in" and is also observed in simulation in conventional studies. As for the simulation result Fig.10-(b), de facto standard5 switches to a new de facto standard2.

The two points that require clarification are that 1) why did the simulation result in Fig.10-(a) differ from one in Fig.10-(b) ?, 2) why did de facto standard switch to another in Fig.10-(b)? It should be clear that the evolutionary learning affects the two points. Agent-based modelers who have ever built a model that had been equipped with learning mechanics do not doubt as to the first point. However, as for the second point, many agent-based modelers must wonder about the market phenomenon. The market phenomenon is one of emergence, and is caused by the essential interactions among economic entities. We call the phenomenon in Fig.10-(b) “lock-in wave”. The reasons for this are technological innovations and evolution of preferences. This can be observed in real markets, like home-video market and game market etc.

We will show an example of mechanism of lock-in wave caused by the evolution of preferences. The consumers’ preferences are affected by the opinion leaders’ ones, when they adopt a standard whose market share is low. The chromosomes in the consumer population drastically change, then the change generates a new trend standard in the product space.

The most important point in this simulation is that this framework can show the lock-in wave that cannot be observed in conventional studies. Although we recognize the importance of corresponding to real market case, the present stage of this study cannot fit the real market data. This framework, however, can represent the various emergent phenomena can been seen in real markets, and has a high potential to explain more interesting phenomena.

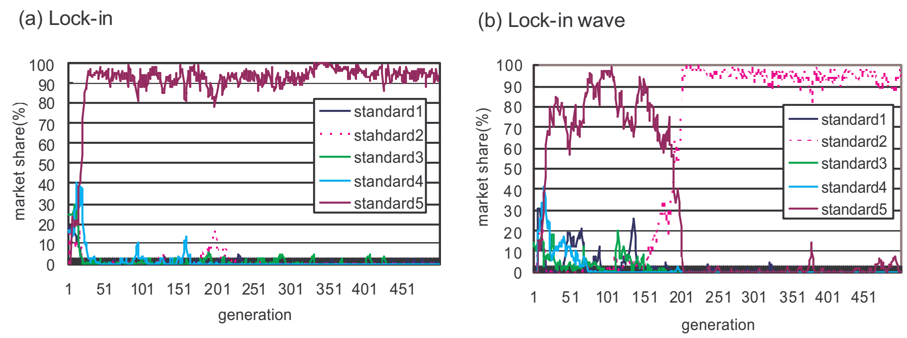


Fig. 10. Emergent phenomena in simulations.

5. Conclusions

In this paper we propose a new framework, called CAMCaT, for market analysis, and suggest that the simulation technologies are required to describe essential micro interactions that correspond to economic actions in real markets.

We should note that the framework has most of essential elements in real markets, the elements which are the micro interactions among economic entities, the decision making of economic entities, the interaction between a firm population and a consumer population, and macro phenomena in the whole market. The market model using the framework can show lock-in wave as an example of emergent phenomena. This observation strongly results from the description of the micro interaction. Thus this framework has ability to understand market phenomena, though it cannot yet explain real market cases. We expect that models using this framework can explain various phenomena in real markets and analyze various market situations.

Finally we should emphasize that agent-based modelers will have to develop simulation technologies for describing interactions in order to develop agent-based modeling. Existing simulation technologies often restricted the modeling of agent activities and learning. In such case some modelers have tended to ignore the partial characteristics of problem situations that they would explain. In addition to this, simplification of modeling has inclined to assist the ignorance. So we will develop the intention-driven simulation technologies to prevent the elimination of important characteristics.

Reference

1. Clayton M. Christensen , "The Innovator's Dilemma, when new technologies cause great firms to fail", Harvard Business School Press, 1997
2. Eric von Hippel , "Democratizing Innovation", MIT Press , 2005
3. H.Deguchi, " Learning Dynamics in Platform Externality,in Applied General Systems Research on Organization",Springer,pp167-176,2003.
4. Jeroen Struben,"Technology Transitions; identifying challenges for hydrogen vehicles" , MIT,2004.
5. John H. Holland , "Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence", Bradford Books, 1992
6. Peter M. Allen, "Knowledge, ignorance and the evolution of complex systems", Frontiers of Evolutionary Economics: Competition, Self-Organization, and Innovation Policy, pp313-350,2003
7. Rohlfs,J.H ,"Bandwagon Effects in High Technology Industries", The MIT Press, 2001
8. Shingo Takahashi, Kotaro Ohori, "Agent-based Model of Coevolutionary Processes of Firms Technologes and Consumer Preferences",NAACSOS Conference, 2005
9. S.Takahashi, "Framework in Agent-based Approach for Analysis of Evolutionary Processes of Consumers' Preferences," J.of the Japan Society for Management Information,Vol.13, pp.1-17,2004.
10. Tackseung Jun, Jeong-Yoo Kim, Beom Jun Kim and M.Y. Choi,"Consumer Referral in a Small World Network", Social Networks, Volume 28, pp232-246, 2006

Session on Economy and Cognition

Welfare stigma allowing for psychological and cultural effects.
An Agent-Based simulation study.

Dalit Contini*, Matteo Richiardi**

*Università di Torino
Dipartimento di Statistica e Matematica Applicata "Diego de Castro"
dalit.contini@unito.it
**Università Politecnica delle Marche
Collegio Carlo Alberto – LABORatorio R. Revelli
m.g.richiardi@univpm.it

Abstract: We investigate the effects of income support on unemployment and welfare dynamics when stigma is attached to welfare provision. Stigma has been modelled in the literature as a cost of entry into welfare. Allowing for psychological factors, we assume that with stigma welfare provision also leads to lower search effectiveness; moreover, we allow for interaction among agents. Carrying out an agent-based simulation study, we find that welfare take-up rates decrease with stigma and welfare spells get longer. Unemployment rates are not monotonically related with the amount of stigma, implying that we can find higher levels of unemployment with stigma than with no stigma.

Keywords: Job-search, Welfare Dependence, Stigma, Agent-Based Modelling.

1. Introduction

The effect of income support policies has been the object of extensive theoretical (OECD 2005) (Rogerson *et al* 2005) and empirical research (Moffitt 1992; Moffitt, 2002; Blank 2002). The focus is on work disincentives. The leading paradigm is rational choice: if the benefit is high enough with respect to wages, individuals choose welfare and stay out of the labour market. The body of work from the empirical literature confirms that transfer programs considerably reduce work effort.

The aim of this paper is to investigate the effects of income support on unemployment and welfare dynamics when social stigma is attached to welfare provision. Stigma is acknowledged as one of the determinants of welfare take-up behaviour[1] (Hernanz *et al* 2004) and it has been modelled as a cost of entry into welfare (Moffitt 1983). By

[1] The take-up rate is defined as the ratio between the number of individuals receiving the benefit and the total number of those who are eligible for it.

providing a disincentive for welfare participation, stigma negatively affects take-up rates: the higher the cost of entering welfare, the lower the propensity to enter welfare.

Our model is more comprehensive. While still embedded in a utility-maximisation framework, individual behaviour is allowed to depend on psychological and cultural factors. Referring to the model for welfare dependence proposed in Contini and Negri (2007), we develop a discrete-time job search model - assuming we are dealing with “weak” individuals, with low skills and low job opportunities - with labour market features taken as given. We let stigma affect preferences by representing a cost of entry into welfare, as other authors do, but in addition we assume that with stigma welfare provision can lead to a reduction of search effectiveness, due to progressive loss of self-confidence of recipients and to unfavourable attitudes of potential employers. Moreover, we allow for interaction among individuals: living in environments where most people rely on welfare can cause preferences to change by reducing the perceived cost of stigma, therefore making the benefit more desirable.

These effects are difficult to investigate empirically, as it is hard to separate the consequences of the specific policies under study from the effects of other policies at work and from macro changes occurring in the meantime. For this reason, we follow a different approach, carrying out an Agent-Based (Tesfatsion 2001) simulation study. The use of simulations is motivated by the fact that, given the complexity of the model, it would be difficult to derive analytically all the relevant results.

We explore the consequences of income support for the poor/unemployed, on welfare take-up rates, unemployment and welfare participation rates, on unemployment and welfare spell duration. Environments with and without stigma are compared. We will show that take-up rates steadily decrease with the amount of stigma, while welfare spells get longer. With respect to unemployment, we highlight two related results:

According to traditional job-search models, stigma – by reducing the work disincentive of welfare provision and enhancing the relative value of work – should reduce unemployment. Allowing for psycho-social effects this conclusion no longer holds: unemployment rates are not monotonically related with the strength of stigma. In many situations we find (*ceteris paribus*) higher levels of unemployment with stigma than with no stigma.

Without stigma higher benefit levels can coexist with lower unemployment and welfare participation rates. Hence, we provide some explanation[2] for the evidence that in Northern Europe, where benefits are quite generous and little stigma is attached to benefit provision because of the universalistic character of the welfare system (Saraceno 2002), unemployment and poverty rates (OECD 2005) and the length of poverty spells (European Commission 2002) are generally lower than in Southern Eu-

[2] Of course, many other explanations are possible. Unemployment rates are shown to be correlated with the rigidity of the employment protection legislation, with union bargaining power, with the strength of active labour market policies (Scarpetta 1996).

rope, where benefit levels are low (OECD 2004; Saraceno 2002) and being in welfare is stigmatized[3].

These results crucially rest on the assumption that employability decreases with elapsed time in welfare with stigma, and it is often reinforced by the interaction effects among individuals.

The paper is structured as follows. In section 2 we illustrate the features of the job-search model, while section 3 is dedicated to the simulation design. Results are discussed in detail in section 4. Conclusions follow.

2. The model

In typical job-search models individual behaviour is based exclusively on rational choice. Individuals are subject to random job offers, that may be accepted or rejected according to the future value of utility associated with the different options. Benefit provision to the unemployed increases the reservation wage: the larger the subsidy and the longer its expected duration, the less individuals are attracted by work, triggering the so-called "welfare trap".

In our model individuals do not operate in a completely rational manner: behaviour is allowed to depend on psychological and cultural factors. In neo-classical economics individuals' preferences are taken as given and the budget line describing available options is only subject to exogenous changes; to acknowledge for psycho-social factors, both these assumptions will be relaxed.

The main features of the model are sketched as follows:

1. The object of individual's decision is search effort. Greater effort increases the probability to find a job, but reduces current utility by reducing time for leisure. Unemployment benefits and social assistance are treated in a unified framework[4]. Stigma may be attached to welfare provision. Following Moffitt (1983), we let stigma represent a fixed cost of entering welfare[5]; as a consequence, not all the eligible will eventually claim the benefit. In this framework, at every point in time, the unemployed has to take the joint decision: 1) whether to search for work; 2) whether to enter welfare.
2. Behaviour may also be affected by psychological and cultural factors (Bane and Ellwood 1994). We claim that these factors should be especially relevant in those environments where stigma is strong. "… Living on public support, which in certain cultural contexts is equated with living on charity, exposes the individual to social disqualification and stigma, imprisoning him in marginal social networks

[3] Unemployment rates for 2003: 8.7% in Italy, 11.3% in Spain, 9.7% in France, 5.4% in Denmark, 5.8% in Sweden, 4.4% in Norway. Poverty rates for 2000: 12.9% in Italy, 11.6% in Spain, 7.0% in France, 4.3% in Denmark, 5.3% in Sweden, 6.3% in Norway (OECD 2005).

[4] Although Contini and Negri (2007) clearly distinguish between unemployment compensations and social assistance, we treat income support policies in a unified framework because we wish to keep the specification simple.

[5] Moffitt also allows for a variable component, depending on the size of the benefit, but this component does not seem to be empirically relevant.

and isolating him – even more than unemployment – from those social contacts which help to gain access to work opportunities. Demoralization and learned helplessness may also take root…" (Contini and Negri 2007), favouring reduction of the effectiveness of job-search. Moreover, prospective employers themselves might be less willing to hire welfare recipients. Thus, stigma may play a role in reducing re-employment probabilities, even when the individual rationally chooses to undertake the job-search.

3. Values and attitudes can be affected when individuals live in situations of socio-economic exclusion and spatial segregation; if the majority of the neighbours rely on welfare rather than work, preferences can change, making work less attractive. Our model allows for interaction among individuals; each individual occupies a cell in a bi-dimensional space, and the behaviour of those in adjacent cells contribute to shaping individual's preferences.

Let $U = f(C, L)$ be the utility function associated with consumption C and leisure L. We operate in a rigid labour market with full time jobs only. People are assumed to consume all their earnings (there are no savings, nor other sources of income), thus, consumption amounts to current income. Income is C_E if employed, C_0 if unemployed with no benefit – taken to be below the poverty threshold - and C_B with income support, where $C_0 < C_B < C_E$. A universalistic policy is considered, so that all the unemployed (who are also poor) are eligible for welfare benefits, which are in principle of unlimited duration.

Standardizing total time to 2, we fix minimum time for leisure L to 1, time for work is 1 and time devoted to job search is either s=0 or s=1. No search on the job is allowed, so that L=1 for the employed and L=2-s for the unemployed. Market wage is always higher than individuals' reservation wage: $f(C_0,2) < f(C_E,1)$. Thus, if no benefits are provided, it is better to work rather than not work.

Present utility for the unemployed is a development of the simple Cobb-Douglas function $U = C^{\alpha} L^{\beta}$. We assume that stigma is the only factor responsible for take-up behaviour[6]. With stigma, since individuals have to choose whether to search for work and whether to claim the benefit, current utility depends on C, L, and A, where A=1 if he is assisted and 0 otherwise. With no stigma, i.e with no disincentives to enter welfare, individuals will never choose A=0. On the other hand, in environments where benefit provision is disqualified, a cost ϕ of entering welfare must be applied. Moffitt (1983) proposes the following model:

$$U(C, L, A) = U(C, L) - \phi A \tag{1}$$

Adapting this specification to the Cobb-Douglas function and allowing for cultural effects in the model - see point (iii) above – according to which preferences change if individuals are "close" to other welfare recipients - we obtain the following function:

[6] According to Hernanz *et al.* (2004), various potential explanations of low-take-up rates for welfare benefits have been addressed in the literature: pecuniary determinants, information costs, administrative costs, social and psychological costs (stigma).

$$U = f(s, A, f) = \left[C_0{}^{\alpha}(2-s)^{\beta}\right]^{1-A}\left[C_B{}^{\alpha}(2-s)^{\beta} - \phi(1-f)\right]^{A} \qquad \textbf{(2)}$$

where f is the proportion of welfare recipients among the person's neighbours. When stigma is present, $\phi>0$ and individuals living close to other welfare recipients will be less affected by stigma: their preferences will change, reducing the value of work. In the extreme case where all neighbours are welfare recipients ($f=1$), no stigma effects are perceived by the unemployed, who are thus more prone to enter welfare and eventually stop searching for a job.

The probability of finding a job is allowed to change with elapsed time in unemployment, as skills tend to become obsolete and social contacts facilitating the match between labour supply and demand loosen (Granovetter 1995). Moreover, we assume that stigma can be the cause of further reduction in employment prospects as time spent in welfare grows longer - see point (ii). The re-employment probability is thus specified as follows:

$$p_t = \gamma_0\left(1-\theta^{U}\right)^{\tau^{U}}\left(1-\theta^{A}\right)^{\tau^{A}} \qquad \textbf{(3)}$$

where γ_0 is the corresponding probability at the beginning of the unemployment spell. Loss of skills developing with time elapsed in unemployment is related to θ^U, while reduction of work opportunities caused by welfare recipiency, occurring when stigma is present, is related to θ^A. τ^U and τ^A are respectively time elapsed in unemployment and in welfare. Notice that $\tau^A \leq \tau^U$, as people can delay welfare entry. With no search there are no chances of receiving job offers, while re-employment allows to recover the original value γ_0.

We assume that the employed are satisfied with their work activity, so that there is no search on the job, but they will loose their current job with probability δ. The unemployed, instead, choose whether to search for work or not. We assume that they evaluate present utility U_0 and the expected utility for two time units ahead; search effort and welfare participation decisions are determined by $\max_{(s_0, A_0)} V_0$, where V_0 is given by:

$$V_0 = U_0 + E[U_1]R + E[U_2]R^2 \qquad \textbf{(4)}$$

where $E[U_t]$ is the expected utility at time t, and $R \in (0,1)$ is a discount factor. Hence:

$$\begin{aligned} V_0 &= U(s_0, A_0, f_0) + \left[U(s_1, A_1, f_1)(1-p_0) + U_E p_0\right]R + \\ &\left\{U(s_2, A_2, f_2)\left[(1-p_0)(1-p_1) + p_0\delta\right] + U_E\left[p_0(1-\delta) + (1-p_0)p_1\right]\right\}R^2 \end{aligned} \qquad \textbf{(5)}$$

where s_t and A_t represent search effort and welfare participation at time t. U_E is the utility of being employed, while p_t is the probability to work at time t+1 given job-search at time t. Notice that p_t itself is a function of s_t. Individuals are assumed to correctly forecast their loss of employability, even when the loss is due to welfare participation[7].

There are 2^6=64 different combinations of values 0 and 1 for (s_0, A_0, s_1, A_1, s_2, A_2). V_0 is evaluated at each combination[8], and the (s_0, A_0) maximising V_0 is taken as the optimal choice for time t=0. In the following time unit, options are evaluated with respect to new current utility and the utility of the two subsequent points in time. Thus, the values of (s_1, A_1, s_2, A_2) maximizing V_0 need not to be equal to the actual choices that will be made at times t=1 and t=2[9].

Because of the progressive loss of employability with time elapsed in unemployment and welfare, job-search can be the optimal choice when employability is still high, but it becomes no longer optimal when employment prospects fall below a certain level (Richiardi and Contini 2006 demonstrate this result with regard to decisions taken with respect to an infinite time horizon). Moreover, if with no stigma everybody takes the benefit from the very beginning, when stigma is present the eligible might not ask for income support at the onset of the unemployment spell, delaying claiming when search effectiveness falls below some threshold.

3. The simulation design

The analytical model is investigated *via* a discrete-time[10] agent-based simulation (Tesfatsion 2001)[11]. The simulation schedule is reported in Figure 1. At time 0 the model is initialized and the world (an IxJ torus grid) is inhabited by agents, each agent having 8 neighbours. A fraction *fractionEmployed* of these agents start as employed, while the others are unemployed. Given that employability decreases over time in unemployment and welfare, infinitely lived agents would inevitably end up in unemployment, which is an absorbing state of the system. To prevent such a deadlock we assume that agents exit the labour market after *maxAge* periods. At the beginning of each period exiting agents are replaced by new agents. Again, with probability *fractionEmployed* each new agent starts as employed, while with probability (1-

[7] This is not necessarily a sensible hypothesis. The issue will be the object of future investigation.

[8] Notice that working condition at time t=2 depends on previous choices and, given that the working condition at time t=3 is not taken into consideration at time t=0, the value of s_2 maximizing V_0 is necessarily 0.

[9] This feature is addressed in the economic literature as "time inconsistency" of choices. In recent years the issue has been object of increasing interest among social scientists (see for example Fang and Silverman, 2004; O' Donoghue and Rabin, 1999).

[10] Discrete-time simulation means that the state of the system is updated (i.e. observed) only at discrete (generally constant) time intervals. No reference is made to the timing of events within a period – see, for example, (Allison, Leinhardt; 1982).

[11] The simulation is written and is build on the open source JAS simulation platform (Sonnessa 2004). The code can be downloaded from http://193.205.134.131/Entra/download/P004473/allegati_doc/Stigma.rar or requested to the authors.

fractionEmployed) he starts as unemployed. Seniority of the agents is randomly initialized at time 0.

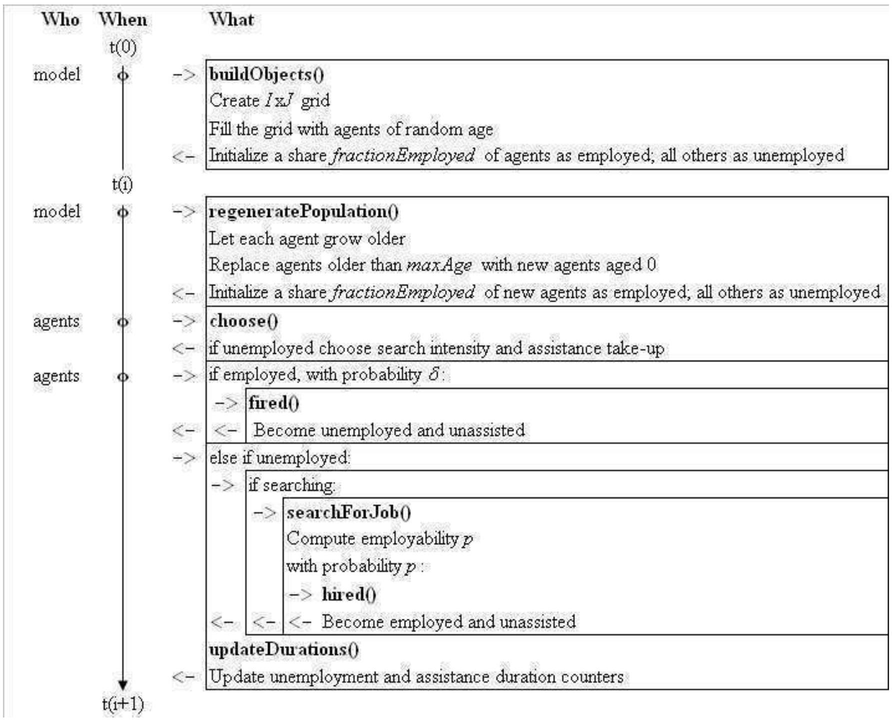


Fig. 1. Simulation schedule

The scheduling of the events within each simulation period should be interpreted as suggested in Figure 2. Employed workers do nothing for the whole period. Unemployed individuals choose their search intensity and whether they are willing to enter welfare at the beginning of the period. Search lasts for the whole period, while benefits are given, if the agent is still unemployed, at the end of the period. Transitions (from unemployment to employment and from employment to unemployment) also take place at the end of the period. As a consequence, employed workers who are fired at the end of a period start the following period with maximum employability γ_0 (their elapsed duration in unemployment and welfare is still 0 at the time when they begin searching for a new job).

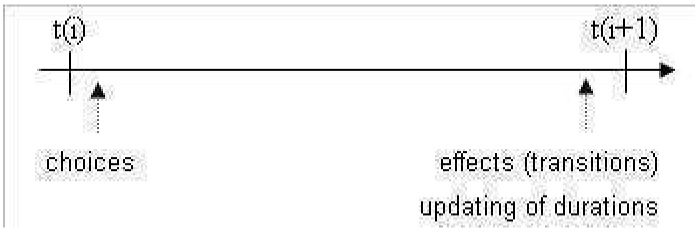


Fig. 2. Continuous interpretation of the discrete time schedule

The speed of convergence to the equilibrium, defined as a regime when all relevant time series look stationary, depends on three key parameters: the equilibrium share of people in assistance, the size of the population and the number of periods individuals stay in the labour market. Given the neighbourhood effect formalised in (1), whenever one individual enters welfare it increases the likelihood that other individuals around him, if unemployed, would also choose to become assisted. Hence, the length of the chain reaction that is triggered depends on how many individuals are at risk (i.e. unemployed) and, for given levels of stigma, on the amount of the benefit.

With 900 simulated individuals, even in the worst case scenarios convergence is obtained within 250 periods, *i.e.* about 2 generations. All the results reported below refer to stationary values, averaged over a large number of periods.

In each period the state of the system can be summarized along two dimensions: unemployment and unemployed behaviour, which in turn depends on search intensity and welfare take-up.

Figure 3 shows a typical simulation outcome: the blue cells in the grid are for employed individuals; red cells are for the unassisted unemployed and yellow cells are for the assisted unemployed. The share of active searchers would complete the description of the system. Note that assisted individuals can cluster together: this is due to the fact that the stigma associated with being in welfare is reduced if the neighbourhood contains a higher proportion of similarly assisted individuals.

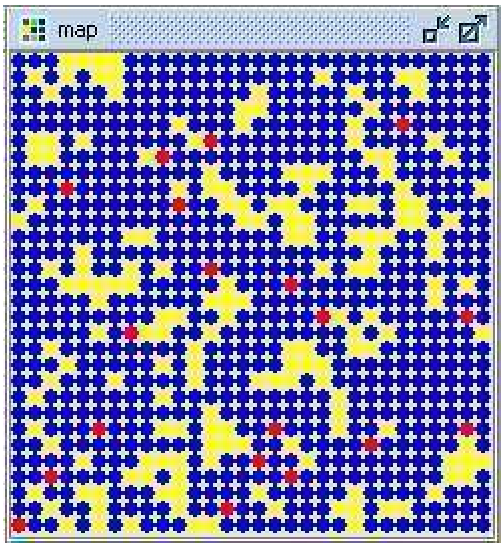


Fig. 3. Simulation outcome

The simulation design has been implemented with parameter values summarised in Table 1. We consider three different environments. One with no stigma effects, the other two characterised by different levels of stigma – which is assumed to affect both the component ϕ in the utility function (1), and the component θ^A in the re-employment probability function (2). Ideally, the NO STIGMA environment should represent Northern Europe countries, while situations with stigma should somehow depict Southern Europe countries. Having no objective way to fix realistic parameters (we are not aware of studies dealing with this issue), we have chosen two alternative combinations: the one labelled as STIGMA 1 has lower values of both ϕ and θ^A than STIGMA2. Thus, the latter represents a situation with stronger stigma. Nevertheless, it is important to note that, given that these values are not empirically founded, we are not able to denote these situations as carrying out "low" or "strong" stigma in absolute terms[12].

Income with no benefit is set at C_0=1 and market wage at C_E=4. Two levels of income support are considered: C_B=1.5 and C_B=2.5, providing 37.5% and 62.5% of market wage respectively. Re-employment probabilities at the beginning of the unemployment spell can take two alternative values: γ_0=0.25 or γ_0=0.4, reflecting situations where individuals can have worse or better job-prospects. On the other side, the employed will loose their job with a probability δ=0.05. Core of the utility function is $U = C^{\alpha} L^{\beta}$, with α=2 and β=0.5[13]. We think of each time unit as one month. For

[12] A calibration exercise is one of our aims for future work.

[13] These values are not empirically founded either. They have been set ad hoc, in order to bring about possibly "reasonable values" for present utility in the different situations under consideration. The consumption component of the function amounts to 16 for employment, 2.25 or 6.25 for unemployment with income sup-

this reason the discount factor is close to 1 (R=0.98). Individuals are followed for 120 time units, thus "life time" is 10 years. All individuals are unemployed at "birth", as *Fraction_Employed* (see section 3) is set to 0.

Environments	NO STIGMA: θ^A=0, ϕ=0
	STIGMA 1 : θ^A =0.1, ϕ=1
	STIGMA 2: θ^A =0.15, ϕ=1.5
Benefit levels	LOW BENEFIT: C_B=1.5
(duration unlimited)	HIGH BENEFIT: C_B=2.5
Income	C_0=1, C_E=4
Job turnover	γ_0=0.25 or γ_0=0.4
	θ^U=0.05
	δ=0.05
Discount factor	R=0.98
Utility function	α=2
	β=0.5
Life length	120 time units
Fraction employed	0

Table 1. Simulation parameter values

4. Results

Cross-section results are summarised in Tables 2a-2d. The following statistics have been computed with respect to the three environments considered (NO STIGMA, STIGMA1, STIGMA2): average percentage[14] across time of unemployed and welfare recipients, percentage of unemployed actively searching for a job among assisted and not assisted, welfare take-up rate, average number of assisted neighbours for the assisted. In order to highlight the role of interaction among agents, simulations where stigma is present but neighbourhood effects are not at work have been carried-out as well.

We have also performed a sensitivity analysis aimed at exploring the effects on the cross-section indicators of the variation of each parameter with a finer grid (holding the other constant). Results related to θ^A and ϕ are reported in Figures 4 and 5.

port, 1 for unemployment with no benefit. Since search effort $s \in \{0,1\}$ and consequently leisure $L \in \{2,1\}$, given consumption, with no searching utility is $\sqrt{2}$ times utility with searching.

[14] In order to limit the complexity of the results, standard errors have not been reported; notice that their values are generally small, so that conclusions are unaffected.

Longitudinal results are described in Tables 3a-3d. We show some features of the distribution of unemployment and welfare spells[15], together with the number of spells in the observation window. In particular, given the heavy skewness of the distribution, we report the 50°, 75° and 90° percentiles (when estimable), the highest non-censored spell length and the estimated survival at that value.

The main results can be sketched as follows.

Welfare

Take-up rates are higher with no stigma while welfare spells are longer with stigma.

1. *Take-up rate*. It is always 100% with no stigma, as welfare entry has no costs. From figures 4 and 5 we can see that the percentage of the eligible entering welfare is sensitive to ϕ - and this is a well known result - but also to θ^A. When θ^A reaches a certain level, the take-up rate starts decreasing. This should be due to the fact that the unemployed correctly anticipate their future loss of employability, and thus, if this loss is strong, they might prefer not to enter welfare[16]. The joint effect of ϕ and θ^A can be observed in tables 2a-2d. For high benefit, take-up rate is always 100%, but with low benefit, rates are lower with stigma, in particular with stronger stigma. Given that neighbourhood effects have an impact on preferences, lowering the perceived cost of welfare entry, take-up rates can rise if these effects are at work (cfr. columns II and III, table 2c).

1. *Welfare spell length.* As expected, because of the progressive loss of employability occurring when assisted, welfare spells are found to be much longer with stigma in all the situations considered, and the tendency is strengthened with increasing level of social disqualification (tables 3a-3d).

Unemployment

As we have pointed out before, job-search models including stigma predict that stigma – acting as a work incentive – reduces unemployment. According to our model, this conclusion is no longer true. In all the situations under consideration, the percentage of unemployed is higher with STIGMA1 than with NO STIGMA (*ceteris paribus*). Moreover, we observe a *lower* percentage of unemployed in some environments with no stigma and *high benefit* than we do in environments with stigma and *low benefit*. For example, with γ_0=0.4, with STIGMA1 and benefit set at 1.5, 26.3% of the simulated population is unemployed; the percentage falls at 14.3% with NO STIGMA and benefit set at 2.5.

[15] Deduced from the Kaplan-Meier estimate of the survival function.

[16] Notice that if individuals were not able to forecast the progressive reduction of employment prospects (see also footnote 8) we should observe a different result.

1. *Unemployment spell length*. From result b) we know that, for those who enter welfare, welfare spells are longer with stigma. On the other hand, not all the unemployed are welfare recipients: a lower number of people enter welfare if there is stigma (see result a). Thus, no simple general result seems to hold. In all the simulations we have carried out, unemployment spells are longer with stigma than with no stigma, but spells do not increase (on the contrary, they can get shorter) as stigma gets stronger.
2. *Proportion of unemployed.* Everybody is unemployed "at birth", but since unemployment spell length is not monotonically related with stigma (result c), also the relation between the proportion of unemployed and stigma cannot be simply determined *a priori*. In all the simulated cases, the stock of unemployed is always higher in environment STIGMA1 than with NO STIGMA, but it can become either even higher or much lower in the environment STIGMA2. From figure 4, we see that the percentage of unemployed follows a reversed-U shape as θ^A increases. The reason for this behaviour seems to be related, once again, to the fact that the unemployed, forecasting future loss of employability if assisted, can decide not to enter welfare and keep searching for work if the anticipated loss is too big (see result a). Notice that, holding constant θ^A, the proportion of unemployed behaves as predicted in the traditional models (decreasing steadily) when ϕ varies.

<u>Welfare (more)</u>

1. *Proportion of assisted (welfare participation rate)*. As we have seen (results a and b) with stigma a lower number of eligible enters welfare, but those who do will stay longer. Moreover, since the size of the eligible population (i.e., the percentage of unemployed) does not depend monotonically on stigma, the proportion of welfare recipients over the whole population does not vary in a simple fashion. The stock of assisted is found to be much lower with no stigma than with stigma when both the level of the benefit and initial employability are relatively high (table 2d). On the other hand, with stronger stigma and low benefit, the size of the assisted population can be very small or even null (see tables 2a and 2c, columns iv and v).

<u>Job-search</u>

1. With no stigma welfare recipients keep searching for a job more often than with stigma. In all the simulated cases, the number of job-searchers among the assisted decreases with the strength of stigma (although figure 4 shows that it is not a monotonic pattern). The job-search behaviour of the unassisted population is not very clear, as that population can be very small in many cases.

<u>Assisted neighbours</u>

1. The average number of assisted neighbours for those who are assisted themselves provides a rough indication of whether the assisted individuals tend to cluster together. Clustering should occur with stigma, when neighbourhood effects are at work. Yet, this number by itself is not very informative: it should be compared with the average number of assisted neighbours corresponding to a completely

random allocation of welfare recipients in the bi-dimensional space. If the proportion of assisted is p, with a constant number of neighbours n, the relevant distribution is that a binomial, thus the average is simply np. Both the observed average number and the expected value under the assumption of randomness are reported. There is no strong evidence of clustering effects, perhaps with the exception of environment STIGMA2, case of low benefit and low employability, neighbourhood effects at work (column v, Table 2a).

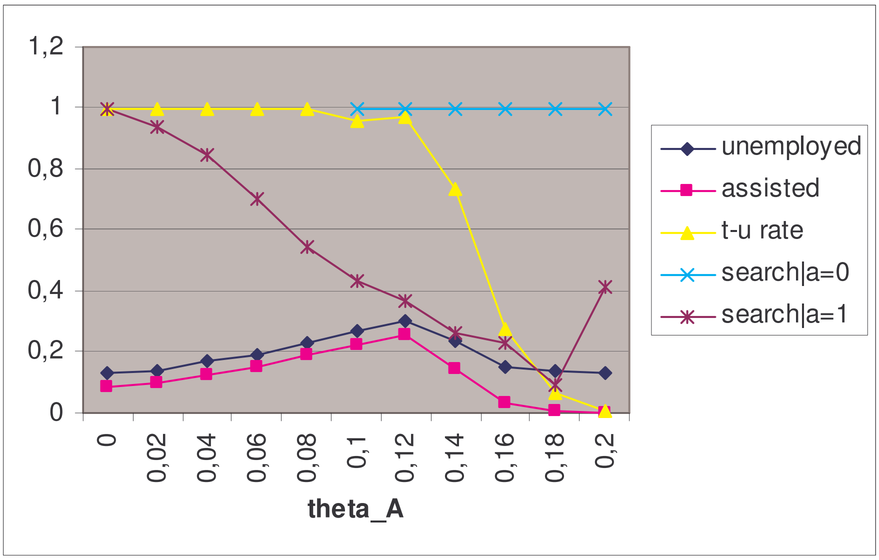


Fig. 4. Simulated results for varying θ^A with ϕ=1, C_B=1.5, γ_0=0.4, with neighbourhood effects

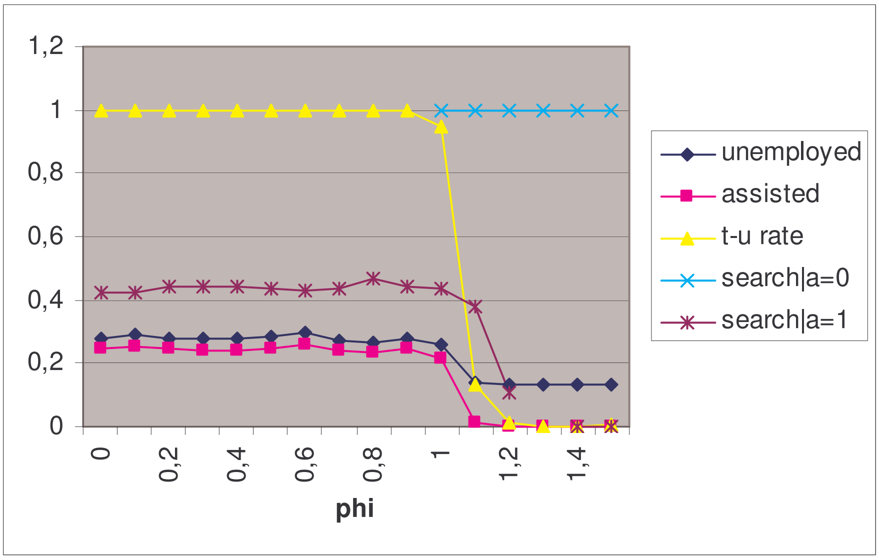


Fig. 5. Simulated results for varying ϕ with θ^A=0.1, C_B=1.5, γ_0=0.4, with neighbourhood effects

5. Conclusions

In traditional job-search models stigma is shown to negatively affect welfare take-up rates; accordingly, unemployment levels tend to decrease with stigma. Thus, although from the point of view of attaining the goal of poverty reduction stigma is considered "a bad", because high take-up rates are among the goal of the policies (Hernanz *et al.* 2004), from the point of view of unemployment reduction stigma is considered "a good".

Under the assumption that social disqualification attached to receiving the benefit not only affects preferences (by representing a cost of welfare entry), but that it leads to a reduction of search effectiveness as well, this mainstream conclusion no longer holds. We show that the unemployment rate is related with the amount of stigma in a non-monotonic fashion: after some level it does decrease with the level of stigma, but "in the beginning" (moving away from the no stigma situation) the unemployment level is shown to increase with stigma.

This result *could* help explaining why unemployment rates, poverty rates and persistence of poverty are lower in Northern Europe than in Southern Europe, even if welfare and unemployment benefits are much more generous there. Other explanations are obviously possible. But we must say *could* also because parameter values were set arbitrarily and not derived from empirical evidence. This means that which values can adequately represent each group of countries – even in a stylized fashion – cannot be assessed yet. This point will be the object of further investigation.

Policy implications are in principle very interesting, although the reduction of stigma, involving well rooted attitudes and beliefs, is not an easy goal to accomplish.

Given that higher unemployment and welfare participation levels are observed when neighbourhood are allowed, the obvious conclusion is that (from many other points of view as well, indeed) spatial segregation and urban ghettoes should be heavily contrasted.

References

Allison P., Leinhardt S. (ed.) (1982) *Discrete time methods for the analysis of event histories*, Jossey-Bass, pp. 61-98.

Bane, M. and Elwood, D. (1994). *Welfare Realities: From Rhetoric to Reform*, Cambridge, MA: Harvard University Press.

Blank, R. (2002). Evaluating Welfare Reform in the United States, *Journal of Economic Literature*, **XL**, 1105-1166.

Contini D., Negri N. (2007) Would declining exit rates from welfare provide evidence of welfare dependence in homogeneous environments?", *European Sociological Review*, 23, pp 21 - 33.

European Commission (2002) European Social Statistics. Income, poverty and Social Exclusion. 2° report.

Fang H. and Silverman D. (2004) Time inconsistency and welfare program participation: evidence from the NSLY, *Cowles Foundation Discussion Paper* n° 1465

Granovetter M. (1995) *Getting a Job: A Study of Contacts and Careers.* University of Chicago Press, Chicago, 2nd edition.

Hernanz V., Malherbet F. and Pellizzari M. (2004) Take-up of welfare benefits in OECD countries: a review of the evidence, *OECD Social Employment and Migration Working Papers,* No 17, OECD Publishing

Moffitt R. (1983) An economic model of welfare stigma, *The American Economic Review*, Vol 73, No 5, pp 1023-1035

Moffitt R. (1992) Incentive effects of the U.S. welfare system: a review, *Journal of Economic Literature*, 30, 1-61.

Moffitt, R. (2002). The Temporary Assistance for Needy Families Program, *NBER Working Paper N° 8749.*

O' Donoghue T., Rabin M. (1999) Doing it now or later? The American Economic Review, Vol 89, n 1, pp 103-124

OECD (2004) Benefit and Wages Indicators, OECD

OECD (2005) Society at Glance. OECD Social Indicators, OECD

Richiardi M., Contini D. (2006) Active and passive policies against poverty with decreasing employability*, Working Paper LABOR-Center for Employment Studies,* No 49

Rogerson R., Shimer R., Wright R. (2005) Search-theoretic models of the labour market: a survey, http:/home.uchicago.edu/~shimer/WP/search-survey.pdf

Saraceno C. (eds.) (2002). *Social Assistance Dynamics in Europe: National and Local Poverty Regimes*, The Policy Press, Bristol.

Scarpetta S. (1996) Assessing the role of labour market policies and institutional setting of unemployment: a cross-country study, *OECD Economic Studies No 26*

Sonnessa M. (2004) "The JAS (java agent-based simulation) library", in Leombruni R., Richiardi M. (eds*), Industry and Labour Dynamics: The Agent-based Computational Economics Approach*, World Scientific Press, Singapore

Tesfatsion L. (2001) Agent-Based computational economics. A brief guide to the literature, in J.Michie (ed.) *Reader's Guide to the Social Sciences*, vol.1, London, Fitzroy-Dearborn

Table 2. Cross section statistics

(2a). Benefit low (C_B=1.5) and low employability (γ_0=0.25)

	(I)	(II)	(III)	(IV)	(V)
	NO STIGMA θ=0, ϕ=0	STIGMA 1 NO NEIGH. EFF. θ^A=0.1, ϕ=1	STIGMA 1 θ^A =0.1, ϕ=1	STIGMA 2 NO NEIGH. EFF. θ^A =0.15, ϕ=1.5	STIGMA 2 θ^A =0.15, ϕ=1.5
% unempl.	25.6	56.6	57.0	22.5	23.2
% assisted	21.9	54.4	54.9	0.9	1.4
take-up rate	100.0	100.0	100.0	5.1	7.2
%search | assist	100.0	24.3	24.1	0.0	0.23
%search | not ass	–	–	–	100.0	100.0
Mean n° assisted neighbours Observed/Random	1.91/1.75	4.44/4.35	4.46/4.35	0.095/0.072	0.17/0.11

(2b). Benefit high (C_B=2.5) and low employability (γ_0=0.25)

	(I)	(II)	(III)	(IV)	(V)
	NO STIGMA θ^A =0, ϕ=0	STIGMA 1 NO NEIGH. EFF. θ^A =0.1, ϕ=1	STIGMA 1 θ^A =0.1, ϕ=1	STIGMA 2 NO NEIGH. EFF. θ^A =0.15, ϕ=1.5	STIGMA 2 θ^A =0.15, ϕ=1.5
% unempl.	59.8	83.3	88.8	83.8	89.2
% assisted	57.8	82.5	88.2	83.0	88.7
take-up rate	100.0	100.0	100.0	100.0	100.0
%search | assist	12.8	3.6	2.2	3.6	2.1
%search | not ass	–	–	–	–	–
Mean n° assisted neighbours Observed/Random	4.65/4..39	6.60/6.27	7.05/6.70	6.64/6.31	7.09/6.74

(2c). Benefit low (C_B=1.5) and high employability (γ_0=0.4)

	(I)	(II)	(III)	(IV)	(V)
	NO STIGMA θ^A =0, ϕ=0	STIGMA 1 NO NEIGH. EFF. θ^A =0.1, ϕ=1	STIGMA 1 θ^A =0.1, ϕ=1	STIGMA 2 NO NEIGH. EFF. θ^A =0.15, ϕ=1.5	STIGMA 2 θ^A =0.15, ϕ=1.5
% unempl.	13.3	14.6	26.3	13.2	13.2
% assisted	9.0	2.2	21.5	0.00	0.00
take-up rate	100.0	21.5	95.2	0.00	0.00
%search | assist	99.0	32.4	43.4	–	–
%search | not ass	–	100.0	100.0	100.0	100.0
Mean n° assisted neighbours Observed/Random	0.92/0.69	0.16/0.17	2.14/1.63	0/0	0/0

(2d). Benefit high (C_B=2.5) and high employability (γ_0=0.4)

	(I)	(II)	(III)	(IV)	(V)
	NO STIGMA $\theta^A=0, \phi=0$	STIGMA 1 NO NEIGH. EFF. $\theta^A=0.1, \phi=1$	STIGMA 1 $\theta^A=0.1, \phi=1$	STIGMA 2 NO NEIGH. EFF. $\theta^A=0.15, \phi=1.5$	STIGMA 2 $\theta^A=0.15, \phi=1.5$
% unempl.	14.3	44.0	44.7	54.7	60.2
% assisted	10.1	41.3	41.9	52.4	58.2
take-up rate	100.0	100.0	100.0	100.0	100.0
%search | assist	84.9	15.5	14.9	10.1	7.6
%search | not ass	–	–	–	–	–
Mean n° assisted neighbours Observed/Random	1.0/0.8	3.4/3.1	3.5/3.2	4.3/4.0	4.7/4.4

Table 3. Longitudinal statistics

(3a). Benefit low (C_B=1.5) and low employability (γ_0=0.25)

	(I)	(II)	(III)	(IV)	(V)
	NO STIGMA θ^A =0, ϕ=0	STIGMA 1 NO NEIGH. EFF. θ^A =0.1, ϕ=1	STIGMA 1 θ^A =0.1, ϕ=1	STIGMA 2 NO NEIGH. EFF. θ^A =0.15, ϕ=1.5	STIGMA 2 θ^A =0.15, ϕ=1.5
			UNEMPLOYMENT SPELLS		
N° spells	21102	14262	14214	21502	21759
50°, 75°, 90° p.	3, 6, 11	3, 10, (n.e)[1]	3, 10, (n.e)[1]	3, 6, 11	3, 6, 11
All cens. from time	22	13	13	45	45
% all cens	2.6	21.0	21.7	0.9	0.7
			WELFARE SPELLS		
N° spells	15325	10530	10448	135	98
50°, 75°, 90° p.	3, 7, 12	4, (n.e), (n.e)	4, (n.e), (n.e)	(n.e.), (n.e), (n.e)	(n.e.), (n.e), (n.e)
highest non cens time (h.n.c..t.)	21	12	12	4	1
survival at h.n.c..t.	3.6	29.2	29.6	97.8	100

[1] Given the number of censored cases, the quantile cannot be estimated.

(3b). Benefit high (C_B=2.5) and low employability (γ_0=0.25)

	(I)	(II)	(III)	(IV)	(V)
	NO STIGMA θ^A =0, ϕ=0	STIGMA 1 NO NEIGH. EFF. θ^A =0.1, ϕ=1	STIGMA 1 θ^A =0.1, ϕ=1	STIGMA 2 NO NEIGH. EFF. θ^A =0.15, ϕ=1.5	STIGMA 2 θ^A =0.15, ϕ=1.5
			UNEMPLOYMENT SPELLS		
N° spells	8557	4505	4670	5988	3937
50°, 75°, 90° p.	3, (n.e.), (n..e.)	(n.e.),(n.e.),(n.e.)	(n.e.),(n.e.),(n.e.)	(n.e.),(n.e.),(n.e.)	(n.e.),(n.e.),(n.e.)
highest n.c. time	6	3	2	3	2
survival at h.n.c.t	26.0	55.0	64.8	54.8	66.5
			WELFARE SPELLS		
N° spells	6291	3429	3640	4549	3051
50°, 75°, 90° p.	3, (n.e.), (n..e.)	(n.e.),(n.e.),(n.e.)	(n.e.),(n.e.),(n.e.)	(n.e.),(n.e.),(n.e.)	(n.e.),(n.e.),(n.e.)
highest non cens time (h.n.c..t.)	5	2	2	2	1
Survival at h.n.c..t.	35.4	72.0	82.4	72.1	85.6

(3c). Benefit low (C_B=1.5) and high employability (γ_0=0.4)

	(I)	(II)	(III)	(IV)	(V)
	NO STIGMA θ^A =0, ϕ=0	STIGMA 1 NO NEIGH. EFF. θ^A =0.1, ϕ=1	STIGMA 1 θ^A =0.1, ϕ=1	STIGMA 2 NO NEIGH. EFF. θ^A =0.15, ϕ=1.5	STIGMA 2 θ^A =0.15, ϕ=1.5
			UNEMPLOYMENT SPELLS		
N° spells	23639	23373	21092	23807	25645
50°, 75°, 90° p.	2, 3, 6	2, 3, 16	2, 4, 8	2, 3, 6	2, 3, 6
Highest n.c. time	30	21	17	–	–
survival at h.n.c.t	0.08	0.74	5.2	–	–
			WELFARE SPELLS		
N° spells	13569	615	10448	1	1
50°, 75°, 90° p.	2, 4, 6	4, (n.e), (n.e)	2, 5, (n.e.)	(n.e.), (n.e), (n.e)	(n.e.), (n.e), (n.e)
highest non cens time (h.n.c..t.)	30	11	15	–	–
survival at h.n.c..t.	0.14	26.5	10.2	–	–

(3d). Benefit high (C_B=2.5) and high employability (γ_0=0.4)

	(I)	(II)	(III)	(IV)	(V)
	NO STIGMA θ^A =0, ϕ=0	STIGMA 1 NO NEIGH. EFF. θ^A =0.1, ϕ=1	STIGMA 1 θ^A =0.1, ϕ=1	STIGMA 2 NO NEIGH. EFF. θ^A =0.15, ϕ=1.5	STIGMA 2 θ^A =0.15, ϕ=1.5
			UNEMPLOYMENT SPELLS		
N° spells	28743	18052	16914	13060	13480
50°, 75°, 90° p.	2, 3, 6	2, 4, (n.e.)	2, 4, (n.e.)	2, 5, (n.e.)	2, 5, (n.e.)
highest non cens time (h.n.c..t.)	15	6	6	5	5
Survival at h.n.c..t.	0.6	14.6	15.5	21.9	25.6
			WELFARE SPELLS		
N° spells	16730	10671	9921	7734	8013
50°, 75°, 90° p.	2, 4, 6	2, 5, (n.e.)	2, (n.e.), (n.e.)	3, (n.e.), (n.e.)	3, (n.e.), (n.e.)
highest non cens time (h.n.c..t.)	14	5	5	4	4
Survival at h.n.c..t.	1.1	24.7	25.5	36.9	42.9

On Emergence of Money in Self-organizing Micro-Macro Network Model

Masaaki Kunigami[1], Masato Kobayashi[2], Satoru Yamadera[2], Takao Terano[2]

[1] Graduate School of Business Science, University of Tsukuba
[2] Computational Intelligence and Systems Science Tokyo Institute of Technology
kunigami@gssm.otsuka.tsukuba.ac.jp, masato@gssm-tsukuba.org,
satoru_yamadera@ybb.ne.jp, terano@dis.titech.ac.jp

Abstract. We are conducting the research on the emergence of the money from the barter economy. This paper presents a new model, which consists of a micro-macro doubly structural network reflecting individual recognitions and social connections among agents. This model will show processes in which particular one of goods attains natures of money as a self-organization of the network. We examine this process by mean-field approximation of dynamics and agent-based simulation.

1 Introduction

A kind of goods called "money" plays a role dominant as the unique medium for exchange in economy and society, even though that has little practical advantage. Researching on mechanism of 'emergence of money' in the society will bring important knowledge for not only historical study on economics but also construction on E-money or local money.

Like Menger and Polani, many economists have discussed function and emergence of money. In addition, recently, there are literatures approaching on this issue by mathematical models (Kiyotaki & Wright[1], Wright [2], Luo [3], Starr [4]) and Agent-based Simulation (ABS) models (Yasutomi [5], Shinohara & Gunji[6], Yamadera & Terano[7]. On the other hand, studies on the natures of social connections among agents (Watts & Strogatz [8], Barabási & Albert [9]), have progressed greatly (Newman [10]) in this decade. Matsuyama et al. [11] research the p2p economy of network between contents-type goods by heterogeneous double-layered network model. Matsuyama et al. [11]

We study on the emergence of money by new model that is composed of a micro-macro doubly structural network and of mutual learning process in the network. In the doubly structural network model, each agent has own micro-level networks each of which represents the corresponding agent's inner recognitions on exchangeability between goods. On the other hand, the macro-level network reflects the social connection among agents. We analyze the contact process in this model by two different approaches i.e. a dynamics with mean-field approximation and a multi-agent simulation.

2 The model

2.1 Doubly Structural Network

We present a model of contact process represents mutual learning on media of exchange between agents on social network. However different from other contact process model, state variables of each agent have own network structure. State variables of each agent represent its inner recognition network on exchangeability of goods. (Fig.1)

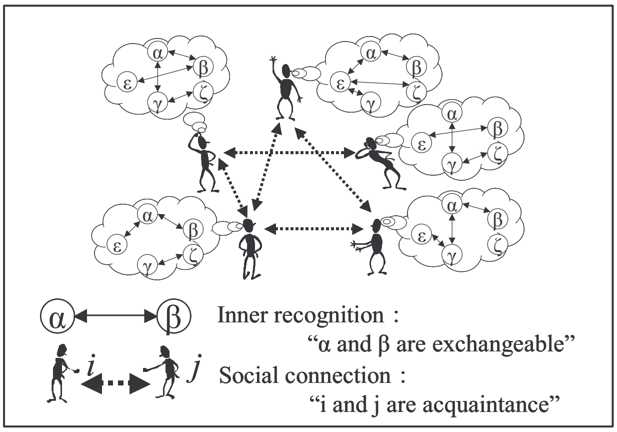


Fig. 1. Doubly Structural Network.

Our model consists of the doubly structural network as follows. At first, we assume that this inter-agent network is given fixedly and expresses the topology of social connection among agents. Node of the social network represents agent and is identified by suffix i or j (i,j=1~N). On the other hand, each agent's inner-network reflects the recognition of the agent, so that each node represents each one of goods that is identified by suffix α, β or γ (1~M). In the inner-network of the i-th agent, the existence of an edge between α and β represents that the i-th agent recognizes the exchangeability of α and β (i.e. agent-i is possible to exchange α and β with anyone). We express this state as (α, β)i=1, and otherwise (α, β)i=0. We mention that these inner-networks of agents change with mutual learning on the inter-agents network.
Different from other social agents models by Epstein & Axtell[12] (sugar-space) or Yamadera & Terano[7], social relations in our model are directly described by inter-agent (macro level) network without using a cellular space structure. Our inter-agent network can represent more general inter-agent relations, and allows us to apply methodologies and outcomes on complex networks to the emergence of money.

2.2 Emergence of Money

Among possible definitions for emergence of money, we focus on the "general acceptability" that is usually defined as exchangeability to any kind of goods of any agent. In the time evolution of our model, if a kind of goods becomes the unique one that has "general acceptability", we can mention that this emerges as money.

In our doubly structural network model, the process in which goods α attains general acceptability is represented by the self-organization in which almost inner-networks become similar star-shaped graph with common hub (The node which connects with almost other nodes) α. (Fig.2)

(Although Starr pointed out that star-shaped network can represent general acceptability of commodity Starr [4], our doubly structural network model can explicitly handle evolutional processes that star-shaped inner networks spread out in the inter-agent network depending upon its graph topology.)

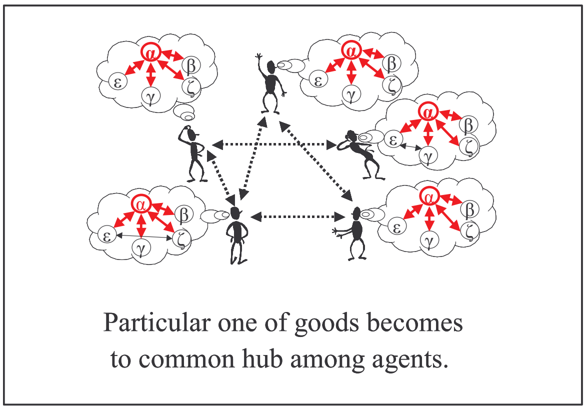


Fig. 2. Emergence of money ~ Self-organization.

2.3 Interaction of Agents

Agents in our model interact each other by the following manners in each time step.

1. Exchange: In the social (inter-agent) network, neighboring agents i and j exchange commodities α and β, if both of them recognize that α and β are exchangeable i.e. (α, β)i=(α, β)j=1. Then this α-β exchange brings reward to i and j with probability P_E.
2. Learning: Learning process of agents consists of the following four ways.
   - Imitation: If an agent i has no recognition of α-β exchangeability (i.e. (α, β)i =0) and if i's neighboring agents j and j' get reward in the α-β exchange, then the agent i will imitate its neighbor's recognition (i.e. (α, β)i=0 → =1) by the probability P_I.
   - Trimming: If an agent 'i' has cycle recognition of exchangeability (ex. (α, β)i=(β, γ)i=(γ, α)i=1), then the agent i will trim its inner-network by cutting randomly one of these cycle edges (i.e. changing one of these three values to 0)

by the probability P_T. Such elimination of circulation means thrift of deliberation and is also consistent with Kiyotaki [1].

We consider that these two processes are essential for 'emergence of money'. In addition, we introduce two more subsidiary processes to avoid local trapping or over-sensitivity on initial conditions.

- Conceiving: Even if an agent i has no recognition of α-β exchangeability (i.e. $(\alpha, \beta)i = 0$), it will happen to conceive this exchangeability (i.e. $(\alpha, \beta)i \to 1$) by the probability P_C.
- Forgetting: Vice versa, even if an agent i has recognition of α-β exchangeability (i.e. $(\alpha, \beta)i = 1$), it will happen to forget this exchangeability (i.e. $(\alpha, \beta)i \to 0$) by the probability P_F. (Despite Japanese Government's endeavors, the exchangeability of Two-thousand yen bill has been becoming almost forgotten in the last decade.)

Although these probabilities will be not state variables but constant data in the model, their values can be dependent on kinds of goods (i.e. $P_E^{(\alpha, \beta)i}$ is not always equal to $P_E^{(\alpha, \gamma)i}$). To simplify the notation, we sometimes omit superscripts$^{(\alpha, \beta)}$ or subscripts$_{\alpha, \beta}$.

2.4 Micro-Macro Interactions

In the last of this section, we emphasize that this model can represent the micro-macro interactions as follows. Adjacency relation and topology of the macro inter-agent network determines interaction between micro inner networks. Similarity between self-organizing inner networks affects the rates of inter-agent interactions.

In this paper, we simplify that the macro inter-agent network is static, but it is natural to extend the macro inter-agent network to dynamic one. In the next paper, we will discuss a micro-macro co-evolutionary model with the doubly structural network.

3 Emergence Scenario in Mean-field Dynamics

3.1 Mean-field Dynamics

To obtain a macro level view of emergence behavior, we derive a dynamics (differential equation system) by a mean-field approximation. In this approximation, states around any agent are approximated by mean of whole agents and nature of social (inter-agent) network is represented by only the average degrees (using constant k) of node. The state variables $x_{\alpha\beta}(t)$ (α=1~M-1, β=α+1~M) represents population ratio of agents that have recognition of α-β exchangeability (inner edge between α and β).

$$\frac{dx_{\alpha\beta}(t)}{dt} = P_C^{(\alpha\beta)} - \left(P_C^{(\alpha\beta)} + P_F^{(\alpha\beta)} + \sum_{\gamma \neq \alpha,\beta} P_T^{(\alpha\beta\gamma)} x_{\alpha\gamma} x_{\beta\gamma} \right) x_{\alpha\beta} \tag{1}$$
$$+ P_E^{(\alpha\beta)} P_I k(k-1) x_{\alpha\beta}^2 - P_E^{(\alpha\beta)} P_I k(k-1) x_{\alpha\beta}^3$$
$$= P_E P_I^{(\alpha\beta)} k(k-1) \Gamma_{\alpha\beta}\left(x_{\alpha\beta}\right).$$

$$\Gamma_{\alpha\beta}\left(x_{\alpha\beta}\right) \equiv b_{\alpha\beta} - a_{\alpha\beta}(\Sigma xx) x_{\alpha\beta} + x_{\alpha\beta}^2 - x_{\alpha\beta}^3 , \tag{2}$$
$$a_{\alpha\beta}(\Sigma xx) \equiv \left(\frac{P_C^{(\alpha\beta)} + P_F^{(\alpha\beta)} + \sum_{\gamma \neq \alpha,\beta} P_T^{(\alpha\beta\gamma)} x_{\alpha\gamma} x_{\beta\gamma}}{P_E P_I^{(\alpha\beta)} k(k-1)} \right), \; b_{\alpha\beta} \equiv \frac{P_C^{(\alpha\beta)}}{P_E P_I^{(\alpha\beta)} k(k-1)} .$$

This mean-field dynamics (1) describes time evolution of these state variables (: population ratios) derived from interaction manners mentioned in previous section. Behavior of the dynamics is determined by the sign (positive or negative) and zero points of RHS, so that we introduce a cubic functions $\Gamma_{\alpha\beta}(x)$ that is normalized function of RHS (2). Of course sign (positive or negative) and zero points of RHS and $\Gamma(x)$ are same. By some manipulations for $\Gamma(x)$, it is shown that all zero points of $\Gamma(x)$ (at least 1 ~ at most 3 points) exist in the region: [0,1].

3.2 Emergence Scenario

Equilibria of our model with mean-field approximation are given by the stable fixed point of the mean-field dynamics, i.e. $\{x \mid \Gamma(x)=0, \Gamma'(x)<0\}$.A state in which almost agents accept only particular one of goods (ex. α) is represented as a state of split zero points i.e. stable zero points of $\Gamma(x)_{\alpha,\mathrm{anyother}}$ are about 1 on the other hand stable zero points of others are about 0. Except for trivial cases in which initial condition of x(t) or fixed points or shapes of $\Gamma(x)$ have distinct difference, this mean-field dynamics has an interesting emergence scenario described by following steps. (Fig.3)

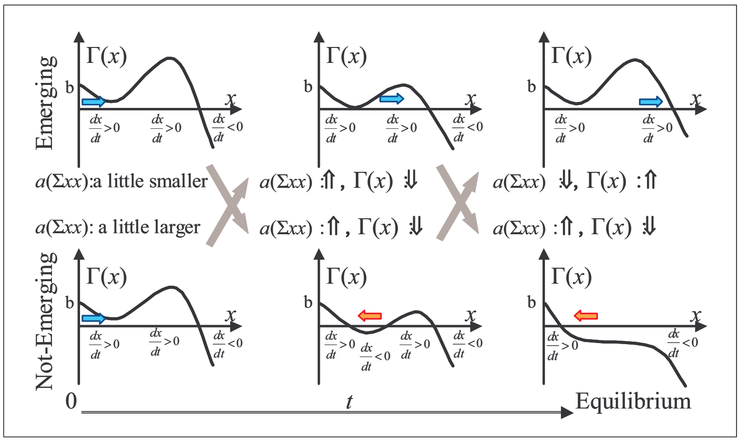


Fig. 3. Process of the emergence scenario.

(i) At first, there are little difference in initial condition of x(t) and in coefficients(shapes and zeros) of Γ(x), so that x(t) start the growth from small values.
(ii) According x(t)s' growth, values of a(Σxx) also increase, so that corresponding Γ(x)s shift downward. Then bottom values of some Γ(x)s go below 0, so that corresponding x(t) are changed to decreasing.
(iii) Decreasing x(t) pushes up others' Γ(x)s through a(Σxx), and these increasing Γ(x)s shift own zeros and corresponding x(t) to right side. Vice versa, increasing x(t) shifts other x and zeros of Γ(x) s to the left. At last, differentiation of zeros (i.e. some equilibrium points go to 1 and others go to 0.) is grown acceleratively.
To work this emergence scenario, it is necessary (not sufficient) that some Γ(x) s have bottoms and 3 zeros. (Fig.4)

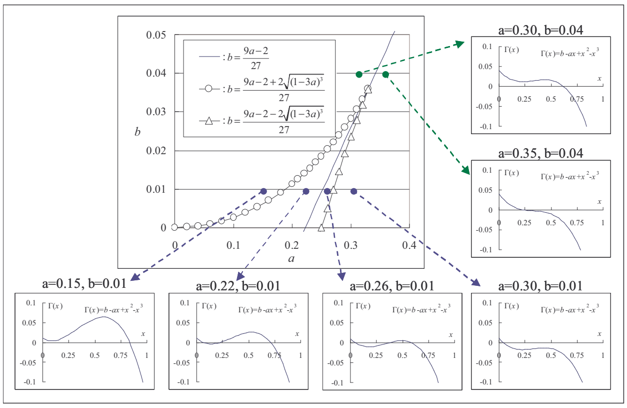


Fig. 4. The region Γ(x) has 3 zeros : between O and Δ.

The numeric example of emergence in the corresponding parameter range is shown in Fig.5

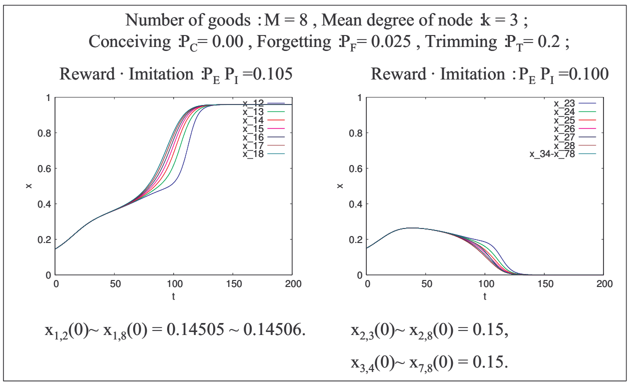


Fig. 5. A numeric example of the emergence. (with small P_C, P_F)

4 Agent-Based Simulation

4.1 Agent Model

The mean-field dynamics is an applicable approach to find analytically the existence and sketchy behavior of emergence scenario. However, mean-field approximation requires strong assumptions that situation around particular agent and natures of the inter-agents networks can be substituted by the global mean of agents and the mean degree of nodes.

Then this approach has little effectiveness to research a locally heterogeneous system or complex networks that have long tailed or specific distributions of node degree.

Here we apply agent-based simulation (ABS) approach to explore emergence of money grounding on micro level interaction without referring the macro information by micro agents.

Over existent models our advantages are to handle a complex inter-agent network explicitly and be able to completely ground micro interaction. This simulation model implements straightforwardly assumptions and situations on the goods and agents in previous section.

4.2 Simulation

We use output matrix M that is total sum of adjacency matrices of inner-networks of recognition. Inside an agent, the status in which a goods α has exchangeability to all other goods is represented by the agent's adjacency matrix that has the value 1 on the α-th row and on the α-th column and otherwise has value 0.

Hence, the status in which anyone recognizes a goods α has exchangeability to all other goods is represented by the output matrix **M** that has the value of total population on the α-th row and on the α-th column and otherwise has value 0.
Samples of our simulation (Fig.6, Fig.7) illustrates that emergence of money is possible with some sets of parameters on different types of social networks. In the subsequent papers, we will discuss more detailed studies on emergence and network.

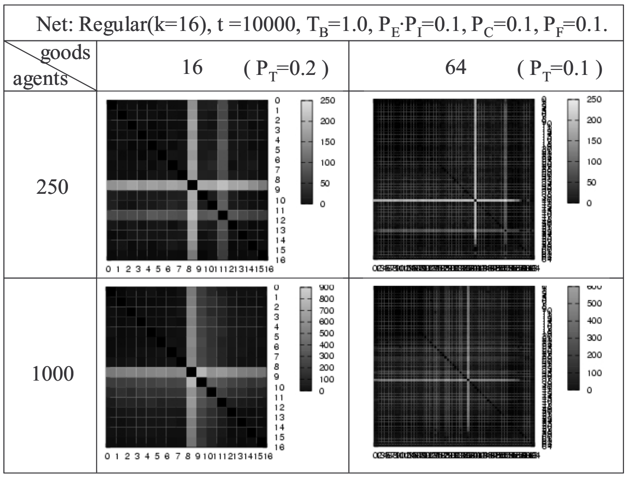


Fig. 6. The ABS can show the emergence with wide ranges of the population and the number of goods.: Social network is Regular(Watts & Strogatz [8]). (The output matrix M. White means large population of agents.)

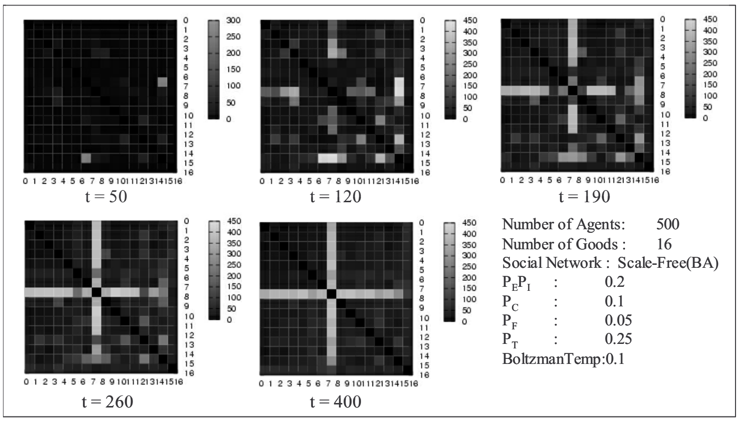


Fig. 7. The ABS can show a time sequence of the emergence.: Social network is Scale-free(Barabási & Albert [9]). (Time evolution of the output matrix M. White means large population of agents.)

5 Summary and Conclusion

In this paper, we present micro-macro doubly structural network model and illustrate the new model is applicable to both analytical and ABS approaches on emergence of money.

The advantages of the presented model are next two points; (i) it can describe a complex network structure of real society or goods, and (ii) it allows dynamical analysis by both differential equations and simulations. Especially by introducing explicitly structured inter-agent network, an important research area has been opened on the relation between the nature of social networks and emergence of money.

In addition, the doubly structural network presents not only a new model of 'emergence of money' but also the new model framework in which each node of the global network is connected to each of interacting local networks that make emergence of some global natures.

Reference

[1] Kiyotaki, N., Wright, R.: On money as a medium of exchange, Journal of Political Economy, vol. 97(1989), pp. 927-54.

[2] Wright, R.: Search, evolution, and money, Journal of Economic Dynamics and Control, vol. 19(1995), pp. 181-206.

[3] Luo, G.,Y.: The evolution of money as a medium of exchange, Journal of Economic Dynamics and Control vol. 23(1999), pp. 415-458.

[4] Starr, R.: Why is there money? Endogenous derivation of ‘Money’ as the most liquid asset: a class of examples, Economic Theory vol. 21(2003), 455–474.

[5], Yasutomi, A,: The emergence and collapse of money, Physica D vol. 82(1995), pp. 180-194.

[6], Shinohara, S., Gunji, Y., P.: Emergence and collapse of money through reciprocity, Applied Mathematics and Computation, vol. 117(2001), pp. 131-150.

[7] Yamadera, S., Terano, T.: Examinig the Myth of Money with Agent-Based Modeling, in Klaus G. Troitzsch ed. “Rpresenting Social Reality” - Pre-Proceedings of ESSA 2005(2005), pp. 258-265.

[8] Watts, D.J., Strogatz, S.H.: Collective dynamics of “small-world” networks, Nature, vol. 393(1998), pp. 440–442.

[9] Barabási, A. L., Albert, R.: Emergence of scaling in random networks, Science, vol. 286(1999), pp. 509–512.

[10] Newman, M. E. J.: The Structure and Function of Complex Networks, SIAM REVIEW, vol. 45(2003), pp. 167–256.

[11] Matuyama.S, Kunigami,M., Terano,T.: Analysis of the hub contents though Agent-based simulation, JSAI2007.

[12] Epstein, J. S., Axtell,R.: Growing Artificial Societies, The Brookings Institution, 1996.

Modelling proximity effects on industrial district competitiveness

Vito Albino[1], Nunzia Carbonara[1], Ilaria Giannoccaro[1]

[1]DIMEG, Politecnico di Bari, viale Japigia 186, 70126 Bari, Italy
v.albino@poliba.it, ncarbonara@poliba.it, ilaria.giannoccaro@poliba.it

Abstract. This paper investigates how proximity affects the Industrial Districts (ID) competitiveness by taking into account four dimensions of proximity, namely geographical, organizational, cognitive and social. The aim is to identify the optimal combinations of proximity that fosters ID competitiveness. The complexity science is adopted as both theoretical and methodological approach. Complexity theory is used to identify whether the proximity matter for the ID competitiveness by analyzing the influence of the proximity on the ID adaptive capacity. Our argument is that the more adaptive the ID, the more the competitive success. Based on the complexity theory, we identify the structural features that allow ID adaptation and their best values. Then, by developing a computational model based on the Business Dynamics, we carry out a simulation analysis to evaluate the influence of the proximity on the values of the ID structural variables. We look for the best combinations of proximity providing results coherent with the optimal theoretical values suggested by the complexity theory.

1 Introduction

The literature on Industrial Districts (IDs) presents several definitions of ID each of them emphasizes different and complementary aspects characterizing IDs (Becattini, 1992; Markusen, 1996; Porter, 1998). Recently, studies in the fields of the economic geography and the regional economics adopt the proximity concept to define and study IDs and put in light how proximity contributes to the innovative performance of IDs (Boschma, 2005). When the proximity concept is used, what is often actually meant is geographical proximity. However, besides geographical proximity, other dimensions of proximity are also investigated, namely cognitive, organizational, and social proximity, and their relevance for the innovative performance of IDs are put in light (Asheim et al., 2004; Filippi and Torre, 2003; Nooteboom, 1999). Foregoing studies present some limitations: some of them focus on the influence of only one dimension of proximity on ID innovation capability (Davenport, 2005; Wuyts et al., 2005); others compare the effects of the geographical proximity with just another proximity dimension (Breschi and Lissoni, 2006; Oerlemans and Meeus, 2005). Thus, they do not consider the complex effects of all the dimensions of proximity on the ID competitiveness. Furthermore, from a methodological point of view, they mainly adopt a case study approach rather than a wide statistical analysis, so that provide outcomes that can not be extended to other contexts.

In this paper we investigate how proximity affects the ID competitiveness by taking into account all the four dimensions of proximity, namely cognitive, geographical, organizational, and social. Our aim is to identify the optimal combinations of proximities that foster ID competitiveness. To this aim, we adopt the complexity science as both a theoretical and a methodological approach.

In fact, first complexity science is used as a conceptual framework to investigate the reasons for the ID competitive success. In particular, by considering the ID as a complex adaptive system (CAS), we identify three structural features of IDs that give them the adaptive capacity and thus their competitive success. The value that the key structural features of the ID should assume to improve the adaptive capacity are also identified based on the complexity theory, i.e. a moderate level of interconnectivity, a high level of heterogeneity, and a moderate level of control.

To investigate the effect of proximity on the ID competitiveness, we analyze the influences of the four dimensions of proximity, namely cognitive, geographical, organizational, and social, on the three ID structural features. If the proximity influences the values of the structural features so as to have them coherent with the identified optimal values assuring the highest adaptive capacity, the ID will be competitive.

To analyze the effect of proximity, we develop a simulation model by using the Business Dynamics (Sterman, 2000). The model represents the complex system of interactions existing among the proximity dimensions and the ID structural features. Such a model is first developed at a conceptual level by identifying cause-effect relations between the proximity and the ID structural features based on the literature. Then, the computational model is developed and simulation carried out.

Simulation analysis is conducted to evaluate the influence of the proximity on the values of the ID structural variables. At the end of simulation we look for the best combinations of proximity providing results coherent with the optimal theoretical values suggested by the complexity science.

The paper is organized as follows. First, the definitions of proximity dimensions are summarized and the ID discussed as a nexus of proximities. The complexity science is briefly presented in Section 3 where the main ID structural features influencing adaptive capacity are identified and their optimal values also defined. Section 4 presents the conceptual model describing the influence of the proximity on the ID structural variables. In Section 5 the computational model based on the Business Dynamics is presented. Finally, the simulation analysis is discussed in Section 6.

2 The proximity dimensions and the Industrial districts

Several definitions of the ID are proposed in the related literature, each of them emphasizes different and complementary aspects characterizing IDs (Becattini, 1992; Markusen, 1996; Porter, 1998) and all of them evidence that the ID is the locus in which firms and individuals are characterized by different dimensions of proximity, namely geographical, organizational, cognitive, and social proximity. In the follow we explain why each proximity dimension characterizes IDs.

Geographical proximity: IDs are formed by co-localized firms, that is firms that are located in the same geographical bounded area, therefore they are characterized by high geographical proximity. For examples, the Italian IDs are mostly concentrated in one or few provinces, covering areas that may range from 100 to 200 squared kilometres. In the U.S., clusters are spread within areas that range from very large areas, is this the case of Silicon Valley, to very narrow areas, as in the case of Hollywood's entertainment cluster.

Organizational proximity: IDs are formed by small and medium sized firms highly specialized on one or more phases of a production process, and integrated through a dense network of inter-firm relationships. The contemporary presence of competition and cooperation is one of the most important feature of the inter-organizational networks within IDs. Furthermore, the recent evolution of some IDs has carried out the hierarchization of the network structure (Corò and Grandinetti, 1999). Therefore, IDs are characterized by a certain level of organizational proximity. Low organizational proximity, when the relationships are basically aimed at market transaction (Carbonara et al., 2002). Medium organizational proximity, when the dense network of inter-firm relationships involves both cooperative and competitive relations. Where the strong competition occurs between firms doing the same things; co-operation and collaboration, especially over technical innovation and design, takes place between firms doing different things, i.e. at different stages of the production process (Brusco, 1992). High organizational proximity, when the IDs is characterized by the presence of hierarchical networks or strong ties, ranging from contractual agreement to mutual co-operation.

Cognitive proximity: IDs are characterized by the agglomeration of firms specialized in a particular field, by a very high workers mobility, by a specialized set of services, including also the training and educational services, tailored to the unique products/industries of the ID, and by an extensive inter-firm networks supporting frequent and repeated knowledge sharing. All these ID features increase the level of cognitive proximity among the actors belonging to an ID. Firms within IDs tend to exhibit similar mental models and competitive behaviour (Pouder and St. John, 1996). However, IDs often include manufacturers of complementary products, companies operating in complementary industries, and public/private research organizations that are sources of new knowledge. Also, IDs are characterized by a strong rivalry among similar firms that motivates them to differentiated each others. The above two ID features act as forces that reduce the cognitive proximity within and ID.

Social proximity: According to Becattini (1992), an ID is a combination of a population of firms and a local community, where community and firms tend to merge. In particular, the local community is characterized by a relatively homogeneous system of values and views and the population of firms is not just an aggregate of productive units. In fact, personal relationships – such as kinship and friendship - between entrepreneurs and other employees operating in different companies are present in the ID. Therefore IDs are generally characterized by a high level of social proximity. Some factors may reduce the social proximity, such as the entry of foreign companies in the ID, the acquisition of ID companies from external companies, generational changes, and the hiring of not local managers.

Once we recognize that IDs are characterized by different dimensions of proximity, it seems particularly interesting to investigate the effect of the proximity dimensions on the ID competitiveness.

3 The complexity science and the Industrial District competitiveness

Complexity science studies CAS and their dynamics. CASs consist in evolving network of heterogeneous, localized and functionally-integrated interacting agents. CASs have adaptive capability and co-evolve with the external environment, modifying it and being modified (Axelrod and Cohen, 1999; Gell-Mann,1994; Lane, 2002). Complexity science offers a number of new insights that can be used to seek new dynamic sources of competitive advantage. In fact, application of complexity science to organization and strategy identifies key conditions that determine the success of firms in changing environments associated with their capacity to self-organize and create a new order, learn and adapt (Levy, 2000; Mitleton-Kelly, 2003). In line with the works above, complexity science is used in our study to identify what conditions of IDs enable them to adapt to external environment. Therefore, the basic assumption of this study is that IDs are CASs, given that they exhibit different properties of CASs, such as the existence of different agents (e.g. firms and institutions), non-linearity, different types of interactions among agents and between agents and the environment, distributed decision making, decentralized information flows, and adaptive capacity (Albino et al., 2005). In what follows assuming that IDs are CASs and based on the CAS theory, the ID structural features that affect the adaptive capacity of ID are defined. Moreover, the value that the key structural properties of ID should assume to improve their adaptive capacity are also identified, i.e. a moderate level of interconnectivity, a high level of heterogeneity, and a moderate level of control.
In so doing, ID competitive advantage is not the result of a set of pre-defined features characterizing IDs, but it is the result of the adaptability process of IDs with the external environment.

3.1 The key ID structural features influencing adaptation

Interconnectivity. CAS theory identifies the number of interconnections within the system as a critical condition for self-organization and emergence. By using the concept of tunable landscape and the *NK* model, Kauffman (1995) demonstrates that the number of interconnections among agents (K) influences the adaptive capacities of the system. As K increases, the rate of adaptation decreases. Therefore, in order to assure the adaptation of the system to the landscape, the value of K should be moderate. This result has been applied in organization studies to modeling organizational change and technological innovation (Levinthal, 1997; Rivkin and Siggelkow, 2002). Referring to IDs, the level of interconnectivity can be associated to the economic and social links among the ID firms. Higher the links among firms, higher the probability that firm behavior is affected by the behavior of the other firms. Thus, applying the result of the complexity to the ID, the level of interconnectivity of the ID structure measured as number of links among ID firms should be moderate to enhance adaptive capacity.

Heterogeneity. Different studies on complexity highlight that variety destroys variety. For example, Ashby (1956) suggests that successful adaptation requires a system to have an internal variety that at least matches environmental variety. Systems having agents with appropriate requisite variety will evolve faster than those without. Dooley (2002) states that one of the main properties of a complex system that supports the evolution is diversity. Such a property is related to the fact that each agent is potentially unique not only in the resources that it holds, but also in terms of the behavioral rules that define how it sees the world and how it reacts. In a complex system diversity is the key towards survival. Without diversity, a complex system converges to a single mode of behavior. Referring to ID firms, the concept of agent heterogeneity can be related to the competitive strategy of firms. This results from the resources that firm possesses and defines the behavior rules and the actions of firms in the competitive environment (Grant, 1998). Thus, to enhance the adaptive capacity, the heterogeneity of ID firms should be high.

Level of control. The governance of a system is another important characteristic influencing CAS self-organization and adaptive behaviors. Le Moigne (1990) observes that CASs are not controlled by a hierarchical command-and-control center but instead manifest a certain form of autonomy. The latter is necessary to allow evolution and adaptation of the system. The level of control in IDs is determined by the governance of the ID's organizational structure. The higher the degree of governance, the higher the level of control exerted by one or more firms on the other ID firms (Storper and Harrison, 1992).

Applying the concepts of complexity theory to the ID, a medium level of control of the ID organizational structure is required to improve the ID adaptive capacity.

4 Proximity and structural features of the industrial districts

4.1 The conceptual model

Based on theoretical and empirical studies presented in the literature we develop a conceptual model that explains the influences of the four dimensions of proximity, namely cognitive, geographical, organizational, and social, on the three structural features of ID.

Geographical proximity. *Geographical proximity increases the level of interconnectivity among firms.* A large body of literature claims that firms that are spatially concentrated have more face-to-face contacts at reasonable costs therefore they can increase the frequency of their relationships keeping low the cost of interactions (Marshall, 1920). Surely, one of the main factor that affect positively on the reduction of the interaction cost is the short distance between co-located firms that keep low the transportation costs (Weber, 1920).

Geographical proximity has a twofold effect on the level of heterogeneity. In fact, it both increases and decreases the level of heterogeneity in a system of firms. Geographical proximity decreases the level of heterogeneity because it favours with different mechanisms the knowledge sharing among firms. In fact, firstly, regular face-to-face contacts existing among co-located firms favour the exchange of tacit knowledge. Secondly, firms in a geographical cluster have access to a highly specialized labour market pooling. The reason for this is that firms can use two sources of workforces: 1) firms can recruit workers from local educational institutions which provide the training that is locally requested, 2) a geographical concentration of technologically related firms creates into a cluster a pool of highly specialized and experienced skills. Another mechanism that guarantees the knowledge transfer within a system of co-located firms is the high mobility of workers that characterized IDs. With this regard, Saxenian (1994) states that engineers in Silicon Valley shifted between firms so frequently that mobility is not only socially acceptable but it is the norm; it is rare for a technical professional to have a career in a single company. Finally, it is widely recognized in the literature that geographical proximity increases the effectiveness of knowledge spillovers, both unintentional and intentional (Torre and Gilly, 2000; Breschi and Lissoni 2001).

The opposite effect that geographical proximity exerts on the level of heterogeneity is due to the strong rivalry among firms in geographical clusters (Porter, 1990), which motivate firms to continuously improve their productivity and innovation performance, so differentiating each others.

Geographical proximity does not have a direct influence on the level of control among firms.

We will show in the next that geographical proximity influences the social proximity and then affects the common culture and the trust, so reducing the transaction costs and in turn decreasing the level of control among firms (Williamson, 1999).

Cognitive proximity. *Cognitive proximity has a twofold effect on the level of interconnectivity. In fact, it both increases and decreases the level of interconnectivity in a system of firms.* If people and/or firms share the same knowledge base and expertise, they can successfully communicate, understand and process the stock of knowledge that is transferred among them. Thus, cognitive proximity increases the level of interconnectivity among firms because it reduces the communication costs facilitating effective and efficacy communication (Boschma and Lambooy, 1999). On the other hand, cognitive proximity may decrease the level of interconnectivity because firms can share few complementary capabilities when they interact with similar firms, so reducing the potential for learning and the possibility to exploit new knowledge (Cantwell and Santangelo, 2002; Nooteboom, 2000). Thus, firms characterized by a small cognitive distance could be not interested in establishing relationships among them.

Cognitive proximity decreases the level of heterogeneity in a system of firms. Cognitive proximity among firms increases their absorptive capacity, that is the capacity of a firm to effectively identify, interpret and exploit the knowledge owned by the other firm (Cohen and Levinthal, 1990). In other words, when the cognitive distance between firms is rather small knowledge spillovers increase and in particular the risk of involuntary knowledge spillovers increases. In such circumstances, the level of firms heterogeneity decreases.

Cognitive proximity decreases the level of control among firms. Cognitive proximity increases the effectiveness and efficiency of knowledge and information transfer process, thus it reduces the uncertainty of communication (Nooteboom, 2000). This in turn decreases the transaction costs, so fostering firms to establish market-based relationships rather than hierarchy-based relationships (Williamson, 1999).

Organizational proximity. *Organizational proximity increases the level of interconnectivity among firms.* According to Torre and Gilly (2000), organizational proximity refers to actors that are close in organizational terms belonging to the same space of relations (networks), i.e. actors are in interactions of various natures. Therefore greater is the organizational proximity in a system of firms, higher is number of relationships among them and greater is the level of interconnectivity.

Organizational proximity decreases the level of heterogeneity in a system of firms. Organisational proximity between agents, that is the existence of a system of formal relations, such as alliances, agreements, partnerships, etc. is a vehicle that enables the transfer and exchange of information and knowledge (Cooke and Morgan, 1998). This, in turn decreases the level of heterogeneity in a system of firms. This is coherent with the other definition of organizational proximity proposed by Torre and Gilly (2000). The authors state that actors close in organizational terms are quite alike, that is, they have the same reference space and share the same knowledge.

Organizational proximity increases the level of control among firms. Organizational proximity can be defined by the rate of autonomy and the degree of control that can be exerted in organizational arrangements. A continuum is assumed from one extreme, i.e. low organizational proximity, meaning no ties between interdependent actors, as in the pure market, to the other extreme, i.e. high organizational proximity, embodied in strong ties, as in a hierarchically organized firm or network. In the middle there are loosely coupled networks, namely weak ties between autonomous entities (Boschma, 2005). Therefore, higher is the organizational proximity higher is the degree of control in a system of firms.

Social proximity. *Social proximity increases the level of interconnectivity among firms.* The notion of social proximity is based on the idea that economic relations are to some extent always embedded in a social context (Granovetter 1985). In other words, firms that are socially proximate are interconnected not only by economic relationships but also by means of friendship and kinship relations. The latter foster frequent face-to-face contacts (Breschi and Lissoni, 2006), so increasing the level of interconnectivity among the firms.

Social proximity decreases the level of heterogeneity in a system of firms. The literature suggests that the more socially embedded are the relationships of a firm, the more is the tacit knowledge shared among them (Boschma, 2005). One reason is that social proximity increase the frequency of face-to-face contacts that facilitate the exchange of tacit knowledge. The other reason is that social proximity includes situations in which agents share a common culture, namely the same system of values, norms, habits, languages. Thus, it ensures effective tacit knowledge transmission.

Social proximity decreases the level of control among firms. A shared system of norms, values, a common culture and language, the system of that trust-based social relationships, characterizing situations with high social proximity are all factors that reduces the risk of opportunistic behaviour. This in turn decreases the transaction costs (Dasgupta, 1998), so fostering firms to establish market-based relationships rather than hierarchy-based relationships (Williamson, 1999)

Influences between proximity dimensions. In the foregoing, it emerges that influences exist among the different dimensions of proximity. The modeling of these influences is coherent with the literature on this topic.

In particular, the conceptual model includes the following influences:

1. Geographical proximity both increases and decreases the cognitive proximity. In fact, on one side, geographical proximity increases the effectiveness of knowledge spillovers, favours the exchange of tacit and codified knowledge, reduces the communication costs, so doing it facilitates the knowledge diffusion and then it reduces the cognitive distances among co-located firms (Freel, 2003; Marshall, 1920). On the other side, the strong rivalry characterizing co-located similar firms leads them to differentiate each others by continuously developing new knowledge that is incorporated into products and processes (Porter, 1990).
2. Geographical proximity increases the social proximity. Spatial proximity facilitates trust relationships based on friendship and kinship and stimulates the development of both a common culture and a shared set of norms and values (Harrison, 1992).
3. Organizational proximity increases the cognitive proximity. According to Nooteboom (2000), strong links among agents (both individuals and organizations) reduce the cognitive distance among them. Conversely, in loosely coupled networks the cognitive maps of the agents have a greater knowledge differentiation.

5 The Business Dynamics model

This Section presents the computational model based on Business Dynamics, which represents the complex system of interactions existing among the proximity dimensions and the ID structural features. Main elements of a Business Dynamics model are: stocks, flows, auxiliary variables and links (Sterman, 2000).

Stocks are accumulations and characterize the state of the system. Inflows and outflows are rates that modify the value of the stock variables. Auxiliary variables are introduced to represent further variables influencing the dynamics. Finally, the links model the influences between the variables.

In the next the computational model is described, first by identifying the variables and then by discussing the block diagrams that model the influences of all the proximity on the three ID structural features. The computational model is developed by using VensimPLE (Sterman, 2000).

5.1 Variables

The model consists in six stock variables, namely the geographical distance, the cognitive proximity, the not shared knowledge, the knowledge spillovers, the organizational proximity, and the number of interconnections.

The *geographical distance* is modelled as a stock with an outflow represented by a rate of approach of firms. It is defined as the average physical distance among firms measured in kilometres.

The *cognitive proximity* is a stock with an outflow represented by the rate of new knowledge creation. It is defined as the common knowledge that firms shared. The rate of new knowledge creation is an inflow for the stock “*Not shared knowledge*” that decreases via the outflow rate of knowledge diffusion. In turn this is an inflow for the stock “*Knowledge spillovers*”. The *organizational proximity* is a stock that increases by the rate of organizational arrangements. It is defined as the ratio between the strong ties among firms and the total ties.

The *number of interconnections* is a stock that increases via the rate of attachment among firms and decreases by the rate of detachment. It is defined as the total number of links among firms.

The other variables presented in the conceptual model are modelled as auxiliary variables.

6 Simulation analysis

Simulation analysis is carried out to analyze the influence of the proximity on the ID structural features. The baseline model is used to simulate the ID creation through the increase of geographical proximity (agglomeration process) and for different values of organizational proximity in order to take into account different types of IDs. According to our model, cognitive and social proximity depend on the geographical and organizational proximity. At the end of simulation we measure the value of the three structural features and we compare these values with the optimal values as suggested by the complexity science. The fit with the optimal values means that proximity matter for the ID competitiveness.

6.1 The baseline model

The baseline model corresponds to the case of a typical ID. The values of the coefficients are assumed by referring to the literature and to qualitative data available thanks to a field research conducted on IDs in Southern Italy (Carbonara, 2002; Carbonara et al, 2002). The baseline model is used to simulate the formation of the ID by decreasing the geographical distance from 2000 to 50 Km. When geographical distance is minor than 500 Km we assume that an ID is formed. Three values of organizational proximity are considered, i.e. 0.04; 0.49; 0.94.

The results of the simulation consist in curves mapping the values of the ID structural features (interconnectivity, heterogeneity, and level of control) as the geographical distance increases. To identify whether the values of the structural features are coherent with the optimal ones, we divide their range into three intervals: *low* from 0 to 0.33, *medium* from 0.34 to 0.66, and *high* from 0.67 to 1.

Results are depicted in Figure 1. Interconnectivity varies with geographical distance. As geographical proximity increases, it first decreases and then increases, but its value always remains at a medium level. Higher levels of interconnectivity are achieved when the organizational proximity is higher. The level of the heterogeneity decreases as the geographical proximity increases. The rate of decrease is influenced by the level of organizational proximity: higher the organizational proximity, higher the rate of decrease. The geographical distances at which the heterogeneity becomes medium are equal to 390, 510, 630 for low, medium, and high values of the organizational proximity, respectively. The level of control decreases as the geographical proximity rises. Moreover, for medium and high level of organizational proximity it assumes very low values. For a low organizational proximity, it assumes the optimal value for the values of the geographical distance between 50 and 980 Km.

Based on these results, the region where all the ID structural variables assume the optimal values is defined by medium levels of geographical proximity (between 390 and 500 Km[1]) and by a low level of organizational proximity (0.04). In this region the cognitive and the social proximity assume a medium values (see Figure 2). Notice that we have carried out other experiments by increasing the organizational proximity to identify the highest value of organizational proximity that is permitted in the system, namely that assure the optimal value for the level of control. Organizational proximity can increase only up to 0.06.

Therefore, IDs characterized by *medium* geographical proximity, *low* organizational proximity, *medium* cognitive and *medium* social proximity are more competitive given that they enhance their adaptive capacity. Based on the above, we also advance the following propositions:

P1. When ID firms become too close, their adaptive capacity will be reduced.

This is confirmed by recent studies that evidence the risk of too much proximity, which determines different types of lock-in: spatial, cognitive, social, and organizational (Boschma, 2005). According to Pouder and St John (1996), when firms become too spatially close they lose their innovative capacity and are less adaptable to new developments. Following Nooteboom (2000), too much cognitive proximity means a lack of sources of novelty and decreases the potential for learning. Uzzi (1997) states that too much social proximity may have adverse impact on learning and innovation. Finally, too much organizational proximity is detrimental because reduces interactive learning and flexibility of the system (Boschma, 2005).

P2. The organizational proximity among ID firms should be low so as increasing the adaptive capacity.

It seems that because in the IDs there is the contemporary presence of geographical, social, and cognitive proximity, when the organizational proximity increases, through for example processes of hierarchization, the entire system loses its ability to learn and create new knowledge, and then its adaptive capacity.

---

[1] Notice that we limit the upper bound to 500 Km because we assume that the maximum value of ID geographical distance is 500Km.

7 Conclusions

The importance of proximity to explain the competitive success of Industrial Districts (ID) is a topic emerging in the literature. What has been put in evidence is that the geographical proximity beside the organizational, the cognitive, and the social proximity enhance the innovation performance of ID firms. While previous studies have mainly focused on the influence of only one dimension of proximity or in a few cases have compared the effects of geographical proximity with just another proximity dimension, we have investigated how all the four dimensions of proximity affect the ID competitiveness. In particular, the aim has been to identify the optimal combinations of proximities that foster ID competitiveness.

This aim has been pursued by using the complexity science as theoretical and methodological approach. According to complexity science, we have identified the structural features that allow the adaptation of the complex adaptive systems (CASs) and their optimal values, namely a medium level of interconnections, a high level of heterogeneity, and a medium level of control. Assuming that IDs are CASs, we have argued that the IDs that would foster their adaptive capacity, should exhibit the optimal values for the identified structural features.

To study whether the proximity matters for ID competitiveness, we have analyzed the influence of proximity on the values of the structural features and thus, as a consequence, on the ID adaptive capacity. If the influence determines values coherent with the optimal ones, we can conclude that proximity matters.

The methodology we used is the Business Dynamics. The model has been first developed by identifying the causal loops among the variables based on the literature and then has been coded by using the Business Dynamics software into stocks, flow, and auxiliary variables and links. The simulation analysis has been then carried out to identify the optimal combinations of proximity. Results show that medium levels of geographical, cognitive and social proximity and low levels of organizational proximity assures the highest adaptive capacity. This result confirms recent studies that evidence the risk of too much proximity that can determine various types of lock-in: spatial, cognitive, social, and organizational (Boschma, 2005).

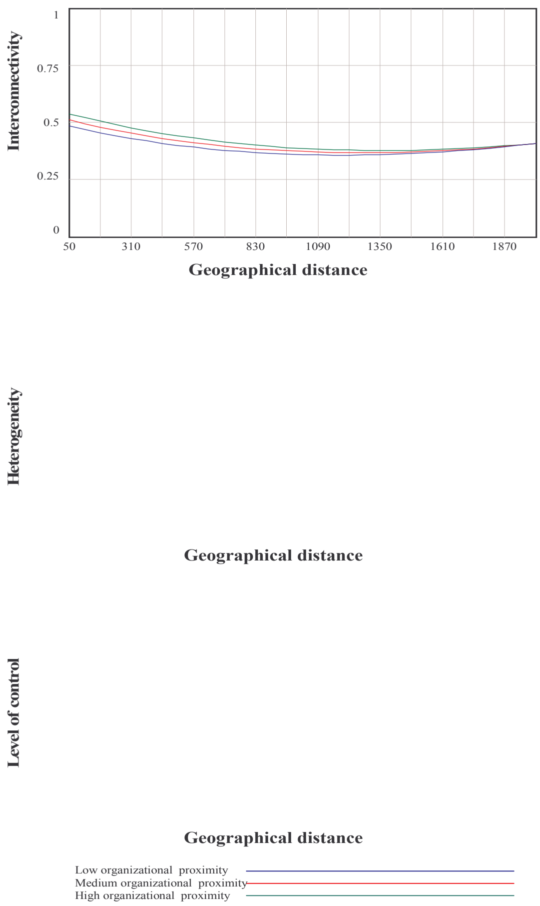


Fig. 1. Results of the baseline model.

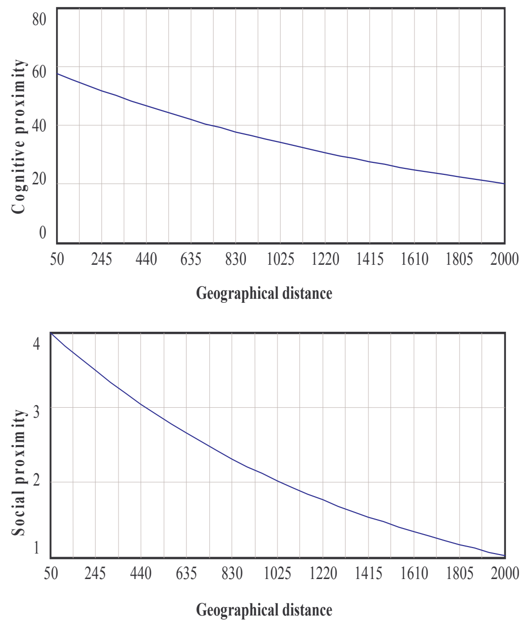


Fig. 2. The values of the cognitive and the social proximity.

References

Albino, V., Carbonara, N., Giannoccaro I., (2005), "Industrial districts as complex adaptive systems: Agent-based models of emergent phenomena", in C. Karlsson, B. Johansson and R. Stough (eds.), *Industrial Clusters and Inter-Firm Networks*, 73-82, Edward Elgar Publishing.

Ashby, W.R., (1956), *An Introduction to Cybernetics*. Chapman & Hall: London.

Asheim, B.T., Coenen, L., Moodysson, J., (2004), "Nodes, networks and proximities: on the knowledge dynamics of the Medicon Valley biotech cluster", European Planning Studies, 12 (7), 1003-1018.

Axelrod, R., Cohen, M.D., (1999), *Harnessing Complexity: Organizational Implications of a Scientific Frontier*. The Free Press: New York.

Becattini, G., (1992), "The Marshallian industrial district as a socio-economic notion", in G. Becattini, F. Pyke and W. Sengenberger (eds.), *Industrial districts and Inter-firm co-operation in Italy,* 37-52, International Institute for Labour Studies: Geneva.
Boschma, R., (2005), "Proximity and innovation: a critical assessment", Regional Studies, 39 (1), 61-75.
Boschma, R. A. , Lambooy, J. G., (1999), "Evolutionary economics and economic geography", Journal of Evolutionary Economics, 9, 411–429.
Breschi, S., Lissoni, F., (2001), "Knowledge spillovers and local innovation systems: a critical survey", Industrial and Corporate Change, 10 (4), 975-1005.
Breschi, S., Lissoni, F., (2006), "Mobility of inventors and the geography of knowledge spillovers. New evidence on US data", Paper presented at the Eiasm Workshop on Complexity and Management, Oxford, 19-20 June.
Brusco, S., (1992), "The idea of the industrial district: Its genesis", in G. Becattini, F. Pyke and W. Sengenberger (eds.), *Industrial districts and Inter-firm co-operation in Italy,* 10-19, International Institute for Labour Studies: Geneva.
Cantwell, J., Santangelo, G. D., (2002), "The new geography of corporate research in Information and Communication Technology (ICT)", Journal of Evolutionary Economics, 12, 163–197.
Carbonara, N., Giannoccaro, I., Pontrandolfo, P., (2002), "Supply chains within industrial districts: a theoretical framework", International Journal of Production Economics, 76 (2), 159-176.
Carbonara, N., (2002), "New models of inter-firm network within industrial district", Entrepreneurship and Regional Development, 14, 229-246.
Cohen, W.M., Levinthal, D.A., (1990), Absorptive capacity: a new perspective on learning and innovation, Administrative Science Quarterly, 35 (1), 128-152.
Cooke, P., Morgan, K., (1998), *The associational Economy. Firms, Regions, and Innovation,* Oxford, Oxford University Press.
Corò, G., Grandinetti, R., (1999), "Evolutionary Patterns of Italian Industrial Districts", Human Systems Management, 18.
Dasgupta, P., (1998), "Economic development and the idea of social capital", in P. Dasgupta and I., Serageldin (eds.), *Social Capital. Integrating the economist's and the sociologist's perspective*, 1-48, World Bank, Washington D.C..
Davenport, S, (2005), "Exploring the role of proximity in SME knowledge-acquisition", Research Policy, 34 (5), 683-701.
Dooley, K.J., (2002), "Organizational Complexity", in M. Warner (ed.), *International Encyclopedia of Business and Management,* Thompson Learning: London.
Filippi, M., Torre, A., (2003), "Local organisations and institutions. How can geographical proximity be activated by collective projects", International Journal of Technology Management, 26, 386–400.
Freel, M. S., (2003), "Sectoral patterns of small firm innovation, networking and proximity", Research Policy, 32, 751–770.
Gell-Mann, M., (1994), *The quark and the jaguar*. Freeman & Co: New York.
Granovetter, M.S., (1985), "Economic action and social structure: the problem of embeddedness", American Journal of Sociology, 91 (3), 481-510.
Grant, R.M., (1998), *Contemporary Strategy Analysis. Concepts, Techniques, Applications*. Blackwell: Oxford.
Harrison, B., (1992), "Industrial districts: old wines in new bottles", Regional Studies, 26, 469–483.
Kauffman, S.A., (1995), *At home in the universe: the search for laws of self-organization and complexity*. Oxford University Press: New York.

Lane, D., (2002), "Complexity and Local Interactions: Towards a Theory of Industrial Districts, Complexity and Industrial Districts", in A.Q. Curzio and M. Fortis (eds.), *Complexity and Industrial Clusters*, Physica-Verlag: Heidelberg.

Le Moigne, J. L., (1990), *La Modélisation des Systèmes Complexes*. Dunod : Paris.

Levinthal, D.A., (1997), "Adaptation on Rugged Landscapes", Management Science, 43, 934-950.

Levy, D.L., (2000), "Applications and Limitations of Complexity Theory in Organization Theory and Strategy", in J. Rabin, G. J. Miller and W.B. Hildreth (eds.), *Handbook of Strategic Management,* Marcel Dekker: New York.

Markusen, A., (1996), "Sticky places in Slippery Space: A typology of Industrial Districts", Economic Geography, 72(3), 293-313.

Marshall, A., (1920), *Principles of Economics*. Macmillan: London.

Mitleton-Kelly, E., (2003), "Ten Principles of Complexity and Enabling Infrastructures", in E. Mitleton-Kelly (ed.), *Complex Systems And Evolutionary Perspectives Of Organisations: The Application Of Complexity Theory To Organisations*, Elsevier.

Nooteboom, B., (2000), *Learning and Innovation in Organizations and Economies,* Oxford, Oxford University Press.

Noteboom, B., (1999), "Innovation and inter-firm linkages: new implications for policy", Research Policy, 28 (8), 793-805.

Oerlemans, L.A.G., Meeus, M.T.H., (2005), "Do organizational and spatial proximity impact on firm performance?", Regional studies, 39 (1), 89-104.

Porter, M., (1990), *The Competitive Advantage of Nations.* Macmillan, London.

Porter, M., (1998), "Clusters and the new economics of competition", Harvard Business Review, 76, 77-90.

Pouder, R., St. John, J. C., (1996), "Hot spots and blind spots: geographical clusters of firms and innovation", Academy of Management Review, 21, 1192–1225.

Rivkin, J.W., Siggelkow, N.J., (2002), "Organizational Sticking Points on NK Landscape", Complexity, 7 (5), 31-43.

Saxenian, A., (1994), *Regional Advantage: Culture and Competition in Silicon Valley and Route 128.* Harvard University Press, Cambridge, MA.

Sterman, J., 2000, *Business Dynamics. System Thinking and Modeling for Complex Work*, Irwin/McGraw-Hill.

Storper, M., Harrison, B., (1992), "Flessibilità, gerarchie e sviluppo regionale: la ristrutturazione organizzativa dei sistemi produttivi e le nuove forme di governance", in F. Belussi and P. Bianchi (eds.), *Nuovi modelli d'impresa gerarchie organizzative e imprese rete*, F. Angeli: Milan.

Torre, A., Gilly, J.P., (2000), "On the analytical dimension of proximity dynamics", Regional Studies, 34 (2), 169-180.

Uzzi, B., (1997), "Social structure and competition in interfirm networks: the paradox of embeddedness", Administrative Science Quarterly, 42, 35–67.

Weber, A., (1920), *Theory of the Location of Industries*. University of Chicago Press, Chicago.

Williamson, O. E., (1999), *The Economics of Transaction Costs.* Edward Elgar, Cheltenham.

Wuyts, S., Colombo, M.G., Dutta, S, Nooteboom, B., (2005), "Empirical tests of optimal cognitive distance", Journal of Economic Behavior & Organization, 58 (2), 277-302.

Session on Opinion and Cultural Dynamics

Drifting to more extreme but balanced attitudes: Multidimensional attitudes and selective exposure

Diemo Urbig[1] and Robin Malitz[2]

[1] Humboldt-Universität zu Berlin, School of Management and Economics and Department of Computer Science, diemo@urbig.org
[2] Humboldt-Universität zu Berlin, Computer- und Medienservice, malitzro@cms.hu-berlin.de

Abstract. We present a model of opinion and attitude dynamics that incorporates multi-dimensionality of a single attitude, i.e. attitudes are composed of evaluations of several features. We include selective exposure that is modeled as a hierarchical version of the bounded confidence model. Individuals adapt their feature evaluations if their attitudes and the respective feature evaluation are sufficiently close. We show that this model can explain a balancing principle, such that features associated with the same object tend to be evaluated similarly. We also show that such individuals can get more extreme despite they might even interact only with agents that have less extreme attitudes.

1 Introduction

An 'attitude is a psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor' (Eagle & Chaiken, 1993). If the 'entity' is a behavior, then an attitude can be considered as an individual's tendency to engage in a specific behavior. Thus, attitudes have been at the heart of models that predict behavior of people. Behaviors of interest are, e.g., buying decisions by consumers, voting behavior, or the adoption of new technologies by firms. Hence, there are many models that describe and explore processes of attitude formation. While some models focus on how single individuals form their opinions in response to specific stimuli, other models focus on social processes of opinion formation. This article focuses on how attitudes evolve in groups of interacting individuals.

Simple models on interacting individuals and resulting changes in attitudes were studied by Festinger (1950), French (1956), Latané et al. (1994), Deffuant et al. (2000), Hegselmann and Krause (2002), Galam (2002), and others. A basic assumption is that interacting and especially communicating individuals mutually affect their attitudes. Regarding communicating individuals there has also been some work on explaining why individuals do not tell others their true attitudes (e.g., Kuran, 1987). Thus, there might be a gap between an individual's attitude, as the mental state, and his expressed opinion as element of communication. As Urbig (2003) argues such differences might also be caused by limitations inherent in languages or mechanisms used for communication. Because we are not concerned with biases of the attitude–opinion relation we use the terms opinion and attitude as synonyms throughout this text. Furthermore, these communication processes are moderated by selective exposure, which describes a mechanism where people tend to maintain their beliefs by exposing themselves to information supporting their beliefs.

Models of attitude and opinion dynamics as introduced by Nowak et al. (1990), Hegselmann and Krause (2002), and Deffuant et al. (2000) represent some of the more simplistic approaches; but they allow for a clear analysis and understanding of the fundamental dynamics. A more complex model of consumer behavior involving attitudes has been introduced by Jager (2000). From my point of view, taking the simple models as starting point and increasing complexity step by step enables us to keep track of the sources of complexity and of specific dynamics and thus allow to include only those elements that have large effects (see Hegselmann, 2004). Most simple models of opinion dynamics in groups of interacting individuals consider an individual's opinion as a one-dimensional value, either binary as good or bad opinion, or with a finite number of stages between the extremes, or as a continuum between extreme values (see Hegselmann & Krause, 2002, Weisbuch et al., 2002). Sometimes such models also include uncertainty about the opinion (see for example Amblard et al., 2003). The idea of selective exposure is included in these models as bounded confidence or relative agreement, i.e. individuals only others if their attitudes towards the object under consideration are sufficiently close.

One-dimensional approaches to attitudes and opinions are beneficial for understanding basic dynamics, but it is questionable whether they can capture the complex dynamics displayed by opinion dynamics in real settings. For instance, when analyzing the diffusion of innovative products or the acceptance of complex programs of political parties it is somewhat heroic to assume that people either agree on all aspects or on no aspect, as it would be implied by one-dimensional approaches. According to consumer theory and psychological attitude theory a one-dimensional approach seems to be inappropriate. According to this research stream an attitude is composed of different *impressions* that themselves are composed of two elements: *beliefs* (also called *cognitions*) about the presence of some attributes and *evaluations* of these attributes (Ajzen, 1991). In fact, we end up with a qualitative distinction between knowledge and evaluations and a multidimensionality regarding a set of attributes related to an entity. This multidimensionality is at the core of most of the reliable approaches to attitude measurement (Ajzen, 2002). While the distinction between beliefs and evaluations might be questioned, even competing approaches to measurement of attitudes keep the multidimensional nature (see Trommsdorff, 1998). Measuring attitudes gets important when linking simulations to real measured data. Altogether, multidimensionality, but also – to a smaller extend – the distinction between evaluations and cognitions seem to be very general though relevant concepts. Although such a compositional approach to attitudes finally results in a one-dimensional attitude, we will see that an explicit modeling of the different dimensions enables a modeling of a more differentiated and from our point of view more realistic communication process.

There are several attempts to include multi-dimensionality into simple models of opinion dynamics. But unfortunately the multi-dimensionality is then interpreted as a set of opinions instead of several elements of the same opinion or attitude. Deffuant et al. (2000), Weisbuch et al. (2002), as well as Fortunato et al. (2005) consider a vectors of binary values as a set of opinions toward different objects. Lorenz (2003) investigates a model of n-dimensional vectors of real values, where the real values represent the amount of money allocated to n projects. In Lorenz's model agents interact and adapt their opinions if the Euclidian distance of these opinions is smaller than a constant, which is a possible extension of the Bounded confidence concept by Hegselmann and Krause (2002) and similarly by Fortunato et al. (2005). These models assume an all-or-nothing strategy for adapting to other's opinions, which means that either all dimensions are adapted or no dimension. This is not plausible from my point of view because individuals might be willing to adapt in one dimension while at the same time they strongly resist any adaptation in another dimension.

Furthermore, for determining the condition for an effective interaction agents evaluate their whole vectors of opinions and not only their own but also the whole vector of their current communication partner. From my point of view this extensive comparison of all elements is a strong weakness of these models because it requires an extensive interaction even with those partners that differ extremely; one needs this extensive comparison just to figure out that they actually differ to such a large degree. It seems more plausible that there is some comparable aggregate or placeholder (aggregated evaluation or signal for group membership in cultural models) that finally determines if individuals get involved into a deeper discussion or reject any discussion and therefore get away being unaffected by each other without having had a deep interaction on the mutual differences.

In this article we want to incrementally develop a model of dynamics of structured attitudes that gets closer to psychological attitude theory, i.e. to the model of Fishbein and Ajzen (Ajzen, 1991). Still we are aware that this will only be an inspiration by psychological theory but not an implementation of the theory on attitudes, given that a consistent theory exists at all. Our general approach, which can be labeled as *hierarchical selective exposure* or *hierarchical bounded confidence* approach, can be described as follows:

> *Having first just a general interaction about the attitude regarding an object under consideration an interaction continues only in case of a sufficient closeness. If they continue they may talk about the different attributes related to the object. At this level they may again be bounded in their predisposition to consider their peer's opinion seriously.*

All together, there are different interpretations of vectors of opinions, but as far as we know there are only two interpretations that refer to structured attitudes or opinions, where the attitude is an aggregate of multiple dimensions that might be adapted independently, namely Urbig and Malitz (2005) and Deffuant and Huet (2007). The first paper is on an agent-based simulation of a preliminary version of the model presented here. The latter article introduces a model that assumes that a 'filter' selects only important features with a higher threshold of importance when the attitude about the feature is incongruent with the global attitude about the object. Furthermore, individuals transmit only features that are congruent with their global attitude. It turns out that in such cases the sequence of perceiving information is highly significant for the overall evolution of attitudes. However, this model deviates in it basic structure significantly from the idea of continuous opinion dynamics.

2 Model

Let $a_{i,t} \in \mathbb{R}_{[0,1]}$ be the attitude of agent $i \in \{1..N\}$ regarding a focal object at time t. The attitude is composed of a set of impressions $e_{i,k,t} \in \mathbb{R}_{[0,1]}$, which are in fact evaluated features comprising an evaluative component and a belief component about the presence of some feature. The attitude is a summary evaluation of all features. We normalize the attitude by dividing it by the number of features.

$$a_{i,t} = \frac{1}{n} \sum_{k=1}^{K} e_{i,k,t} \tag{1}$$

In every step t an agent i interacts with a randomly chosen agent j. If they both adapt their attitudes and thus adapt on at least one feature evaluation, then they are considered as interacting effectively. The condition of interacting effectively on feature k is indicated by $ef_k(i,j) \in \{0,1\}$, which gives 1 for effective interaction and 0 otherwise. This function is symmetric, $ef_k(i,j) = ef_k(j,i)$, such that if one adapts also the other adapts. A convergence parameter $0 < \mu_k \leq 0.5$ describes how much agents adapt . For now let it be specific for the different features k. If $\mu_k = 0.5$ then agents' effective interaction on a feature leads to an agreement in the respective feature evaluations. The smaller μ_k is, the less quickly they converge in their feature evaluations. Given two randomly chosen agents i and j, equation 2 defines how agent i updates her feature evaluations and thus indirectly her attitude. The rule is symmetric for agent j. Note that an agent's attitude only changes if the two agents' attitudes are sufficiently close and if their feature evaluations are sufficiently close.

$$\begin{aligned} e_{i,k,t+1} &= e_{i,k,t} - ef_k(i,j) \cdot \mu_k \cdot (e_{i,k,t} - e_{j,k,t}) \\ \text{with } ef_k(i,j) &= \begin{cases} 1\ if\ |a_{i,t} - a_{j,t}| < \varepsilon_a & \wedge \quad |e_{i,k,t} - e_{j,k,t}| < \varepsilon_f \\ 0\ if\ |a_{i,t} - a_{j,t}| \geq \varepsilon_a & \vee \quad |e_{i,k,t} - e_{j,k,t}| \geq \varepsilon_f \end{cases} \end{aligned} \tag{2}$$

3 Get different by getting more alike

A first analysis of this model in an agent-based simulation has shown that there seems to be tendency that agents can get more extreme attitudes although they communicate with agents that have on average less extreme attitudes (Urbig &

Malitz, 2005). However, there was no systematic analysis of this effect. Therefore, we now show that the process described in our model converges, keeps average attitudes and average feature evaluations constant, but attitudes may get more distant through an interaction.

Consider two interacting agents, agent 1 with attitude $a_{1,t}$ and agent 2 with attitude $a_{2,t}$. Without loss of generality let agent 2 have the more optimistic attitude, i.e. $a_{1,t} < a_{2,t}$. Let C contain those features, on which the two agents adapt, i.e. $\forall 1 \leq k \leq K : |e_{1,k,t} - e_{2,k,t}| < \varepsilon_f$.

Because both agents' updating processes are characterized by the same convergence parameter μ_k, the mean of the corresponding impressions keep constant, i.e. $\frac{1}{2}\left(e_{1,k,t+1} + e_{2,k,t+1}\right)$ equals $\frac{1}{2}(e_{1,k,t+1} + \mu_k \cdot (e_{2,k,t} - e_{1,k,t}) + e_{1,k,t+1} + \mu_k \cdot (e_{1,k,t} - e_{2,k,t}))$, which finally translates into $\frac{1}{2}\left(e_{1,k,t} + e_{2,k,t}\right)$. Because the attitude is the average of impressions also the average of all agents' attitudes keeps constant. Although agents can get closer and can get more distant, these dynamics stabilize, which means that they do not move back and forwards for ever.[3]

The distance of two agents' corresponding impressions never increases due to our definition of updating. However, the distance in attitudes can increase. This is not surprising if the updating process was not symmetric for two interacting agents, but in our model it is. One can easily show that the changes in attitudes of two interacting agents have the opposite sign (proof available on request). An increasing distance thus implies that the change in one agents attitude needs to have the opposite sign of the difference between own and the other agent's attitude. The change in attitude of agent i interacting with agent j is given as $\frac{1}{n}\sum_{k\in C}\mu_k\left(e_{j,k,t} - e_{i,k,t}\right)$. Because we assume that $a_{1,t} < a_{2,t}$, we need $0 > \frac{1}{n}\sum_{k\in C}\mu_k\left(e_{2,k,t} - e_{1,k,t}\right)$ for having opposite signs. This implies that the two agents adapt on a feature where the order of their respective impressions is inversed compared to the order of their attitudes. Thus, the move of agent 1 has to have a different sign that the difference between the own and the other agent's attitude.

Our further analysis will focus on a special case. For this we assume that the convergence parameter is equal for all features, i.e. $\mu_k = \mu$. We also look at the case of two features since it is much easier to illustrate than case with more features. For these two instantiations we can derive more precise criteria when agents get more distant while they interact: If the convergence parameter μ is equal for all features, i.e. $\mu_k = \mu$, then agents can only get more extreme if they do not compromise on at least one feature. This is easily show by contradiction:

[3] If agents adapt their feature evaluation towards another agent's evaluation, then they assign the weight $1 - \mu_k$ to their own and μ_k to the other's feature evaluation. Since agents are defined symmetrically, either both interaction partners adapt their evaluations or none of them. Therefore, (1) every agent always assign his own evaluation a weight larger than zero, (2) if an agent assigns a positive weight to another agent's evaluations then this happen vice versa, and (3) all positive weights are always above or equal to $1 - \mu_k$. Hence, according to the stabilization theorem by Lorenz (2005) the process stabilizes in the feature evaluation space. Because the attitude space is just a mapping it stabilizes there as well.

Without loss of generality we assume that the second agent's attitude is larger the first agent's attitude, i.e. $a_{2,t} > a_{1,t}$. Let the convergence parameter be equal for all features, i.e. $\mu_k = \mu$. If all feature evaluations are affected, then we can rewrite the above stated condition as follows $0 < -\sum_{k \in C} (e_{2,k,t} - e_{1,k,t}) = -\sum_{k=1}^{K} (e_{2,k,t} - e_{1,k,t}) = -\sum_{k=1}^{K} e_{2,k,t} + \sum_{k=1}^{K} e_{1,k,t}$. This leads to $a_{1,t} > a_{2,t}$, which contradicts our assumption about the initial relation between these two agents' attitudes. From our analysis so far we can also conclude that if there are only two features, then two agents can only get more distant if they adapt on only one of the two features and this displays an inverse order compared to the attitudes.

4 Drifting to more extreme though balanced attitudes

Getting more distant by compromising is just a possibility and this effect might be wiped out in a population of interacting agents. Therefore we now look at the population level.

For analyzing our model in groups of interacting individuals there are several methods that could be applied. One is agent-based simulation that simulates populations of single agents, i.e. in every step one knows the attitudes of every single agent. We already did agent-based simulation on a similar model as introduced here and have recognized that exploring the full parameter space requires huge computational efforts (see Urbig & Malitz, 2005). Therefore we decided to numerically solve rate equations for analyzing the dynamics (Lorenz, 2006, calls this approach Interactive Markov Chains). Thus, we indirectly assume a population of infinite many agents, i.e. $N = \infty$, and in every step every agent gets a randomly chosen interaction partner and all could adapt in every step. The deviations have an impact, but it is rather small (see for instance for other models of the same class Deffuant & Huet, 2007 and compare Deffuant et al., 2000 with Ben-Naim et al., 2003).

We focus on a two-dimensional model. We define a discrete space of equidistant impressions or feature evaluations. They are labeled from 1 to R. The initial probability mass located in the two-dimensional feature evaluation space at point i, j is given by $F_0(i, j)$ with $1 \leq i, j \leq R$. The two-dimensional feature evaluation space is mapped into the one-dimensional attitude space. Because the feature evaluations are discrete, also the attitude space is discrete with $2 \cdot R - 1$ levels. The probability mass in this space is given by $A_0(i)$ with $1 \leq i \leq 2R - 1$. The mapping from the feature evaluation space onto the attitude space is given by (3). It represents an averaging of the feature evaluations.

$$A_t(i) = \begin{cases} \sum_{x=1}^{i} F_t(x, i - x + 1) & \text{if } i \leq R \\ \sum_{x=i}^{R} F_t(x, R - x + 1) & \text{if } i > R \end{cases} \tag{3}$$

The left part of Figure 1 illustrates the mapping. For a better understanding we plot the isoquants for the attitudes; along these lines agents have the same attitude. The attitude increases the more one gets to the upper right corner.

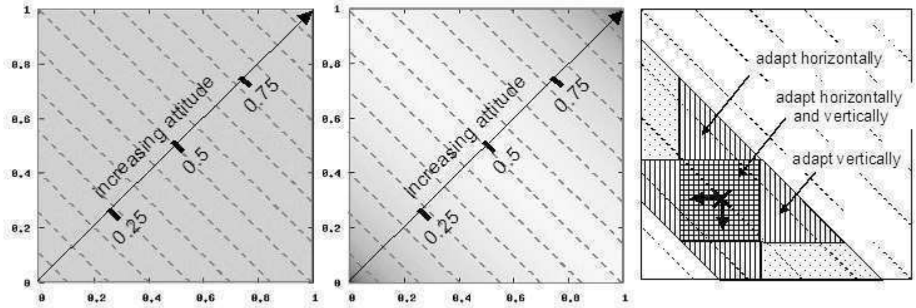


Fig. 1. Left: equally distributed feature evaluations, Center: equally distributed attitudes, Right: an agent's tendency to get more extreme

On can easily see that there are more agents, i.e. a larger probability mass, for intermediate attitudes. The equally distributed feature evaluations translate into a triangular distribution in attitudes.

If we want to compare our model with one-dimensional models that assume initially equally distributed attitudes, then we have to make an equivalent assumption on our initial distribution. We thus manipulate the initial distribution in the feature evaluation space such that at the attitude level agents are equally distributed. The required distribution is illustrated in the middle plot of Figure 1. We need a distribution that emphasizes those areas that imply more extreme attitudes. Our analysis first considers equally distributed feature evaluations to be compatible with traditional models of multidimensional opinions and later compares it with the second case of equally distributed attitudes.

We now define the rate equation for our model. Remember, we assume two dimensions and agreement on all features where agents interact effectively, i.e $K = 2$ and $\mu_k = \mu = 0.5$. We assumed that feature evaluations are discrete. Thus, if two agents compromise they might want to compromise on a value that is exactly in between of two or four of the available values. In such cases we assume that they assign randomly to the respective values. Due to the geometry of the space, only three cases can happen: (1) agents' new feature evaluations can be located at a single cell; (2) they are located on a border between two cells; (3) they are located at a corner, where four cells meet. In the first case we assume that the probability is 1.0 to move into this cell. If it is on a border, then the probability is 0.5 to get into either of two corresponding two cells. For cases when agents would be located at corners the probability is 0.25 to get into one of the four cells. Given this rule let us define two additional functions. Function $d^+(i, j, x_1, y_1, x_2, y_2) \in \{0.25, 0.5, 1.0\}$ gives the probability that an agent maintains feature evaluations according to the cell (x_1, y_1) and moves to cell (i, j) due to an interaction with an agent from cell (x_2, y_2). Furthermore, function $d^-(i, j, x, y) \in \{0.0, 1.0\}$ gives the probability that an agent from cell (i, j) moves somewhere due to an interaction with an agent from cell (x, y). Based on these two functions we now define a rate equation that describes how the distribution of agents within the space of feature evaluations changes from one step to another.

$$\begin{aligned} f_{t+1}(i,j) &= f_t(i,j) + D_t^+(i,j) - D_t^-(i,j) \qquad\qquad \text{with} && (4)\\ D_t^-(i,j) &= \sum_{x=1}^{R}\sum_{y=1}^{R} F_t(x,y)\cdot F_t(i,j)\cdot d^-(i,j,x,y) \\ D_t^+(i,j) &= \sum_{x_1=1}^{R}\sum_{y_1=1}^{R}\sum_{x_2=1}^{R}\sum_{y_2=1}^{R} F_t(x_1,y_1)\cdot F_t(x_2,y_2)\cdot d^+(i,j,x_1,y_1,x_2,y_2) \end{aligned}$$

The sequential calculation of the rate equation is implemented in Java. It either stops after 500 steps or if the sum of all changes is smaller than 10^{-6}. We also calculated arrows that roughly indicate the drifts within the multidimensional space. Figures 2 and 3 illustrate the results of our analysis.

Figure 2 illustrates the distribution in the feature evaluation space for three steps and three different settings. We look at the second setting, where there is in fact no selective exposure at the attitude level. It gets closest to previously analyzed multidimensional dynamics. Because the condition at the attitude level is not restricting the dynamics at all, the dynamics are almost equal to those of two combined one-dimensional dynamics, see for instance the discussion by Fortunato et al. (2005). As Fortunato et al. (2005), we also observe the formation of four equally sized major clusters. Interesting to note, in this scenario agents might get more extreme, but on the population level this effect is wiped out.

If we look at the setting where the selective exposure is restricting the dynamics at both levels, we can indeed observe a drift to feature evaluations, where both features tend to be evaluated similarly and to more extreme attitudes. The difference to the previous case gets obvious in Figure 2. Obviously the selective exposure at the attitude level makes a difference. Right part of Figure 1 illustrates the reason. Agents from the dotted, squared, and striped areas succeed with respect to selective exposure at the attitude level. These restrictions together with the general limits at zero and one create the situation that there is a larger probability to be affected by a more extreme agent at a specific feature evaluation than to be affected by a more moderate agent.

In case that the selective exposure is not restricting the dynamics at the feature evaluation level, we end up with a dynamics that is close to a one-dimensional dynamics. Due to the selective exposure of 0.25 we might expect two clusters, but because the initial distribution is a triangular distribution there is a stronger tendency to form only one large central cluster. However, we observe two small minorities.

While Figure 2 only looks at the distribution in the feature evaluation space and only considers equally distributed initial feature evaluations, Figure 3 plots the evolution of distributions in the attitude space and additionally plots the results for settings, where the initial distribution of attribute evaluations are adjusted such that the initial attitudes are equally distributed.

For the case without a binding restriction at the feature evaluation level we observe a typical evolution of the attitude given that the initial attitudes are

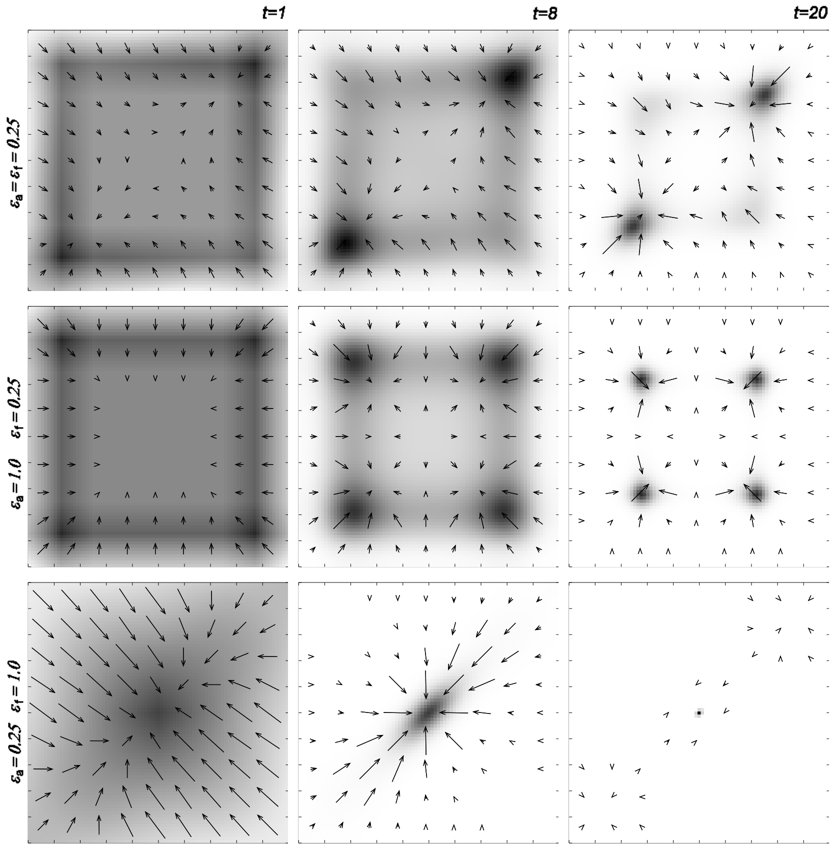


Fig. 2. Evolution of the distribution of attitudes in different scenarios: A: $\varepsilon_a = \varepsilon_f = 0.25$, B: $\varepsilon_a = 1.0$, $\varepsilon_f = 0.25$, and C: $\varepsilon_a = 0.25$, $\varepsilon_f = 1.0$

equally distributed. For the case of no binding restriction on the attitude level we hypothesize that if equally distributed attitudes are assumed, we always end up with equally sized majorities.

In our last analysis (see Figure 4) we take a closer look at the case that the selective exposure is equal on both levels, at the attitude level and at the feature evaluation level, i.e. $\varepsilon_a = \varepsilon_f = \varepsilon$. We vary ε between 0.1 and 0.8. This gives us a much better insight regarding the range of the parameter for selective exposure, where polarization or consensus can be expected. For creating the plot we did some transformations. For each distribution we normalized the values by dividing all cells by the maximum value of all cells. From Ben-Naim et al. (2003) we know that minorities are typically $3 \cdot 10^{-4}$, while majorities are much larger. We use the threshold 0.001 to discriminate between majorities and minorities. Doing this one should keep in mind that majorities can be rather small. And in fact, in some scenarios they get rather small, e.g. the central cluster or majorities

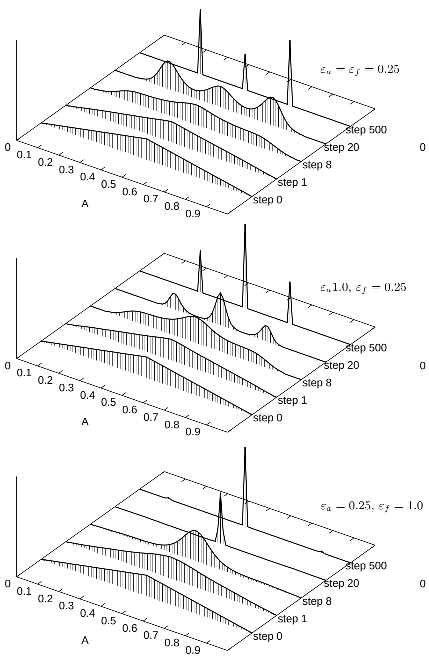


Fig. 3. Evolution of the distribution of attitudes in different scenarios. Left column represents cases where initial evaluations of features are equally distributed and thus attitudes follow a triangular distribution, while second column represents cases with equally distributed initial attitudes.

that are close to the extremes, where already at the beginning of the dynamics there are only small probability masses due to he triangular distribution. Since Figure 4 marks every majority with a dot, these very small majorities look like large majorities. In Figure 4 we observe that for the scenario where selective exposure works on all levels, polarization indeed appears for larger parameter values, $\varepsilon > 0.25$. If we consider equally distributed initial attitudes we even get polarization for some $\varepsilon > 0.35$.

5 Conclusion

In this paper we extended simple models of the social formation of attitude by the aspect of multidimensionality of single attitudes. We introduced the idea of hierarchical selective exposure, which means that when selecting interaction partners and issues first select at a rather general level and in case of selection they step-wise go down in the hierarchy to more specific partners or issues. This model removes a rather heroic assumption of previous models of multidimensional opinion dynamic models, which assume that before selecting a partner both partners have perfect knowledge about all of the other's beliefs and atti-

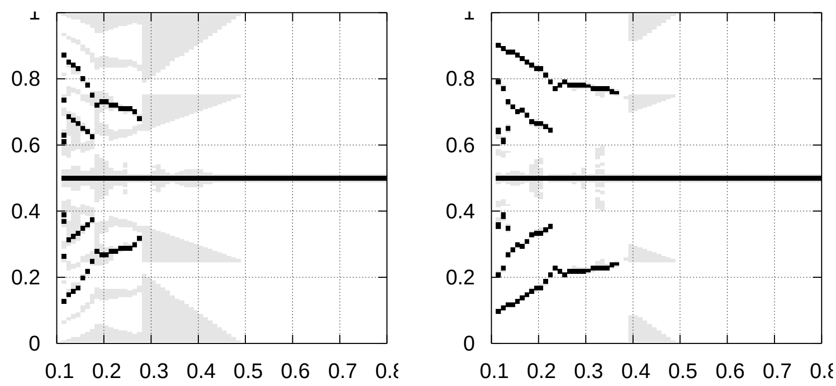


Fig. 4. Bifurcation diagrams for varying $\varepsilon_a = \varepsilon_f$ from 0.1 to 0.8. Left: initially equally distributed feature evaluation, Right: initially equally distributed attitudes

tudes. We have demonstrated that our extension can explain a systematic drift to more extreme attitudes. Polarization that was previously usually not expected for parameter of selective exposure beyond 0.25 can occur systematically for values way beyond this threshold.

For future work one could split the concept of impressions regarding features into beliefs about the presence of features and evaluations of these features. Agents might interact regarding both of these aspects. Agents could affect their beliefs without necessarily affecting their values assigned to specific attributes. Furthermore, applying this idea to more complex interaction rules, e.g. relative agreement model, would probably be of interest.

Acknowledgments. Many thanks to Sylvie Huet and Guillaume Deffuant for their very helpful suggestions regarding the model and its analysis.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and Human Decision Processes*, *50*, 179–211.

Ajzen, I. (2002). Attitudes. In R. F. Ballesteros (Ed.), *Encyclopedia of psychological assessment* (Vol. 1, pp. 110–115). London: Sage Publications.

Amblard, F., Weisbuch, G., & Deffuant, G. (2003). *The drift to single extreme appears only beyond a critical connectivity of the social networks.*

Ben-Naim, E., Krapivsky, P., & Redner, S. (2003). Bifurcations and patterns in compromise processes. *Physica D*, *183*(190).

Ben-Naim, E., Krapivsky, P., Vazquez, F., & Redner, S. (2003). Unity and discord in opinion dynamics. *Physica A*, *330*, 99–106.

Deffuant, G., & Huet, S. (2007). Propagation effects of filtering incongruent information. *Journal of Business Research*, *60*(8), 816–825.

Deffuant, G., Neau, D., Amblard, F., & Weisbuch, G. (2000). Mixing beliefs among interacting agents. *Advances in Complex Systems*, *3*, 87–98.

Eagle, A. H., & Chaiken, S. (1993). *The psychology of attitude.* Orlando, FL: Harcourt Brace Jovanovich.

Festinger, L. (1950). Informal social communication. *Psychological Review*, *57*, 271–282.

Fortunato, S., Latora, V., Pluchino, A., & Rapisarda, A. (2005). Vector opinion dynamics in a bounded confidence consensus model. *International Journal of Modern Physics C.*

French, J. R. P. (1956). A formal theory of social power. *Psychological Review*, *63*(3), 181–194.

Galam, S. (2002). Minority opinion spreading in random geometry. *The European Physical Journal B*, *25*, 403–406.

Hegselmann, R. (2004). Opinion dynamics: Insights by radically simplifying models. In D. Gilles (Ed.), *Laws and models in science.* London.

Hegselmann, R., & Krause, U. (2002). Opinion dynamics and bounded confidence models: Analysis, and simulation. *Journal of Artifical Societies and Social Simulation (JASSS)*, *5*(3).

Jager, W. (2000). *Modelling consumer behaviour.* Unpublished doctoral dissertation, University of Groningen, Centre for Environmental and Traffic psychology, Kurt Lewin Instituut, Groningen. (<http://www.ub.rug.nl/eldoc/dis/ppsw/w.jager/>)

Kuran, T. (1987, September). Preference falsification, policy continuity and collective conservatism. *The Economic Journal*, *97*(387), 642–665.

Latané, B., Nowak, A., & Liu, J. H. (1994). Measuring emergent social phenomena: dynamism, polarization, and clustering as order parameters of social systems. *Behavioral Science*, *39*, 1-23.

Lorenz, J. (2003, May). *Multidimensional opinion dynamics when confidence changes.*

Lorenz, J. (2005). A stabilization theorem for continuous opinion dynamics. *Physica A*, *355*(1), 217–223.

Lorenz, J. (2006). Consensus strikes back in the hegselmann-krause model of continuous opinion dynamics under bounded confidence. *Journal of Artificial Societies and Social Simulation*, *9*(1). (<http://jasss.soc.surrey.ac.uk/9/1/8.html>)

Nowak, A., Szamrej, J., & Latané, B. (1990). From private attitude to public opinion: A dynamics theory of social power. *Psychological Review*, *97*(3), 362–376.

Trommsdorff, V. (1998). *Konsumentenverhalten* (third ed.). Stuttgart; Berlin; Köln: Kohlhammer.

Urbig, D. (2003). Attitude dynamics with limited verbalisation capabilities. *Journal of Artificial Societies and Social Simulation (JASSS)*, *6*(1).

Urbig, D., & Malitz, R. (2005, Sep). Dynamics of structured attitudes and opinions. In *Representing social reality. Pre-proceedings of the third conference of the european social simulation association (ESSA '05)* (pp. 206–212).

Weisbuch, G., Deffuant, G., Amblard, F., & Nadal, J. P. (2002). Meet, discuss, and segregate! *Complexity*, *7*(3).

Effects of Mass Media and Opinion Exchange on Extremist Group Formation

Steven Butler[1] and Joanna J. Bryson[1]

Department of Computer Science, University of Bath, BA2 7AY. UK

Abstract. Contemporary communication technologies are thought to facilitate the growth of the small autonomous terrorist groups indicative of 'modern' terrorism. In this study, the Animal Liberation Front provides an example of a culturally distinct organisation from which an extremist force—the Animal Rights Militia—emerges to pose the threat of violence. Agent-based modelling is used to simulate this emergence, and subsequent dynamics, under a variety of conditions. The simulation model not only implements local opinion exchange within a population, but also the polarising effect of mass media. Results show the significance of mass media, of limits on the cell size that are independent of the policing efforts, and the impact of societies with rapid population turn over, such as are found near universities.

Keywords. Agent-based simulation of social phenomena, cultural dynamics, opinion dynamics.

1 Introduction

The autonomous nature of small, terrorist networks or *cells* is indicative of 'modern' terrorism—a terrorist movement that does not need a hierarchical command structure, but instead simply the dissemination of goals, guidelines and ideologies. Modern communication technologies, the Internet in particular, easily facilitate such cultural transactions. This is a trend that can only become more widespread, and so it is important for us to understand how these cells may develop and the factors that may aid or hinder their growth.

In this study, we attempt to provide an insight into the proliferation of extremist violence from within a culturally distinct group. We create a model, the TerrorPlex, which simulates the emergence of terrorist groups and the subsequent population interaction dynamics on the individual and group level. A range of factors are explored within the model, including the impact of mass media, the nature of cell structures (particularly size limits) and the impact of population turnover, such as is found around universities.

The group chosen for this study is the Animal Liberation Front (ALF), a trans-global collective of individuals largely united by an opposition to animal testing. The ALF is not in itself a violent movement, but the assumption can be made that the Animal Rights Militia (ARM), which *has* claimed responsibility for numerous acts of violence, consists of extremist ALF members intent on recruiting into secretive groups other ALF members who exhibit an extremist bent. This arrangement is thought to be typical of terrorist organisations, many of which might be seen as a more significant threat than the ARM. However, the ALF/ARM are relatively accessible for study, and the ARM do meet technical definitions of a terrorist group. The aims and rules of the ALF are readily available[1], enabling individuals and small groups to follow the ALF agenda whilst retaining autonomy. Guidelines even recommended size limits for activist groups.

This degree of transparency can be logically credited to the relative mildness of ARM actions. In Great Britain, Animal Rights activists are not responsible for attacks involving high-grade explosives and resulting in the death of targeted persons. From a sociological perspective, this is unlikely to change (see Taylor [10]) due largely to the emotional values that usually provide the impetus for involvement with the animal rights movement or similar issues.

Although one may face opposition when attempting to label the ARM as 'terrorist', their *modus operandi* is similar to that of the autonomous networks that are indicative of modern terrorism. We may define any sub-national group as terrorist if it performs premeditated actions that are:

a) Politically or ideologically motivated,
b) Targeted against non-combatants,
c) Violent or destructive, and
d) Executed in a clandestine manner.

The ARM is clearly encapsulated by this definition and it is hoped that the results of this study may be extended to other groups which—while operationally similar—have a tendency for more extreme violence.

Another culturally distinguishable group modelled in the TerrorPlex simulation is Pro-Test. This is the pro-animal-testing movement which emerged during 2006 in response to the campaigns waged by the ALF. In our model, any agent 'in favour' of animal testing is identified as Pro-Test, broadly reflecting the similar autonomy to be found within the ALF.

To better understand the dynamics underlying the growth of terrorist organisations, we chose to build an Agent-Based Model (ABM). After a review of ABM frameworks available, the Java distribution of RePast [9] was selected. The basic principles for the agent interaction were derived from the published ALF guidelines and other externally documented factors pertaining to clandestine and terrorist groups.

This paper provides a brief description of the simulation model, comprising of a simply definable environment and agent attributes. We shall then cover functionality of the particularly pertinent features of opinion exchange and recruitment, before detailing how a mass media influence is implemented. We then review key results, including circumstances in which the ALF 'self annihilates'. Readers at any time seeking further clarity and a more broad range of recorded experiments are directed to Butler [3]. Code for the TerrorPlex is available on demand from the web or by request from the authors.

[1] The Animal Liberation Primer by Animal Liberation Front Information Services is available from: http://www.animalliberationfront.com/ALFront/ALFPrime.htm [Accessed 21 June 2007].

2 The TerrorPlex Model

We present an ABM by first specifying the environment, then the agents' characteristics or *state*, and finally the agents' behaviour [2].

2.1 Environment

During runtime, the model is viewed as a two-dimensional grid, but is in fact a three-dimensional torus, or 'doughnut' shape. The world inhabited by the agents is not the conventional physical-proximity model used to such great effect by pioneers such as Epstein and Axtell [5]; it is a metaphysical domain intended to acknowledge the pervasiveness of modern communication technologies. The metaphysical arrangement represents an environment where each agent has a number of neighbours with which it communicates regularly. The proximity is not necessarily representative of a physical closeness, but rather a psychological familiarity. In this way, the simple model accommodates a range of social interactions, with neighbouring agents potentially representing school friends, work colleagues, family and household members, as well as Internet chat room or notice board associates.

The effect of the Internet in particular is highly contested as either a moderating or polarising force. It was noted that, whilst it would be difficult to identify cases where the Internet has resulted in an individual turning to violence, it would be impossible to gauge how many have been turned from violence by the same medium. We handle the impact of mass media differently from the social network, as will be discussed below under "agent behaviour".

2.2 Agent Characteristics

Agents within the simulation are objects. As such, they are endowed with a number of attributes which constitute something akin to a 'personality', plus other attributes used for housekeeping functions. Personality attributes are randomly generated numeric values that fall between 1 and 100. These are identified as Law; an agent's inherent respect for law and order, and Pacifist, which represents the agent's pacifism. We may call these *fixed parameters*, since once set, these values remain constant for the duration of an agent's lifespan. Law and Pacifist play an important role when used in conjunction with a measure of personal isolation for discerning an agent's suitability for recruitment into an extremist group. Broadly then, an agent's tendency towards violent action can be calculated like so:

$$Tendency = \frac{Isolation}{Law \times Pacifism} \tag{1}$$

What we refer to as 'housekeeping' attributes are necessary for controlling the simulation and include, for example, those which identify whether an agent is an ALF member, describe an agent's location, and maintain a record of any groups to which an agent belongs. In this model, an agent's position is another example of a fixed parameter, whereas details of membership can be classed as *dynamic*, since these may vary throughout the course of a simulation.

Table 1. Agent attributes

Attribute	Type
x	integer
y	integer
law	integer
pacifist	integer
isolation	integer
myGroup	integer
lifetime	integer
myDuration	integer
isProTest	boolean
isALF	boolean
isActive	boolean
CIS	int array

Table 2. Cultural thresholds

No. of 1s	Cultural Type	Colour
< 2	Pro–Test	Blue
2 - 7	Neutral	Green
> 7	ALF	Red

Cultural Identity To describe the views of an agent in relation to animal testing, we employed a Cultural Identity String, or CIS. The CIS is an integer array consisting of a binary pattern, which is at the core of opinion exchange within the model. For each agent, this array is randomly generated when the agent first enters the simulation. Three conditions may arise from the content of this full dynamic parameter. With respect to the animal testing debate, an agent which has a large majority of 0s in its CIS is designated as Pro-Test, a large majority of 1s means ALF, and a mixture of values within upper and lower bounds indicate that an agent is Neutral.

2.3 Agent Behaviour

On each cycle of the TerrorPlex simulation, agents engage on a course of behaviour dependent on their cultural identity. Behaviour includes the exchange of opinion and, for some agents, recruitment. For example, ALF agents maintain their cultural identity by not updating their own opinion in accordance with their neighbours. However, if their tendency towards violent action increases to a pre-defined threshold then they become activist agents. Agents undergo an increased propensity for violence if they become more isolated or if the *terror gradient* increases. The terror gradient is calculated periodically and is a measure of the difference in the prevalence of terrorist activity between two points in time. Neutral agents possess a more substantial behaviour repertoire. They may assimilate the cultural identity of neighbours, incorporate mass media opinion and potentially develop an affiliation with Pro-Test or the ALF.

Opinion Exchange Opinion exchange within the model is achieved on a local basis with net effects propagated—where applicable—throughout the agent population. The mechanism for these exchanges is known as tag-flipping, described by Epstein and Axtell [5]. An agent eligible for interaction compares the value held at a randomly selected location within its CIS with the corresponding value in each eligible neighbour's CIS. If the majority of neighbouring agents share the same value, then no change is made.

However, if most neighbours have a different value for the particular CIS location, then the agent flips its own value to match the neighbouring majority.

Deffuant et al. [4] provide a model for opinion exchange between agents in which a threshold governs whether an exchange is made. If two agents have an opinion that differs by a margin exceeding the threshold, then no exchange of opinion is possible. They call this ‘bounded confidence’. A threshold, per se, is not used here but rather the social divisions or cultural thresholds that are implied by the makeup of agent’s CIS string. These divisions are essential aspects of our model. Further to dictating an agent’s behaviour, they may act as barriers to communication, so that neutral agents may—given certain parameters—identify ALF and Pro-Test agents as beyond the scope of their cultural interactions. Similarly, ARM agents are *always* omitted from opinion exchanges as the secrecy of such individuals would be closely guarded.

Recruitment Gould [6] states that the process whereby individuals become involved in movements through people whom they know, is one of the “most frequently cited facts about social ties and activism”. It will come as no surprise then, that the recruitment of ALF agents into activist groups by ARM agents plays a pivotal role in our simulation. A very small proportion of activist agents come to exist in the model under their own volition (see above). ALF agents whose tendency towards violent action reaches a pre-defined ‘conversion’ threshold become ARM agents. These individuals are responsible for recruiting ALF agents into their clandestine groups, and in turn the recruited become the recruiters. A more complete model could provide mechanisms for group norms and polarisation, but that is beyond the scope of the present work[2].It should be noted at this point, that although ARM agents must actively recruit ALF agents, there is no such requirement for an agent to become culturally Pro-Test or ALF. This is justifiable since Pro-Test members and (non-ARM) ALF sympathisers do not operate illegally, and so are not obliged to exercise the degree of secrecy that the ARM must do to avoid prosecution.

For an agent to be recruited into an ARM group, several conditions must be met:

1. There must be communication (physical or virtual) between both parties.
2. There must exist mutual trust between these two, although not necessarily between the new member and all existing group members.
3. The person being recruited must be known to share the same ideologies as those of the group.
4. Based upon the person’s ideologies and other attributes, it should be deemed likely that they would wish to become part of the group.
5. The group must want a new recruit.

These rules are met within the simulation as follows:

Communication may exist between agents who are classed as ‘neighbouring’ in the radial catchment area within the social space defined by the environment and described

[2] Berry et al. [1] present an interesting paper in which entire cliques are eligible for conversion to terror when their ‘disgruntlement’—calculated as an average over all members—exceeds some threshold.

earlier. Parameters for the catchment—either von Neumann or Moore neighbourhood methods (see below) and the length of the radius—are controlled through the interface.

Mutual trust is established between two agents after having been neighbours for a pre-defined number of simulation cycles. This may appear to be a rather crude approach to a very complex notion, but one must recall the nature of the simulation space; agent proximity represents an emotional connection. Therefore, it should not be too unreasonable to assume that a prolonged period of emotional familiarity may engender a degree of trust between two parties.

As human beings are adept at assessing others' values through contextual, conversational and visual clues, ARM agents are able to identify those neighbours who share ALF ideologies by direct access to the variable state.

The likelihood that an agent could be recruited is represented by their tendency towards violent action. This is a personal attribute, but if we have already said that mutual trust exists, then we can suppose that such knowledge has been accumulated by the recruiting party. An existing ARM agent will access the '*tendency*' value of a potential recruit and if it falls below a set threshold then the ALF agent is eligible for recruiting.

ARM agents accept suitable recruits until the membership quota of their cell (if any) is met.

2.4 Mass Media

With careful definition of the agent space, we have implicitly managed of one aspect of modernity—communication technologies—by not restricting ourselves to an environment of physical proximity. Also to be considered is another phenomenon unique to recent history; mass media. The effects of mass media cannot be handled as conveniently as modern communication devices, but demand an explicit solution. Indeed, without such an approach, we would be unable to exercise any convenient control over the behaviour of mass media within our model and thus analyse recorded variations.

Mckeown and Sheehy [8] build upon the bounded confidence model [4] by introducing a second mechanism for mass communication. This mechanism involves the interaction of a fixed opinion, representing a media source.

Because our study concerns the two opposing forces of ALF and Pro-Test, it is sensible that we incorporate two corresponding media opinion sources. These are essentially two CIS elements, but deliberately and statically biased in opposite directions. The TerrorPlex simulation by default asserts a probability of 0.7 that the binary values will be 0 for one media source, and a probability of 0.3 for the other source. Wherever possible, randomising functions are used. When the mass media functions are enabled, an agent will incorporate one of these media sources into the tag-flipping exercise as though it were a neighbouring agent. Whilst a neutral agent will liberally select at random a media source to assimilate, ALF and Pro-Test agents will not be so indiscriminate. These agents will ignore the media source which does not represent their cultural identity, thus facilitating the reinforcement of ideologies that is found to act within groups.

To add realism to our model, we have employed an innovative approach to mass media representation by harnessing the power of feedback. To paraphrase Wardlaw [11], when terrorism becomes institutionalised, the perceived significance of terrorist acts is reduced. Therefore, it may follow that the effect of media bias upon an individual will be less if violent acts are routinely reported. Indeed, if violence becomes commonplace, it may not be reported at all. With this in mind, a ‘terror gradient’ has been implemented whereby the media presence in an agent’s tag-flipping routine is weighted in accordance with the relative change in terrorist activity from one period to another. A significant change in activity will herald a more significant effect upon the routine. For our purposes, we shall say that there is a direct correlation between the number of activist agents and the prevalence of terrorist acts.

2.5 Simulation Execution

At run-time, an initial representation of the model is created. Data channels are defined for both real-time graphical output and the recording of simulation data for analysis. These channels are updated upon each cycle. Elements recorded to disk are the number of ARM, ALF, Pro-Test and Neutral agents, plus tag-flipping activity, the number of ARM groups and the terror gradient.

Every aspect of the model has a default value within the code. However, the majority of these may be changed through the user interface. Before execution then, the user can easily change parameters which include environment size, population size, whether mass media is present, and the range of an agent’s view. One may also express a preference for either a von Neumann or Moore neighbourhood method [7]. The von Neumann method searches for neighbouring agents in the North, South, East and West directions only, so that given a range of one, the maximum number of neighbours would be four. Moore, however, includes the diagonal directions North-East, South-East, South-West and North-West, potentially yielding eight neighbours when given the same range.

Table 3. TerrorPlex simulation parameters

Parameter	Type	Entry
Population	Integer	Direct
GridSize	Integer	Direct
RestrictGroupSize	Boolean	Check–box
MassMedia	Boolean	Check–box
NeighbourMethod	Integer	Drop–down List
Range	Integer	Direct
IgnoreOtherCultures	Boolean	Check–box
Respawn	Boolean	Check–box
StepLimit	Integer	Direct

The three-dimensional torus environment is generated and projected onto the screen as a grid, and this is populated with a randomly positioned agent population. Each agent object is instantiated with its attributes as described earlier.

A schedule is defined which asserts that each agent shall undergo one of several behavioural routines (also described earlier) once every cycle. Which routine is selected will depend upon the cultural identity of an agent and any relevant parameters. The program does not iterate through a sequential list of agents, but rather a list that is randomised at the beginning of each cycle. The schedule also ensures that the terror gradient is calculated at regular intervals.

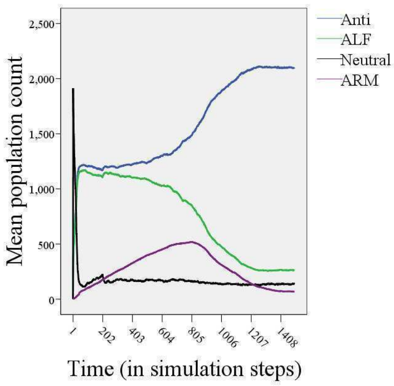
3 Results

The experiments we conduct provide a substantial dataset supporting interesting real-time observations. The most noteworthy results can be seen in tests for which the *Respawn* option is active. Respawn regularly replenishes the agent population by removing those who have reached their assigned lifespan and reintroducing to the simulation an equal number of new agents. Respawned agents are created with their attributes being randomly-assigned under the same rules as at simulation startup. When this option is not in use, the population stagnates after relatively few iterations of the simulation and does not allow us to enjoy the dynamics of changeable social interactions (see Figure 1(d)). Notably, in human societies we also see more social change and sometimes radicalism come from areas such as university towns where there is a steady but not overly rapid turn over of population. If turn over is too rapid, trust cannot be established.

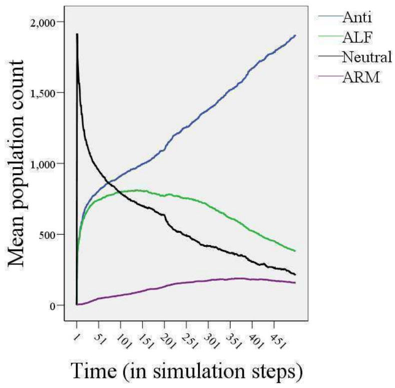
Some expected simulation behaviours were recorded, that include an excellent example of clustering. This occurs where neutral agents are free to interact with ALF and Pro-Test agents, but the 'respawn' option is disengaged. Under these conditions and an absence of mass media presence, we see that the three cultures achieve a similar abundance within the population in distinctive clusters. In examples where the mass media mechanism is applied, clustering is still evident, but the number of neutral agents will dwindle until none remain.

Tests also provided a few surprises. It was shown that placing limits on the size of activist groups was absolutely necessary for avoiding a saturation of activist agents within the ALF community. Emergent terror cells in the clustering examples mentioned above are very sparse when the default group size restriction of 5 is imposed. If the restriction is lifted however, ARM membership increases to include nearly the entire ALF contingent. We had originally assumed that the terror cells would be limited in size by the inherent difficulties involved in recruiting suitable candidates. In real societies, cell size is also limited by policing, since the larger the number of members, the higher the probability that the cell is discovered. Adding this probability of discovery into the TerrorPlex would be interesting further work.

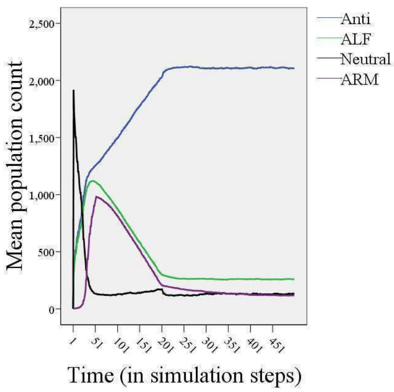
The polarising effect of mass media can be observed throughout the tests as expected. Polarisation of opinion is evident without mass media, but at a much-reduced rate. An unforeseen phenomenon though, is that of a self-annihilating ALF. What we witness is the initial growth of both ALF and Pro-Test groups. This may also be accompanied —in the case of unrestricted group size—by a proportional increase in ARM activity or, as we see with low-value size restrictions, a steady increase in ARM numbers. In either situation, and for other restriction values between these two extremes, there



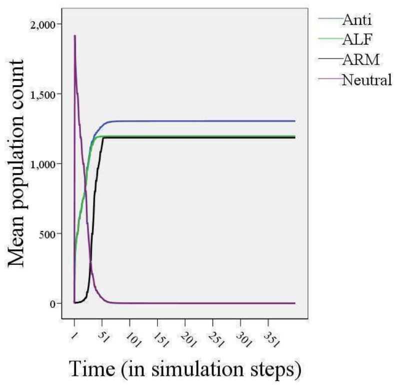
(a) mass media, respawn, cell limit of 5



(b) no mass media, respawn, cell limit of 5



(c) mass media, respawn, no cell limit



(d) mass media, no respawn, no cell limit

Fig. 1. Plots showing means for 15 runs in each condition. *Anti* here represents Pro-Test. *Mass media* determines whether the two mass media influences are available to agents (see main text.) In the *Respawn* condition, each agent is randomly assigned a lifespan of between 0-200 cycles, otherwise agents enter the simulation only at its beginning. *Cell limit* indicates the size at which terrorist cells no longer recruit even if compatible agents are found.

is a clearly observable phenomenon (see Figure 1). After the initial polarisation into ALF and Pro-Test groups, there is a period of 'power sharing' between the two cultural identities. This gives way to a divergence of fortunes however, where ALF membership diminishes suddenly, with Pro-Test agents increasing in number at an equal rate to gain almost complete consensus. This reduction in ALF members is closely related to ARM membership and, in fact, is easily explainable by the stipulations that we have made with respect to the transfer of cultural identity.

For reasons of personal security, ARM activists exist outside of the normal social context. As such, they are omitted from the cultural view of other agents. A 'neutral' agent, when assimilating its neighbours' culture, will not involve an ARM activist, but *will* involve an ALF member. If there is a situation where the ARM population has increased to a majority of the ALF population, then an agent involved in cultural exchange will have mostly other neutral agents and Pro-Test agents to engage with. This imbalance quickly leads to a simulation space dominated by Pro-Test agents. The proliferation of ARM activists has effectively led to the annihilation of the entire ALF movement.

As mentioned, tests have also demonstrated the importance of size restrictions upon ARM groups, showing that they have a tendency to balloon in size where restrictions are not in place. Herein lies a surprising and interesting relationship: The sustainability of the ALF population actually increases with the application of more stringent restrictions on ARM group size. Far from aiding the growth of the ALF, the presence of ARM agents causes a decline by increasing the probability of media attention, but concurrently removing culturally influential ALF members (those that have now turned to activism). In the TerrorPlex simulation, the ALF population is only able to compete with Pro-Test if the number of activist agents is kept low.

4 Conclusion

Beginning with just a simple definition of agent behaviour, we have been able to create a fertile simulation environment which enables us to explore emerging culturally complex scenarios for modelling terrorist recruitment and other sorts of ideological social affiliative behaviour.

Local opinion exchange—enhanced with the concept of bounded confidence—has been developed with the addition of polarised mass media sources. This mass media influence has been further advanced with a dynamic terror gradient which enables cultural changes within the agent population to be fed back into the model.

The element of feedback has, in turn, led to some fascinating results that include an example of how autonomous terrorist networks can, through their own prosperity, cause a cultural imbalance which leads to the eventual demise of not just the terrorist faction but the entire movement.

References

[1] Berry, N., Ko, T., Moy, T., Smrcka, J., Turnley, J., and Wu, B. (2004). Emergent Clique Formation in Terrorist Recruitment. In *Agent Organizations: Theory and*

Practice, Session on Organisational Models. The AAAI Press, Menlo Park, California. American Association for Artificial Intelligence workshop.

[2] Bryson, J. J., Yasushi, A., and Lehmann, H. (2007). Agent-based models as scientific methodology: A case study analysing primate social behaviour. *Philosophical Transactions of the Royal Society, B — Biology*, 362(1485). in press.

[3] Butler, S. (2006). A simulated exploration into the growth of modern terrorist networks. Technical Report CSBU–2006–18, University of Bath, UK. Honours Undergraduate Dissertation, with Joanna J. Bryson.

[4] Deffuant, G., Neau, D., Amblard, F., and Weisbuch, G. (2000). Mixing Beliefs Among Interacting Agents. *Advances in Complex Systems*, 3:87–98.

[5] Epstein, J. M. and Axtell, R. (1996). *Growing Artificial Societies*. The Brookings Institution, Washington, DC.

[6] Gould, R. V. (2003). Why do Networks Matter? Rationalist and Structuralist Interpretations. In Diani, M. and McAdam, D., editors, *Social Movements and Networks*, pages 233 – 257. Oxford University Press.

[7] Hegselmann, R. and Flache, A. (1998). Understanding complex social dynamics: A plea for cellular automata based modelling. *Journal of Artificial Societies and Social Simulation*, 1(3). Available from: http://www.soc.surrey.ac.uk/JASSS/1/3/1.html.

[8] Mckeown, G. and Sheehy, N. (2006). Mass Media and Polarisation Processes in the Bounded Confidence Model of Opinion Dynamics. *Journal of Artificial Societies and Social Simulation*, 9(1). Available from: http://jasss.soc.surrey.ac.uk/9/1/11.html.

[9] North, Collier, Vos, Najlis, and Maciorowski (2004). Repast Revolution: An Overview of New Repast Developments. In *Social Dynamics: Interaction, Reflexivity and Emergence*.

[10] Taylor, B. (1998). Religion, Violence and Radical Environmentalism: From Earth First! to the Unabomber to the Earth Liberation Front. *Terrorism and Political Violence*, 10(4):1–42.

[11] Wardlaw, G. (1982). *Political Terrorism*. Cambridge University Press, Cambridge, UK. Australian Institute of Criminology.

Information feedback and mass media effects in cultural dynamics

J.C. González-Avella[1],M. G. Cosenza[2], K. Klemm[3], V. M. Eguíluz[1] and M. San Miguel[1]

[1] IFISC (CSIC-UIB) Campus Universitat Illes Balears, E-07122 Palma de Mallorca, Spain

[2] Centro de Física Fundamental, Universidad de los Andes, Mérida, Mérida 5251, Venezuela

[3] Bioinformatics, Department of Computer Science, University of Leipzig, 04107 Leipzig, Germany

Abstract. We study the effects of different forms of information feedback associated with mass media on an agent-agent based model of the dynamics of cultural dissemination. Two mechanisms of information feedback are investigated: (i) direct mass media influence, where local or global mass media act as an additional element in the network of interactions of each agent, and (ii) indirect mass media influence, where global media acts as a filter of the influence of the existing network of interactions of each agent. Our results substantiate previous findings showing that cultural diversity builds-up by increasing the strength of the mass media influence. We find that this occurs independently of the mechanisms of action (direct or indirect) of the mass media message. However, through an analysis of the full range of parameters measuring cultural diversity, we establish that the enhancement of cultural diversity produced by interaction with mass media only occurs for strong enough mass media messages. A main different result is that weak mass media messages, in combination with agent-agent interaction, are efficient in producing cultural homogeneity. Moreover, the homogenizing effect of weak mass media messages are more efficient for direct local mass media messages than for global mass media messages or indirect global mass media influences.[1]

1 Introduction

In an influential paper, Robert Axelrod [1] addressed the question

> if people tend to become more alike in their beliefs, attitudes, and behavior when they interact, why do not all differences eventually disappear?.

To investigate this problem, Axelrod introduced an agent-based model to explore mechanisms of competition between the tendency towards globalization and the

[1] This article was published in the Journal of Artificial Societies and Social Simulation: http://jasss.soc.surrey.ac.uk/10/3/9.html

persistence of cultural diversity (see the applet http://ifisc.uib.es/research_topics /socio/culture.html). Culture in this model is defined as a set of individual attributes subject to social influence. The state of an agent is described by a set of F cultural features, each with q possible values or traits. The parameter q gives a measure of initial culture diversity for a random distribution of trait values. The local interaction between neighboring agents follows two basic social principles. The probability of interaction is determined by a principle of homophily: it is proportional to the cultural overlap between the agents, that is to the amount of cultural similarities (number of features) that they share. The result of the interaction is determined by a principle of social influence, so that similarity is enhanced when interaction occurs. In a typical dynamical evolution, the system freezes in a multicultural state with coexisting spatial domains of different cultures, illustrating how a simple mechanism of local convergence can lead to global polarization. Later systematic studies of Axelrod's model have identified a globalization-polarization transition depending on the value of q for a fixed F [2–6]. There is a threshold value q_c such that for $q < q_c$ globalization (cultural homogeneity) occurs, while for $q > q_c$ a persistent multicultural state (polarization or cultural diversity) is reached.

Several extensions of this model have been investigated, some of them already suggested in Axelrod's paper. For example, it has been shown that frozen polarized configurations are not robust against cultural drift modelled as random perturbations [3]. Other extensions include the consideration of quantitative instead of qualitative values for the cultural traits [7], the extension of the model to continuous values of the cultural traits and the inclusion of heterophobic interactions [8], the simulation of technology assimilation [9], the consideration of specific historical contexts [10], or the effect of a fixed external cultural influence [11]. A number of works have addressed the issue of the role of the social network of interactions. The existence of long range links of interaction in a social network with a small world structure naturally promotes cultural globalization [12]. However, a study of the increase in the range of interaction [13] suggests that an increase in communication promotes the emergence of a global culture, but it can also function to maintain areas of cultural uniqueness. A related question is the co-evolution of the social network of interaction and the dynamics of cultural changes, so that possible social interaction are not fixed from the outset [14]. Within this general context of the studies of the effects of different forms of social interactions, Shibanai et al.[15] considered the process of information feedback into the social system. Such feedback mechanism is one of the functions of mass media. From a general perspective, this question can be addressed as the competition of local agent interactions with mechanisms of global coupling [16]. Shibanai et al. concluded that

> mass media, contrary to lay beliefs of their strong uniforming power, would rather contribute to creating differences in the long run.

This result is certainly surprising since mass media are believed to be powerful instruments to influence people's attitudes and opinions to homogenize society.

Conclusions in reference [15] are based on the analysis for a single value of the parameter q. In this paper we consider the full range of values of q when addressing the general question of the effects of different forms of information feedback on cultural dynamics in the framework of Axelrod's model. We consider information feedback mechanisms as different types of mass media influences acting on a social system. This extension of Axelrod's model was referred to as "public education and broadcasting" [1]. Our aim is to explore mechanisms, and their efficiency, by which mass media modifies processes of cultural dynamics based on local agent interactions. We investigate two main mechanisms that differ on how the information feedback affects the agents [15]: (i) direct mass media influence, and (ii) indirect mass media influence. In the first case, mass media acts as an additional element in the network of interactions of each agent. As applications of direct mass media influence we study models of global and local mass media and compare their effects on the system. In the second feedback mechanism, mass media acts as a moderator or filter of the influence of the existing network of interactions of each agent. The assumption of this second model is that mass media have indirect influential power which reinforces the effect of personal networks. In particular, we focus on a model of indirect influence of global mass media.

We deal with states of the agents and mass media influences described by vectors whose components can take discrete values. We consider mass media influences that originate endogenously but can act either directly or indirectly on the system. In the first case, the agent-agent interaction and the interaction of the agents with the mass media are two independent processes and both are based on the same homophily and social influence principles of Axelrod's model. This scheme constitutes a model for a social system interacting with global or local mass media that represent endogenous cultural influences or plurality information feedback at different levels. In the case of indirect action, agent-agent interactions is not independent of the state of global mass media, the interaction being reinforced when the states of the agent and mass media message are connected.

Our results support those of Shibanai et al. [15] showing that cultural diversity builds-up by increasing the probability of interaction of the agents with the mass media. This occurs independently of the form of action (either direct or indirect) of the mass media message. However, we find that cultural diversity is only promoted by strong enough mass media messages, while weak mass media messages, in combination with agent-agent interaction, are efficient in producing cultural homogenization. Moreover, this homogenizing effect in the way to globalization turns out to be much more efficiently implemented by direct local mass media messages, as for example local or regional TV (narrowcast) than by direct global mass media messages as worldwide TV channels (broadcast), or by an indirect influence of global mass media as in a filtering process of local interactions.

The model of direct mass media influence, including the description of the local and global mass media, is presented in Section II. This section shows the effects of the direct action of these two types of mass media. We describe different effects occurring in the culturally homogeneous state of the system ($q < q_c$) as well as in the multicultural state ($q > q_c$). Section III describes the model of indirect global mass media influence and the effects of this mechanism in the collective states of the agents. Section IV contains the conclusions of our work, including a comparison of the effects of the different models of information feedback that we have considered.

2 Models of direct mass media influence

We consider a system of N agents at the nodes of a two-dimensional regular lattice. The state c_i of agent i is defined by a vector of F components $\sigma_i = (\sigma_{i1}, \sigma_{i2}, \ldots, \sigma_{iF})$. In Axelrod's model, the F components of c_i correspond to the cultural features (language, religion, etc.) describing the F-dimensional culture of agent i. Each component σ_{if} can take any of the q values in the set $\{0, 1, \ldots, q-1\}$ (called cultural traits in Axelrod's model). As an initial condition, each agent is randomly and independently assigned one of the q^F possible state vectors with uniform probability. We introduce a vector M representing the mass media message, with components $(\mu_{i1}, \mu_{i2}, \ldots, \mu_{iF})$. We distinguish two types of direct mass media influences, schematically shown in Fig. 1.

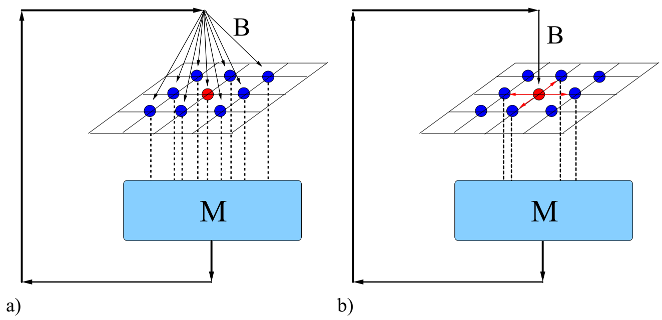


Fig. 1. Diagrams representing two types of direct, endogenous mass media influences acting on the system. a) Global mass media. b) Local mass media.

(i) The *global mass media* is spatially uniform and may vary in time. Each μ_{if} is assigned the most abundant value exhibited by the f-th component of the state vectors of al the agents, which we denote by μ_f. If the maximally abundant value is not unique, one of the possibilities is chosen at random with equal probability. In this type of influence the same global information is feedback to each agent at any given time. The components of M may change as the system evolves. In the context of cultural models, this represents a global mass media influence shared

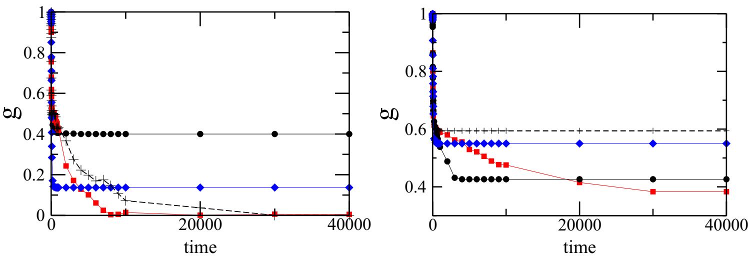


Fig. 2. Evolution of g in a system subject to a global mass media message for different values of the probability B, with fixed $F = 5$. Time is measured in number of events per site. System size $N = 50 \times 50$. Left: $q = 10$; $B = 0$ (crosses); $B = 0.0005$ (squares); $B = 0.15$ (diamonds); $B = 0.6$ (circles). Right: $q = 30$; $B = 0$ (crosses); $B = 0.0005$ (squares); $B = 0.005$ (circles); $B = 0.1$ (diamonds).

identically by all the agents and containing the most predominant trait in each cultural feature present in a society (a "global cultural trend") [15].

(ii) The *local mass media*, is spatially non-uniform and non-constant on time. Here the component μ_{if} is assigned the most frequent value of the f-th component of the state vectors of the agents belonging to the von Neumann neighborhood of agent i. If there are two or more maximally abundant values of component f one of these is chosen at random with equal probability. This type of influence can be interpreted as a local mass media conveying the "local cultural trend" of its neighborhood to each agent in a social system.

Systems subject to either local or global mass media describe social systems with endogenous cultural influences. Cultural influences generated endogenously represent a plurality information feedback, which is one of the functions of mass media [15], but this can occur at a global ("broadcast") or at a local ("narrowcast") level.

The strength of the coupling to the mass media is controlled by a parameter $B \in [0, 1]$ that measures the probability of interaction of the agents with the mass media message. We shall assume that B is uniform, i.e., the mass media influence reaches all the agents with the same probability. The parameter B can be interpreted as the probability that the mass media message attracts the attention of the agents in the social system. The parameter B takes into account factors of the mass media influence that can be varied, such as its amplitude, frequency, attractiveness, etc. At any given time, we assume that any agent can either interact with the mass media message or with other agents in the system. Each agent in the network possesses a probability B of interacting with the message M and a probability $(1 - B)$ of interacting with its neighbors.

Formally, we treat the mass media acting on each agent i as an additional neighbor of i with whom an interaction is possible. The mass media message is represented as an additional agent $\phi(i)$ such that $\sigma_{\phi(i)f} = \mu_{if}$ in the following definition of the dynamics. The dynamical evolution proceeds iterating the following steps:

(1) Select at random an agent i on the lattice (called active agent).

(2) Select the source of interaction j. With probability B set $j = \phi(i)$ as an interaction with the mass media vector. Otherwise, choose agent j at random among the four nearest neighbors of i on the network.

(3) Calculate the cultural overlap (number of shared features) $l(i,j) = \sum_{f=1}^{F} \delta_{\sigma_{if},\sigma_{jf}}$. If $0 < l(i,j) < F$, sites i and j interact with probability $l(i,j)/F$. In case of interaction, choose h randomly such that $\sigma_{ih} \neq \sigma_{jh}$ and set $\sigma_{ih} = \sigma_{jh}$.

(4) Update the mass media vector M if required. Resume at (1).

The role of mass media can be seen as an additional neighbor of each agent i with whom an interaction can take place. This "fifth" neighbor competes with the other four neighbors to influence an active agent i. B is the probability of capturing the attention of any agent to interact with either the global or the local mass media message M. The effect of that interaction on each agent depends on the specific cultural overlap between the agent and the message M. The total probability to interact with the mass media is Bp_{iM}, where $p_{iM} = l(i,M)/F$, and $l(i,M)$ is the overlap between agent i and the message. On the other hand, the active agent i has a probability $\frac{1}{4}(1-B)p_{ij}$ to interact with a randomly chosen neighbor j. Therefore, the parameter B weights the influence of the mass media with respect to the influence of the neighbors on an agent i. This dynamics is based on the same principles of Axelrod's model, namely homophily or similarity and social influence. But now the agents interact either between them or with M. The competition of these interactions in measured by B.

To characterize the collective final state reached by the dynamics, we calculate the average fraction of cultural domains $g = \langle N_g \rangle / N$. A cultural domain is a set of contiguous sites with identical cultural traits. N_g is the number of cultural domains formed in the final state of the dynamics for a given realization of initial conditions. A culturally homogeneous state is characterized by values $g \to 0$. When the system settles into a culturally diverse state we have $\langle N_g \rangle \gg 1$. Our numerical results in this work are based on averages over 50 realizations for systems of size $N = 50 \times 50$, and $F = 5$.

Figure 2 shows the average fraction of cultural domains g as a function of time under the direct action of global mass media, for two values of the number of traits q with $F = 5$, and for different values of the probability B. Fig. 2 (left): for small values of q the system reaches a culturally homogeneous state in absence of mass media influence ($B = 0$) and also for small values of B. When the probability B increases, we see that global mass media is able to induce cultural diversity in the system. This result agrees with the results obtained by Shibanai et al. [15] about the ability of global mass media to induce cultural diversity in the system ($g \neq 0$). However we observe in Fig. 2 (right) that, for larger values of q, where the system would be in a culturally polarized state with no mass media influence ($B = 0$), small values of the probability B can reduce the number of cultural groups. Effects similar to those shown in Fig. 2

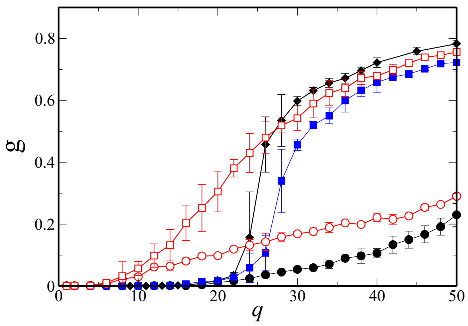


Fig. 3. Asymptotic value of the fraction of cultural domains g as a function of q, for different values of the probability B and for different types of mass media influences. $B = 0$ (diamonds); $B = 0.0005$ (solid squares, direct global mass media); $B = 0.6$ (empty squares, direct global mass media); $B = 0.0005$ (solid circles, direct local mass media); $B = 0.1$ (empty circles, direct local mass media).

are observed when a local mass media message interacts with the system. We therefore conclude that mass media can increase or decrease cultural diversity depending on the value of the parameter q.

A global picture of direct mass media influences is provided by Figure 3, which shows the asymptotic value of g for long time as a function of q, with $F = 5$, for different values of the probability B. In absence of any mass media influences ($B = 0$) there is a threshold value of the number traits $q_c \approx 25$, such that for $q < q_c$ dynamical evolution always leads to one of the possible q^F homogeneous states ($g \to 0$). On the other hand, for values of $q > q_c$, the behavior of the system changes and it settles into a multicultural state when $B = 0$. When the probability B is increased, the threshold value of q decreases. There is a value $q_c(B)$ below which the system still reaches in its evolution a homogeneous cultural state ($g \to 0$) under the action of any of these mass media messages.

Figure 4 shows the spatial configurations of the final states of the system subject to a global mass media message, when $q < q_c$. In the absence of mass media influence, i.e. $B = 0$, the system settles into any of the possible q^F homogeneous cultural states. When the probability B is increased, the system is driven towards a homogeneous cultural state that depends on the evolution of the global mass media from the initial conditions. Thus, for small values of B, global mass media contributes to maintain a globalized state in the system, as one may expect. However, there is a value of the strength B above which there is no convergence to a homogeneous state, but a state of cultural diversity emerges in the system. A similar behavior is observed when the system is under the influence of local mass media. These results describe for $q < q_c$ a transition at a threshold value of B from a culturally homogeneous state to a state of cultural diversity characterized by an increasing number of cultural domains as B is in-

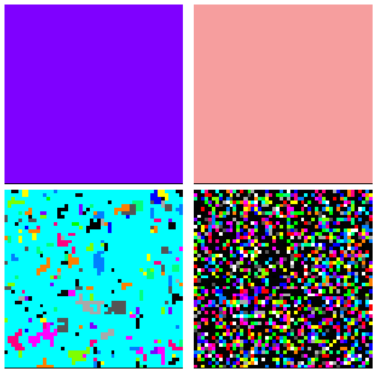


Fig. 4. Asymptotic cultural configurations for different values of the probability B for a direct global mass media influence, for $F = 5$, $q = 10$, and $N = 50 \times 50$. Top left: $B = 0$; top right: $B = 0.01$; bottom left: $B = 0.1$; bottom right: $B = 0.9$.

creased. Thus, we find the counterintuitive result that for $q < q_c$, above some threshold value of the probability of interaction, mass media induces cultural diversity in a situation in which the system would be culturally homogeneous under the effect alone of local interactions among the agents.

Figure 5 shows q_c as a function of B. The threshold value q_c for each type of media decreases with increasing q for $q < q_c$. In each case, the threshold curve q_c versus B in Fig. 4 separates the region of cultural diversity from the region where homogeneous, monocultural states occur on the space of parameters (B, q). For values of B above this curve, the interaction with the mass media dominates over the local interactions among the individual agents in the system. Consequently, agents whose states exhibit a greater overlap with the state of the mass media have more probability to converge to that state. This process contributes to the differentiation of states between neighboring agents and to the formation of multiple cultural domains in the system for large enough values of the probability B.

When there are no mass media influences ($B = 0$), the system always freezes into culturally polarized states for $q > q_c$. Figure 3 shows that the effect of mass media for $q > q_c$ depends on the magnitude of B. For the two types of mass media messages that we consider, small values of B produce a drop in g to values below the reference line corresponding to its value when $B = 0$. Thus, the limit $B \to 0$ does not recover the behavior of the model with only local agent-agent interactions. The fact that for small values of B the interaction with mass media promotes cultural homogeneity is related to the non-stable nature of the inhomogeneous states in Axelrod's model. When the probability of interaction B is very small, the action of mass media can be seen as a sufficient perturbation that allows the system to escape from polarized states with frozen dynamics. The role of mass media in this situation is similar to that of cultural drift [3].

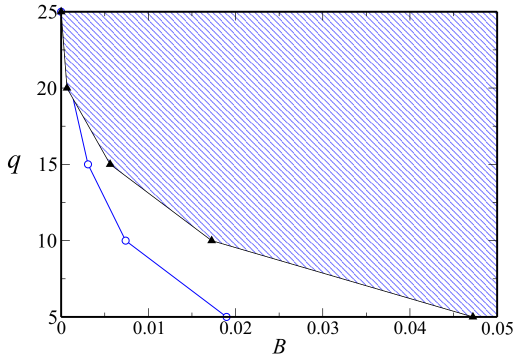


Fig. 5. Threshold boundaries $q_c(B)$ vs. B for $q < q_c$ corresponding to the global and mass media. Each line separates the region of cultural diversity (above the line, in grey) from the region of a global culture (below the line) for direct global (circles) and local (triangles) mass media influences.

The drop in the asymptotic values of g for small values of B from the reference value ($B = 0$) that is observed for the local mass media in Fig. 3 is much more pronounced than the corresponding drops for global mass media. This can be understood in terms of a greater efficiency of a local mass media influence as a nonuniform perturbation that allows the system to escape from a frozen inhomogeneous configuration. Increasing the value of B results, in both types of mass media, in an enhancement of cultural polarization in the system, but the local mass media always keeps the amount of cultural diversity, as measured by g, below the value obtained for $B = 0$. Information feedback at the regional level is more efficient in promoting cultural homogenization. Figure 6 shows configurations reached under the influence of the direct local and global mass media in the multicultural region $q > q_c$. A smaller number of cultural domains (smaller g) that for $B = 0$ are observed for global mass media field. But this number in much smaller under the action of local mass media field. In addition, the time to reach the final state is much longer for a local than for a global field (see Movie 1). Local mass media does not lead to an early frozen state, leaving room for agent-agent interactions that result in a culturally homogenized state.

3 Model of indirect global mass media influence: the filtering of local interactions

In this section we analyze a model of global information feedback where the global mass media acts as a moderator or filter of the local influence of neighbors, as proposed by Shibanai et al. [15]. In the original Axelrod's model one feature with different traits for two neighboring agents is chosen, and the trait of the active agent is changed to that of the neighbor. This is modified in the model of indirect global mass media influence analyzed here, taking into account the agreement of the chosen trait of the neighbor and that of the global mass media or *the plurality* of the population. If the trait of the neighbor is concordant with the dominant one, that is, the same as that of the global mass media message M, the feature of the active agent will be changed to that of the neighbor. But if the feature of the neighbor is different from that of the global mass media message M, then, with probability R the active agent will not change. Thus, this model assumes that agents are more likely to adopt a trait from those neighbors that are concordant with the plurality.

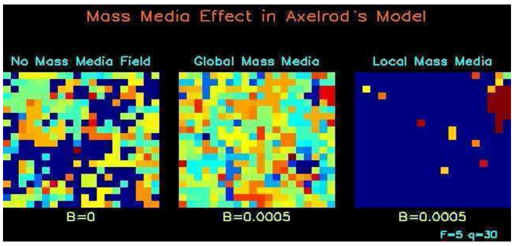


Fig. 6. Cultural configurations for different values of the probability B for different mass media influences in the multicultural region, for $F = 5$, $q = 30$, and $N = 50 \times 50$. Left: no mass media $B = 0$; center: Global mass media with $B = 0.0005$; right: Local mass media with $B = 0.0005$.

We use the definition of a uniform global mass media as in Section II, $M = (\mu_{i1}, \mu_{i2}, \ldots, \mu_{iF})$. The dynamical evolution of the filter model can be described in terms of the following iterative steps:

(1) Select at random an agent i on the lattice (active agent).

(2) Select at random one agent j among the four neighbors of i.

(3) Calculate the overlap $l(i, j)$. If $0 < l(i, j) < F$, sites i and j interact with probability $p_{ij} = l(i, j)/F$. In case of interaction, choose h randomly such that $\sigma_{ih} \neq \sigma_{jh}$. If $\sigma_{jh} = \mu_h$, then set $\sigma_{ih} = \sigma_{jh}$; otherwise with probability R the state of agent i does not change and with probability $1 - R$ set $\sigma_{ih} = \sigma_{jh}$

(4) Update the global mass media vector M if required. Resume at (1).

Figure 7 shows a diagram of the filter model. The parameter R describes the intensity of the filtering effect of the global mass media on the local interactions. The case $R = 0$ corresponds to the original Axelrod's model, while $R = 1$ implies that cultural interaction only causes a change if the chosen trait of the neighbor was equal to that of the global mass media. The overall probability of interaction between an active agent i and a chosen neighbor j is $p_{ij}(1 - R)$ if the chosen trait of j is different from that corresponding to M, and p_{ij} if the chosen trait is equal to that corresponding to M.

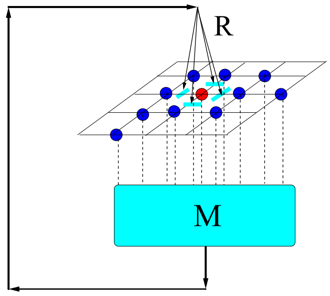


Fig. 7. Diagram representing the filter model.

Figure 8 shows the average fraction of cultural domains g as a function of time in the global mass media filter model, for two values of q with $F = 5$, and for different values of the filtering probability R. In Fig. 8 (left), when $q < q_c$ the system reaches a homogeneous state for $R = 0$ and also for small values of R. However, when the probability R increases, the filtering influence of the global mass media can induce cultural diversity. Our results for $q < q_c$ support the results obtained by Shibanai et al. [15] about the ability of the filtering process to induce cultural diversity in the same fashion as the model with direct global mass media influence. But comparison with Fig. 2 (left) shows that direct interaction with global mass media is more efficient in promoting cultural diversity than the filtering mechanism of agreement with the global plurality. The analysis of reference [15] was restricted to a single value of $q < q_c$. We have also explored values of $q > q_c$, where the system would be in a heterogeneous cultural state in absence of any filtering ($R = 0$). For these values of q we find (Fig. 8, right) that the filtering mechanism has no appreciable effects for small R, in contrast with the case of direct global mass media influence where for small values of the probability of interaction B with the media message, the number of cultural groups is reduced as a consequence of this interaction.

A systematic analysis of the filtering effect for different values of q is summarized in Figure 9 which shows the asymptotic value for long times of the average fraction of cultural domains g as a function of q, with $F = 5$, for different values of the filtering probability R. When no filtering acts on the system ($R = 0$) the behavior is that of the original Axelrod's model and also coincides with the direct mass media models for $B = 0$.

The effects of the filtering process in the culturally homogeneous region, i.e., for parameter values $q < q_c$, is similar to that of a direct influence of endogenous mass media. When the probability R is increased, the threshold value of q decreases. There is a value $q_c(R)$ below which the system still reaches a homogeneous cultural state under the influence of the filter. An increase in R for parameters $q < q_c(R)$ leads to cultural diversity. Thus, both mechanisms of feedback information, either direct or indirect, promote multiculturality in the region of parameters where globalization prevails in the absence of any feedback. The similar behavior found for the three types of mass media influences considered here suggests that the phenomenon of mass media-induced diversity should be robust in this region of parameters, regardless of the type of feedback mechanism at work.

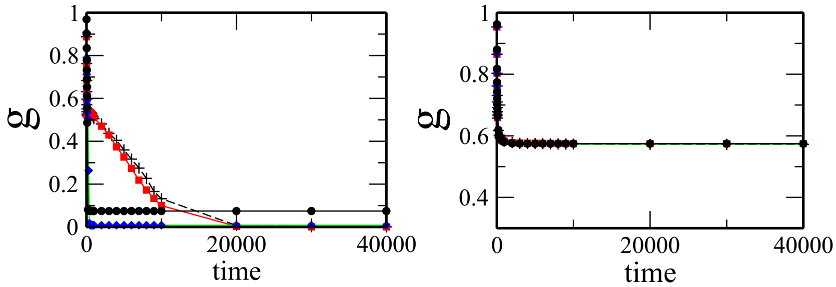


Fig. 8. Time evolution of the average fraction of cultural domains g in the filter model for different values of the probability R, with fixed $F = 5$. Time is measured in number of events per site. System size $N = 50 \times 50$. Left: $q = 10$; $R = 0$ (crosses); $R = 0.0005$ (squares); $R = 0.15$ (diamonds); $R = 0.6$ (circles). Right: $q = 30$; $R = 0$ (crosses); $R = 0.0005$ (squares); $R = 0.005$ (circles); $R = 0.1$ (diamonds).

However, in the region of parameters $q > q_c$ where multiculturality occurs for $R = 0$ or $B = 0$, the behavior of the filter model differs from those of the direct mass media influence. The filtering mechanism has little effect for values of the probability $R < 1$. As $R \to 1$ there is a small decrease in the number of cultural groups formed in the system. But at $R = 1$ a discontinuity appears: the fraction of cultural groups g jumps from a value close to the one for $R = 0$ to a value close to $g = 1$ corresponding to maximum multiculturality (number of cultural groups equal to the number of agents in the system). The case $R = 1$ corresponds to an extreme restriction on the dynamics, when no adoption of cultural features from neighbors is allowed unless the state of the neighbor coincides with the one of the global mass media. Since we are considering random initial conditions, when q is large enough, the probability that the features of any agent coincide with those of the global mass media message M is quite small, making the convergence to globalization, i.e., a common state with the media, very unlikely. As a consequence, the random multicultural state subsists in the system and manifests itself as a maximum value of g. The small probability of interaction with the global mass media for large values of q when $R = 1$ is also reflected in the very long convergence time needed to reach the final multicultural state as compared with the convergence time for $R < 1$.

We note that the case $B = 1$ in the model of direct global mass media influence is less restrictive than the condition $R = 1$ in the filter model. Although local agent-agent interactions produce negligible effects in both cases, in the model of direct influence, an agent can still interact with the global mass media when there is some cultural overlap between the agent and the mass media message.

4 Conclusions

We have studied the effects of different forms of information feedback on a social system in the framework of Axelrod's model for cultural dissemination. Two basic mechanisms of information feedback have been interpreted as models of direct or indirect mass media influence on the system [15]. For the direct information feedback mechanism we have considered the cases of global mass media and local mass media influences [16]. The dynamics of direct interaction with the mass media is based on the similarity or overlap between vector states. Although their origin is different, at the local level the direct global and local mass media influences act in the same manner, as a "fifth" effective neighbor whose specific source is not essential. Consequently they produce similar effects in the system. First, we find the nontrivial result that direct mass media can enhance cultural diversity, in agreement with the results of [15]: for parameter values for which the system reaches a global culture due to the local interaction among the agents, there is a threshold value of the probability of interaction B with mass media. For values of B above this threshold cultural diversity emerges. This happens because there is a competition between the consequences of the similarity rule applied to the agent-agent interactions, and applied to the agent-mass media interaction. This competition leads to the formation of cultural domains and to polarized states. Secondly, we find another surprising effect: for parameter values for which the dynamics based on the local interaction among the agents produces a frozen multicultural configuration, very weak directly interacting mass media produces cultural homogeneity. This range of parameters was not explored in Ref. [15], ad we find here that mass media interaction for $q > q_c$ produces the opposite effect of the one described before. That is, it favors cultural homogenization.The limit $B \to 0$ is discontinuous and the homogenizing effect for $B << 1$ occurs because the interaction with the mass media acts as a perturbation on the non stable multicultural configurations with frozen dynamics that appear for $B = 0$. As a general result for any value of q, we find that for small values of B, the interaction with mass media promotes cultural homogenization: for $q < q_c$ this interaction preserves homogeneity, while for $q > q_c$ it causes a drop in the degree of cultural diversity in the system as measured by the average fraction of cultural domain g. We have calculated, for the two types of direct mass media influences considered, the corresponding boundary in the space of parameters (B, q) that separates globalization and polarization states. The spatially nonuniform local mass media has a greater ordering effect than the uniform global mass media in the regime $q > q_c$. The range of values of B for which globalization occurs for $q < q_c$ is also larger for the nonuniform local mass media.

The effect of the indirect global mass media influence as a filtering process in the culturally homogeneous region, i.e., for $q < q_c$, is similar to that caused by a direct influence of mass media. For small values of the filtering probability R the system reaches a culturally homogeneous state. For values of R greater than a threshold value the system converges to a state of cultural diversity. Thus, both mechanisms of feedback information, either direct or indirect, promote multicul-

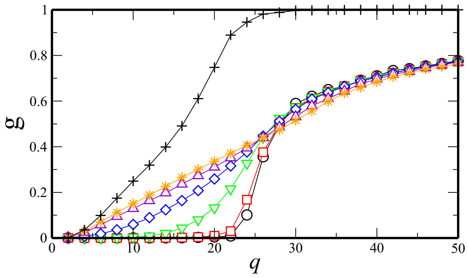


Fig. 9. Average fraction of cultural domains g as a function of q, for different values of the probability R for the filter model. $R = 0$ (circles); $R = 0.01$ (squares); $R = 0.1$ (triangles down); $R = 0.5$ (diamonds); $R = 0.9$ (triangles up); $R = 0.99$ (stars); $R = 1.0$ (plus signs)

turality in a region of parameters where it would not be present in the absence of any feedback. In the region of parameters $q > q_c$ where multiculturality occurs for either $B = 0$ or $R = 0$, the filtering mechanism has, for values of the probability $R < 1$, a very weak effect in comparison to the one caused by a direct mass media influences: there is only a small decrease in the number of cultural groups formed. However, when the extreme restriction $R = 1$ is imposed, the number of cultural groups jumps discontinuously to a value corresponding to maximum multiculturality.

The similarity of the behaviors of the different types of mass media influences considered here suggests that the phenomenon of mass media-induced diversity should be robust, regardless of the type of feedback mechanism. In spite of the differences mentioned between direct and indirect information feedback processes as well as between uniform and nonuniform mass media influences, it is remarkable that the collective behavior of the agents displays analogous phenomenology for the three types of mass media considered.

Generally speaking, our analysis unveils the delicate compromise between direct agent-agent interactions and feedback processes. Mass media reflects local or global cultural trends created by local agent-agent interaction, but mass media information is processed by agent interactions, while the agent-mass media interaction is conditioned by the overlap of the cultural features of the agent and the mass media message. We have analyzed the effect of different forms of mass media for the full range of the parameter q that measure an initial cultural diversity. Our results indicate qualitatively different effects when globalization ($q < q_c$) or polarization ($q > q_c$) would prevail when no mass media feedback is taken in account. In summary, we find that, when the probability of interacting with the mass media is sufficiently large, mass media actually contribute to cultural diversity in a social system, independently of the nature of the media. But direct mass media influences are found to be efficient in promoting cultural homogeneity in conditions of weak broadcast of a message, so that local interactions among individuals can be still effective in constructing some cultural overlap with the mass media message. Strong media messages do not lead to cultural homogenization because agent-agent interaction becomes inefficient. These results identify the power of being subtle in mass media massages. In addition, direct local mass media appear to be more effective in promoting uniformity in comparison to direct global broadcasts, which identifies the importance of local media (feedback at regional levels) in the cultural globalization path.

Acknowledgments

J. C. G-A., V. M. E. and M. SM acknowledge financial support from MEC (Spain) through project CONOCE2 (FIS2004-00953). M. G. C. acknowledges support from FONACIT (Venezuela) under grant F-2002000426. K K. K. acknowledges support from DFG Bioinformatics Initiative BIZ-6/1-2 and from Deutscher Akademischer Austausch Dienst (DAAD).

References

1. R. Axelrod, J. of Conflict Resolution **41**, 203 (1997).
2. C. Castellano, M. Marsili, and A. Vespignani, Phys. Rev. Lett. **85**, 3536 (2000).
3. K. Klemm, V. M. Eguíluz, R. Toral, and M. San Miguel, Phys. Rev. E **67**, 045101(R) (2003).
4. M. San Miguel, V. M. Eguíluz, R. Toral, and K. Klemm, Computing in Science & Engineering **7**, 67 (2005).
5. K. Klemm, V. Eguíluz, R. Toral, and M. San Miguel, Physica A, **327**, 1 (2003).
6. K. Klemm, V. Eguíluz, R. Toral, and M. San Miguel, J. Econ. Dyn. Control, **29**, 321 (2005).
7. A. Flache and M. Macy, Los Alamos Arxiv:physics/0604201 (2006).
8. M.W. Macy, J. Kitts, A. Flace and S. Bernard, in Dynamic Social Network Modeling and Analysis. EDs. R. Breiger et al., National Academies Press (Washington, 2003).
9. L. Leydesdorff (2001), Journal of Artificial Societies and Social Simulation **4**, no. 3, http://www.soc.surrey.ac.uk/JASSS/4/3/5.html.
10. R. Bhavnani (2003), Journal of Artificial Societies and Social Simulation **6**, no. 4, http://jasss.soc.surrey.ac.uk/6/4/1.html.
11. J. C. González-Avella, M. G. Cosenza, and K. Tucci, Phys. Rev. E **72**, 065102(R) (2005).
12. K. Klemm, V. M. Eguíluz, R. Toral, and M. San Miguel, Phys. Rev. E **67**, 026120 (2003).
13. J. Greig, J. of Conflict Resolution **46**, 225 (2002).
14. D. Centola, J. C. González-Avella, V. M. Eguíluz and M. San Miguel, arxiv:physics/0609213.
15. Y. Shibanai, S. Yasuno, and I. Ishiguro, J. Conflict Resolution **45**, 80 (2001).
16. J. C. González-Avella, V. M. Eguíluz, M. G. Cosenza, K. Klemm, J.L. Herrera and M. San Miguel, Phys. Rev. E **73**, 046119 (2006).

Session on Epistemological Issues

Why do Social Geographers Have Problems in Applying Agent-Based Geosimulation?

Andreas Koch

Department of Geography and Geology, University of Salzburg, Hellbrunnerstr. 34, 5020 Salzburg, Austria
Andreas.Koch@sbg.ac.at

Abstract. The paper indicates an epistemological incompatibility between contemporary theories of space in social geography and methodological approaches in geosimulation. There is, however, on both sides the potential to relax extreme positions and then to benefit from geosimulation in social geography, if the idea of constructionism is added to the positivistic paradigm of geosimulation. This leads to the idea that reality is not given objectively and that modelling techniques, however, are able to represent realities as a result of social communication.

Introduction

The application of agent-based geosimulation methods in social geographical studies is not yet widely accepted. Certain relevance can be revealed in urban geography [1, 2] or transportation [3]. In these fields of research, the modeling of processes of spatial spreading and allocation is central rather than modeling explicitly socio-spatial phenomena. This reluctance is to a certain degree determined by apparently incompatible epistemological attitudes between theoretical approaches in social geography and methodological approaches in the sphere of modeling and simulation. The article aims to overcome this incompatibility by arguing that, on the one hand, social geography has to abandon its claim of a one-sided theoretical reference to space, and on the other hand, geosimulation has to revise its positivistic epistemology. According to Benenson and Torrens [1], geosimulation, understood as agent-based modeling technique, is defined as follows: "Geosimulation is concerned with the design and construction of object-based high-resolution spatial models, using these models to explore ideas and hypotheses about how spatial systems operate, […] to support object-based simulation, and applying simulation to solving real problems in geographic contexts". An appropriate starting point is Axelrod's [4] general notion of simulation as a third way of doing science – a composition of induction and deduction: "Like deduction, it starts with a set of explicit assumptions. But unlike deduction, it does not prove theorems. […] Unlike typical induction, however, the simulated data come from a rigorously specified set of rules rather than direct measurement of the real world". An example of modeling socio-spatial interactions in an agent-based simulation model will be presented during the conference.

The Paradigm of Social Geography

Contemporary social geographical approaches concerning the relationship between space and society take the view that space is exclusively the result of social constructions. Space emerges through individual actions in social contexts. Consequently, space can be interpreted as a social category, which expresses itself exclusively through language – i.e. as landscape, region or place – and, therefore, cannot be visualized by models. This asymmetry of constructionist capability (not construction itself) can be restored to symmetry by applying a system theoretical approach [5]. As well as social systems, which are able to generate social realities of space and time through communication, we have to take into account spatial systems with independent properties of creating spatial realities. Simultaneously, both types of systems mutually depend on one another insofar as the social construction of space induces the spatial construction of society - by creating spatial references, social systems are embedded into these spatial references. The connection of systems is thereby realized by structural coupling. By applying this approach it is possible to imagine space as a construction *and* a given social reality, being able then to model it.

Geosimulation

While social geography, with its theories of space, makes it difficult to benefit from geosimulation as a method of representing socio-spatial processes of interaction, geosimulation similarly has problems with integrating contemporary theories of space due to its strong emphasis of a positivistic paradigm. With its epistemological approach of representing reality or truth as objective, the scientific community will not readily accept the findings of social geographers. My intention here is not to indulge a passion of postmodernism, but to account for a justified critique of a traditional concept. The application of agent-based geosimulation can contribute to improving knowledge about socio-spatial phenomena by complementary methods of analysis and phenomenology and by explicitly integrating processes into investigations.

Benenson and Torrens [1] opt explicitly for a positivistic understanding of modeling: “Here is our manifesto: We say yes to (a) positivism; […] (c) we consider reductionism as an initial point of any research; […] (e) we believe that practice and experiment are criteria of a theory, what might seem logical can nonetheless be wrong”. To criticize positivistic thinking means foremost to criticize an inductive way of reasoning. One proceeds on the assumption that studies and results based on simulation models can be generalized; what is represented by a model correlates with the objective truth, one even admits that the model is a simplification of reality. “Two motivations for a geosimulation modeling paradigm can be considered […]: *Pragmatic*, which uses the automata approach as a tool without any reference to the theoretical issues that have emerged from urban, social, and general system theory, and *paradigmatic*, which aims at investigation of cities as sociospatial systems”. Undoubtedly, a big advantage of multi-agent simulation is to examine and explore real-world problems in an experimental way; however, without any relation to theoretical issues it seems to be a poor endeavor. Complementary to induction, deduction is an important way of reasoning.

Simulation and modeling do meet two requirements: representation and communication. Both requirements serve to make relationships visible – on different levels. This means to accept multiple forms of representation and to make interactions between modelers, models and users visible. Due to different knowledge, experiences, and intentions of actors being involved in creating and using simulation models, a different perception on what a model represents in communications can be concluded. And this in turn means that a simulation model does not have an objective truth or reality, but that we have to take a synthesis of multiple constructions of subjective truths into account. Such a synthesis towards a model might then be able to achieve a realistic quality within the scientific community.

Against this background, several statements of Benenson and Torrens (p. 249) about the epistemology of agent-based geosimulation should be modified. For instance they wrote: "*The concept of agents makes it possible for a model to directly reflect human behavior*". Apart from the fact that no model, arguably, is able directly to reflect human behavior, we have to examine to what extent human behavior is able to be formalized appropriately and, hence, is able to be translated into a modeling language. This depends crucially on the theoretical premises and conceptualization, respectively. Consequently, the following statement seems to be a problem as well: "*Agent-based models do not demand comprehensive knowledge of the phenomena being studied*". Whether this is possible or not depends on the complexity of the phenomenon being examined and, therefore, it depends on the theoretical context. This is also true for this statement: "*Formalization of behavioral rules reveals gaps in our knowledge of urban processes*". This conclusion can easily be drawn; it is, however, a debatable point whether it can be done solely by a variation of the model parameters, without any feedback to the theoretical approaches in mind as they argue: "The ability to vary formal representation of agent behavior often reveals gaps in our knowledge, and MAS [multi-agent simulation, A.K.] models reveal which behavioral parameters can make a more significant contribution to our understanding of urban residential dynamics, or the dynamics of any urban system".

References

1. Benenson, I., Torrens, P.M.: Geosimulation. Automata-based modeling of urban phenomena. Chichester (2004)
2. Batty, M., Xie, Y., Sun, Z.: Modeling Urban Dynamics through GIS-based cellular automata. In: Computers, Environment and Urban Systems 23. (1999) 205-233
3. Helbing, D., Keltsch, J., Molnar, P.: Modelling the evolution of human trail systems. In: Nature 388 (1997) 47-50
4. Axelrod, R.: Advancing the Art of Simulation in the Social Sciences. In: Conte, C., Hegselmann, R., Terna, P. (eds.): Simulating Social Phenomena. Lecture Notes in Economics and Mathematical Systems, Vol. 456. Springer-Verlag, Berlin Heidelberg New York (1997) 21-40
5. Koch, A.: Autopoietic Spatial Systems. The significance of actor network theory and system theory for the development of a system theoretical approach of space. In: Social Geography, vol. 1. (2005) 5-14

The Challenge of Context Permeability in Social Simulation

Luis Antunes[1], João Balsa[1] Paulo Urbano[2], and Helder Coelho[1]

[1] GUESS/Universidade de Lisboa, Portugal
{xarax,jbalsa,hcoelho}@di.fc.ul.pt
[2] Universidade de Lisboa, Portugal
pub@di.fc.ul.pt

Abstract. In multi-agent-based social simulations the connections between agents are usually represented in a simple network structure. This is done either for the purpose of simplifying the huge complexity of the social world, or because of the features of available tools. Facing the rich complex set of social relations an agent engages in, we propose to go a step further in the modelling by considering a multitude of concomitant social relations. This setting can be seen in a simulation as a n-dimensional scenario where each (n-1)-dimension surface represents a different social relation.
Agents belong to distinct contexts in these multiple relations. So, in models defined over this new social structure, besides interactions between agents, it is possible to define interactions between different contexts, allowing information to flow between them. In this paper, we propose the concept of "context permeability" in a social topology; illustrate the concept with a simple example; then elaborate on various alternatives for materialising this concept; and finally present some comparative results of simulations of a typical case: social consensus.

1 Introduction

In some multi-agent-based simulations (MABSs), agents are located in virtual spaces where dimensionality is irrelevant. Typically, all agents can participate in interactions with all other agents, and the notion of physical space plays no role in the simulation. In other MABSs (e.g. the study of innovation in technological parks [14]), the notion of geographic place is important to establish relevant notions (neighbour, distance, obstacle, etc.) that provide a filter to what kind of actions and interactions the agents can engage in. Usually, simulation environments provide bi-dimensional grids (more recently tri-) that represent the target environment in such a way that can be visualised in a computer screen. It is a thought courtesy made for the eyes and the mind of the observer/analist of the simulation, who can easily and readily capture phenomena with his/her own eyes, that would otherwise be difficult to define formally and detect analytically [1].

The problem is that most times the kind of social relations being studied is not really adapted to this representational support: a continuous grid, marked by an array of discrete slots. Many times, this simple relation representation is used as a simplifying metaphor of the entire complex set of social relations the agent is involved in. In the real social world, agents belong simultaneously to several relations with other agents and/or institutions, and these relations may be of different kind and quality. To expect the complexity of social relations to be collapsed into a single relation depicted in a bi-dimensional space is overly simplistic and may jeopardise the quality of simulation results and undermine the confidence in the derived conclusions and their applicability.

So, we propose to take one step further, and consider multiple concomitant social relations, which provide the agents with the possibility of entering different social contexts. A context is something like your family, your work colleagues, etc. Naturally, these contexts will have points in common, which provide what we call context permeability. Our conjecture is that this feature of the social world is of extreme importance for the dissemination of phenomena, and societal adaptiveness.

In the rest of this paper, we explore the ideas of multiple social relations, and elaborate on possible representations for this concept. We discuss the base structure of the topological notion to support multiple relations and contexts. We then illustrate our ideas with a simple example, social consensus, and present the results of simulations that show how context permeability can quantitatively and qualitatively change the achievement of consensus in a socially structured society.

2 Multiple Social Relations

Obviously, humans and other social agents engage in several relations simultaneously, and many of those relations carry with it a context of intervention that cannot be neglected in most social phenomena, much less in social simulation.

One interesting theme that has arisen in the literature is the application of multiple social networks (so called 'multi-plex relations') to uses in the Semantic Web. Matsuo et al [10] explain how the knowledge about multiple social relations can be elicited from several sources to generate an intricate and complex picture of the sociality of a group of agents. In this paper, three different social relations are presented, taken from a set of real data collected over four different conferences: *knows*, *collaborates*, and *meets*. Each of those relations is unveiled by using distinct methods, but for our present purposes, we will focus on the structures these relations show (cf. figure 1). Besides, the authors do not really say how the integration of those three relations could be done for purposeful use in the Semantic Web.

These multiple social relations illustrate the kind of potential we can expect from endowing our simulations with a similar schema. However, we must be aware that added complexity can arise from such representations. While we want to emphasise the infra-structure and not focus on specific applications, let us examine some examples.

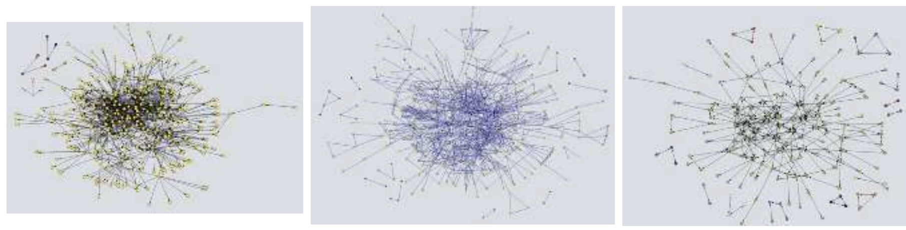


Fig. 1. Multiple social relations: *knows*, *collaborates*, and *meets*. From [10].

The structures found in the real world are often far from theoretical predictions and even unsupported assumptions used in MABS. For instance, the 'small-world' type of network connections is often adopted for the mathematical properties it exhibits and the theoretical and practical tools it allows. Nevertheless, as argued in [13], that structure may be more difficult to find in reality that practitioners suspect, as the conditions and constraints that end up yielding it may be rather atypical. In a case such as the one depicted in figure 1c (relation *meets*), we find several small clusters of agents which are isolated from the main group. Should this relation be used to allow contacts among agents, those isolated groups would never be reached (something that seldom occurs in small-world networks, where a path exists between any given two agents). Our proposal of considering simultaneously several contexts may allow for those groups to be accessed, since the nature of the other contexts will provide increased opportunities for liaisons.

Another reason for considering multiple contexts in separate planes is the unbalanced nature of some relations. Consider relation *knows* in figure 1. While it can be expected that I know the President of the Portuguese Republic (PPR), it can hardly be reasonable that he knows me. So, this relation *knows* is, for certain applications, assymetrical, and can be decomposed in two other relations, that could be called *knows* and *is-known-by*. Again, when considering explicitly the representation for these relations, we deepen the realism adopted in our simulations, which can be useful for the simulation purposes. This approach is adopted in [8], where the authors emphasise the need for using multiple social relations, and describe several techniques for their representation.

The PPR example also gives us another example of the need to maintain representations of multiple simultaneous relations: drawing a paralell with the notorious Erdős number [3]. Imagine that I want to obtain the telephone number of the PPR. Assuming benevolence from all involved, I would call my father, who

[3] Hungarian Paul Erdős was a travelled and prolific mathematician who wrote hundreds of scientific papers with hundreds of co-authors. His Erdős number is 0, whereas the Erdős number of his co-authors is 1. The Erdős number of his co-authors' co-authors is 2 and so on and so forth. An analogue concept was defined for cinema and is known as the Bacon number. Both have been used to exemplify the notion of small-world network structure.

knows someone that surely has the number of the former PPR, which certainly has the number of the present PPR. Let's say I am a PPR4. While the co-author relation used to calculate the Erdős number is well defined and objective, in the PPR example the problem is far more complex. Not only the PPR would have far more trouble to uncover my phone number, due to the assymetry of relations, but we use a multitude of different relations and, moreover, of knowledge about those relations.

A similar point can be made in a network of social actors when you try to establish founded trustworthiness [10], a very important concept in the applications of Multi-Agent Systems in the industry [3, 5]. When information starts being gossiped away in a epidemic manner, its spreading can be greatly enhanced by starting focuses at a distance, due to the simultaneous presence of a defamating agent in several contexts. A repeated lie can so easily be transformed in truth for all relevant purposes, because the targeted agent is helpless to fight the snowball effect.

3 Two Models of Multiple Contexts

We can look at the topologies used in social simulation from two different perspectives. On one hand, we can focus on the problem at hand (target phenomenon) and observe which kind of structure emerges from the problem description and corresponding simulation. This is the kind of approach used in [13], where the authors concentrate on the individual *reasons* and collective conditions that can bring about several topological structures already found in an important collection of typical situations(for instance [16, 4]).

In this paper we take the alternative view. What we are interested is in knowing which kind of representational infra-structure can provide the adequate setting for those complex topological structures to appear, and which kind of restrictions are imposed by the chosen infra-structure. In particular, we want to know how the simulation "geography" can influence the propagation of social concepts through the simulation.

In this first take on the problem, we will focus our attention in two special types of such topological representations: multiple contexts with separate sets of agents and possibility of agent transference; and multiple contexts with common agents but different connections (social networks). Our conjecture is that phenomena will be significantly qualitatively different in both these approaches.

3.1 "Beam Me Up, Scottie."

One important category of topological context permeability in a setting where agents live in separate worlds, and there is the possibility of transference from one world to another. These worlds can represent different relations: an agent can belong to social relations, but possibly not simultaneously. In the simulations, it is as if an agent has focus on one particular relation he is engaged in, and then for whatever reason it can change focus. In figure 2 we represent this scenario. (Hyper-)planes on this (hyper-)cube represent different settings in which relations are placed. Agents can move from one plane to another, and this represents the shift from one context to another.

The features to model in this category are:

- The number of relevant worlds to consider;
- The connectedness of agents in each of the worlds;
- The geographical/social/etc. obstacles to connections between agents and how these are modelled;
- How/when/where/why/to-where transferences of agents occur;
- Which transformations agents undergo during transference.

At his point of our research, we will claim the importance of this view by considering a simplistic version of such a setting. We consider two worlds, in which all agents are placed in an unrestricted uniform grid. Agents can move anywhere on this grid. In each of the worlds there is one special square part of the grid such that when an agent steps inside the square it will get transported to a corresponding place in the other world. See figure 2.

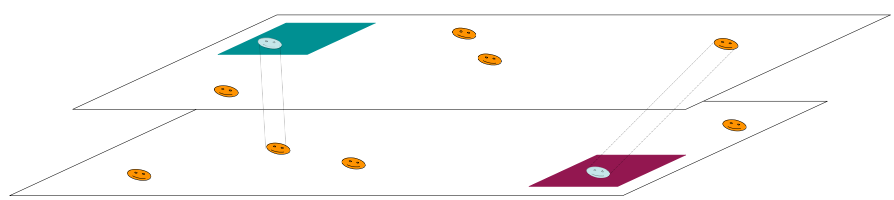


Fig. 2. Separate sets of agents in parallel worlds.

The framework we have just described is not unlike the one in [6], if we correspond our planes to their concept of *groups*. We can even consider that when an agent changes from one plane to another, it can fulfil a different *role*. However, in our case we have an augmented dynamism, as our model emphasises explicitly the policies of transfer between planes. Additionally, their concept of *world* can be corresponded reasonably to our *context*. Ferber et al. consider two types of worlds: organisations and physical worlds, which in our case would correspond to classes of different relevant contexts. In the different planes we can consider mutually exclusive (but not incommunicable) sets of agents, so allowing us to represent a wide range of complex social situations.

Arguably, another important infra-structural feature of the simulations is the update regime of the agents. This is usually left for the simulation scheduler to decide, and has been subject to long debate. Approaches such as Swarm [11] have left that issue to the discretion of the developer, whereas others such as NetLogo [17] or RePast [12] provide default update and scheduler regimes. In the case of NetLogo, it is difficult to envisage a way of changing the default scheduler, but this drawback is easily overtaken by the advantages of quick prototyping it offers. For the moment, we will leave this discussion as it is, and we concentrate on how the static features of the topological solution influence the dynamism of the simulation.

3.2 "Who Told You *That*?!"

In most interesting applications, an agent in immersed in a complex social world where it engages in a multitude of relations with other agents (and even aggregate agents such as institutions or informal groups). In this section we consider another category of context permeability, in which agents can belong simultaneously to several relations. As can be seen in figure 3, now the planes in the cube represent different relational settings, and the agents can engage in those relations with different partners in different settings.

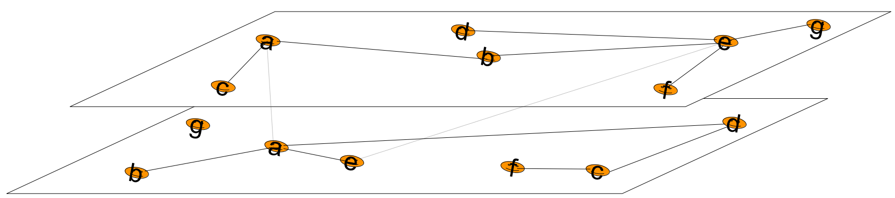


Fig. 3. Different relations for the same agents.

The same agents populate the several views on social reality that are included in the simulation. So, in the course of the simulation, agents will have to face several contexts constantly, either simultaneously, or at rhythms imposed by the different settings. The key features to model in this category are:

- The number of relevant worlds to consider;
- The connectedness of agents in each of the worlds;
- The geographical/social/etc. obstacles to connections between agents and how these are modelled;
- The pace at which each world evolves;
- The kind of interactions agents can undergo in each world.

In this case, the main objective of the framework is to include the possibility of finer grained modelling. For instance, we can use this multiple-planed composition only to provide one of the planes of the previous section, as we depict in figure 4.

Our move from simple to multiple social relations is similar to the one that Flache and Hegselmann made when considering irregular grids in cellular automata approaches [7]. In their case, they were concerned with the shape of

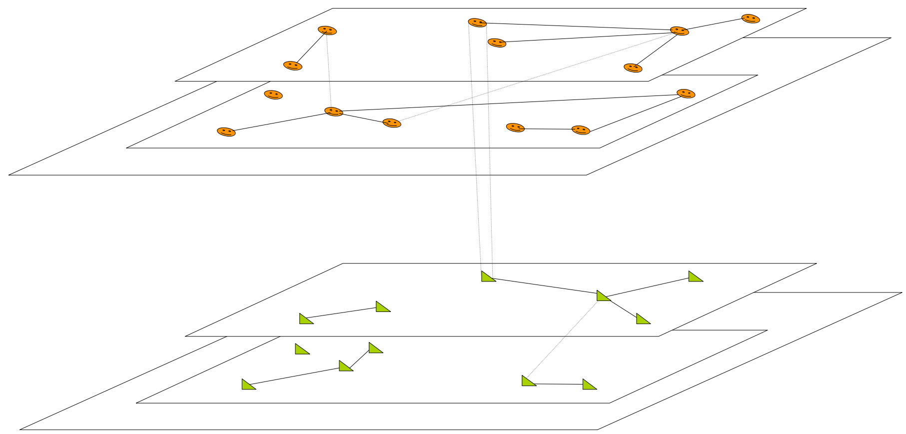


Fig. 4. Composition of modelling approaches.

neighbourhoods, accurately pointing that neighbourhoods such as von Neuman's or Moore's imposed a far too restrictive constraint on the modeller. Flache and Hegselmann's analysis is focused on the robustness of results when changing neighbourhood shapes, whereas in our case we are concerned with the expressiveness of the modelling effort by considering multiple simultaneous relations.

4 Representation and Visualisation

In Multi-Agent-Based Social Simulation it is important not to unbalance the trade-off between representation and visualisation. A problem can be so rich that the required representation paraphernalia is extremely complex. However, especially due to the exploratory character of many simulations, it is important that results can be manageable by the experimenter and provoke interesting insights. It is frequent that the simulation run takes a few minutes and then the analysis of results several weeks (see for instance [2]). An important feature of manageability is visualisation. As we argued before, many concepts in a simulation can only be 'discovered' *a posteriori*, and due to the fuzziness of its nature can hardly be defined and automatically detected. A visual probe of the experiment can be most useful for the experimenter to capture such phenomena and to think about them.

The danger of balancing simulations towards visualisation is that it can constrain the representation and make the designer use overly simplistic concepts. What can be visualised in necessarily much simpler than what can be represented, since it must all be adapted for screen observation (cf. for instance [8]). This is especially important when choosing the simulation geographical space and geometrical laws. Visualisation can help analysis, and the role of the experimenter/observer cannot be neglected [1], but we mustn't constrain our designs by the features we predict we can observe.

We propose to use the 3D features of modern platforms (e.g. NetLogo [17]) and visualise the multitude of social relations as a cube, such as depicted in figures 2 and 3. Even if the visualisation needs can still drive us towards simplistic representations (as can, for instance, the management of complexity in representations, or else we run the risk of making a model too complex to provide any enlightenment on the problem at stake), we can have each plan in this cube with different topologies, geographies, rules, geometries, constraints, etc.

For the example we built for illustration, we use a sliced 3D cube. In every slice (square piece of plan) we have a uniform geography with Euclidean geometry.

5 Example: Consensus Formation

The formation of consensus has been studied in communities of artificial agents [9, 15]. To illustrate the notion of context permeability we run a set of experiments related to this example.

The first series of experiments were concerned with a simple consensus game where agents have to make a collective (but essentially arbitrary) choice. Examples of similar choices are the side of the street you drive on, or the particular word you assign to a given meaning. It is important to have consensus, but the actual content of the consensual choice is irrelevant. So, as the subject to be agreed on we picked just the 'colour' attribute of the agent.

In our game, randomly selected pairs of agents would interact until total consensus was achieved, that is, all agents have the same colour. In each interaction, one of the two agents would choose to adopt the colour of the other. While in the description any two agents can be selected, for the NetLogo simulation we let the agents walk randomly in the grid and consider that a selection of a pair coincides with two agents walking into the same patch. We made several tests to ensure that this is by no means different from selecting randomly the two agents from a common pool.

We then proceeded to repeat this experiment with the multiple context approach (same agents, different planes, translated easily into squares at different heights in a 3D NetLogo cube). We executed 1000 runs of each simulation, and recorded the number of individual meetings necessary to achieve total consensus. Our findings are summarised in table 1.

It seems apparent that there is no significant difference between the one-dimensional and the multi-dimensional approaches, even when the number of agents (n) grows. Now, this game can be seen as a multiple particle Brownian movement game, and when n grows, the time until consensus increases quadratically. This fact is due to the arbitrary nature of each of the decisions. And this also explains why the multiple-plane approach renders small or no improvements. We are arguing in favour of context in a setting where apparently context

Table 1. Average number of meetings necessary to achieve total consensus.

No. Agents	100	200
1 plane	7295	26845
2 planes	6996	25885
3 planes	6918	25355
4 planes	6913	26569
5 planes	6839	28137

is absolutely irrelevant. The game is too simple, and too *staccato* for context to be useful.

Hence, in the next series of experiments, we explored a slightly more complex game, in which agents do not change sides arbitrarily, but rather *remember* the total of meetings they had with each side, and adopt the side that has the highest frequency. This is called the 'majority' game. When deployed in a single surface, this game typically shows the following behaviour: after a random initialisation, for some time both sides keep quite similar proportions. Then suddenly, one of the sides becomes decisively stronger, and proceeds quickly towards colouring all the agents.

When n grows, the total number of meetings to achieve global consensus grows logarithmically ($O(n \, log \, n)$), and so does the the initial period of instability. Table 2 shows what happens when we run the game in a multiple-plane setting. We defined this notion of instability by the bounding of the proportions between both sides of the choice in a 20% band. That is, we consider that the instability period is over as soon as the difference between the proportion of sides is higher than 20%. The winning side by that time is invariably the final winner of the game.

Table 2. Majority game: averages (over 5000 simulation runs) of the number of meetings necessary to achieve total consensus, the number of such meetings that took place after the instability phase, and the percentage duration of the instability phase over total time.

Agts.	50			100			200		
	Meetings	After inst.	Inst.	Meetings	After inst.	Inst.	Meetings	After inst.	Inst.
1 pl.	1680	1222	27%	4910	3510	29%	13666	9479	30%
2 pl.	1924	1266	34%	4933	3383	31%	11857	8261	30%
3 pl.	2569	1590	38%	6314	4170	34%	14132	9844	30%

When we pass from one plane to two planes, we see that the total number of meetings required to achieve consensus is of the same magnitude (worsens slightly when we have 50 or 100 agents, and sensibly improves when we consider 200 agents). When we pass to three planes, we find that the total number of meetings is quite higher. Further experiments with four and five planes showed that this tendency is kept. It is interesting to note, however, that convergence to consensus after the initial instability period is quite faster when we consider multiple planes. This can be seen in absolute numbers, as the pace of growth of the number of meetings after instability ends is very small, and significantly smaller than the growth in total number of meetings. To reinforce this, we can look at figure 5. In less populated worlds, the weight of the instability in the total number of meetings can be quite high. But as the number of agents grows, the instability phase tends to keep a constant proportion.

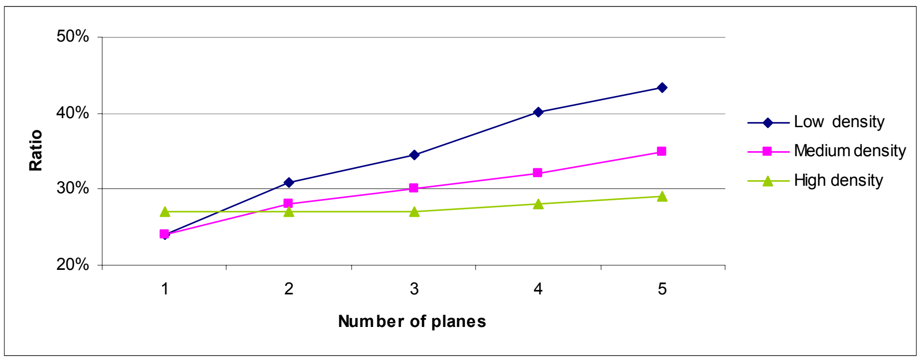


Fig. 5. Evolution, for different number of planes, of the ratio convergence period over total time, with different agent densities.

While admitting that more experimentation will be required, our simulations show us that the multiple context approach yields results that are qualitatively different from the ones obtained by considering a single context. Some of the numbers still need further analysis, and more parameter exploration. All these experiments were conducted a very simple setting, where the key notions for connectedness are based on geographical space and Euclidean geometry. After we set the basics of the multiple context framework, we will need to further explore the structure of social relations and its consequences for the simulation.

6 Conclusions and Prospective Work

Agents engage in a simultaneous multitude of social relations, all of which can be relevant for a given problem. Current literature focus on this issue from the standpoint of the origin (and respective conditions) of such relations. In this paper we proposed to expand this focus by tackling the use of such multiple relations in social simulation. For that purpose, we need to contemplate a balance between representation and visualisation. Two categories of environments were considered, one with the same agents in multiple relations, and another with distinct physical environments where agents can be placed. Our experiments considered only the first of these categories. It is clear that multiple contexts will provide solutions that were impossible in a single one, for instance, to access agents in communities that are isolated for some given connection (e.g. computers inside a firewall), but can be addressed through other connections (in the same e.g., secure protocols). Our preliminary results show that the benefits from this approach can also come either from enlarging the problem space search phase (enhancing exploration of solutions before a premature convergence) or from using preliminary search to reduce the weight of subsequent solution approximations.

Future work will concentrate first on completing the exploration of the present set of experiments, in order to obtain more solid results, as some of the numbers remain still unexplained. Then we will proceed to propose a more general framework to consider multiple social relations and respective contexts in a social simulation setting. In particular, we will survey the literature to build an ontology of simple social relations and their representations, to propose then a more complex ontology that encompasses multiple simultaneous relations. For each of these relations and combinations of relations we will then describe the most common and useful notions of context (neighbourhoods, distances, connectedness, and other properties).

Acknowledgements

We wish to express our thanks to the anonymous reviewers for their comments and useful suggestions.

References

1. L. Antunes, H. Coelho, J. Balsa, and A. Respício. e*plore v.0: Principia for strategic exploration of social simulaiton experiments design space. In *Proc. of the First Congress on Social Simulation WCSS 2006*, Kyoto, Japan, 2006.
2. R. Axelrod. A model of the emergence of new political actors. In N. Gilbert and R. Conte, editors, *Artificial societies: the computer simulation of social life*. UCL Press, London, 1995.
3. C. Castelfranchi, S. Barber, J. Sabater-Mir, and M. Singh, editors. *Trust in Agent Societies*. AAMAS Workshop, Utrecht, The Netherlands, 2005.
4. J. Delgado. Emergence of social conventions in complex networks. *Artif. Intell.*, 141(1/2):171–185, 2002.
5. R. Falcone, S. Barber, J. Sabater-Mir, and M. Singh, editors. *Trust in Agent Societies*. AAMAS Workshop, Hakodate, Japan, 2006.
6. J. Ferber, F. Michel, and J.-A. Báez-Barranco. AGRE: Integrating Environments with Organizations. In D. Weyns, H. V. D. Parunak, and F. Michel, editors, *Environments for Multi-Agent Systems*, volume 3374 of *LNAI*, pages 48–56, 2005.
7. A. Flache and R. Hegselmann. Do irregular grids make a difference? relaxing the spatial regularity assumption in cellular models of social dynamics. *J. Artificial Societies and Social Simulation*, 4(4), 2001.
8. R. A. Hanneman and M. Riddle. *Introduction to Social Network Methods*. `faculty.ucr.edu/ hanneman`, University of California, Riverside, 2005.

9. F. Kaplan. *The emergence of a lexicon in a population of autonomous agents (in French)*. PhD thesis, Université de Paris 6, 2000.
10. Y. Matsuo, M. Hamasaki, Y. Nakamura, T. Nishimura, K. Hasida, H. Takeda, J. Mori, D. Bollegala, and M. Ishizuka. Spinning multiple social networks for semantic web. In *Proceedings, The Twenty-First National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference, July 16-20, 2006, Boston, Massachusetts, USA*. AAAI Press, 2006.
11. N. Minar, R. Burkhart, C. Langton, and M. Askenazi. The swarm simulation system: a toolkit for building multi-agent simulations. Technical Report 96-06-042, Santa Fe Institute, 1996.
12. M. J. North, N. T. Collier, and J. R. Vos. Experiences Creating Three Implementations of the Repast Agent Modeling Toolkit. *ACM Transactions on Modeling and Computer Simulation*, 16(1):1–25, January 2006.
13. J. M. Pujol, A. Flache, J. Delgado, and R. Sangüesa. How can social networks ever become complex? modelling the emergence of complex networks from local social exchanges. *Journal of Artificial Societies and Social Simulation*, 8(4), 2005.
14. M. Schilperoord. *Complexity in Foresight, experiences with INTERSECTIONS: an agent-based simulation workbench to help achieve adaptiveness in strategic planning*. PhD thesis, Erasmus University, Roterdam, 2005.
15. P. Urbano, L. Antunes, and H. Coelho. Application of consensus games to the selection of metrical conventions. In M. Mohammadian, editor, *Proceedings of International Conference on Computational Intelligence for Modelling Control and Automation (CIMCA 2003)*, Vienna, Austria, February 2003.
16. D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393:440–442, April 1998.
17. U. Wilensky. NetLogo, 1999. Center for Connected Learning and Computer-Based Modeling, Northwestern University. Evanston, IL. http://ccl.northwestern.edu/netlogo/.

Morphogenesis of epistemic networks: a case study

Camille Roth

CRESS, Department of Sociology, University of Surrey, Guildford, GU2 7XH, United Kingdom
c.roth@surrey.ac.uk

Abstract. Agents producing and exchanging knowledge are forming as a whole a socio-semantic complex system. We argue that several significant aspects of the structure of a knowledge community are primarily produced by the co-evolution between agents and concepts, i.e. the evolution of an epistemic network. Focusing on a particular community of scientists working on a well-defined topic, we micro-found various stylized facts regarding its structure by exhibiting processes at the level of agents accounting for the emergence of epistemic community structure. After assessing the empirical interaction and growth processes, and assuming that agents and concepts are co-evolving, we successfully propose a morphogenesis model rebuilding relevant high-level stylized facts.

Introduction

Agents producing, manipulating, exchanging knowledge are forming as a whole a socio-semantic complex system: they are fully immerged in flows of information on which they can have an impact and leave their footprints at the same time. The massive availability of informational content and the potential for extensive interactivity has recently made the focus slip from single "groups of knowledge" to the entire "society of knowledge", in a *networked* fashion, calling for the use of new methods and the characterization of new phenomena, with knowledge being distributed and appraised on a more horizontal basis.

Understanding the structural aspects of these communities relate more broadly to a recent issue in social science, social network formation modeling, involving several disciplines from graph theory (computer science and statistical physics), mathematical sociology to economics [1–3]. Most of the recent interest has stemmed from the empirical observation that real social network structure strongly differs from that of uniform random graphs *a la* Erdős-Rényi (ER) [4], suggesting that agents interact non-randomly with respect to heterogenous preferences for interacting with other agents. While this fact was already well-documented in social science [5, 6], general network models had been limited for long to ER-like random graphs [7–9]. Subsequently, much work has been devoted to determining novel non-uniform interaction and growth mechanisms reconstructing complex network structures consistent with the real world, through a rich set of statistical parameters [10]. On the whole, this amounts to find the solution of a reverse problem: given such an evolving system, what kind of (possibly minimal) dynamics rebuild its structure? In other words, we look for a valid network morphogenesis model for the real-world structure.

We focus here on a particular socio-semantic complex system, a scientific community working on a well-defined topic, and we make the following assumption: *modeling* *interactions at the level of agents who co-evolve with the concepts they manipulate is sufficient to carry the micro-founded reconstruction of this complex system.* More precisely, we will rebuild several aspects of the structure of such a community by introducing a co-evolutionary framework based on a social network, a semantic network and a socio-semantic network; as such an epistemic network made of agents, concepts, and relationships between all of them. We will then show that dynamics at the level of this epistemic network are sufficient to reproduce several stylized facts of interest. To achieve such a morphogenesis model, we first build tools enabling the estimation of interaction and growth mechanisms from past empirical data. Then, assuming that agents and concepts are co-evolving, we successfully reconstruct a real-world scientific community structure for a relevant *selection* of features.

1 Networks

(Social) network morphogenesis models. Networks (or graphs) are omnipresent in the real world: from the lowest levels of physical interaction to higher levels of description such as biology, sociology, economics and linguistics. For long however, network appraisal had been restricted to theoretical approaches in graph theory and small scale empirical studies on a case-by-case basis; while network models were mostly limited to the seminal work of Erdős-Rényi [4] (ER), which was assumed to be realistic for most purposes. In this respect, the recent availability of increasingly larger computational capabilities has made possible the use of quantitative methods on large networks, which yielded surprising results, often contradictory with those provided by ER models. This consequently precipited an unprecedented interest in networks [10–12]. Three statistical parameters in particular appeared to provide an enormous insight on the topological structure of networks: (i) clustering coefficient (the proportion of neighbors of a node who are also connected to each other, averaged over the whole network), (ii) average distance (i.e. the length of the shortest path between two nodes, averaged over all pairs of nodes), (iii) degree distribution (the degree (or the connectivity) of a node is basically the number of nodes this node is connected to).

Several recent works suggested morphogenesis models matching empirical data on these statistical parameters, contradicting and eventually replacing the ER model [13–16]. More specifically, Barabasi & Albert (BA) [16] insisted on the point that such topology could be due to two very particular phenomena that models were so far unable to take into account: network growth, and preferential attachment of nodes to other nodes. They thus pioneered the use of these two features to successfully rebuild a scale-free degree distribution. In their network formation model, new nodes arrive at a constant rate and attach to already-existing nodes with a likeliness linearly proportional to their degree. This model was a great success and has been widely spread and reused. As a consequence, the term “preferential attachment” has been often understood as degree-related preferential attachment only, in reference to BA’s work. Since then, many other authors introduced network morphogenesis models with diverse modes of preferential link creation depending on various node properties (attractiveness [17], common neighbors [18], fitness [19], centrality [20], hidden variables and “types” [21], bipartite structure [22], etc.) and various linking mechanisms (stochastic copying of links [23], competitive trade-off and optimization heuristics [20, 24], payoff-biased network reconfiguration [25], group formation [26], to cite a few). On the other side, growth processes (if any) were often reduced to the regular addition of nodes which attach to older nodes — sometimes growth is absent and studies are focused on the evolution of links only.

The idea is usually to exhibit high-level statistical parameters and suggest low-level network processes, in order to deduce the former from the latter. Obviously, after selecting a set of relevant stylized facts to be explained, model design consists of two subtasks: defining the way agents are bound to interact with each other, as well as specifying how the network grows. However, even in recent papers, hypotheses on such mechanisms are often arbitrary and rarely empirically checked. This attitude is still convenient for normative models but is rather questionable for descriptive models. Here, we therefore endeavor to (i) exhibit high-level stylized facts characteristic of epistemic networks, (ii) point out relevant low-level features that may account for these high-level facts, (iii) design measurement tools to appraise these low-level features, and (iv) design a reconstruction model based on the *observed* low-level dynamics.

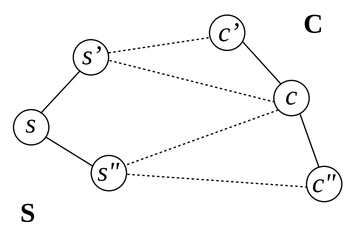


Fig. 1. Sample epistemic network with $\mathbf{S} = \{s, s', s''\}$, $\mathbf{C} = \{c, c', c''\}$, and relations $\mathcal{R}^{\mathbf{S}}$, $\mathcal{R}^{\mathbf{C}}$ (solid lines) and $\mathcal{R}$ (dashed lines).

Epistemic networks and empirical setting. We first introduce the objects we deal with: we distinguish a *social network* (linking agents), a *semantic network* (linking concepts) and a socio-semantic network (linking agents to concepts). Nodes in the social network $\mathcal{S}$ are agents, and links represent co-occurrence of two agents in an event. Thus $\mathcal{S} = (\mathbf{S}, \mathcal{R}^{\mathbf{S}})$, $\mathbf{S}$ is the set of agents and $\mathcal{R}^{\mathbf{S}}$ the set of undirected links. The semantic network $\mathcal{C} = (\mathbf{C}, \mathcal{R}^{\mathbf{C}})$ is the network of co-occurrences of concepts within events: $\mathbf{C}$ denotes the concept set, $\mathcal{R}^{\mathbf{C}}$ denotes links between concepts. The socio-semantic network $\mathbf{G_{SC}}$ is made of agents of $\mathbf{S}$, concepts of $\mathbf{C}$, and links between them, $\mathcal{R}^{\mathbf{SC}} = \mathcal{R}$, denoting usage of concepts by agents (undirected links for co-occurrence of agents and concepts). An *epistemic network* is thus given by these three networks, key for providing an account of the reciprocal influence and co-evolution of authors and concepts (Fig. 1). This is not to be confused with a bipartite network and its various projections: while the socio-semantic graph is indeed bipartite, social and semantic networks are not projections of the latter.

Translated in this framework, events are articles, agents are their authors, and concepts are made of expert-selected abstract words. We considered the community of embryologists working on the model animal *"zebrafish,"* during the period 1997–2004. Our main source of data is `MedLine`, a US National Library of Medicine reference database. The dataset contains around $10,000$ authors, $6,000$ articles and 70 concepts, adopting a weak linguistic assumptions by assuming that a lemmatized term corresponds to a concept. We restrict the dictionary to the 70 most used and significant words in the community selected with the help of a domain expert, in order to avoid rhetorical and neutral terms ("stop-words"). These concepts are given *a priori*: in the semantic network, only links appear, not nodes.

2 High-level features

We first endeavor to describe a few high-level statistical parameters particularly appropriate for epistemic networks. While we could have looked at many single-network parameters (such as assortativity [27], giant component size [28], single-network communities [29, 30], etc.), we focused instead on features specific to this epistemic network (thus, mostly bipartite parameters) — many results and models are already available for most traditional statistical features.

Degree distributions. In an epistemic network, ties appear in the social, semantic, and socio-semantic networks; hence, four degree distributions are of interest (Fig. 5):

1. *degrees k for the social network*: this distribution has been extensively studied in the literature [31, 32], and is traditionally said to follow a power-law, although only the tail actually does; some may suggest that this distribution follows a log-normal [33] or q-exponential law [34].
2. *degrees k_c for the semantic network*: since there are only 70 concepts the data are very sparse, we considered cumulated distributions (see exact definitions on Fig. 5) — all concepts are being progressively connected to each other.
3. *degrees from agents to concepts ($k_{a\to c}$)*: following a power-law; few agents use many concepts, many agents use few concepts.
4. *degrees from concepts to agents ($k_{c\to a}$)*: few concepts being used by a lot of agents, and most concepts being used by an average number of agents.

Clustering. The clustering coefficient is another valuable parameter [13]. It is basically a measure of the transitivity in one-mode networks, expressing how neighbors of a given node are connected to each other ("friends of friends are friends"). This coefficient is usually found to be very high in empirical social networks when compared to typical random networks such as those produced by ER, BA models. Along with degree distribution, this stylized fact has been the target of many more recent models [18, 35].

We use the local clustering coefficient, $c_3(i)$, measuring the proportion of neighbors of node i who are connected together: $c_3(i) = \frac{\text{[number of pairs of connected neighbors]}}{k_i \cdot (k_i - 1)/2}$ with k_i degree of i. This coefficient is close to 1 here and decreases rather slowly with node degree (Fig. 6). Yet, networks built with an underlying event structure are bound to exhibit a high coefficient [36, 37], thus a poorly informative criterion. By contrast, a *bipartite clustering coefficient* counting the proportion of diamonds [38] is a meaningful measure of how two agents connected to a same concept are likely to be connected to other concepts (as such a very local kind of structural equivalence): c_4 is the proportion of common neighbors among the neighbors of a node — in other words, are two agents connected to a same concept likely to be connected to other concepts? This coefficient appears to be one order of magnitude larger compared to that measured in scale-free random networks: pairs of agents linking together to certain concepts share other concepts abnormally often (Fig. 6).

Epistemic community structure. A key high-level stylized fact characteristic of epistemic networks is the particular distribution of epistemic communities (ECs) as *groups of agents using jointly the same concepts*, or maximal bipartite cliques in the socio-semantic network [39, 40]. An adequate epistemic network model should ultimately yield the same EC profile as in the real-world, which shows a significantly larger proportion of high-size ECs — see Fig. 7.

Besides, just as we observed the bipartite clustering between agents and concepts, we may want to know whether agents in the network are semantically close to each other. Likewise, and more specifically, in which manner are they semantically close to their social neighborhood? To this end, we need to introduce a semantic distance, i.e. a function of a dyad of agents which (i) decreases with the number of shared concepts, (ii) increases with the number of distinct concepts, (iii) equals 1 when there is no concept in common, 0 when all concepts are identical. Given $(s, s') \in \mathbf{S}^2$, and $s^\wedge$ the set of concepts s is linked to, we suggest the following metric semantic distance $\delta(s, s') \in [0; 1]$, based on the classical Jaccard coefficient [41], such that $\delta(s, s') = \frac{|(s^\wedge \setminus s'^\wedge) \cup (s'^\wedge \setminus s^\wedge)|}{|s^\wedge \cup s'^\wedge|}$. As δ takes real values in $[0, 1]$ we discretize δ, using a uniform partition of $[0, 1[$ in $I - 1$ intervals, to which we add the singleton $\{1\}$. We thus define a new discrete distance d taking values in $\mathcal{D} = \{d_1, d_2, ..., d_I\}$ such that: $\mathcal{D} = \{[0, \frac{1}{I-1}[,\ [\frac{1}{I-1}, \frac{2}{I-1}[, ...[\frac{I-2}{I-1}, 1[, \{1\}\}$. Then, we look at the distribution of semantic distances in the network, both on a global scale (by computing the distribution for all pairs of agents) and on a more local scale (by carrying the computation for pairs of already-connected agents only). Results on Fig. 7 suggest that while similar nodes are usually rare in the network, the picture is radically different when considering the social neighborhood: acquaintances are at a strongly closer distance.

3 Low-level dynamics

3.1 Measuring interaction behavior

Designing a credible social network morphogenesis model requires to understand both low-level interaction and growing mechanisms. We therefore first show how to design such low-level dynamics from empirical data. If the observed empirical structure diverges from the ER uniform random model, this suggests that interactions are not occurring totally at random but, rather, are directed by preferences on agents. In other words, there is prefential attachment from some agents to some other kinds of agents.

Formally, preferential attachment (PA) is the likeliness for a node to be involved in an interaction with another node with respect to node properties. Existing *quantitative* estimations of PA and subsequent validations of modeling assumptions are quite rare, and are often either related to the classical degree-related PA [32, 33], or considering PA as a scalar quantity, using direct mean calculation, econometric estimation approaches or Markovian models [28, 42, 43]. We here use a unified framework where properties are neither strictly based on social network topology nor reduced to single scalar quantities, while appraising how distinct properties correlatively influence PA.

We distinguish (i) single node properties, or *monadic* properties (such as degree, age, etc.) from (ii) dyadic properties (social distance, dissimilarity, etc.). When dealing with monadic properties indeed, we seek to know the propension of some kinds of nodes to be involved in an interaction. On the contrary when dealing with dyads, we seek to know the propension for an interaction to occur preferentially with some kinds of couples. We assume the influence on PA of a given monadic property m can be described by a function f of m, the *interaction propension*, independent of the distribution of agents of kind m: $f(m)$ is simply the conditional probability $P(\mathrm{L}|m)$ that an agent of kind m receives a link L. Thus, it is $f(m)$ times more probable that an agent of kind m is involved into an interaction. For instance, the classical degree-based PA used in BA and subsequent models is an assumption on f equivalent to $f(k) \propto k$. We may estimate f through $\hat{f}(m) = \frac{\nu(m)}{P(m)}$ if $P(m) > 0$, 0 otherwise, where $\nu(m)$ is the expectancy of new link extremities attached to nodes of property m along a period, and $P(m)$ typically denotes the distribution of nodes of type m. We adopt a dyadic viewpoint whenever a property has no meaning for a single node, such as proximity, similarity — or distances in general. Similarly, we assume the existence of an essential dyadic interaction behavior embedded into $g(d)$ for a given dyadic property d defined on couples of agents, corresponding to the conditional probability $P(\mathrm{L}|d)$. Again, g is estimated with $\hat{g}(d) = \frac{\nu(d)}{P(d)}$.

The PA behavior embedded in $\hat{f}$ (or $\hat{g}$) can be used to shape modeling hypotheses, either by taking the empirically estimated function, or by stylizing the trend of $\hat{f}$ (or $\hat{g}$) to allow analytic solutions. When considering a property which enjoys an underlying natural order, it is also useful to examine the cumulative propension $\hat{F}(m_i) = \sum_{m'=m_1}^{m_i} \hat{f}(m')$ as an estimation of the integral of f, especially with noisy data. Besides, when considering a collection of properties one must make sure that they are uncorrelated: for instance, node degrees may depend on age. If two distinct properties p and p' are independent, the distribution of nodes of kind p in the subset of nodes of kind p' does not depend on p', i.e. the quantity $\frac{P(p|p')}{P(p)}$ theoretically equals 1.

3.2 Empirical PA

Using these tools, we examine PA based on (i) a monadic property (node degree) and (ii) on a dyadic property (semantic distance d, rendering homophily). We first consider the node degree k as property m: we compute the real slope $\hat{f}(k)$ of the degree-related PA and empirically roughly verify the classical assumption "$f(k) \propto k$" (Fig. 2). This precise result is not new and tallies with existing studies on degree-related PA [44, 45].

We also assess the extent to which agents are "homophilic" (they prefer to interact with similar agents) by using the semantic distance introduced in Sec. 2. Empirical results on Fig. 2 show that while agents favor interactions with slightly different agents, they still very strongly prefer similar agents. Besides, the exponential trend of $\hat{g}$ suggests

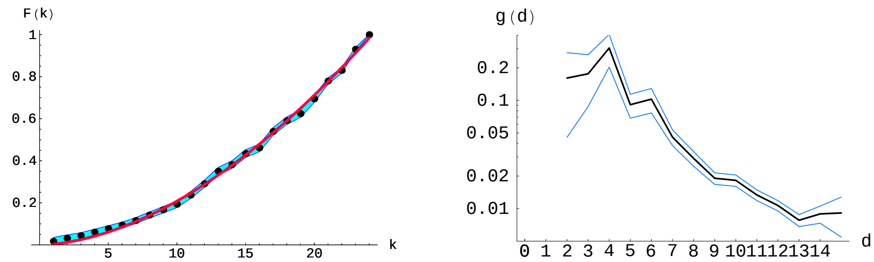


Fig. 2. *Left:* Cumulated propension $\hat{F}$. Dots represent empirical values, the solid line is the best non-linear fit for $\hat{F} \sim k^{1.83}$ (i.e. $\hat{f} \sim k^{0.83}$), and the gray area is the confidence interval. ***Right:*** Homophilic interaction propension $\hat{g}$ with respect to a semantic distance $d \in \{0, ..., 15\}$ (thick solid line) and confidence interval for $p < .05$ (thin lines).

that homophily is even more influent than connectedness. This fiercely advocates the use of semantic content for modeling such networks, while showing that simple non-structural properties may strongly shape interaction behavior in some networks. As underlined above, we check if the two properties are independent, i.e. whether or not a node of low degree is more or less likely to be at a larger semantic distance of other nodes. Here, there is no correlation between degree and semantic distance.

We finally examine if concepts are preferentially chosen: are well-connected concepts used more often, thus ‘interacting’ with even more authors? It turns out that concepts appear proportionally to their socio-semantic degree (i.e. the number of agents who use them) which reflects their popularity.

3.3 Growth- and event-related parameters

These features yield an essential insight on how local interactions occur. Now, in order to complete the description of the way the network grows, studying how events are structured in terms of both authors and concepts is also a crucial information. Regularly, new articles are produced, involving on one side a certain number of authors who have already authored a paper (old nodes) and possibly a fraction of new authors (new nodes), and on the other side, concepts that the authors bring in as well as new concepts.

Network growth. The first step is to determine the raw network growth, in terms of new nodes. How many new events appear, how many new articles are written during each period? Articles gather existing authors as well as new authors around concepts. Since we consider the set of concepts to be fixed *a priori*, new nodes appear in the social network only. The evolution of the size of the social network N_t depends on the number of new nodes per period ΔN_t, with $N_{t+1} = N_t + \Delta N_t$. In turn, there is a strong link between ΔN_t and the number of articles n_t, depending on the fraction of new authors per article. The growth of both ΔN_t and n_t is roughly linear with time: we can approximate the evolution of n by $n_{t+1} = n_t + n_+$, for a given arithmetic growth rate of n_+; every period the number of new articles increases by n_+. In our case, $n_+ \simeq 96$ ($\sigma \simeq 28$). ΔN and n seem to be linearly correlated, suggesting that the proportion of new authors in all articles is stable across periods.

Size of events. This leads us to study how articles are structured: in particular, how many agents are gathered in an event, and how many of them are new nodes? As shown on Fig. 3, the distribution of the number of agents per article appears to follow roughly a geometric distribution. On the other hand, the weight of new authors within articles obeys a distribution centered around three modes $\{0, 0.5, 1\}$, suggesting that in most cases either (i) authors are all new, (ii) they are all old, or (iii) half are new & half are old. Since this proportion is stable across periods, n_t is a good indicator of network growth: new articles appear and pull new authors into the network — on average, articles gather 4.4 authors, among which 55% are new, thus $.55 \times 4.4 = 2.42$ new authors, which is close to the coefficient of the best linear fit of ΔN with respect to n: $\Delta N \sim 2.25n$.

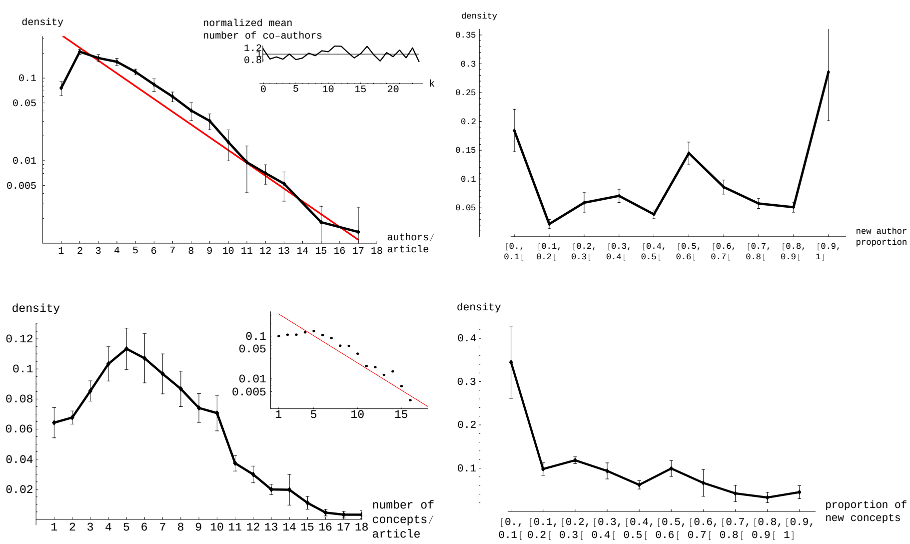


Fig. 3. *Top, left:* Distribution of the size of events, averaged on 8 periods 97-04, with confidence intervals for $p < .05$. The mean number of authors is 4.4 ($\sigma = 3.1$), and the best non-linear fit is $\propto \exp^{-\mu n}$ with $\mu = .36 \pm .06$ (straight line). The inset shows the mean number of coauthors with respect to degree k, relatively to the global mean number of co-authors: in case of independence, this ratio equals 1. *Top, right:* Proportion of new authors with respect to total authors, averaged on 7 periods (98–04) — the mean proportion is 0.55, but $\sigma = .33$ because of the tri-modal distribution. *Bottom, left:* Distributions of concepts per article — mean: 6.5, $\sigma = 3.6$. In the inset, the solid line represent the best exponential fit, $\propto e^{-\mu n}$ with $\mu = 0.29$. *Bottom, right:* Distribution of the proportion of new concepts that none of the agents anteriorly used — only for articles where there is at least one old agent. The mean is .32, with $\sigma = .28$.

Since the size of the network is increased by ΔN in a period, and ΔN here shows a linear behavior, N exhibits a quadratic growth. The fact that the number of articles per period linearly increases is however proper to the evolution of *this* empirical situation. The evolution of n and N is a consequence of this — this is obviously not the case for all networks: if for instance this field of research were to be abandoned, we would have a decrease of articles, not a linear growth.

Exchange of concepts. Knowing the structure of articles, and how authors are gathered, we now investigate how concepts are used. The distribution of the number of concepts is plotted on Fig. 3, and could be accurately approximated by a geometric distribution. Besides, while old authors bring a certain proportion of their concepts, some concepts are used for the first time: they do not belong to the intension of authors. The distribution of the proportion of new concepts — *new to the authors* — also shown on Fig. 3, makes it possible to distinguish concepts chosen within the intension of authors, from new, unused ones. It has a single mode 0, but is on the whole relatively flat.

4 Epistemic network morphogenesis model

Design. Using empirically-measured low-level parameters (article composition, interaction preferences) we design a model that rebuilds a high-level structure compatible with real-world stylized facts (degree and semantic distance distributions, bipartite clustering, EC structure). Three key modeling features are implemented: (i) event-based network growth, (ii) co-evolution between agents and concepts, and (iii) realistic low-level descriptions, especially for interactions. Events are articles, made of agents (more or less active depending on their degree k, and gathering preferentially with respect to their interests) and concepts (more or less popular, depending on their degree $k_{\mathrm{c}\to\mathrm{a}}$). Our low-level dynamics, or model of a coevolving epistemic network, consists in:

1. *Creating and defining events.* n_t articles are created at each period: $n_{t+1} = n_t + n_+$. Author set and concept set sizes follow geometric laws, with observed means.
2. *Choosing authors.* Because of the tri-modal distribution articles feature either only new authors, either only old authors, or equally old and new authors, all equiprobably. If there is at least one old agent, an 'initiator' is randomly chosen proportionally to her social network degree k; then, other old agents of degree k' are picked according to $P(\mathrm{L}|k', d) = P(\mathrm{L}|k')P(\mathrm{L}|d)$, d is the semantic distance to the initiator. Finally new nodes are created.
3. *Choosing concepts.* New concepts (i.e. such that no old agent uses) are a fixed proportion of the article concept set. Other concepts are chosen from the concept set of authors. All are chosen randomly proportionally to their degree $k_{\mathrm{c}\to\mathrm{a}}$.
4. *Updating the network,* once author and concept sets are defined (Fig. 4).

Results. We ran the model for 8 periods $t \in \{1, \cdots, 8\}$, starting with an empty epistemic network — in other words, the morphogenesis starts *from scratch*. Obviously, periods correspond to years. One hundred new articles were to appear during the first period, with a growth rate of 100 articles per period per period: $n_1 = 100$, $n_+ = 100$.

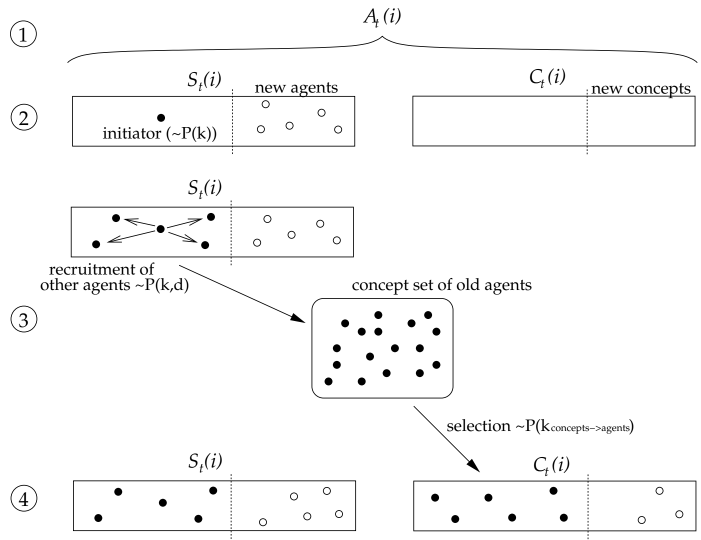


Fig. 4. Modeling an event by specifying contents of article i, $A_t(i) = (S_t(i), C_t(i))$, author and concept sets. The numbered steps indicated here follow those of the model description in Sec. 4.

We focus on networks obtained after simulations are completed for 8 periods, and we have a satisfying adequation for every stylized fact, both in shape and in magnitude:

– *Rebuilding network size.* Simulated networks contain 10982 agents on average ($\sigma = 215$, for fifteen runs), agreeing with empirical data.
– *Rebuilding degree distributions.* Results for all four degree distributions are shown on Fig. 5, indicating a very good fit — in particular, power-law tails have a similar exponent, with a shape which fits a log-normal distribution similar to that of the empirical case.
– *Rebuilding clustering coefficients.* Clustering coefficients are accurately reproduced, as shown on Fig. 6.
– *Rebuilding epistemic community structure.* ECs have been computed (see Fig. 7) and distributions of EC sizes are close to those of the real network. Semantic distances are also correctly rebuilt, see Fig. 7.

Discussion. Hence, epistemic communities are produced by the co-evolution of agents and concepts. Not only is the high-level structure accurately reconstructed by our model, but low-level dynamics are consistent as well — this is a not a minor point: rebuilding high-level phenomena remains dubious if the low-level dynamics is incorrect. Truthfulness of descriptions must reach the higher level *as well as* the lower level. In any case, we may still wonder what weight some of our hypotheses bear towards the apparition of high-level phenomena: is our model a minimal model *as regards the stylized facts we selected*? In particular, consider basic event-based models for social networks —

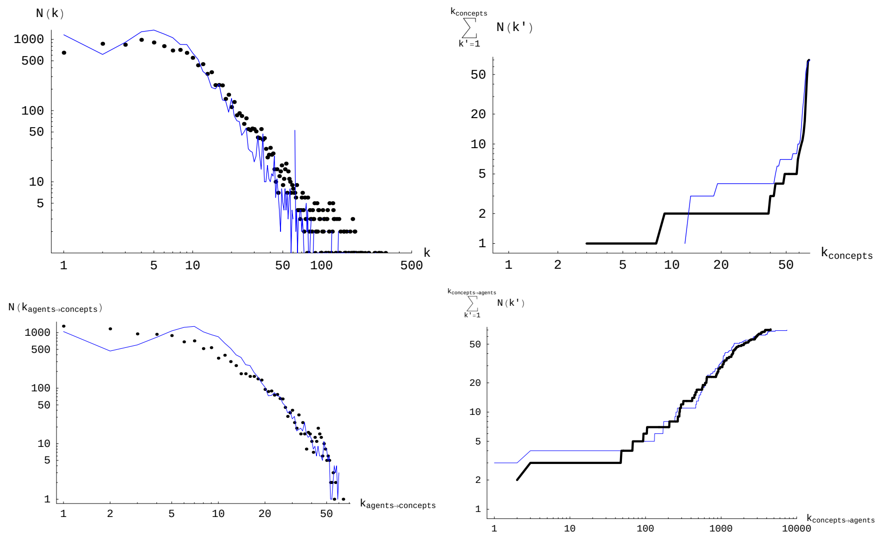


Fig. 5. Social (top-left), semantic (top-right) and socio-semantic (bottom) degree distributions. Simulation results (dots or thick line) acceptably fit empirical data (thin line). *x-axes* correspond to degrees in a given network, while *y-axes* represent the number or cumulated number of nodes having a given degree — both are log-scale.

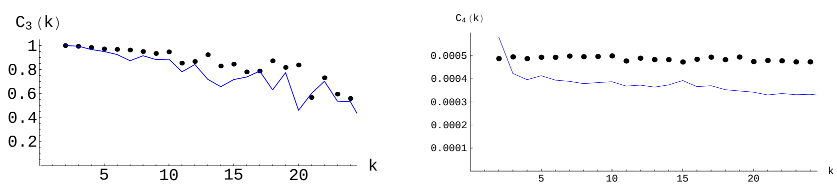


Fig. 6. $c_3(k)$ (left) and $c_4(k)$, simulated (dots) vs. empirical values (solid line). See Sec. 2 for details on these coefficients.

which have become popular very recently among a few other authors as well [22, 26, 28] — that simply rest on n-adic events instead of dyadic interactions and that do not even specify any kind of PA. Yet, these models lead to scale-free distributions and high one-mode clustering coefficients. These results suggest that PA is not required to rebuild degree distributions and c_3, by contrast to dyadic-interaction-based models (such as BA model).

Recall that our model features (i) event-based modeling, (ii-a) degree-related preferential attachment (or activity) for the choice of agents and (ii-b) for concepts, and (iii) homophily of agents. Are the high-level stylized facts still reproduced if we loosen some of these hypotheses? Since many combinations of simplified models are envisageable, we only examine what happens when relaxing one hypothesis at a time. Yet,

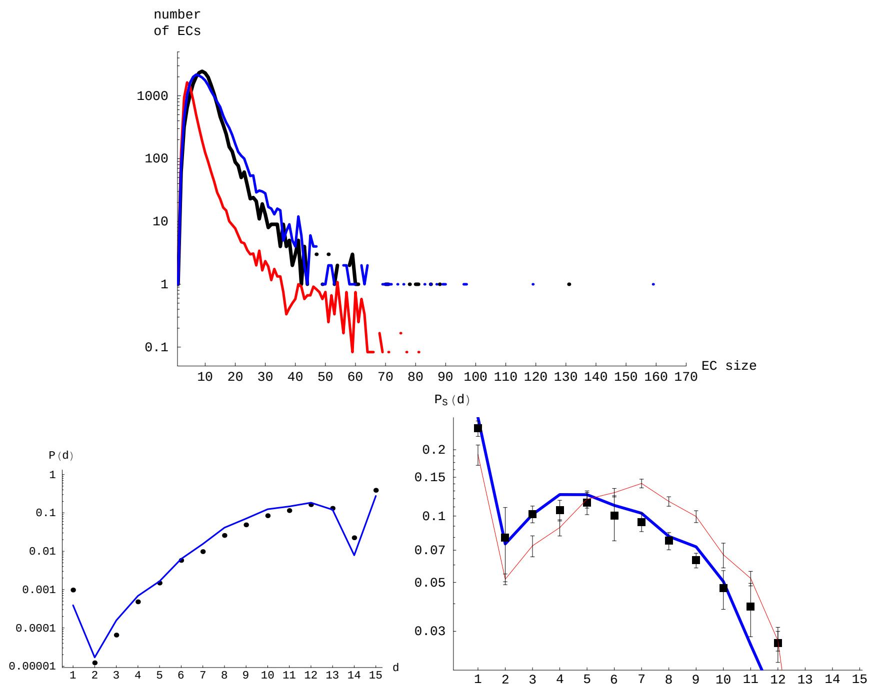


Fig. 7. *Top:* Number of ECs with respect to agent set sizes, in GLs computed for samples of 250 agents. Simulation results (thick black line, above) fit the empirical data (thin blue line, above), compared to a random "rewired" cases where degree distributions on from agents to concepts and from concepts to agents are conserved: as expected, they contain significantly less ECs, by one order of magnitude (thin red line, below). *Bottom, left:* Simulated mean distribution of semantic distances on the whole graph (dots) compared to original empirical data (line). *Bottom, right:* Same quantities, but computed only for the social neighborhood of each agent. Note the thin solid line, representing simulations not using homophily.

at least one high-level fact is not accurately reproduced when relaxing any feature of our model (event-based modeling, degree-related preferential attachment (or activity) for the choice of agents and for concepts, or homophily of agents).[1]

Conclusion

We investigated the formation of the emerging "zebrafish" scientific community and assumed that we could micro-found the evolution of the structure of this social complex system by modeling agents co-evolving with concepts. Therefore, we introduced tools to estimate low-level interaction and growth processes from past data. Only thereafter could we hope for a *realistic*, *descriptive* model. The final success of the reconstruction

[1] Comprehensive details about these results are omitted because of length restrictions.

gives credit to our hypothesis that the structure of knowledge communities is at least produced by the co-evolution of agents and concepts. *In fine*, we argued for an empirical stance when designing models: even if the model reproduces the desired stylized facts, it is essential to know whether the alleged low-level dynamics is empirically grounded.

References

1. Skyrms, B., Pemantle, R.: A dynamic model of social network formation. PNAS **97**(16) (2000) 9340–9346
2. Albert, R., Barabási, A.L.: Statistical mechanics of complex networks. Reviews of Modern Physics **74** (2002) 47–97
3. Cohendet, P., Kirman, A., Zimmermann, J.B.: Emergence, formation et dynamique des réseaux – modèles de la morphogenèse. Revue d'Economie Industrielle **103**(2-3) (2003) 15–42
4. Erdős, P., Rényi, A.: On random graphs. Publicationes Mathematicae **6** (1959) 290–297
5. Touhey, J.C.: Situated identities, attitude similarity, and interpersonal attraction. Sociometry **37** (1974) 363–374
6. McPherson, M., Smith-Lovin, L., Cook, J.M.: Birds of a feather: Homophily in social networks. Annual Review of Sociology **27** (2001) 415–444
7. May, R.K.: Will a large complex system be stable? Nature **238**(413–414) (1972)
8. Barbour, A., Mollison, D.: Epidemics and random graphs. In Gabriel, J.P., Lefevre, C., Picard, P., eds.: Stochastic Processes in Epidemic Theory. Lecture Notes in Biomaths, 86. Springer (1990) 86–89
9. Wasserman, S., Faust, K.: Social Network Analysis: Methods and Applications. Cambridge University Press, Cambridge (1994)
10. Dorogovtsev, S.N., Mendes, J.F.F.: Evolution of Networks — From Biological Nets to the Internet and WWW. Oxford: Oxford University Press (2003)
11. Barabási, A.L.: Linked: The new science of Networks. Cambridge, Mass.: Perseus Publishing (2002)
12. Newman, M.E.J.: The structure and function of complex networks. SIAM Review **45**(2) (2003) 167–256
13. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. Nature **393** (1998) 440–442
14. Redner, S.: How popular is your paper? An empirical study of the citation distribution. European Phys. Journal B **4**(131–134) (1998)
15. Faloutsos, M., Faloutsos, P., Faloutsos, C.: On power-law relationships of the Internet topology. Computer Communication Review **29**(4) (1999) 251–262
16. Barabási, A.L., Albert, R.: Emergence of scaling in random networks. Science **286** (1999) 509–512
17. Dorogovtsev, S.N., Mendes, J.F.F., Samukhin, A.N.: Structure of growing networks with preferential linking. Physical Review Letters **85**(21) (2000) 4633–4636
18. Jin, E.M., Girvan, M., Newman, M.E.J.: The structure of growing social networks. Physical Review E **64**(4) (2001) 046132
19. Caldarelli, G., Capocci, A., Rios, P.D.L., Munoz, M.A.: Scale-free networks from varying vertex intrinsic fitness. Physical Review Letters **89**(25) (2002) 258702
20. Fabrikant, A., Koutsoupias, E., Papadimitriou, C.H.: Heuristically optimized trade-offs: A new paradigm for power laws in the internet. In: ICALP '02: Proceedings of the 29th International Colloquium on Automata, Languages and Programming, London, UK, Springer-Verlag (2002) 110–122
21. Boguna, M., Pastor-Satorras, R.: Class of correlated random networks with hidden variables. Physical Review E **68** (2003) 036112
22. Peltomaki, M., Alava, M.: Correlations in bipartite collaboration networks. arXiv e-print archive **physics** (2005) 0508027
23. Kumar, R., Raghavan, P., Rajagopalan, S., Sivakumar, D., Tomkins, A., Upfal, E.: Stochastic models for the web graph. In:

IEEE 41st Annual Symposium on Foundations of Computer Science (FOCS). (2000) 57
24. Berger, N., Borgs, C., Chayes, J., D'Souza, R., Kleinberg, R.: Competition-induced preferential attachment. In: Proceedings of the 31st International Colloquium on Automata, Languages and Programming. (2004) 208–221
25. Carayol, N., Roux, P.: Micro-grounded models of complex network formation. Cahiers d'Interactions Localisées **1** (2004) 49–69
26. Ramasco, J.J., Dorogovtsev, S.N., Pastor-Satorras, R.: Self-organization of collaboration networks. Physical Review E **70** (2004) 036106
27. Newman, M.E.J., Park, J.: Why social networks are different from other types of networks. Physical Review E **68**(036122) (2003)
28. Guimera, R., Uzzi, B., Spiro, J., Amaral, L.A.N.: Team assembly mechanisms determine collaboration network structure and team performance. Science **308** (2005) 697–702
29. Girvan, M., Newman, M.E.J.: Community structure in social and biological networks. PNAS **99** (2002) 7821–7826
30. Latapy, M., Pons, P.: Computing communities in large networks using random walks. arXiv e-print archive (2004) 0412568
31. Newman, M.E.J.: Scientific collaboration networks. I. Network construction and fundamental results. Physical Review E **64** (2001) 016131
32. Barabási, A.L., Jeong, H., Ravasz, E., Neda, Z., Vicsek, T., Schubert, A.: Evolution of the social network of scientific collaborations. Physica A **311** (2002) 590–614
33. Redner, S.: Citation statistics from 110 years of physical review. Physics Today **58** (2005) 49–54
34. White, D.R., Kejzar, N., Tsallis, C., Farmer, D., White, S.D.: A generative model for feedback networks. Physical Review E **73** (2006) 016119
35. Ravasz, E., Barabási, A.L.: Hierarchical organization in complex networks. Physical Review E **67** (2003) 026112
36. Newman, M.E.J., Strogatz, S., Watts, D.: Random graphs with arbitrary degree distributions and their applications. Physical Review E **64**(026118) (2001)
37. Guillaume, J.L., Latapy, M.: Bipartite structure of all complex networks. Information Processing Letters **90**(5) (2004) 215–221
38. Lind, P.G., Gonzalez, M.C., Herrmann, H.J.: Cycles and clustering in bipartite networks. Physical Review E **72** (2005) 056127
39. Roth, C., Bourgine, P.: Epistemic communities: Description and hierarchic categorization. Mathematical Population Studies **12**(2) (2005) 107–130
40. Roth, C., Bourgine, P.: Lattice-based dynamic and overlapping taxonomies: The case of epistemic communities. Scientometrics **69**(2) (2006) 429–447
41. Batagelj, V., Bren, M.: Comparing resemblance measures. Journal of Classification **12**(1) (1995) 73–90
42. Snijders, T.A.: The statistical evaluation of social networks dynamics. Sociological Methodology **31** (2001) 361–395
43. Powell, W.W., White, D.R., Koput, K.W., Owen-Smith, J.: Network dynamics and field evolution: The growth of interorganizational collaboration in the life sciences. American Journal of Sociology **110**(4) (2005) 1132–1205
44. Newman, M.E.J.: Clustering and preferential attachment in growing networks. Physical Review Letters E **64** (2001) 025102
45. Jeong, H., Néda, Z., Barabási, A.L.: Measuring preferential attachment for evolving networks. Europhysics Letters **61**(4) (2003) 567–572

Session on Cognitive Agents and Social Behaviour

A Model of Mental Model Formation in a Social Context

Umberto Gostoli

Dipartimento di Economia
Università Politecnica delle Marche
Piazzale R. Martelli, 8
60121 Ancona – Italy
u.gostoli@univpm.it

Abstract. This paper presents a model of learning in a context of a relatively large population interacting through random bilateral matching to play a bilateral game in strategic form. While the theory of learning in games commonly assumes that the players can observe only the strategies chosen by their opponents, in this paper is introduced the additional assumption that the players are characterized by phenotypic traits observable by the other players with whom they interact. The extension of the traditional framework allows to introduce a more sophisticated and cognitively plausible expectations' formation model than the ones proposed so far. In particular, this paper introduces a new model of the induction process through which the agents build mental models that take the form of lexicographically structured decision trees.

Keywords: Theory of Learning in Games, Categorization, Social Stereotyping, Fast and Frugal Heuristic Theory, Self-Organization, Data Mining.

1 Introduction

Any dynamic or repeated game is characterized by a particular information structure that defines what the players know before the game starts and what they can observe during the stages, or periods, of the game. Moreover, any game is, implicitly or explicitly, characterized by a certain information processing algorithm that defines and represents the players' cognitive skills. The equilibria reached, as well as the dynamics that leads to them, depend on the information structure of the game and on the decision process through which the players compute the information at their disposal.

It is commonly assumed that the players know the payoff matrix of the game, can observe in every period the action of their opponent and, moreover, that all the players know that the other players have the same prior knowledge and observational skills, that all the players know that the other players know and so on. These are the so-called prefect and common knowledge assumptions that come from the traditional static game theory, whose formal application to the structure of the game can lead to technical and philosophical problems (Binmore, 1997).

While the information available to the players has been considered by the means of formal assumptions, less attention has been paid to the explicit formal definition of the cognitive process through which this information is computed by the players to reach a decision in order to the action to undertake. However, as behavioural game theory shows, the differences between the outcomes of experiments and the predictions of game theory are often due to the unrealistic computational power and rationality assumptions of the latter (Camerer, 2003).

The theory of learning in games is the branch of the literature that introduced models of players less than perfectly rational and that explicitly formalized, together with the information structure of the game, the decision process that underlies the player's action. While game theory tells us which Nash equilibria a particular game has, the theory of learning in games gives us models that determine the path through which a certain equilibrium is reached. In these models, the equilibrium is the outcome of a process in which less than fully rational players grope for optimality over time (Fudenberg and Levine, 1999).

The approaches to learning in games can be divided in two big classes: reinforcing and forecasting learning models. The first class includes models where the players choose a strategy on the basis of its past performance. The imitation models belong to this class: the players are endowed with the skill to asses their neighbours' success and to associate this performance to a particular strategy, unsophisticated but not trivial cognitive skills. The second class is represented by models where the players are endowed with the cognitive skills to develop a forecasting model of the behaviour of other players and, so, to choose the strategy that is the best reply to their opponent's expected action.

The *fictitious play* models belong to this last class. In general, in these models players are endowed with the basic skill to observe and memorize their opponent's action. This cognitive skill allows the agents to keep track of the relative frequencies with which each strategy is played. These relative frequencies represent the players' expectation about the strategies' distribution of the population.

In the model proposed in this paper the information structure of the game is extended to include the players' skill to observe their opponent's phenotype, represented by a string of three binary attributes. Given the information processing skills that characterize the fictitious play models, this extension would allow the players to develop expectations conditional on their opponents' phenotypes.

Of course, this extension is not new in the field of the theory of learning in games. Previous papers have introduced models where the agents are endowed with a visible tag composed by one binary attribute (Axtell, Epstein and Peyton Young, 1999) or by three binary attributes (Hoffmann, 2006). These works demonstrated that the agents' capacity to develop expectations conditional on their opponents' type leads to the formation of social classes. In other words, in these models we can observe the endogenous emergence of social stereotyping.

The main difference between the papers mentioned above and this one is the cognitive process through which the information at the agent's disposal is processed. While in the former papers the agents develop conditional probabilities over opponents' play conditional on the particular type they are playing against, according to the classical fictitious play algorithm, in this paper is proposed a model of information processing algorithm inspired by a recently introduced decision process theory: the Fast and Frugal Heuristic Theory (FFHT). The FFHT proposes a lexicographic decision process, that is, a process through which the attributes are looked up in a particular order of validity by the means of a decision tree. Figure 1 shows a lexicographic decision tree with which is possible to classify instances, identified by three binary attributes, in the two classes A and B. In front of the situation 0-1-1, the agent would classify it as A: the third attribute (identified with the number 2) has value 1, so it has to consider the second node, containing the first attribute (identified with the number 0). Its decision process does not need to proceed further because, being this attribute's value equal to 0, the decision tree ends with a leaf, containing in this case the class A. The model presented in this paper is based on the assumption that the agent's mental model is structured as a decision tree.

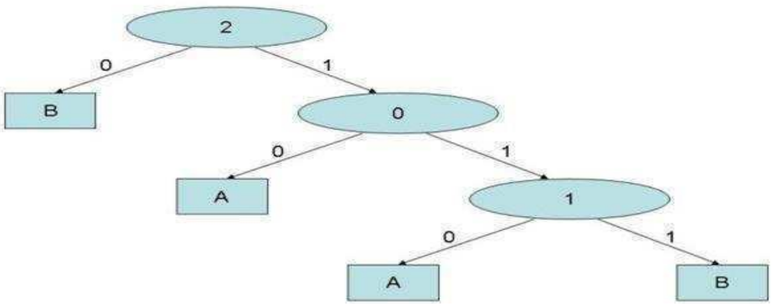


Figure 1

The FFHT is based on the bounded rationality paradigm: because of mind's limitations, real decision makers have to use approximate methods to handle most tasks (Simon, 1955). The function of these methods, that we will call *heuristics*, is to make reasonable, adaptive inferences about the social and physical world given limited time and knowledge. In this situations, indeed, real agents cannot perform the cognitively demanding multiple linear regression, but, more plausibly, they look up only few available cues in a fixed order of validity (Gigerenzer and Goldstain, 1996).

The cues relevant in the decision-making process and the order in which they have to be considered can be learned in three main ways: they could be genetically coded, they can be learned through cultural transmission or they can be learned from the agent's experience. In the model presented in this paper, the number and the kind of cues the agents can detect are genetically determined but the decision tree by which they are hierarchically ordered to form their mental model is the result of an induction process that the agents perform on their experiences' database.

According to FFHT, in an evolutionary context, decision-making strategies are selected for their accuracy, frugality and speed, measures that relate the decision process to the environment in which the decision has to be taken. In fact, the FFHT introduces a new concept of rationality called *ecological rationality*: a heuristics, or mental model, is rational from an ecological point of view if it allows the agents to take effective decisions in the environment in which the agents live. This means that in order for a given heuristics to be rational it has to match the information structure of the environment in which that particular heuristics is used.

In a social context, however, the environment of each agent is represented by other agents. So, the mental model of each agent is rational from an ecological point of view if it matches the mental model of the other agents with whom it interacts. The main aim of this paper is to find out if a social system formed by agents that make heuristic inferences about their opponent's behaviour can reach an equilibrium, that is, a situation where the mental model that each agent develops through its experiences allows it to make ecologically rational decisions given the mental models developed by the other agents with whom it interacts. The model presented in this paper is a model of *co-adaptation* of mental models with which we try to find out if the system ever reaches the steady state in which each agent's mental model is confirmed by, and because of, the other agents' mental models.

2 The Model

Being this an agent-based model, it has to be described at two levels: the population and the agent level. The former specifies the kind of interaction that takes place among the agents who compose the population, while the latter specifies the decision process that take place inside each single agent. The model presented in this paper considers a population of *N* agents that are in each period randomly paired to play the *Stag Hunt Game*, whose payoff matrix is shown in Figure 2. This game has two Nash equilibria: the socially optimal equilibrium S – S and the risk-dominant equilibrium H – H. While in the evolutionary game theory the agent has no choice but to play the strategy to which it is genetically or culturally associated, in this model the agent has to take in each period a decision about which strategy to play. Of course, this decision is based on the expectation the agent has about its opponent's action: it will play S if it expects its opponent to play S, it will play H otherwise.

Table 1: the Stag Hunt Game

	S	H
S	2,2	0,1
H	1,0	1,1

To understand how expectations form and evolve, we have to describe the model at the level of the agent. Regarding the prior knowledge of the agents, this model adopt the limited knowledge paradigm that characterizes the models adopted by the theory of learning in games: the agent knows only its own payoff matrix and does not know what its opponent gets from the interaction.

Each agent is characterized by phenotypic traits represented by three binary attributes. This means that the population is composed by eight different phenotypes: 000, 001, 010, 011, 100, 101, 110, 111. Moreover, each agent can observe the phenotypic traits of its opponent, together with the action its opponent performs. So, after each game, the phenotypic traits and the action of the opponent is stored in the agent's memory M, its size being a parameter of the model. An example of an agent's experiences database could be the matrix shown in Figure 3. This matrix shows that in its last game, the agent met an opponent whose phenotype was 001 and who played strategy S. Of course its own phenotype and action have been stored at the same time in the experiences database of its opponent.

We call *s* the agents' memory size, that is, the number of experiences the agent can store in its memory. When the number of periods $p = s$, the agent's memory is full, new experiences are stored according to the FIFO system: the oldest experience in the database is discarded to make room for the last experience.

Having defined the agent's gathering and storing cognitive skills, we come to the information processing algorithm that underlies the agent's decision. We call *l* the number of periods that constitutes the agent's *learning period*, its length being another parameter of the model. For the first *l* periods the agent has not yet developed a mental model that allows it to forecast its opponent's move so it makes a random forecast and chooses the action that is the best response to this forecast.

Table 2: an agent's experiences database

Attribute 1	Attribute 2	Attribute 3	Strategy
0	1	0	S
1	0	0	S
1	1	1	H
...	...	...	...
1	0	1	H
1	1	0	S
0	0	0	H
...	...	...	...
0	1	1	H
0	1	1	H
0	0	1	S

At the end of the learning period the agent builds, through an induction process performed on its experiences database, a mental model that allows it to forecast, from its opponent's phenotype, the strategy its opponent will play. An example of an agent's mental model can be the decision tree of Figure 1, interpreting the letters A and B in the rectangles as, respectively, the two strategies S and H. If an agent with this mental model meets an opponent whose phenotype is, for example, 011, the agent's forecast of its opponent's move would be S. The mental model represents, in fact, a set of hierarchically organized behavioural rules of the kind *If/Then* that maps each phenotype to a strategy to be chosen when meeting an agent with that phenotype.

From the cognitive point of view, we have to distinguish two different processes: the *induction process* through which the agent develops its mental model on the basis of its experiences, and the *decision process,* that takes place when the agent forecasts, with its mental model, its opponent's move. In other words, with the induction process the agent makes the mental model, with the decision process it puts it into use. While the decision process takes place in each match, the induction process can also be performed once every given number of periods r, this number being another parameter of the model.

While the decision process is quite straightforward, it is necessary to specify in a detailed way the induction process through which the agent develops its mental model from the database of its experiences. In this model, this cognitive process is modelled by the means of data mining techniques. The best way to describe the information processing algorithm that represents the agent's induction process is to look closely at an example. Let's consider the database of Table 4, containing 20 instances. The database is formed by three independent variables that define each instance (V1, V2, V3) and one dependent variable representing the class to which the instance belongs (C). The first and the third independent variables are binary variables: they can have either value 0 or 1. The second independent variable can have three values: 0, 1 or 2. The instances can belong either to class 0 or to class 1.

Table 3: a database

	V1	V2	V3	C
1	0	2	1	1
2	0	1	0	1
3	0	2	0	1
4	1	2	1	0
5	0	1	1	0
6	1	0	1	0
7	1	2	1	1
8	0	0	0	1
9	0	0	0	0
10	1	0	1	1
11	1	1	0	0
12	1	1	1	0
13	1	1	0	0
14	1	2	0	1
15	0	1	0	1
16	1	1	1	0
17	0	1	0	1
18	0	1	0	1
19	0	2	0	1
20	1	0	1	1

The first step of the data mining process is to find out the first independent variable of the decision tree. In order to do so, we have to compute the *average information value* for each independent variable. According to the information theory, the information value of a database is 1 minus the *entropy* of the database. The database's entropy is defined as the number of bits required to specify the class of an instance given that it belongs to the database. The entropy value can go from 0, if we the information that an instance belongs to the database is all we need to classify the instance, to 1, if the fact that the instance belongs to the database do not gives us any useful information about its class. If we call N the total number of instances in a database composed by a variable that can be classified in r classes and n_i the number of instances belonging to class i, the entropy E of a database is given by (1):

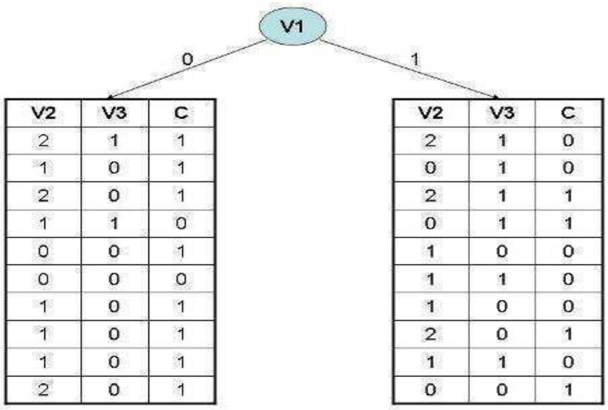
$$E = -\frac{n_0}{N}\log_r\frac{n_0}{N} - \frac{n_1}{N}\log_r\frac{n_1}{N} - \ldots - \frac{n_i}{N}\log_r\frac{n_i}{N} - \ldots - \frac{n_r}{N}\log_r\frac{n_r}{N} \qquad (1)$$

The entropy of the dependent variable column C is 0.971. It represents the additional number of bits we need in order to classify an instance given that it belongs to the database. Therefore, the information value of column C, *info(C)*, is 0.029 (1 – 0.971).

If column C was composed by the same number of 0's and 1's, the database's entropy would be 1 and the its information value 0. This makes sense because, intuitively, the fact that an instance belongs to the database, in this case, would not give us any useful information regarding its class. Conversely, if the C column was composed, for example, entirely by 1's, the database's entropy would be 0 and its information value 1: the fact that an instance belongs to the database would give us all the information we need to classify the instance as an instance of class 1.

With these basic information theory concepts we are now able to compute the average information value of each independent variable. Let's suppose that the most informative variable is variable V1. The next step is to find out the variable with the highest average information value for *each* branch, excluding variable V1 that has been already chosen as root of the decision tree. This means building from the original database, a subset for each branch: a subset of instances where V1 has value 0 (right branch) and a subset where V1 has value 1 (left branch).

For each of them we will compute the average information value of V2 and V3, choosing for each node the variable with the highest average information value (Figure 2).



V2	V3	C
2	1	1
1	0	1
2	0	1
1	1	0
0	0	1
0	0	0
1	0	1
1	0	1
1	0	1
2	0	1

V2	V3	C
2	1	0
0	1	0
2	1	1
0	1	1
1	0	0
1	1	0
1	0	0
2	0	1
1	1	0
0	0	1

Figure 2

The decision tree we get at the end of this data mining algorithm will be the decision tree that maximizes the number of instances that are correctly classified or, in other words, that minimizes the number of the exceptions. This decision tree can also be defined as the theory that best describes the database or, alternatively, the most likely theory given the database.

From a general point of view, the model described above is a feed-back process: the agents' experiences determine the agents' mental model which determine the agents' behaviour which determines the agents' experiences and so on. The main aim of the simulations, whose results will be presented in the following section, is to find out if this process ever reaches the steady state where the agents' behaviour produces experiences that confirm the agents' mental model.

3 Simulations' Results

In this section are presented the results of a particular simulation that will serve us as example, results that are, however, representative of the outcome of simulations based on the social and cognitive model presented above. In fact, making the agents in their learning period random forecasts, different simulations will produce results that are different in their details, but, nevertheless, we can see that they share general statistical features.

In the simulation below, the population is formed by 1000 agents who, in each period, are randomly paired to form 500 couples to play a one-shot stag hunt game. At their birth the agents are assigned a randomly generated phenotype and a memory, which is randomly drawn from a normal density function having average 40 and variance 4. The learning period for each agent is ¼ of her memory. This means that each agent will form a mental model after its memory has been filled for ¼ of its size. Then, the mental model will then be updated after every game ($r = 1$).

The graph in Figure 3 shows the *forecasting performance* of the population of agents, measured as the number of agents that, in each period, forecast correctly their opponent's choice.

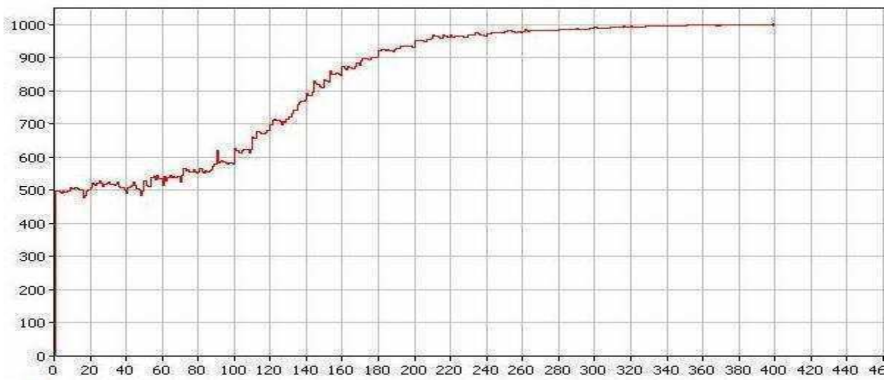


Figure 3: forecasting performance

From the beginning of the game till the end of the learning period this performance will be around 50%, that means that half of the population has a correct expectation about the opponent's behaviour, because of the random guess the agents make during the learning period. Then, as the agents develop and update their mental model, we expect this performance to increase until it reaches, eventually, the equilibrium, a state where 100% of the agents have correct expectations about their opponent's strategy. The graph of Figure 2 shows the path of the average forecasting performance in the particular simulation we are considering: as we can see, at around period 50, the forecasting ability of the agents begins to grow at increasing speed until period 180 when about 90% of the agents guess correctly their opponent's choice. Then, the performance growth tends to slow down, until it reaches 100% around period 350.

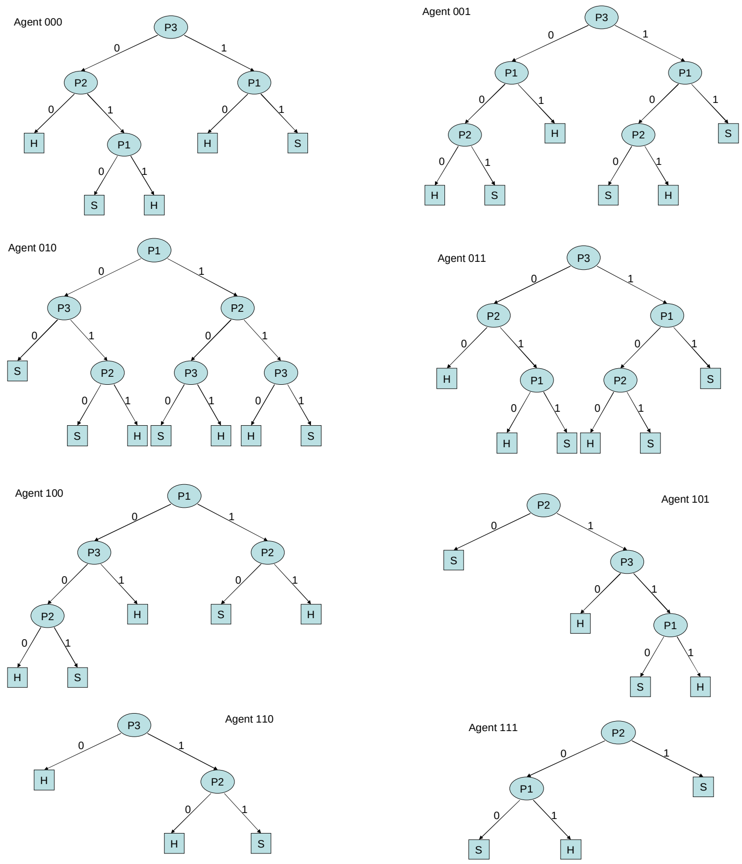


Figure 4: the mental models of eight randomly chosen agents

Even if the exact timing of the forecasting performance growth changes from one simulation to the other, the common feature of all simulations is that it always reaches 100%. This means that the system always reaches the mental model equilibrium, where the expectations of each agent of the population consistently match the other agents' expectations. After period 350, the experiences the agents make are consistent with their mental model and, consequently, the evolution of their mental models reaches a steady state. At this point, for each agent its mental model is the 'true' mental model because it allows it to forecast perfectly its opponent's behaviour in every game.

The mental models that characterize the equilibrium are a peculiar character of each simulation because they depend on stochastic fluctuations that take place during the learning period, when each agent makes random forecasts. The eight decision trees shown in Figure 4 represent the mental models of a sample of eight agents, one for each phenotype, after the system has reached the mental model equilibrium.

To see how these mental models determine the agents' behaviour, let's consider the decision tree of the agent whose phenotype is 000. It tells us that the fist thing this agent looks at is the third phenotypic trait (P3) of its opponent. If the value of this phenotypic trait is 0, then it looks to the second phenotypic trait (P2) and, if it is 0, the decision process terminates: the agent will forecast that its opponent will play H. If the value of P3 is 1, this agent only needs to look at the first phenotypic trait of its opponent (P1): if it is 0 then the forecast will be H, otherwise the forecast will be S. If the agent 000 meets the agent 001, we can see that the outcome will be H – H, so, in this case, we can say that the players' expectations confirm each other, as we should expect at the equilibrium.

With the mental models shown above, the outcome of the 36 possible matches is shown in Table 4.

000 – 000 → H - H	010 – 101 → H - H
000 – 001 → H - H	010 – 110 → H - H
000 – 010 → S - S	010 – 111 → S - S
000 – 011 → H - H	011 – 011 → S - S
000 – 100 → H -H	011 – 100 → H - H
000 – 101 → S - S	011 – 101 → S - S
000 – 110 → H - H	011 – 110 → S - S
000 – 111 → S - S	011 – 111 → S - S
001 – 001 → S - S	100 – 100 → S - S
001 – 010 → S - S	100 – 101 → S - S
001 – 011 → H - H	100 – 110 → H - H
001 – 100 → H - H	100 – 111 → H - H
001 – 101 → S - S	101 – 101 → S - S
001 – 110 → H - H	101 – 110 → H - H
001 – 111 → S - S	101 – 111 → H - H
010 – 010 → S - S	110 – 110 → H - H
010 – 011 → H - H	110 – 111 → S - S
010 – 100 → S - S	111 – 111 → S - S

Table 4

First of all, we can notice that all the matches show a perfect correspondence that characterizes, in every match, the two players' expectations. In other words, all the mental models are consistently confirmed by the outcome of the interaction: we have reached the mental model equilibrium. Secondly, 19 of the 36 matches are characterized by the S – S strategy equilibrium, while the remaining 17 matches are characterized by the H – H equilibrium. This means that, at the population level, in each period around 53% of the agents play the strategy S and around 47% play the strategy H. This represents the strategy equilibrium that determine the overall efficiency reached by the system in this particular simulation. While in all the simulations the system reaches the perfect internal coherence among the mental models of the agents, the level of efficiency reached by the population is path-dependent and so varies from one simulation to the other. Third, we can notice that this equilibrium is characterized by a particular average payoff distribution: while the agent with the phenotype 111 plays the strategy S in 6 of the eight possible matches it can have, with an average payoff of 1.75, the agent with the phenotype 110 plays S only in 2 matches, with an average payoff of 1.22. Again, the average payoff distribution among the agents' types depends on the particular simulation run.

Simulations show that these results are robust relatively to the memory's size and the length of the learning period: varying these parameters we observe only small differences in the number of periods the system takes to reach the equilibrium. Simulations show that increasing the number of the cues does lengthen the period of co-adaptation, but the system reaches in any case the equilibrium.

4 Conclusions

While the specific results of the simulation presented in the previous section depend on variables and parameters determined through stochastic algorithms and, consequently, change from one simulation to the other, we can point out some general characteristics of the dynamics of the social and cognitive model presented in this paper.

First of all, the mental model equilibrium, and the consequent strategy equilibrium, are reached in *all* the simulations, after around 300 periods. This means that, given the assumptions of the model, the emergence of a social stereotyping system, that is, a set of socially formed and evolved beliefs that tend to confirm and strengthen each other, is statistically a very likely phenomenon. These beliefs are not true from an *absolute* point of view: at the beginning of the simulation the phenotypes attached to each agent do not have any influence on the agent's behaviour. However, by playing repeatedly the game, the agents develop beliefs about their opponents' behaviour that become true from a *social* point of view: they are true because the agent's opponent holds beliefs that make them true. In other words, the model presented in this paper suggests a mechanism for the endogenous emergence of social conventions, defined as the equilibrium reached by a system of beliefs that evolves until it reaches a state of internal coherence.

We have to notice that, even if the particular social convention that the system develops depends on stochastic events, or historical accidents, nevertheless, once it gets established it is in the interest of each single agent to follow it. In other words, even if would be in the interest of each agent to change the convention to reach the socially optimal equilibrium, the situation where all the agents choose S, an agent that would decide not to follow the convention would be worse-off in a system where the other agents follow it. This makes the convention a steady state from which it is almost impossible to escape, unless subsets of the population decide collectively to change it.

Secondly, the strategy equilibrium that characterizes the social convention is not the social optimum but it is not the worst social outcome, represented by the risk-dominant equilibrium, neither. In fact, the simulations show that the proportion of interactions S – S goes from 40 to 60% of the total. However, we have seen that each equilibrium is characterized also by an unequal average payoff distribution among the various phenotypes. This fact would have important consequences in an evolutionary setting, where the phenotypes having the higher average payoff would tend to grow in the population, a dynamic that would lead to the socially optimal outcome, where all the agents play Stag.

Finally, from a general point of view, these simulations show that, if we give up the perfect and common knowledge paradigm that characterizes the classical game theory to embrace the bounded rationality paradigm of the theory of learning in games, the dynamics and the equilibrium eventually reached by the system depend crucially on the assumptions about the cognitive skills of the agents, cognitive skills that, in order to build models that have some positive or normative value, need to be empirically justified. One of the aim of this paper has been the proposal of an agent that is one step closer to the cognitive sophistication of real agents than the agent that have populated the fictitious play models so far. The additional assumption adopted in this paper is that the agents' behaviour is based on the mental model that takes the form of a decision tree through which the agents analyse the input in a lexicographic way.

References

1. Axelrod, R.: The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration. Princeton University Press, Princeton (1997)
2. Axtell, R. L., Epstein, J. M., and Young, P. H.: The Emergence of Economic Classes in an Agent-Based Bargaining Model. In: Durlauf, S., Young, H. P. (eds): Social Dynamics. MIT Press, Cambridge, Mass. (2001)
3. Binmore, K.: Modeling Rational Players - Part I. Economics and Philosophy: 3 (1987) 179-214.
4. Camerer, C.: Behavioural Game Theory. . Princeton University Press, Princeton (2003)
5. Fudenberg, D., and Levine, D. K.: A Theory of Learning in Games. MIT Press, Cambridge, Mass. (1998)
6. Gigerenzer, G., and Goldstain, D. G.: Reasoning the Fast and Frugal Way: Models of Bounded Rationality. Psychological Review 103 (1996) 650-669
7. Hoffmann, R.: The Cognitive Origins of Social Stratification. Computational Economics 28 (2006) 233-249
8. Holland, J. H., Holyoak, K. J., Nisbett, R., and Thagard, P. R.: Induction: Processes of Inference, Learning and Discovery. Press, Cambridge, Mass. (1986)
9. Shannon, C.: A Mathematical Theory of Communication. Bell System Technical Journal: 27 (1948) 379-423, 623-656
10. Skyrms, B.: The Stag Hunt and the Evolution of Social Structure. Cambridge University Press, Cambridge (2004)
11. Vanderschraaf, P.: Learning and Coordination: Inductive Deliberation, Equilibrium and Convention. Routledge, London (2001)

Modelling Crowd dynamics

Influence factors related to the probability of a riot

Nanda Wijermans, René Jorna[1], Wander Jager[1], and Tony van Vliet[2]

[1] University of Groningen, Landleven 5, 9700 AV Groningen, The Netherlands
[2] Research institute TNO, Kampweg 5, 3769 ZG Soesterberg, The Netherlands

Abstract. This research aims for more understanding of behavior in crowds and riots. We state that crowd behavior can only be understood by studying the individuals and their interactions, a multi-level study. By selecting influence factors that influence the probability of a riot and translating this on how the individual is influenced, we have taken our first step in modeling and simulating human crowd behavior.

1 Introduction

Crowd behaviour is the behaviour shown by a large amount of individuals gathered on one physical location at a certain point in time. When some members in such a collective express violent/aggressive behaviour, we generally adress the crowd phenomenon as a *riot*. In the Netherlands for instance, riots regularly happen around soccer matches. Once in a while a demonstration gets out of hand and the day the Dutch celebrate the birthday of their queen has ended more than once in a violent encounter with the police. It might seem that there are many riots going on, however the contrary is true. Often large numbers of people gather at the same physical location at the same time without any problems. Daily commuters traveling to work by public transportation gather on stations, large-scale festivals and regular Saturdays in shopping centers also attract many people, to name a few examples. As crowds are so common, what makes some turn into a riot while most remain calm? Or to be more specific, what factors affect the probability of a crowd situation to turn into a riot? This study focuses on answering this question, in which this paper describes the first step, a conceptual model of crowd/collective behaviour. This model shows the relevant influence factors in a crowd context in relation to violent/aggressive behaviour. It emphasises the multi-level study and the focus on the individual that generates the behaviour being influences by all levels. The explicit/implemented model in a simulation of crowd and riot behaviour represents the following step.

Crowd behaviour is a complex and dynamic phenomenon, where a multitude of factors influences behaviours of the individuals in a crowd. This makes it difficult to diagnose a situation as being a high risk to turn into a riot. In dealing with crowds, the task for the riot police is very difficult as they are expected to act proportionally in a given situation. This requires knowledge of the ongoing processes within a crowd and the influence factors on the individual and thus crowd behaviour. Unfortunately, only limited knowledge is available concerning the relation between micro (individual) and macro (group) levels of crowd behaviour, as current research is limited in moving beyond observational studies towards experiments. These limitations are caused by the constraints of performing real/empirical experiments with crowds in terms of methodological rigor on the one hand and the safety of subjects on the other. Due to the multitude of variables that influence each other, it is virtually impossible to keep certain variables constant whilst manipulating others. The lack of control raises an ethical issue as the safety of subjects cannot be guaranteed while studying the turnover point of a crowd to a riot. Altogether, the lack of systematic knowledge about crowd behaviour and its underlying processes prevents professionals in practice, e.g. riot police, to improve crowd management on the basis of present-day scientific knowledge.

Although knowledge of the dynamic processes in crowds is still in a rudimentary stage, crowd research has advanced a lot in the last 15 years by numerous observational studies (McPhail, 1991; Adang, 1998). These studies had a major impact as they falsified the dominating intuitive crowd theories with empirical evidence. Major assumptions made in these intuitive theories were the existence of a group mind and shared predispositions. The idea of a group mind refers to losing cognitive control when entering a crowd, which makes people behave at will of this crowd mind (LeBon, 1895). The other perspective explained crowd behaviour by claiming that similar dispositions of crowd members result in similar behaviour (Allport, 1924; Miller & Dollard, 1941). The main insights of these observational studies proved that crowd/riot behaviour is a part of human behaviour that arises on individual level and is situation driven. Another insight was that crowd behaviour is not uniform as individuals do not behave in the same way at the same time. It appears that for instance, in soccer riots, only a small part of a crowd (10 %) actually shows violent/aggressive behaviour (Adang, 1998). Besides falsifying dominating theories, this research also gave insight into the complexity of crowd behaviour and thus the difficulty of studying this phenomenon.

In this research we aim at gaining better insight in crowd behaviour by studying the interaction and influence processes between individuals in a crowd using a multi-agent simulation model. By relating these underlying processes with the behavioural patterns shown on group level, more knowledge can be gained about the process leading towards the emergence of a riot. This contributes both to a better estimation of the probability of a riot, as well as an understanding of the effect of interferences on these processes. By developing a model that represents an individual in a crowd, we are able to develop a simulation that enables us to perform experiments with arbitrary crowd scenarios. Apart from this flexibility that overcomes the limitations of real experiments, simulations also serve as a methodological tool to develop theories that enable us to study domains that are not approachable in real/empirical experiments (Helmhout, 2006).

This paper represents the first stage in this research: developing a model of an individual in a crowd. In doing so, relevant influence factors in a crowd and related to aggressive/violent behaviour will be discussed in §2. As all influences on crowd behaviour go via the individual, §3 shortly addresses the various physiological and functional aspect of indivuals as human information processing systems. In this paragraph we also discuss some hypotheses that indicates how the mentioned variables are intervene. In paragraph §4 we give conclusions.

2 Influences on the probability of a riot

This section describes the influence factors that are related to the emergence of riot behaviour, but also those influences that are inherently present in a crowd context, e.g. human density. We do so by addressing the relevant environmental (inter-individual) as well as the internal (intra-individual) influences that play a role in the probability of a riot.

2.1 The environment

The environment represents the external world in which an individual is situated. In a crowd context this environment consists of people and other physical objects, from which physical and social influences originate. The distinction between both types of influence is based on the level tehy influence directly. Which is mostly physiology for the physical influences, for instance by directly increasing the arousal level when a high level of noise is observed. Whereas observing behaviour (social influence) directly influences the functional level.

The physical environment. In a crowd, individuals are gathered on the same physical location. Being situated in an environment, the physical surrounding influences human behaviour through their perception of physical information via the bodily senses, e.g. touch, sight, hearing, taste, smell. For instance in moving, visual/tactile information is used to avoid objects, e.g. walls or other people. In a crowd, avoiding collisions is a basic mechanism in which human density influences walking/moving behaviour. Human density, i.e. ρ_{human}, is an inevitable characteristic of crowds as it involves the co-presence of other individuals. Apart from human density, several other relations between physical factors and the occurrence of riots have been found, for instance temperature (weather), noise and scent. We will discuss them all shortly, however for an extensive overview we refer to Krahé and van de Sande (Krahé, 2001; van de Sande, 2006) .

Several general influence factors evolve around weather conditions, i.e. temperature. Two types of temperature-riot relations are distinguished, a contributing and a inhibiting effect. Hot temperature is related to an increased amount of riots and of crime rate. As it appears, there are more riots in summer than in winter (Anderson & Anderson, 1998). Whereas the other temperature-related relation is the dampening effect of bad weather (cold, windy, foggy, rainy), on aggression in crowds. Lastly, some influence factors worth mentioning are noise and (unpleasant) scent that appear to behave as an intensifier in an already tensed situation, which implies that they work as a moderator on an already aggressive mood (Geen & O'Neal, 1969; Rotten, Barry, Milligan, & Fitzpatrick, 1979). Although both noise and scent are relevant phenomena in crowds, we only use noise as their effects are similar.

In sum, we mentioned influence factors restraining actions (density), improving and dampening influences (temperature) and moderators on an aggressive mood (noise, scent). In Figure 1, the selected physical influence factors, i.e. human density (ρ_{human}), temperature ($^{\circ}C$) and noise, related to the probability of aggression are visualised.

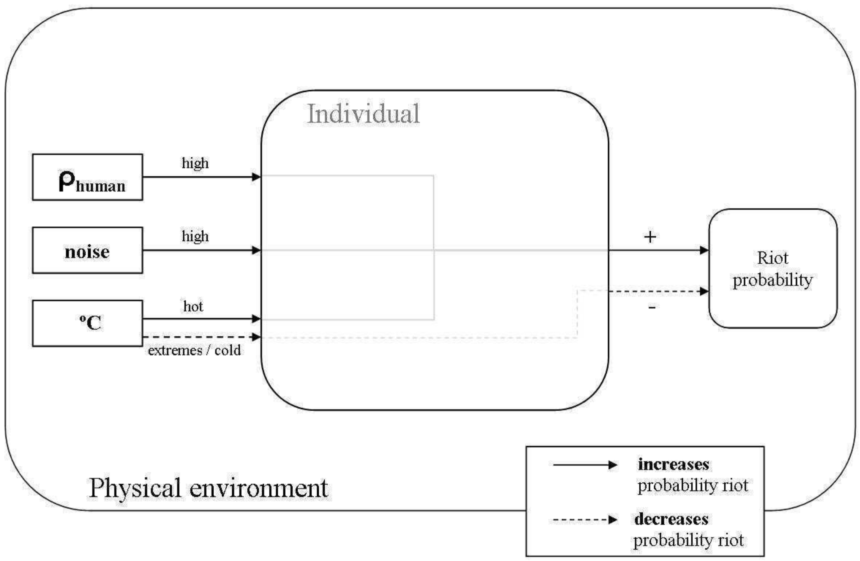


Fig. 1. Relevant physical environment factors related to aggressive/violent behaviour in a crowd

The social environment (inter-individual). A crowd is a group phenomenon that emphasises a social setting. The social environment mainly involves influences by perceiving behaviours of other individuals. In what way these perceived behaviours influence ones own behaviour depends on the connections that exists between the perceiving and observed individuals. For our crowd context we selected three social characteristics that are assumed to play an important role in crowds and thus to turn a crowd into a riot: in/out-group perception and the presence of friends and leaders.

In/out-group perception relates to an typical riot setting, i.e. a *two-block setting*, which is composed of two (or more) opposing groups (van de Sande, 2006). This distinction between *them* and *us* gives rise to certain group processes. For instance, when a clear out-group is present, the membership of the in-group becomes more salient, which gives rise to the so-called *ingroup-outgroup bias*. This bias reflects the tendency of group members to selectively favor the in-group, but to derogate the out-group (Forsyth, 2006). Apart from studying the interactions within a group, also the interplay between groups has practical relevance. For instance, studying the interplay between a rioting group and the riot police gives insight in the effect of different strategies to prevent or deescalate riots.

Friendship indicates the unique relations an individual has with its social surrounding. For example, some people are emotionally more close, due to friendship, club-membership or blood-bond. This bond has an effect on how a person is influenced by the other, which exceeds the rough distinction of in- and out-group effects. Furthermore, it is usually the case that people attend crowds with friends/family/acquaintances that makes an in-group being composed out of smaller subgroups that are closer connected.

Leadership implies that some people exert a larger influence on others than average, which is being perceived and accepted by most (sub)group members. Opinion leaders, according to Glock and Nicosia (1964), are a source of social pressure and social support that influences the decision making process of others. Whereas the effects of leadership have not experimentally been explored in a crowd and riot context, it seems plausible that opinion leaders play a more central/influencing role during riots.

Through interacting an individual is influenced. This is represented in terms of a social network that describes the connections/ties an individual has as well as characteristics it consists of (Wasserman & Faust, 1994). In our crowd context the characteristics of a social network consist of ties that represent in-group/out-group membership, friendship and leadership. This is visualised in Figure 2. The *existence of a relation* in terms of a pattern gives rise to the division of an in and out group. Friendship can be characterised externally in terms of *frequency and direction* of interaction, and leadership in terms of the *amount of in- and outgoing connections* someone has.

2.2 The individual - intra-individual -

Besides the external influences (physical and social environment), behaviour is also determined by internal (physiological and mental) factors. To be more precise, one's bodily & mental states and constraints play an essential role in determining behaviour at the individual and consequently at the group level.

Physiological. Being embodied (having a human body) implies having behavioural and cognitive constraints. Not only does our autonomy define our perceptual range and behaviour options (hearing 20-20.000 Hz, not capable of flying, etc.), it also has impact on mental processes. In a crowd setting physiological measures such as arousal and energy are assumed to play an important

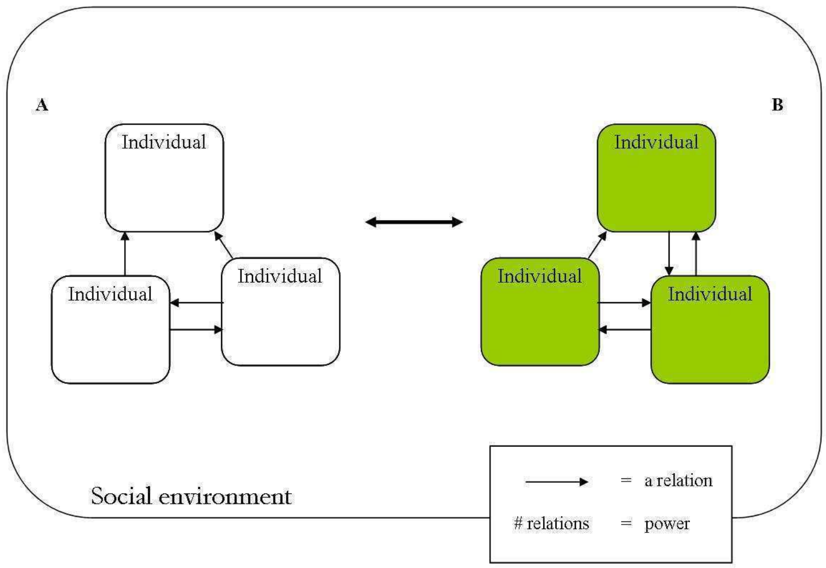


Fig. 2. The relevant influence factor of the social environment is the social network of an individual that describes if the existence of a relation, group individuals and the power someone has.

role in the course of a crowd gathering, but also factors such as alcohol and drugs influencing behaviour via physiology are very relevant in crowd situations.

Arousal is a powerful physiological state as it is related to an important basic mechanism of the sympathetic nervous system: *fight or flight*, which prepares us to act fast in a situation of threat (Baron & Richardson, 1994). Several theories of aggression are based on the notion that arousal is closely related to aggression (Berkowitz, 1981, 1988; Zillmann, 1988). This implies that all non-calm crowds, e.g. festivals, demonstrations, soccer matches, which elicit above average arousal levels, are potential riot situations. This makes the inclusion of arousal in our model very relevant.

Energy is an other important physiological measure that relates to the bodily resources an individual has. Energy is a necessary resource for life (behave, think, etc.). A low energy level may drive people toward the decision to leave the crowd as to get food, drinks or to rest, and so to restore the energy level. Both arousal and energy correspond with primitive/basic behaviours that provide in short term survival mechanisms (fight/flight, food search) that allow humans to respond to a subsistence/life-threatening situation.

As physiological measures can strongly influence behaviour, also factors influencing physiology are important to address here. In many crowd contexts the use of alcohol and drugs is a common phenomenon. Alcohol consumption is often linked with the occurrence of group violence. Even though there is no causal relation between alcohol usage and aggression, evidence shows that a little amount of alcohol leads to increased aggressive behaviour (Baron & Richardson, 1994; Russel, 1993). The role of drugs on the other hand, is probably less important than one would expect. Mostly, the effect of drugs (apart from alcohol) in relation to aggression is not that well explored (van de Sande, 2006). However, the effects of marihuana or combined party drugs, e.g. XTC, LSD, etc., do not appear to effect the course of a riot (van de Sande, 2006). In case of marihuana it even seems to decrease the probability of violent behaviour (Baron & Richardson, 1994) as it makes an individual less aware of its social surrounding. In sum, although many physiological factors mentioned are relevant for a crowd/riot context, the relation between alcohol/drugs and aggressive behaviour has only been proven to correlate respectively positively/negatively, but remains unexplained at the mental level.

Functional. The functional part of an individual is where the behavioural preprocessing takes place. In this, we are interested in those mental processes, characteristics or content types that are involved in the preprocessing of aggressive/violent behaviour. Before addressing some relevant concepts, we will shortly describe the view we have on the cognitive system.

In describing the cognitive system, a cognitive architecture provides a blueprint of the organisation of the human mind in a fixed structure. The cognitive architectures like ACT-R (Anderson & Lebiere, 1998), Soar (Newell, 1990), and Clarion (Sun, 2003) describe functional components that cooperate to process information and to result in behaviour. In this description two other concepts play an important role: representations and processes. A mental representation contains information or other representations. These representations are manipulated by the mental processes. Figure 3 gives an overview of the components and relevant processes that correspond with the hybrid cognitive architecture ACT-R (Anderson & Lebiere, 1998; Helmhout, 2006).

On a more abstract level three main components are normally be distinguished with regard to the human information processing system: perception, cognition (mainly memory) and action. Each have their own substructures and ways of interacting with the other components. Memory in general represents the knowledge an individual has. This includes knowledge about performing an action, such as throwing a rock (procedural knowledge), knowledge of facts, such as knowing what it is you can use rock throwing for (declarative knowledge), and other forms of knowledge. Furthermore, goals are also represented as memory elements. Perception and action on the other hand are the components that allow interaction between the internal and external world, by receiving input via the bodily sensors, i.e. perceiving, and output by behaving, i.e. acting.

To select relevant functional influence factors, a certain content of memory, e.g. knowledge and functions, we need to address the distinction between the way we talk and the way we model. When observing behaviour we tend to talk in terms of motivation, attitude, identity, and self to relate the behaviour we observe with internal settings. It is a way in which we talk about our inner system. Because we want to model intra-individual in according to the general view in AI, we need to translate these terms by operationalise them in terms of representation and mental processes. In including relevant influence factors we mainly focus on mechanisms that contribute to moods, tendencies or actual behaviour of aggression. We selected saliency, motivation, and insecurity to play an important role in the probability of showing aggressive behaviour, which need to be translated in terms of content of the cognitive architecture.

Saliency relates to the phenomenon where something, such as norms or identity, is in focus and therefore has more influence than other norms/identities. (Kallgren, Reno, & Cialdini, 2000; Mullen, Migdal, & Rozell, 2003). We relate this with the *activation* level of the memory elements, which implies that in showing aggressive/violence the elements representing these behaviours have to be highly activated. This activation causes a dynamical ranking of behaviour based on perception or an internal source of a motivation or goal.

Crowd behaviour is very diverse, and thus are the underlying goals/motivations. This makes the dynamics between goals and thus the resulting behaviour essential within crowds. With motivation we mean: "[...] *Motivation is a modulating and coordinating influence on the direction, vigor, and composition of behaviour. This influence arises from a wide variety of internal, environmental, and social sources and is manifested at many levels of behavioural and neural organization*" (Wilson & Keil, 1999). To have a simple representation of driving forces for behaviour we use the concepts of *needs* (Max-Neef, 1993), which represent a fixed set of abstract goals that can be fulfilled with satisfiers/explicit goals. The set of needs we found relevant for crowd behaviour are: *Subsistence (biological)*, *Safety*, *Belonging* and *Autonomy*. Where subsistence is need for self-preservation in terms of energy, safety relates to a need to feel safe, belonging to a group and autonomy to remain self sufficient. These needs are represented as memory elements that are constantly addressed and updated.

The last influence factor we want to mention is insecurity. Insecurity relates to a situation where no behavioural rules, i.e. norms, are present but also not known in terms of experience. For instance, when in a supermarket standing in line waiting on your turn to pay, you know exactly how you are supposed to behave. Most crowd situations are always insecure in some way, as it is not exactly clear what will happen and what is the norm how to behave. For instance, when joining a demonstration there are no clear rules of conduct for a demonstration. So, if there are not really strong associations with behaviours it might explain why some behaviours are very surprising. This factor is also represented in terms of memory elements. Insecurity is a measure of the strength of associations in a certain situation. The absence of norms and expectancies makes the behaviour selection more depending on activation, as no good differentiator in comparing behaviour is present.

In sum, we described the relevant intra-individual influence factors that are physiological (energy and arousal) and the functional (reformulated concepts of: activation, motivation representations, insecurity). The cognitive system as well as its moderating influences are shown in Figure 3.

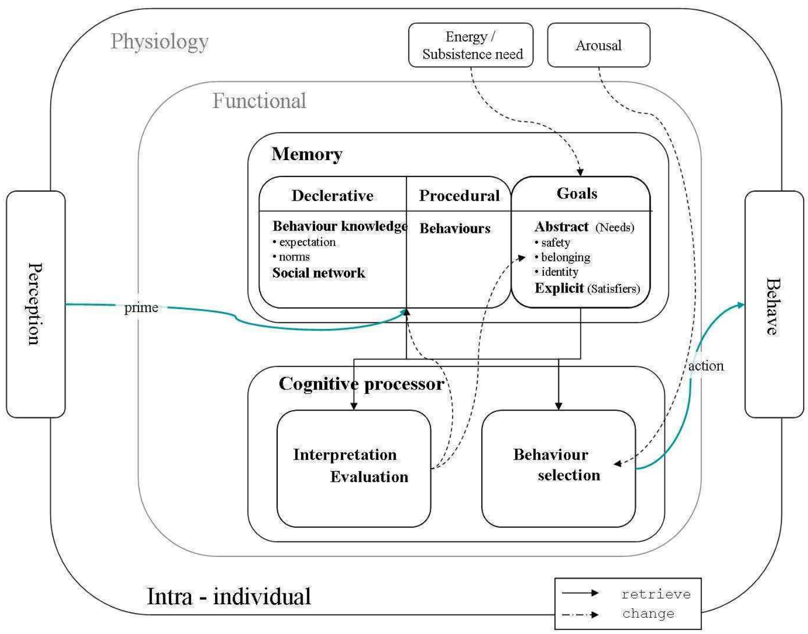


Fig. 3. Internal processes and components of an individual

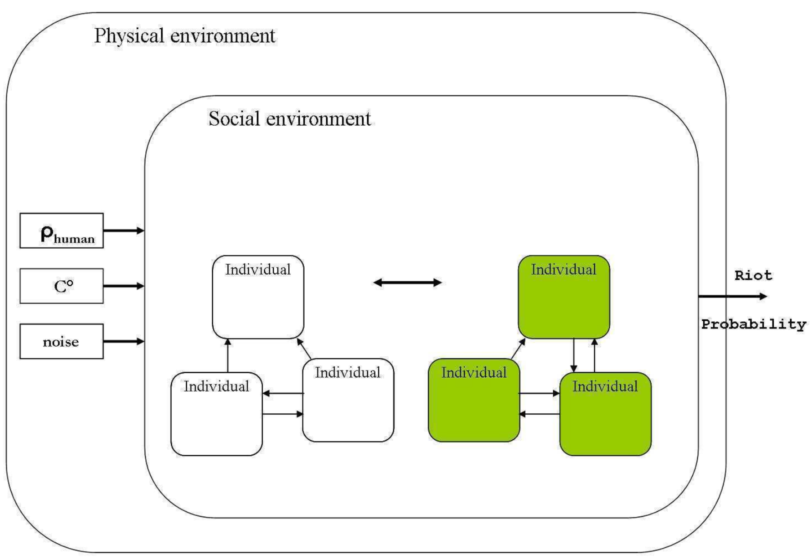


Fig. 4. Overview multi-level influence factors that are related with aggressive, violent or riot behaviour.

This section addressed the known influence factors that are related to aggressive/violent behaviour that are applicable or specifically related to a crowd context. In Figure 4 we give the overall view on the influences that are related to the probability of a riot. The physical influences that count for all individuals as well as the social environment, which depends on the social network a person individually has. All these influences go via the individual where they influence behaviour pre-processing. Altogether, the behaviours shown on group level give rise to a riot or not.

3 Influence factors on individual level

In modeling crowd behavior, we want to develop a model of an individual in a crowd. As we mentioned earlier, crowd behavior arises from the behaviors of its individuals, see Figure 4. This implies that all the influence factors we mentioned in §2, influence the probability of a *riot* via the individual. The model choices are represented by the translation of each factor on the individual level. In this short paper we can not discuss all the hypotheses we formulated for the various influence factors at the various levels. We will discuss three hypotheses concerning arousal (physiology), leadership (social) and needs (functional).

We assume arousal to play an important role in the turn of a riot. The arousal level is related to the functioning of the behavior selection process. This means that when an aggressive/violent behavior is salient and arousal is high, the evaluation and comparing process is inhibited due to time, which increases the probability of showing aggressive behavior.

Hypothesis 1: High levels of arousal increases the probability of turning a crowd into a riot by 'impairing' the behavior selection process.

A leader has a major influence on people who perceive him/her as such. This makes the behavior or (dis)approval of a leader a dominant influence in the behavior selection process. These normative boundaries set by a leader can play a crucial role whether a riot will arise or not.

Hypothesis 2: If a leader is engaged in aggressive behavior, the likelihood of a riot will increase.

The dynamics between dominance in needs give rise to different behaviors. We expect that this dominance of needs is related to a higher activation of a subset of behaviors, where need dominance depends on the situation. In a crowd situation where some people show aggressive behavior high dominance of the need *belonging to a group* in the surrounding individuals will give rise to the 'contagion' process/pattern of violent/aggressive behavior.

Hypothesis 3: If some individual show aggressive behavior, high dominance of the need *belonging to a group* in the surrounding individuals, will increase the probability of a riot.

To move on to the next stage of our research all the relations will be formalised in a fixed structure that represents the cognitive system of an individual. This involves a framework that allows for modeling the internal processes, as well as for the interaction with a social and physical environment. A framework suitable for our domain is an integration of a cognitive architecture with a multi-agent systems framework, in which the first takes care of the internal processes and embodiment related to cognitive capabilities and the latter with the social interaction. In doing so we will explore the possibilities to use RBOT (Helmhout, 2006), which is a cognitive agent based social simulation system that implements ACT-R in Java and allow for social interaction. Only after the formalising and implementing the model into a simulation of crowd behaviour, experiments can be performed to explore the dynamics of crowds/riots and test the efficiency of intervening strategies.

4 Conclusion

In the approach of our simulation research of which we here only presented the conceptual structure we emphasised four design elements. In the first place we showed the necessary multi-level relationships in crowd and riot behavior. The physical, the social, as well as the cognitive level are important in the interdependencies and influence factors. In the second place we included the cognitive plausibility of the individual agents even in crowds in our model. In the third place the various levels are integrated via the cognitive agent itself. Finally, the conceptual structure is constructed out of modular components. This modularity allows for extending components on each level within the overall multi-level structure. All in the light of gaining understanding of crowd and specifically riot behavior.

Bibliography

Adang, O. (1998). *Hooligans, autonomen, agenten. Geweld en politie-optreden in relsituaties*. Samson.

Allport, F. H. (1924). *Social Psychology*. Houghton Mifflin.

Anderson, C. A., & Anderson, K. B. (1998). *Human aggression: Theories, research and implications for social policy.*, chap. Temperature and aggression: Paradox, controversy, and a (fairly) clear picture., pp. 247 – 298. San Diego, Academic Press.

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Lawrence Erlbaum Associates.

Baron, R. A., & Richardson, D. R. (1994). *Human Aggression*. Plenum Press.

Berkowitz, L. (1981). *Multidisciplinary approaches to aggression research*, chap. The concept of aggression, pp. 3 – 15. Orlando FL : Academic Press.

Berkowitz, L. (1988). Frustrations, appraisals, and aversively stimulated aggression. *Aggressive Behavior*, *14*, 3 – 11.

Forsyth, R. D. (2006). *Group Dynamics* (4 edition). Thomson Wadsworth.

Geen, R. G., & O'Neal, E. C. (1969). Activation of cue-elicited aggression by general arousal. *Journal of Personality and Social Psychology*, *11*, 289 – 292.

Helmhout, M. (2006). *The Social Cognitive Actor, A multi-actor simulation of organizations*. Ph.D. thesis, University of Groningen.

Kallgren, C. A., Reno, R. R., & Cialdini, R. B. (2000). A Focus Theory of Normative Conduct: When Norms Do and Do Not Affect Behavior. *Society for Personality and Social psychology*, *26*(8), 1002–1012.

Krahé, B. (2001). *The social psychology of aggression*. Psychology Press.

LeBon, G. (1895). *La Psychologie des Foules*. Alcan.

Max-Neef, M. (1993). *Real-life economics: Understanding wealth creation*, chap. Development and human needs, p. 197.

McPhail, C. (1991). *The myth of the madding crowd*. Aldine de gruyter.

Miller, N., & Dollard, J. (1941). *Social Learning and Imitation*. Yale University Press.

Mullen, B., Migdal, M. J., & Rozell, D. (2003). Self-Awareness, Deindividuation, and Social Identity: Unraveling Theoretical Paradoxes by Filling Empirical Lacunae. *Society for Personality and Social Psychology*, *29*(9), 1071–1081.

Newell, A. (1990). *Unified theories of cognition*, Vol. 3. Harvard University Press.

Rotten, J. F., Barry, T., Milligan, M., & Fitzpatrick, M. (1979). The air pollution experience and physical aggression. *Journal of Applied Social Psychology*, *9*, 387 – 412.

Russel, G. W. (1993). *The social psychology of sport*. Springer.

Sun, R. (2003). A Tutorial on CLARION 5.0. Tech. rep..

van de Sande, H. (2006). *On Crowds*.

Wasserman, S., & Faust, K. (1994). *Social Network analysis: methods and applications.* Cambridge University Press.

Wilson, R. A., & Keil, F. C. (Eds.). (1999). *The MIT Encyclopedia of The Cognitive Science*. The MIT Press.

Zillmann, D. (1988). Cognitive-excitation interdependencies in aggressive behavior. *Aggressive Behavior*, *14*, 51 – 64.

Does Cognition (Really and Always) Matter? The *Vexata Quaestio* of the Micro-Foundations of Agent-Based Models from a Sociological Viewpoint[♣]

Flaminio Squazzoni*

*Department of Social Sciences, University of Brescia, Via San Faustino 74/B, 25122 Brescia, Italy, squazzon@eco.unibs.it

Abstract. In the social simulation community, there is a vivid dispute between supporters and opponents of the relevance of cognition for the analysis of macro social phenomena. This paper aims at illustrating advantages and disadvantages in the cognitive simplification/sophistication of models, enriching the debate in social simulation with insights from outside the community and embedding the debate on sound methodological issues. The paper argues that cognitive complication of sociological models should not be viewed as a price to be paid to the realism of the model, and that social simulations should not be always coupled with cognitive foundations to have a real explanatory power.

Keywords: cognition, social phenomena, micro-foundations, social simulations.

1 Introduction

In the social simulation community, there is a vivid dispute between supporters and opponents of the relevance of agents' cognition for modelling and understanding the emergence of macro social phenomena [8] [17]. Does cognition (really and always) matter for understanding social phenomena? Should we build sophisticated cognitive models at micro level to find satisfactory theoretical explanations of social

---

♣ I would like to acknowledge my grateful thanks to Riccardo Boero and Marco Castellani, two colleagues and friends of mine, for many useful remarks on a preliminary version of this paper, and for the enriching exchange of ideas we had in these last years. Moreover, I would like to express my gratitude to two ESSA conference anonymous referees for some precious comments and suggestions that allowed me to improve the quality of this paper. The usual disclaimer applies.

phenomena? These are two of the most important disputed questions in social simulation.

This paper attempts to illustrate advantages and disadvantages of the cognitive simplification and sophistication of social simulation models, with the aim to enrich the debate in social simulation with insights from outside the community. As a matter of fact, when social simulation scholars deal with the role of cognition, they seem frequently more careful to technical and modelling details than to substantive and theoretical issues. Opening the attention to contributions from outside the community is a way to bring the debate back to substantive and theoretical issues, as well as to methodological ones.

The first section summarises the most representative approaches in the social simulation community. A particular attention is paid on the difference between what cognitive scientists/artificial intelligence scholars doing social simulations, on one side, and sociologists/political scientists/economists doing social simulations, on the other, think and actually do. The difference concerns both the analytical purposes of the models and the methodological practices.

The second section summarises the sociological approach to cognition. Sociologists could be distinguished in three categories: cultural sociologists, who claim the sociological foundations of cognition and cultural facts and, often, the need for ethnographic and qualitative methods to go deep in the analysis of individual knowledge; quantitative sociologists, who are not interested in individual behaviour and in micro details but just on the measurement and the analysis of social facts; mathematical and computational sociologists, or, in general, formalised sociologists who do not believe in the need for sophisticated cognitive foundations of sociological models, given the common idea that sociologists should be satisfied by taking into account the results and not the generative processes of the agents' cognition, and that sociology should have the purpose to understand aggregative interaction processes and their properties [6] [18] [20].

The third section attempts to establish the dispute on some methodological foundations. For this purpose, three methodological practices in social simulation are compared: the ***KISS*** principle [2], the ***kitchen sink*** principle, and the ***complication*** principle. From a standard sociological viewpoint, most of sophisticated cognitive models in social simulation seem often to fall into the second category. In general, KISS principle should be viewed as the preferable standard practice, because of the recognised advantage of simplification, abstraction, and transparency of models for the inter-subjective control and collaboration, and scientific progress. The complication principle is suitable just in particular cases, when theoretical and/or empirical evidences and reasons justify to enrich some components of the model (cognitive level being one of them).

The paper ends up arguing that the level of complication of cognitive foundations of sociological models should be related to the nature of the *explananda* only on the base of functional reasons, taking into account this following good methodological practice: to start with relevant empirical phenomena to be addressed with simple models and general theories, and to further complicate the models just if simple theories do not work in explaining the phenomena, and not because of the presumed complexity of the empirical reality, so to avoid the kitchen sink risk. For the purpose of doing analytical research, models should not be intended as a 1 to 1 (or even a 4 to

1, as it happens in some case with cognitive models) mapping of the complexity of the reality, but as a means to abstract, dissect and simplify among the many factors, variables and mechanisms potentially and empirically involved in generating a given phenomenon.

In this perspective, the question of the analytical levels on which to build and divide a theoretical model (micro-meso-macro) can not be confused with how the reality concretely is, as though analytical levels were ontological objects of the social reality [6]. The methodological and modelling practice of identifying analytical levels of models, and, in this case, of differentiating between cognition (micro) and social (macro) level, is commonplace in science, not only in social simulation. But, the idea to proceed in the analysis by permanently assigning autonomous epistemological existence to the levels, and by taking into account their mutual interaction every time, should be justified only if such a difference adds explanatory power to the model. Otherwise, such a methodological practice can come to damage the heuristic value of the analysis [17].

2 Cognition or Not? A Social Simulation Viewpoint

The relevance of cognition is far from being commonly recognised in social simulation. Cognitive scientists and artificial intelligence practitioners doing social simulations strongly emphasise the cognitive premises and consequences of social emergence, although the results of most of agent-based models in social simulation seem to support the contrary argument. As a matter of fact, the overall evidence of social simulations is that "contra interpretivists, objective structure can emerge, and the existence of those structures can constraint individual agents (via changes in patterns of local interactions), even when agents have no internal representations", that is no relevant cognitive features [28].

For the sake of analysis, four different approaches to the modelling of agents' behaviour can be identified in social simulation (even if, in the reality, they are often blended), with a various degree of thickness of cognitive foundations involved: the "folk psychology" approach; the bounded/adaptive rationality approach; the social cognition approach; the computational logic/artificial intelligence approach.

The "folk psychology" approach, sometimes called "commonsense psychology" or "naïve theory of action", is a quite common means to modelling individual action in sociology and psychology, as well as in part of artificial intelligence [24]. Among the four approaches, this first is the less cognitivistic one. It is based on the undeniable evidence that behaviour explanations are a everyday human activity and that we, as humans, interact with our peers and evaluate their actions by assessing their intentions depending on their (supposed) beliefs and desires [4]. This is the reason why folk psychology is also viewed as a simple and intuitive general theory of the individual behaviour. In psychology and sociology, it is called Belief-Desires-Opportunities (BDO) theory and has been originally established by Jon Elster [12]. In social simulation and in artificial intelligence, it is called Belief-Desired-Intentions (BDI) framework, which is viewed as a computationally tractable model of cognition, easy to understand and quite general.

The assumption behind this theory is that modelling individual action as a set of beliefs, desires and opportunities is a way to make individual action intelligible and thereby explainable [20]. Moreover, one could select a particular ideal-type of explanations of individual action, for instance an opportunity-based explanation rather than a desire-based, or connect these building blocks to formally express simple and general behavioural patterns, such as adaptive preferences, counter-adaptive preferences or wishful thinking, which can be used as causal explanations of individual behaviour. BDO theory is confronted or put together with other alternatives, such as rational choice models, adaptive rationality or learning models. As we see in the next section, sociologists make use of BDO theory to abstract away any sophisticated cognitive features of individual action so to get a suitable model for sociological purposes.

Originally introduced by Herbert Simon with the concept of "bounded rationality" [30], and further developed by James March [25] and, more recently, by Gigerenzer and Selten with their notion of "adaptive rationality" [15], the second approach emphasises the importance of internal cognitive processes of decision making and embeds the cognition in the individual mind, with the purpose of understanding how decision is undertaken in complex environments. Most of agent-based models in computational economics, as well as in evolutionary economics, are inspired by a simplified and reduced version of these cognitive models (local adaptation vs global outcomes), with the aim to leave behind or to complicate respectively the standard assumption of the rational choice model and the standard economics [14] [33]. In this case, the focus is not on the purpose of the individual action, but on cognitive procedures, such as selective attention, time, memory, computational capacity, local information, and so on.

The social cognition approach, which has been put forward by social and cognitive psychologists, such as Zerubavel [35] and Kunda [21], is focused on the attempt to find behaviour explanations not only appealing to psychological bases of individual action or cognitive processes, but rather based on the sense-making that is made possible on the base of social interactions and knowledge. The attention is on how people use social mental categories to persuade each other, communicate or build and exploit social images and reputation. Examples of a similar approach in social simulation can be found in the social influence model by Latané [22], or, with a different analytic purpose and a more sophisticated cognitive modelling, in the reputation model by Conte and Paolucci [9]. Here, the cognitive elements of social action are not conceived as pure procedural internal building blocks of individual decision, as in Simon's view, but as substantive social components of the action's meaningfulness.

Last but not least, there is what we called the computational/logic approach, which should be considered as a third way between folk psychology and social cognition. In this case, scholars have a close interest in using formal logic and computing as an appropriate analogy of cognition, with an emphasis on commonsense reasoning, the structure of human mind and social agency [7] [26] [32].

Social scientists doing simulations see the cognitive notion of "agent" in a different way, either as a pragmatic modelling-oriented problem, more than a theoretical substantive issue, or as a sophistication that would only complicate a model without a strong theoretical reason to do it. Moreover, since theoretical assumptions behind cognitive models suggested by cognitive scientists doing social simulation are *ad-hoc* solutions with extreme theoretical and paradigmatic diversity, the unavoidable consequence for social scientists doing social simulation is that “the choice of a cognitive model […] will probably depend on pragmatic issues such as how easy they are to obtain, previous experience with them and their use in related research” [17], more than for good theoretical and explanatory reasons.

Looking upon the debate, it is easy to realise that there is a clear difference between what cognitive scientists/artificial intelligence scholars doing social simulations, on one side, and sociologists/political scientists/economists doing social simulations, on the other, are actually doing. The difference concerns both the analytical purposes of the models and the assumed methodological practices.

Most of computational sociologists or economists argue that modelling agents should be modelling no more than individual decision rules [13]. This is because resources and attention of a model should be put on the explanatory power of interactions, and on how decisions aggregate to give rise to empirical macro patterns and regularities, which are the ultimate *explananda*. Computational economists, who represent the larger community in social simulation by now, are mostly interested in demonstrating that the invisible hand does not require intelligent fingers. Given such a purpose, the consequence is that their simulation models are built to relax the standard assumption of mainstream economic theory on the absolute rationality of agents and their homogeneity, including agents (heterogeneous to a certain extent) following simple behavioural rules. Game theorists and scholars, who are engaged in the field of cooperation and social dilemmas, do about the same. Most of computational sociologists believe that sociology should explain particular features of aggregate behaviour and collective patterns, and not features of individual action [19] [6].

On the opposite, cognitive scientists doing social simulations strongly argue that, under each agent’s decision, there is a causal cognitive process that can influence, and be influenced by social factors, with the consequence that modelling agents without cognitive features can give rise to strong epistemological mistakes and to unsatisfactory explanations of a social phenomenon [7]. As Sun has recently outlined, in his interesting attempt to combine social simulation with cognitive sciences, models of social phenomena need for cognitive foundations because the understanding of aggregate phenomena is possible only if we have a ”better understanding” of individual cognition, given that “cognitive models may provide a more realistic basis for understanding multi-agent interaction, by embodying realistic constraints, capabilities, and tendencies of individual agents in their interaction with their environments” [32]. This is the same argument put forward by Pylyshyn [27], when he argued that explanations of individual behaviour that do not involve cognitive terms run the risk of misattributing the reasons for that behaviour.

On this need for deepening the micro-foundations, cognitive scientists meet artificial intelligence practitioners and theorists, who are working on the conceptualisation and the modelling implementation of theories of “agency” [26]. Unfortunately, at least from a social science viewpoint, these last seem mostly interested in generating taxonomies to pinpoint differences between types of agents, meta-rules and rules of behaviour, and to define ontological and/or modelling frameworks that should be used for the first purpose. The construction of computational objects is mostly not related to the explanatory purpose on which a sociologist or a social scientist is called to take into account, that is to explain a given macro empirical social phenomena starting from theoretical assumptions on agents' behaviour and interactions. As a consequence, technical issues in the modelling of agents are sometimes confused with theoretical issues, with the result to have complicated models, methods, and tools without a real explanatory power.

But, although cognitive scientists call attention to the aim of realism of models, sociologists and economists argue for the need to abstract away the micro details of agents' behaviour to concentrate on micro-macro components of models. The level of in-depth examination of individual action to be established in sociological models is instrumental to the analytic purpose of explaining aggregate social phenomena and not a analytic purpose in itself.

3 What Sociologists Think and Do (or Should Do)

With reference to cognitive issues, one can roughly distinguish three types of sociologists:

1) cultural/qualitative sociologists, who claim the sociological foundations of cognition and cultural facts and, often, the need for ethnographic and qualitative methods to go deep in the analysis of individual knowledge;
2) quantitative sociologists, who are not interested in individual behaviour and in micro details, but just on the measurement and the analysis of social facts;
3) standard mathematical, computational or formalised sociologists, who do not believe in the need for sophisticated cognitive foundations of sociological models, given the common idea that sociologists should be satisfied by taking into account the results and not the generative processes of the agents' cognition and that sociology should have the purpose to understand particular features of aggregate behaviour and collective patterns and not features of individual behaviour [6] [20].

As said before, it is quite evident that sociologists do not deal with the challenge of explaining individual action [6]. They would aim at explaining social structures and collective phenomena, and, in particular, the objective nature of these, as quantitative and mathematical sociologists would clearly argue [18]. An example is the famous Schelling model of segregation and its many variants [8] [16], where a population of neighbouring agents is assumed to be divided in two groups. It is quite undeniable that the reason of such a division should make reference to relevant socio-cognitive basic processes such as social identity and identification. But, the Schelling's analytic purpose was not to explain why people were divided in two groups, or why they felt to belong to a particular group, and so on. The model was created to explain the unintended aggregate consequences of individual action, taking for granted all the basic cognitive processes that could be the causes of such action. In a word, the explanatory purpose was to explain the emergence of social macro dynamics out of interactions, not the feature of individual action [17].

Vice-versa, sociologists, who are interested in deepening the relation between objective and subjective nature of social institutions or structures, strongly believe in the primacy of qualitative methods, and are not interested in computational and mathematical approaches to sociology. In these cases, cognition undoubtedly matters for sociological analyses, but is necessarily investigated through discursive and narrative methods, or even through other formalised tools.

Computational and mathematical sociologists, as well as formalised sociologists, who believe in the added value of simplification and parsimony for the scientific enterprise, argue that most of the computational resources of a sociological model should be addressed to the explanatory power of interactions, with the aim to understand how individual decisions aggregate giving rise to macro patterns and regularities, which are the ultimate sociological *explananda* [6] [10] [19].

This last argument was suggested by Coleman, with his famous example of the "boat" (see figure 1), who correctly emphasised the complexity trade-offs between the different components of a theory in sociology. Since the sociological purpose is to understand how individual actions combine to generate a macro outcome, in order to explain synchronic and diachronic changes or features at a macro level, a sociological theory should contain four basic components: a macro situation *X* (a given empirical macro-regularity, possibly measured in terms of a social indicator, a variable, and so on, which is a set of possibilities/constraints upon individual action); a set of individual actions (explained in terms of a rational action, for instance, inside a BDO framework); an interaction structure that connects individual actions; a micro-macro link able to explain how interaction produces a macro outcome *Y*.

As a consequence, to allow greater complexity to the point 3) and 4), which are typically of greater sociological interest, "one must keep the action component as simple as possible by abstracting away all elements not considered crucial" [20]. Sociologists who opted for folk psychology theory, a BDO framework or a rational action theory as micro-foundations of their models, did it intentionally for such instrumental and theoretical reasons [6].

There are two further theoretical arguments that corroborate such a position, at least in sociology [17]. The first one is that the understanding of a social phenomenon at a macro level does not always require the assumption of behavioural heterogeneity at a micro level, with the consequence that the attention to endogenous cognitive features of behaviour would be a pointless exercise, or a complication of the model without a theoretical purpose to do it. Moreover, even if there would be empirical and/or theoretical evidence that behavioural heterogeneity matters in the case of interest, this should not automatically imply the recourse to cognitive features and sophistications, given the possibility of pulling up the micro-foundations of a model to the behavioural/action level. In most cases, the reason for complicating at the micro level depends on the need for exploring the space of possible behavioural parameters in more exhaustive way [10], and not necessarily on the need for plunging into the cognitive level. If this is true, then the problem of the micro-foundations of sociological models is more a quest of 'horizontal exploration' (among alternative micro-foundations), than a quest of 'vertical in depth-examination' (between behavioural and cognitive levels of assumptions).

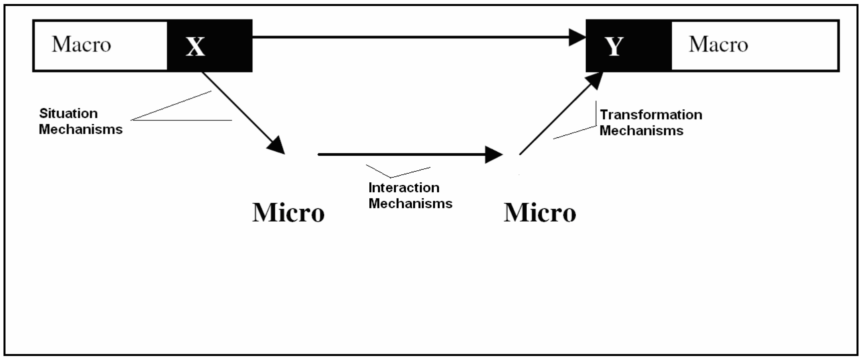


Fig. 1. A Representation of the so-called "Coleman Boat".

The second one concerns the so-called argument of the "multiple realizability" [28], which stands for the evidence that a wide range of possible micro behavioural mechanisms might come to produce the same macro phenomenon at the level of sociological interest, so that a general sociological explanation should abstract away the micro details and the concrete generative micro mechanisms of the macro realisation. As Nigel Gilbert rightly outlined, the reason why Schelling's model of segregation is a good model lies in its theoretical generality in sociological terms, and precisely in its capacity to demonstrate the role of unintended macro consequences of the households' migration, beyond any supposed argument about the heterogeneous and differentiated underlying motivations at the household level [17]. Anyway, this does not impede that a further scientific goal should be to explore alternative theoretical assumptions as regards to Schelling's theoretical starting point [16], as well as to carefully verify the assumptions of the model in particular empirical contexts, as other scholars more recently did [5].

At this point, a computational, mathematical or formalised sociologist would ask why bother with cognition. According to an analytical and formalised sociological approach [6], [20] [18], an appropriate micro-foundation of a sociological model should be a plausible theory of human action, based on empirical-grounded assumptions "known to be at least roughly correct in the real-world setting that we are analyzing" [20], expressed in an abstracted and precise theoretical framework, that is as simple as possible. The theoretical abstraction allows avoiding accounts of individual action that are based on "thick" descriptions, given the epistemological assumption that "neither the psychology for the phenomenology of human action is a focus of analytic attention: no attempt is made to capture the full diversity of the cognitive or motivational aspects of action nor of the nuances of its subjective meanings" [18].

This position definitively depends on the idea, mentioned before, that the primary analytic concern of sociology is to elucidate the macro-micro-macro link, that is to show the ways in which a number of individual actions, given some social constraints and a *ex-ante* macro situation, come together so as to generate macro-social phenomena of interest, as well as how such macro-phenomena can come to affect individual action. The sociological in depth-examination of individual action is instrumentally viewed as a means to understand macro phenomena, and not an analytic purpose in itself.

This is the reason why the concepts of individual that is appropriate to cognitive sciences, where both analytic and explanatory purposes deal with individual phenomena and the higher level (to be explained) is cognition, and the concepts of individual that is appropriate to sociology, where analytic purpose is to explain macro level phenomena, differ very much [23]. Following Coleman [6], Stinchcombe [31], Hedström [20] and Golthorpe [18] in their footsteps, the point is that the relevant advantages of formalisation, simplification and abstraction of models, as well as their explanatory power, are lost if the lower-level processes, which are introduced to account for higher-level social regularities (to be explained), require themselves complex (often *ad-hoc*) explanations.

4 Methodological Issues: KISS, Kitchen Sink and Complication

The debate on cognition in social simulation and in sociology cannot be abstracted away from methodological issues and foundations. This is because, as said before, adding cognition to a sociological model is adding complication at a micro level and, as a consequence, an unavoidable trade-off between simple/transparent explanations and complicated models.

It is worth noting that deciding within such as trade-off is difficult, and, ultimately, a matter of scientific perspectives and tastes. From one hand, it is manifest that computer simulations, in particular agent-based models, allow to include more complications and details in the modelling, as compared to what scholars usually can do with mathematical and standard analytical models. On the other, the fact that a tool is particularly flexible and allows to easily express complicated models is not a good methodological and theoretical reason to play it on without measure. In fact, if this was the criterion, the resulting paradox would be that doing simplified and transparent models in social simulation could be implicitly perceived as driving a race car doing no more than 15 miles per hour.

To give a more structured look at such debate, one can compare three methodological practices that are currently confronting in social simulation: the ***KISS*** principle [1], the ***kitchen sink*** principle and the ***complication*** principle.

The ***KISS principle*** stands for the motto "Keep It Simple Stupid", and it is the standard practice in social simulation: it means that a given social phenomenon should be understood in theoretical terms by abstracting away irrelevant details and empirical richness and simplifying assumptions and modelling components at the most. This allows for an easy understanding of what is going on in the model, as well as an effective inter-subjective control and falsification, which are synonymous of the scientific progress. From a sociological viewpoint, the KISS principle is the opposite of what might be called a cognitive foundation. Actually, it is not a case that most of simulation models in sociology and in economics are micro-founded on simple behavioural rules, with very simple bounded rationality processes at a micro level (usually, some meta-rule that allows rules' up-to-dating, selection and change over time). Moreover, it is not a case that all of the most famous and successful theoretical and/or simulation models are kept very simple at the micro level. In sociology, examples are Schelling segregation model [29], Granovetter's model of collective phenomena [19], or Axelrod's model of cooperation [1], just to name a few. The epistemological and methodological advantages of KISS approach are: the simple intuition behind the model, the wide range of application of the model, the theoretical generality of the explanation, the possibility given to colleagues to enter the model in detail, to use it, re-use it, extend it, test it, and control it. As a matter of fact, it is more congruent for the scientific success of a discipline to proceed step by step in a cumulative and collective way. Simplification, abstraction, generality and parsimony are the main features to be able to proceed in such way. This is what KISS models allow to do.

The ***kitchen sink principle*** is that a model should map the richness, details, and features of the reality as close as possible, with the result that "you simply open the fridge and take everything in it to put in the kitchen sink and prepare the soup" [1]. The assumption is that the model of a phenomenon should be able to include all the factors, variables and objects that are called to play a role in the genesis or emergence of the phenomenon, following both theoretical or empirical inspiration. It is a common attitude in empirical case-studies, or in qualitative and ethnographic studies, that has been imported also in formalised social science. As said before, such attitude is not so uncommon also in social simulation, given the belief that the advantages of computer simulations in respect to standard analytical models is that you can add a lot of details and complications, and see what happens.

As said before, the argument of the higher level of flexibility of computational tools as compared with mathematical or other formalisation tools is a out-of-place argument. As Hedström correctly pointed out, "since there are hardly any constraints on the analyses that can be performed, it sometimes leads analysts to specify simulation models that are almost as complex and difficult to comprehend as the real-life processes that they are supposed to help us understand. To be useful, a model must be realistic yet sufficiently simple and transparent to further our understanding of the key mechanisms at work. *If it is not, then we may as well use real data and simply describe the patterns observed in them* [italics mine]" [20].

The ***complication*** principle is a variant of the so-called KIDS principle [11]. It is of paramount importance for the cognitive perspectives on social simulation. KIDS stands for "Keep It Descriptive Stupid" and means that one should start with a descriptive model, which may be quite complex, to simplify just at the next step and if such a need is justified. KIDS approach overturns KISS practice. If KISS practice starts from a strong belief in the value of theoretical abstraction, KIDS starts with a descriptive adequacy, mostly related to the belief that adding empirical knowledge, also common-sense descriptions, is better than relying on theoretical abstractions. In the KIDS view, social simulation models should start with a wide range exploration of factors and components as candidates for being considered relevant features of the model. In the KISS view, social simulation models should start with theoretical abstraction and simplification. KIDS emphasises the *in the course of* and *ex-post* value of the model, even contemplating the involvement of stakeholders in the

[1] I owe this definition to Gérard Weisbuch.

modelling, as empirical knowledge bearers. KISS proudly emphasises the value of *ex-ante* theory used to build the model.

Here, we suggest that, if KISS is the best way to start with a model, because it emphasises the role of theory and the added value of simplification, there can be sound reasons to do a further complication step, when models and simulation results force analysts to come back to theoretical assumptions and to complicate the starting points. Such reasons can be of two types, theoretical and empirical ones. On the second type of reasons, there is a vibrant methodological debate on empirical calibration and validation of models that is currently ongoing in social simulation, which is impossible to summarise in a few lines here (i.e. [3] [34]). On the first type of reasons, the assumption here is that even if the KISS approach is in principle the best one to do social science, there can be relevant cases in which one is forced to add step by step further levels of complication/sophistication in the theoretical assumptions behind the model, because simplified starting points cannot satisfactorily explain the phenomenon the analyst is trying to explain. The complexification of a model can be therefore motivated by sound theoretical and empirical reasons. But, the point is that such a complexification does not mean necessarily a sophistication of the cognitive level of the model, but, let's say, a complication of the interaction structure or of other model's relevant components. The choice of the analytical level involved in the complexification depends on the analytical purpose of the analyst, for instance, on empirical data or evidence available or on pure theoretical reasons.

Whatever this methodological debate is of interest or not here, the conclusion is that, from a sociology viewpoint, most of sophisticated cognitive models often fall into the kitchen sink methodological practice**.** The complication of the cognitive level seems often a price to be paid in advance to theoretical richness or empirical realism of the model. We argue that, in general, KISS principle should be viewed as the preferable standard practice, because of the recognised advantage of simplification, abstraction, transparency of models for inter-subjective control and collaboration. The complication principle is suitable just in particular cases, when theoretical and/or empirical evidence and reasons justify the needs to enrich some of the model's components (cognitive level is one of them).

5 Concluding Remarks

In this paper, we have suggested some theoretical and methodological reasons to defend the idea that social simulations do not always need to be coupled to cognitive micro-foundations to have real explanatory power for the analysis of social phenomena, and that such a coupling should be carefully advocated just if there is a clear evidence that it can add explanatory power to the theory behind the model. On the contrary, the formalized model loses the unequivocal advantage achieved from simplification and abstraction, without a concrete increasing in its theoretical and explanatory sociological value. Therefore, the sophistication of the cognitive level of a model should not be viewed as a price to be paid to the realism of a sociological model, as a means to achieve a good isomorphism between model and reality, or as a way to honour a supposed appropriate interdisciplinary style of doing social science.

As we have argued, such a position seems to be confirmed on the basis of sound methodological and substantive reasons. On the methodological side, it is worth outlining that the historical experience in the usage of formalized models in the social sciences seems unequivocally to prove the pre-eminence of simplified, abstracted and usable models upon overcomplicated, specific and *ad hoc* ones [1] [13]. Although it is manifest that social simulation models, in particular agent-based models, allow the scholar to include more complications and details in the modelling of agents, interactions and environments, as regards to standard analytic models, this fact cannot constitute *per sé* a self-justified reason for the need to introduce complicated cognitive foundations in sociological analyses. In fact, if this complication would be introduced without a sound theoretical and explanatory reason, the consequence would be to stuff a sociological model with layers, details, pieces of *ad hoc* theories and assumptions, as well as with a set of complicated micro-macro-micro feedbacks, with the result of building a model more complex than the phenomenon it is called to explain. As a matter of fact, at a certain level of model's complication, the cost benefit ratio of formalizing a model, as well as its explanatory advantage as regards to standard qualitative methods and analyses, significantly decrease.

Thus, if this is true, there are good reasons to argue that social scientists, which are seriously interested in taking advantage of the usage of formalised theoretical models, should leave the analysis of complicated and sophisticated cognitive phenomena to cognitive scientists, at the same time leaving the challenge to explore the deep (and interesting!) link between cognitive and sociological analytical levels to other approaches and research traditions (i.e. ethnography, qualitative studies, and case-studies). Since the *explanandum* of a sociological model is not the individual behaviour but an aggregate phenomenon, and given the methodological and pragmatic constraints posed by the usage of formalized models, a good practice in the case of a sociological model would be to put effort and attention in the modelling and the understanding of interactions among agents and how individual actions aggregate at the macro level. However, it is of paramount importance that sociological models require the explicit generative reference to individual action and interaction to avoid the risk of reifying and evoking macro categorical concepts, such as "class", "culture", or "system", which are often used as explanatory categories in macro sociology in a tautological way. This is the reason why social simulations (agent-based models) are a powerful and irreplaceable means to formalize sociological generative models.

References

1. Axelrod R:, The Complexity of Cooperation. Agent-Based Models of Competition and Collaboration. Princeton, Princeton University Press (1997)
2. Axelrod R: Advancing the Art of Simulation in the Social Sciences. Complexity. 3 (1998) 16–22
3. Boero R. and Squazzoni F: Does Empirical Embeddedness Matter? Methodological Issues on Agent-Based Models for Analytical Social Science, in «Journal of Artificial Societies and Social Simulation. 8 4 (2005) <http:jasss.soc.surrey.ac.uk/8/4/6.html>

4. Braddon-Mitchell D. and Jackson F.: Philosophy of Mind and Cognition. Oxford, Blackwell (1996)
5. Bruch E. and Mare R. D.: Neighborhood Choice and Neighborhood Change. American Journal of Sociology. 112 (2006) 667-709
6. Coleman J.: Foundations of Social Theory. Cambridge, Massachusetts, The Belknap Press of Harvard University Press (1990)
7. Conte R. and Castelfranchi C.: Cognitive and Social Action, London. University College of London Press (1995)
8. Conte R., Edmonds B., Scott M., Sawyer R. K.: Sociology and Social Theory in Agent-Based Social Simulation: A Symposium. Computational and Mathematical Organization Theory. 7 (2001) 183-205
9. Conte R. and Paolucci M.: Reputation in Artificial Societies. Social Beliefs for Social Order. Kluwer Academic Publishers Dordrecht (2002)
10. de Marchi S.: Computational and Mathematical Modelling in the Social Science. Cambridge, Massachusetts, Cambridge University Press (2005)
11. Edmonds B. and Moss S.: From KISS to KIDS- An 'Anti-Simplistic' Modelling Approach. In Davidsson P. et al. (Eds.), Multi Agent Based Simulation 2004. Lecture Notes in Artificial Intelligence. Berlin Heidelberg, Springer Verlag. 3415 (2005) 130-144
12. Elster J.: Ulysses and the Sirens: Studies in Rationality and Irrationality. Cambridge, Cambridge University Press (1979)
13. Epstein J.: Generative Social Science. Studies in Agent-Based Computational Modeling. Princeton, Princeton University Press (2007)
14. Epstein J. and Axtell R. : Growing Artificial Societies. Social Science from the Bottom-Up. Cambridge, MA, MIT Press (1996)
15. Gigerenzer G. and Selten R. (Eds.): Bounded Rationality: The Adaptive Toolbox. London, The MIT Press (2001)
16. Gilbert N.: Varieties of Emergence. In Sallach D. (ed), Social Agents: Ecology, Exchange, and Evolution. Agent 2002 Conference. University of Chicago and Argonne National Laboratory (2002) 41-56
17. Gilbert N.: When Does Social Simulation Need Cognitive Models?. In R. Sun (Ed.), Cognition and multi-agent interaction: From cognitive modeling to social simulation. Cambridge, Cambridge University Press (2005)
18. Goldthorpe J. H.: On Sociology. Volume One: Critique and Program. Second Edition, Stanford, California, Stanford University Press (2007)
19. Granovetter M.: Threshold Models of Collective Behavior. American Journal of Sociology. 83 (1978) 481-510
20. Hedström P.: Dissecting the Social. On the Principles of Analytical Sociology. Cambridge, Cambridge University Press (2005)
21. Kunda Z.: Social Cognition. Making Sense of People. Cambridge, Massachusetts, The MIT Press (1999)
22. Latané B.: Pressures to Uniformity and the Evolution of Cultural Norms. Modeling Dynamics Social Impact. In Hulin C. E Ilgen D. (Eds.), Computational Modeling of Behavior in Organizations. The Third Scientific Discipline. American Psychological Association, Washington DC (2000) 189-215
23. Lindenberg S.: Homo Socio-Economicus: The Emergence of a Generalized Model of Man in the Social Sciences. Journal of Institutional and Theoretical Economics. 146 (1990) 727-748
24. Malle B. F.: How the Mind Explains Behavior. Folk Explanations, Meaning, and Social Interaction. Cambridge, Massachusetts, The MIT Press (2004)
25. March J.: A Primer on Decision Making. New York, Free Press (1994)
26. Nehaniv C. L. (Ed): Computation for Metaphors, Analogy and Agents. Berlin Heidelberg, Springer Verlag (1999)

27. Pylyshyn Z.: Computation and Cognition: Toward a Foundation for Cognitive Science. Cambridge, Massachusetts, The MIT Press (1984)
28. Sawyer R. K.: Social Emergence. Societies as Complex Systems. Cambridge, Cambridge University Press (2005)
29. Schelling T. C.: Micromotives and Macrobehavior. W.W. Norton & Co. Ltd, London (1978)
30. Simon H.: Models of Bounded Rationality. II Volumes. Cambridge, Massachusetts, The MIT Press (1982)
31. Stinchcombe A. L. : The Conditions of Fruitfulness of Theorizing about Mechanisms in Social Science. In Sorensen A. B. and Spilerman S. (Eds.), Social Theory and Social Policy: Essays in Honor of James S. Coleman. Westport, Praeger (1993)
32. Sun R.: Cognitive Science Meets Multi-Agent Systems: A Prolegomenon. Philosophical Psychology. 14 1 (2001) 5-28
33. Testfatsion L.: Agent-Based Computational Economics: Growing Economies From the Bottom-Up. Artificial Life, 8 1 (2002) 55-82
34. Windrum P. Fagiolo G. and Moneta A.: Empirical Validation of Agent-Based Models: Alternatives and Prospects. Journal of Artificial Societies and Social Simulation 10 2 (2007) <http:jasss.soc.surrey.a.uk/10/2/8.html>
35. Zerubavel E.: Social Mindscapes. An Invitation to Cognitive Sociology. Cambridge, Massachusetts, Harvard University Press (1998)

Balancing Internal and External Cognitive Connectivity in Young Enterprises to Explore and Exploit Inter-Organizational Relationships

Michael Beier[1]

[1] University of Cologne
Chair for Business Administration, Corporate Development and Organization
Albertus-Magnus-Platz
50923 Cologne, Germany
beier@wiso.uni-koeln.de

Abstract. This study analyzes the complex interplay of internal and external cognitive connectivity of founding teams and their inter-organizational relationships. For this purpose a comprehensive model has been formulated, formalized, and simulated. The results support young enterprises to increase their connectivity to potential external partners, while ensuring a necessary minimum of internal connectivity, to explore and exploit intended inter-organizational relationships.

Keywords: Entrepreneurship, Inter-Organizational Relationships, Social Cognition.

1 Introduction

Young enterprises often need to rely on external resources and capabilities [1]. To gain access to these complementary resources young enterprises have to cooperate with diverse external partners from domains different to their own [2]. Adequate levels of shared languages and systems of meaning are precondition for long-lasting and beneficial inter-organizational relationships [3]. One major success factor for young enterprises is a high degree of similarity of their identities and cognitive schemata within the founding team. This internal cognitive connectivity is beneficial in early stages of founding [4], but can evolve to a cognitive lock-in [5], limiting the connectivity to potential external partners [6] in further stages. Therefore, it is a crucial but difficult task for young enterprises to balance their internal and external cognitive connectivity to explore and exploit the potential benefits available via inter-organizational relationships.

So far, little is known about the factors and mechanisms influencing the interrelated development of internal and external cognitive connectivity of young enterprises and their inter-organizational relationships over time [7]. To generate deeper insights, a comprehensive model integrating recent theoretical and empirical results concerning attributes and processes on the micro-level of individual actors within founding teams and potential external partners has been formulated. In a further step this model has been formalized and simulated to analyze the linkages on the macro level between the input parameters and the outcomes in terms of cognitive connectivity, inter-organizational relationships, and network success

2 Model and Simulation

The interrelated processes of the development of cognitive connectivity of a focal founding team and of the development of its inter-organizational relationships constitute the core of the model.

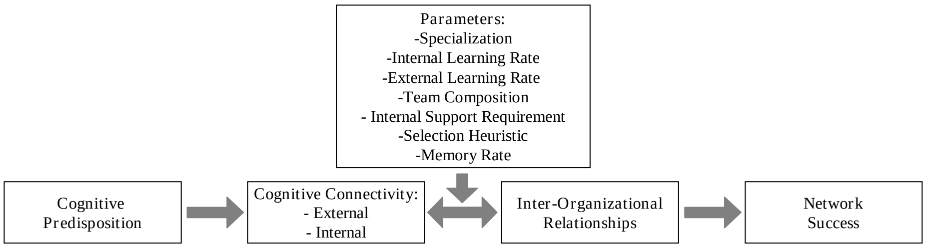


Fig. 1. Overview of the model

Cognitive predispositions of internal and external agents are conceptualized as sets of cognitive schemata [8] which are formalized as numerical vectors. Each schemata of an external agent represents its individual requirements on a relationship to the focal founding team. Each schemata of an internal agent represents its perception of the requirements of an external partner it interacts with. Each internal and external agent belongs to a class of expertise. Within these classes the similarity of cognitive schemata is higher than between the classes. *Cognitive connectivity* is defined by the deviation of two respective schemata, but can also be calculated as aggregated measure between sets of agents. Interacting with an unknown actor or in a new context, actors rely on their already acquired schemata as default for expectations on the new situation [9]. Internal agents add a copy of their schema with the lowest deviation to the interaction partner's schema to their repository as a perception pattern for the acquired knowledge. Experiences during partnerships lead to the adaptation of these schemata over time: For each dimension of each schema vector the value is modified, so that the deviation to the corresponding value of the partners' vector decreases per period according to the respective learning rate.

Parameters which are covered by the simulation model are *specialization* of relationship management within the founding team, the *internal and external learning rate* of internal agents, *team composition*, *internal support requirements*, the *selection heuristic* applied for external partner evaluation as well as the *memory rate* of focal agents, which defines how long not applied cognitive schemata are available within the repository. *Network success* is measured as the degree to what different classes of resources are available to a founding team by *inter-organizational relationships*.

Initially, the simulation generates by random generator a respective number of cognitive classes defined by sets of intervals concerning the dimensions of schema vectors as well as cognitive schemata within these classes for all simulated internal and external agents. In each period of simulation each internal agent chooses an external partner, adds the respective schemata for external and internal interactions to its repository, and adapts all schemata of ongoing partnerships. External agents evaluate the deviations to abort the respective relationship if it exceeds their tolerance. The schema deviations between internal agents influence their support within the founding team. By computing these processes of external and internal agents on the micro level the simulation allows for the analysis of network dynamics on the macro level in terms of external and internal cognitive connectivity, inter-organizational relationships and network success.

3 Results and Implications

The core contribution of this study is to develop a deeper understanding how the presented parameters influence the complex interplay of internal and external cognitive connectivity of founding teams and their intra- and inter-organizational relationships. Young enterprises are supported to increase their connectivity to potential external partners, while ensuring a necessary minimum of internal connectivity, to explore and exploit intended inter-organizational relationships.

References

1. Lechner, C., Dowling, M.: Firm Networks: External Relationships as Sources for the Growth and Competitiveness of Entrepreneurial Firms. In: Entrepreneurship and Regional Development, Vol. 15. (2003) 1-26
2. Hargadon, A., Sutton, R.I.: Technology Brokering and Innovation in a Product Development Firm. In: Administrative Science Quarterly, Vol. 42. (1997) 717-749
3. Nahapiet, J., Ghoshal, S.: Social Capital, Intellectual Capital, and the Organizational Advantage. In: Academy of Management Review, Vol. 23. (1998) 242-266
4. Oliver, A.L., Liebeskind, J.P.: Three Levels of Networking for Sourcing Intellectual Capital in Biotechnology: Implications for Studying Interorganizational Networks. In: International Studies of Management and Organization, Vol. 27. (1998) 76-103
5. Grabher, G.: The Weakness of Strong Ties. The Lock-In of Regional Development in the Ruhr Area. In: G. Graber (ed.): The Embedded Firm. On the Socioeconomics of Industrial Networks. Routledge, London, UK (1993) 255-277
6. Maurer, I., Ebers, M.: Dynamics of Social Capital and their Performance Implications: Lessons from Biotechnology Start-ups. In: Administrative Science Quarterly, Vol. 51. (2006) 262-292
7. Baum, J.A.C., Calabrese, T., Silverman, B.S.: Don't Go It Alone: Alliance Network Composition and Startups' Performance in Canadian Biotechnology. In: Strategic Management Journal, Vol. 21. (2000) 267-294
8. Rumelhart, D.E.: Schemata and the Cognitive System. In: Wyer Jr., R.S., Srull, T.K. (eds.): Handbook of Social Cognition, Vol. 1., Lawrence Erlbaum, Hillsdale, NJ (1984) 161-189
9. Neisser, U.: Kognitive Psychologie, Klett, Stuttgart (1974)

Session on Reputation and Communication

Un modèle multi-agents pour évaluer le rôle des réseaux dialogiques sur la dynamique de l'innovation en agriculture

Marie Houdart[1,4], Muriel Bonin[1], François Bousquet[2], Patrick Rio[3]

[1] Cirad, UMR Tetis, Station de Neufchâteau, Ste Marie, 97130 Capesterre-Belle-Eau, Guadeloupe
[2] CIRAD Green TA 60/15 Campus de Baillarguet 34398 Montpellier Cedex 5 France
[3] INRA, UMR LAMETA, 34060 Montpellier Cedex 2
[4] Cemagref, UMR Metafort, 24 avenue des Landais, BP 50085, 63172 Aubière
marie.houdart@cemagref.fr
{muriel.bonin, bousquet}@cirad.fr
rio@ensam.inra.fr

Résumé. A partir de résultats d'enquêtes en exploitation agricole et de référentiels théoriques sur réseaux et innovation, l'innovation est formalisée dans le modèle multi-agents comme un processus d'apprentissage individuel et collectif. Le modèle permet de simuler la distribution des systèmes de culture et la vitesse de mise en œuvre des systèmes de culture innovants en prenant en compte leur dynamique naturelle, l'émergence et le processus de construction de l'innovation, les réseaux de dialogue entre exploitants. Deux paramètres sont testés : le nombre et la force des liens entre agents. Il apparaît que plus les liens sont nombreux, plus les échanges d'information sont nombreux, quelle que soit la force des liens. En conséquence, la quantité de liens introduit une grande différence dans la mise en œuvre de systèmes de culture innovants. Ceci pose de nouvelles questions à la formalisation des réseaux de dialogue de même qu'à la construction de dispositifs de terrain.

Mots clefs : innovation, réseaux dialogiques, apprentissage, SMA, Guadeloupe

1. Introduction

Les fonctionnalités des SMA pour la gestion des ressources renouvelables en liaison avec les dynamiques sociales et spatiales sont nombreuses [1] [2]. Les SMA recouvrent notamment une utilité majeure quant à la mise en question des données d'enquêtes sociologiques (étape de formalisation, test des dynamiques, confrontation entre actions individuelles et collectives). C'est un outil pertinent pour confronter ces données d'enquête à la théorie [3].

En Guadeloupe, île française des Petites Antilles, une recherche a porté sur la capacité d'un groupe d'agriculteurs à mettre en œuvre des innovations agro-écologiques dans un contexte de crise environnementale et économique [4]. Une analyse basée sur l'étude des réseaux dialogiques entre agriculteurs a permis de mettre en valeur un lien entre structure du réseau et innovations stratégiques.

En conséquence, l'objet du modèle présenté dans ce papier est d'offrir une représentation multi-agents de la théorie de la construction de l'innovation pour évaluer la façon dont les réseaux de dialogue influent sur la dynamique de l'innovation.

L'action des réseaux de dialogue sur la dynamique de l'innovation est peu prise en compte dans les SMA : c'est l'approche que nous retenons et qui fait l'objet d'une première section. Dans la seconde section, nous présentons les bases conceptuelles du modèle, entre traduction d'une théorie sociologique et traduction des résultats de données d'enquête. La troisième section est consacrée à la description de la structure du modèle et aux résultats de la simulation.

2. Réseaux sociaux et innovation dans les SMA : l'intégration des réseaux de dialogue

D'un point de vue théorique, les courants qui analysent la relation entre réseaux sociaux et innovation s'inspirent fréquemment de l'évolutionnisme [5]. Leur prolongement a orienté des approches diffusionnistes de l'innovation [6] dans lesquelles la progression de l'innovation, en agriculture notamment, est perçue comme une épidémie : la pénétration de l'innovation se fait selon une courbe en S sur laquelle on distingue les innovants, les adoptants précoces, la majorité précoce, la majorité tardive et les retardataires [7]. D'une façon générale, la prise en compte de cette théorie diffusionniste revient à postuler que le rôle des agents est en partie stable du fait de la stabilité même de la structure du réseau. Cela amène à considérer que les données qui circulent ont intrinsèquement du sens et à peu près le même pour tout le monde.

Aussi plusieurs sociologues vont-ils à l'encontre de cette théorie "diffusionniste" [8] [9]. Par le développement de la théorie des réseaux d'acteurs, ils insistent sur la nécessité de prendre en compte le processus de construction de l'innovation. Cette théorie repose sur l'idée que l'agriculteur est un être social : il participe à des collectifs d'échanges techniques, des réseaux, des organisations ; il est influencé et influence ; les décisions qu'il prend dépendent pour partie des groupes sociaux auxquels il participe [10] [11] [12]. L'émergence de conceptions et de pratiques hybrides, résultant de l'interaction dialogique que les agriculteurs entretiennent avec d'autres, est prise en compte. La forme que peut prendre l'espace social a alors des effets sur la dynamique de changement pour la mise en œuvre des pratiques agricoles [13].

Les SMA constituent un outil pertinent pour simuler les effets des réseaux de dialogue sur la dynamique de changement pour la mise en œuvre de pratiques innovantes. Le principal apport des SMA réside en effet dans la possibilité de prendre en compte à la fois les dynamiques spatiales et les dynamiques sociales. L'approche générale consiste à modéliser des agents qui perçoivent un environnement et agissent sur celui-ci après avoir délibéré [14] [15] [16] [17].

Utilisant les principes de l'auto-organisation, de la confrontation entre analyse de comportements individuels et collectifs et celui de phénomènes émergents propres aux SMA, plusieurs auteurs modélisent la dynamique de l'innovation selon la théorie diffusionniste [18] [19] [20] [21]. En soi, les réseaux de dialogues ne sont pas pris en compte pour aller dans le sens de la théorie de la construction de l'innovation, malgré l'intégration de la notion de réseaux d'acteurs. Plusieurs modèles développent cependant la question de l'apprentissage dans la mise en œuvre de l'innovation, se rapprochant ainsi des concepts fondamentaux de la théorie de la construction de l'innovation [22] [23] [24]. D'autres notions sont afférentes alors à celle d'innovation dans ces modèles : celles d'apprentissage, par expériences individuelle et collective, et de confiance [25] [26].

3. Principes théoriques du modèle élaboré : traduction d'une théorie et traduction de données de terrain

Construit de façon à évaluer le rôle des réseaux de dialogue sur la dynamique de l'innovation, le modèle repose sur trois principes : une définition de l'innovation résultant des données d'enquête (auprès de 18 exploitants agricoles d'une petite zone rurale du Sud de la Guadeloupe) ; une conceptualisation de la théorie de la construction de l'innovation ; une conceptualisation des réseaux de dialogue en fonction des données d'enquête (Tab. 1).

Tableau 1. Conceptualisation des principes théoriques et des données d'enquête pour la modélisation.

Principes théoriques retenus	Analyse empirique (données d'enquête)	Conceptualisation pour la modélisation
Innovation	L'innovation spontanée apparaît face à une crise économique et prend la forme d'un changement d'orientation productive de l'exploitation	Innovation = changement de système de culture plus performant (marge brute)
Construction de l'innovation (vs diffusion)		Notion de Discours sur l'Innovation (ID)
Principe de l'innovation endogène, spontanée		Mise en œuvre de l'innovation par un agent pionnier (IDemergent)
Discours sur l'Innovation = objet hybride, muable, apprentissage collectif		Transmission de l'ID d'un agent à l'autre (IDtransmis)
Regard sur expérience individuelle, apprentissage individuel		Interprétation de l'ID reçu par un agent pionier ou non (IDtraduit)
Réseaux de dialogue	Forme des réseaux	Nombre de liens
	Modes de communication différents : fréquence des échanges et confiance/ influence	Force des liens

3.1. L’innovation : un changement de système de culture

En agriculture, l’innovation peut être considérée comme un changement technique ou stratégique, provoqué de manière spontanée (endogène) ou impulsée par des leviers d’action (exogène).

En Guadeloupe, les résultats du travail mené sur les réseaux de dialogue entre exploitants agricoles d’une même zone ont montré que les changements opérés de façon spontanée par les exploitants agricoles avaient pour origine le contexte de crise économique touchant notamment le secteur de la banane. La diversité des solutions adoptées en termes de systèmes de production (diversification, renforcement de la production bananière, agri-tourisme) de même que d’organisation spatiale des activités agricoles (échanges de parcelles), nous a amené à considérer l’innovation comme un changement d’orientation productive de l’exploitation agricole.

Dans le modèle, la mise en œuvre d’une innovation correspond alors à la mise en œuvre d’un système de culture dit innovant, plus performant. Cette performance est évaluée à l’échelle de l’exploitation par le calcul, à chaque pas de temps, de la marge brute. Cette dernière dépend uniquement, dans la version du modèle présentée ici, de l’âge du système de culture.

3.2. Le Discours sur l’Innovation (ID) : théorie de la construction de l’innovation

Ces changements de systèmes de culture se font à la suite d’échanges entre les agents du modèle, les exploitants. En reprenant les travaux de [8] et [9], nous posons dans notre modèle :

- qu'en même temps qu'une innovation spontanée est proposée, un Discours sur cette Innovation (ID) émerge (émergence),
- que cet ID vise à construire la position sociale de l’innovation, en termes de qualification-déqualification-requalification, ce qui implique une forme d'apprentissage collectif[1] (transmission),
- que cet ID se transforme au fur et à mesure que les agents non-pionniers adoptent l'Innovation, ce qui implique une forme d'apprentissage individuel (traduction).

On voit ainsi ce que notre analyse met en jeu. La mise en œuvre de l'innovation n'étant pas séparable du discours qui se fait à son propos (ID), l'importance relative des paramètres qui contrôlent cette construction du discours constitue l'enjeu majeur de notre interrogation.

---

1 On notera qu'un réseau non verbal (voir ce que le voisin fait sur son exploitation) peut parfaitement participer de l’ID. Dans le cas de changements stratégiques, la visibilité suffit à construire l’ID alors que la mise en œuvre d’une innovation technique nécessiterait des apprentissages effectifs ou une délégation de compétences (mimétisme ou recours aux techniciens).

3.3. Structure du réseau : une déconstruction-reconstruction des données d'enquête

Deux paramètres sont naturellement candidats à l'évaluation du contrôle du discours [27] :

- un paramètre traduisant la sociabilité de chaque agent,
- un paramètre traduisant la familiarité que les agents ont entre eux.

L'analyse des données recueillies en Guadeloupe a permis d'identifier plusieurs groupes de communication : des groupes de communauté d'orientation productive, au sein desquels les échanges sont fréquents et les stratégies similaires ; des groupes animés par des convictions politiques mais n'impliquant aucune communication sur les stratégies et les systèmes de culture en particulier mais au sein desquels chacun sait ce que l'autre fait ; des groupes présentant des stratégies similaires par le jeu d'influence d'un ou deux agents ; enfin, d'une façon générale, l'information passe beaucoup par l'observation des changements visibles opérés sur les parcelles des uns et des autres en raison de la proximité des exploitations, de leur situation sur un même axe de communication et de l'ouverture du paysage. Au regard de ces données d'enquête, les paramètres de contrôle du discours peuvent se mesurer en termes de nombre de liens d'une part, de force des liens d'autre part.

Un gradient de quantité de liens est représenté par trois types de groupes (Fig. 1) :

- un groupe unique au sein duquel tous les agents sont reliés aux autres.
- groupes reliés : tous les agents ne sont pas reliés les uns aux autres mais la connectivité est de 4. Chaque agent connaît 4 autres agents, qui eux-mêmes en connaissent 3 autres. Chaque groupe est relié aux autres par un agent.
- groupes non reliés : groupes dispersés, entre lesquels n'existe aucun lien mais au sein desquels tous les agents sont reliés les uns aux autres.

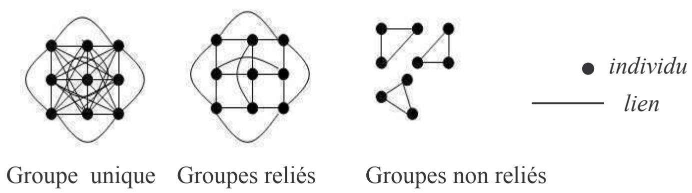


Fig. 1. Trois types de groupes traduisant un gradient de quantité de liens dans le modèle.

Les liens peuvent avoir différentes forces, en fonction de la confiance que les agents accordent aux propos des autres d'une part, de la fréquence de ces échanges d'autre part.

Dans ce modèle, nous retenons trois formes de communication :

- communication essentiellement basée sur l'observation, en raison de la proximité spatiale des parcelles, du passage répété d'un exploitant devant les parcelles d'un autre : la fréquence des échanges est faible, voire nulle. La force du lien est considérée comme faible.

- communication basée sur des échanges réguliers mais sur une confiance faible dans la mesure où les échanges ne portent pas directement sur la question de l'innovation. La force du lien est alors considérée comme moyenne.
- communication basée sur des échanges quotidiens, avec une confiance forte. La force du lien est considérée comme forte.

3.4. Hypothèse

En prenant en compte les deux paramètres retenus de la structure des réseaux que sont le nombre de liens et la force des liens, un point de vue pragmatique sur la question nous amène à émettre l'hypothèse que :

- pour un même nombre de liens, le nombre d'ID construits et transmis augmente avec la force des liens;
- si la force des liens est similaire pour toutes les relations existantes, le nombre d'ID construits et transmis augmente avec la quantité de liens.

Au final, en nous basant sur la théorie des réseaux, nous pouvons émettre l'hypothèse que plus il y a de liens entre individus et que ces liens sont forts, plus l'innovation (adaptée aux spécificités individuelles) est fréquente.

4. Le modèle

4.1. Objectifs

L'objectif du modèle est de simuler la distribution des systèmes de culture sur un territoire et la vitesse de mise en œuvre des systèmes de culture innovants en prenant en compte :

- la dynamique naturelle des systèmes de culture (deux types de Système de Culture : SdeCa ou SdeCb) : vieillissement et diminution de marge brute associée,
- l'émergence et le processus de construction de l'innovation (dynamique de l'ID)
- les réseaux de dialogue entre exploitants, à travers deux paramètres que sont le nombre et la force des liens.

Selon les principes théoriques implémentés dans le modèle, la dynamique de l'ID qui détermine en partie celle des systèmes de culture, est liée aux réseaux dialogiques par le nombre et la force des liens entre agents.

L'objectif assigné aux simulations est donc de mesurer comment ces deux paramètres caractéristiques des réseaux sociaux affectent *in fine* la mise en œuvre de systèmes de culture innovants, en tenant compte des conditions dans lesquelles le discours sur l'innovation transmet une innovation (émergence de l'ID, transmission, traduction).

4.2. Structure

La formalisation des bases conceptuelles pour l'intégration dans le modèle peut correspondre à la création d'agents, d'attributs ou encore de méthodes d'évolution de l'agent (Tab. 2).

Tableau 2. Formalisation des principes théoriques et des données d'enquête pour la modélisation.

Conceptualisation pour la modélisation	Formalisation dans le modèle
Innovation = changement de système de culture	Attribut de l'agent exploitant SdeCa et SdeCaInnovant SdeCb et SdeCbInnovant
Notion de Discours autour de l'Innovation (ID)	ID = agent
Emergence de l'ID	IDemergent = attribut de l'agent ID
ID transmis	IDtransmis = attribut de l'agent ID
ID traduit	IDtraduit = attribute de l'agent ID
Nombre de liens	Définition donnée dans les méthodes d'initialisation de l'agent exploitant
Force des liens	Définition donnée par la probabilité d'envoi des ID dans une des méthodes de l'agent exploitant

Le modèle repose sur une structure simple et générique (Fig. 2). Il est composé :

- d'entités spatiales élémentaires, les cellules, et d'entités agrégées, les exploitations, chacune appartenant à un exploitant.
- d'agents de type localisé communiquant : les exploitants.
- d'objets localisés créés au cours de la simulation : les ID.
- de messages qui permettent la visualisation de la communication entre agents : les MessageID.

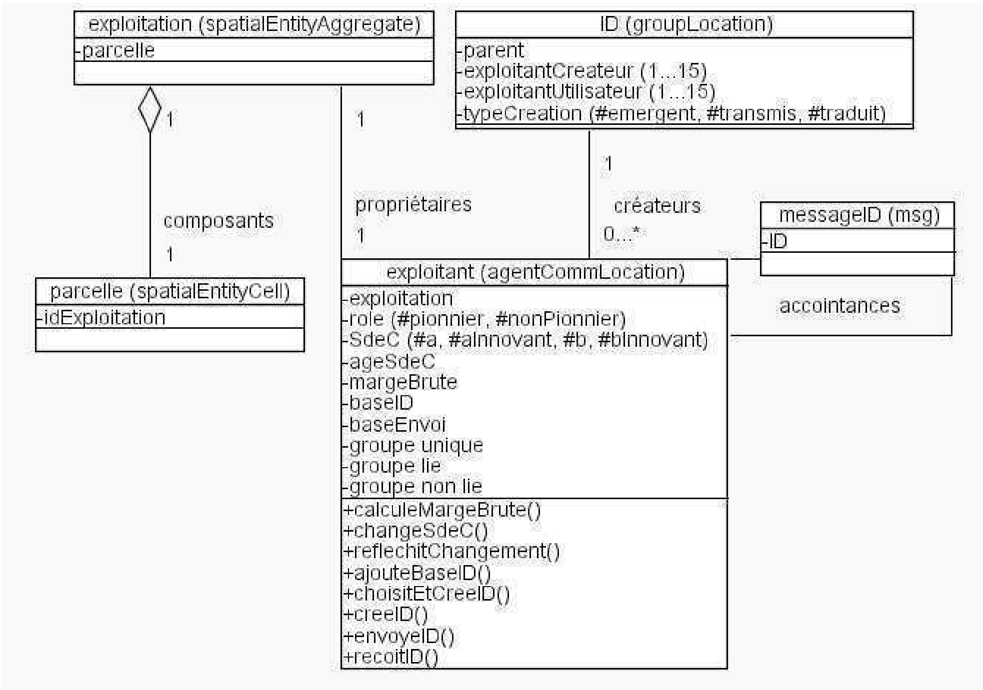


Fig. 2. Diagramme de classe du modèle.

Les agents "exploitant" sont au nombre de 15. Chacun possède une cellule de la grille spatiale, correspondant à une exploitation (Fig. 3).

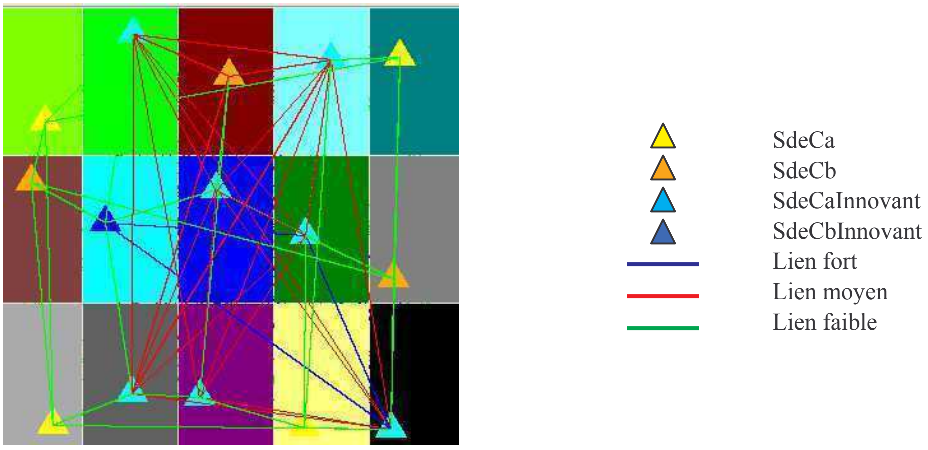


Fig. 3. Illustration de la grille spatiale au 8ème pas de temps d'une simulation incluant les différentes forces de liens.

Dans la mesure où une parcelle équivaut à une exploitation gérée par un exploitant, toutes les règles concernant le système de culture, de même que la marge brute, sont fixées au niveau de l'exploitant.

Deux systèmes de culture (SdeC) sont intégrés (SdeCa et SdeCb), de façon à ce que chacun ait une marge brute différente pour un âge similaire. Chaque système de culture peut évoluer vers un système de culture innovant : SdeCa devient SdeCaInnovant ; SdeCb devient SdeCbInnovant. Pour un âge similaire, la marge brute de ces systèmes innovants est plus élevée que celle du système initial.

Les SdeC sont répartis de façon aléatoire : chaque exploitant a 50% de probabilité de choisir SdeCa ou SdeCb.

Les SdeC ont un âge calculé aléatoirement à l'initialisation, compris entre 0 et 4.

L'exploitant a 1 chance sur 5 d'être pionnier.

Les liens entre exploitants sont définis au niveau des méthodes d'initialisation de l'exploitant. Ce sont ces liens qui déterminent la présence de groupe unique, groupes reliés ou groupes non reliés dans les différents scenarii.

La force des liens est traduite dans le modèle par une probabilité d'envoi des ID définie au niveau de l'exploitant : y sont fixées les probabilités d'envoi des ID. Chaque agent envoie la liste de ses ID à l'ensemble des membres de son réseau selon une probabilité d'occurrence qui peut varier. Plus les liens sont forts, plus la probabilité d'envoi des ID est élevée. Trois cas sont intégrés dans le modèle : le lien fort correspond à une probabilité d'envoi de 1, le lien moyen à une probabilité de 0.5, le lien faible à une probabilité de 0.2.

4.3. Déroulement de la simulation et démarche d'analyse

La simulation se déroule de la façon suivante (Fig. 4) : à chaque pas de temps, les SdeC vieillissent, la marge brute est ainsi calculée. Si elle passe en dessous d'un certain seuil, l'exploitant envisage de changer de stratégie sur l'ensemble de son exploitation. Si l'exploitant est pionnier, il change de système de culture, sinon, il observe sa boîte de ID. Si cette boîte est vide, il ne se passe rien ; si elle contient des ID, l'exploitant retient l'un d'eux au hasard, change de système de culture et la mise en œuvre de ce système de culture innovant implique la création d'un nouvel ID, de type IDtraduit. Le changement de système de culture provoque l'envoi de l'ID créé.

Deux paramètres clefs sont modifiés au cours des simulations de façon à évaluer leur importance respective et, au final, discuter autour des configurations "idéales" à donner aux réseaux de dialogue :

- le nombre de liens : dans un premier scénario, tous les agents appartiennent à un groupe unique ; dans un second scenario, les agents sont reliés par des groupes reliés ; dans le troisième scenario, 5 groupes indépendants relient chacun 3 agents.
- la force des liens est définie par la probabilité avec laquelle les ID sont envoyés. Trois scenarii sont envisagés : probabilité de 1, de 0.5 et de 0.2.

En croisant ces variations de paramètres, 9 scenarii sont testés.

Chaque simulation de scenario est lancée 30 fois sur 50 pas de temps. La moyenne des résultats de ces 30 itérations est enregistrée. Les sondes analysées sont :

- la quantité d'IDtransmis, d'IDemergent et d'IDtraduit

– le pourcentage de cellules en SdeCa, en SdeCaInnovant, en SdeCb et en SdeCbInnovant.

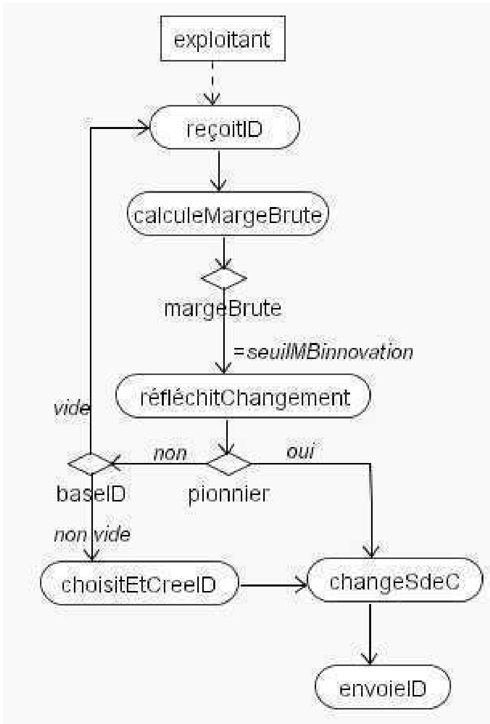
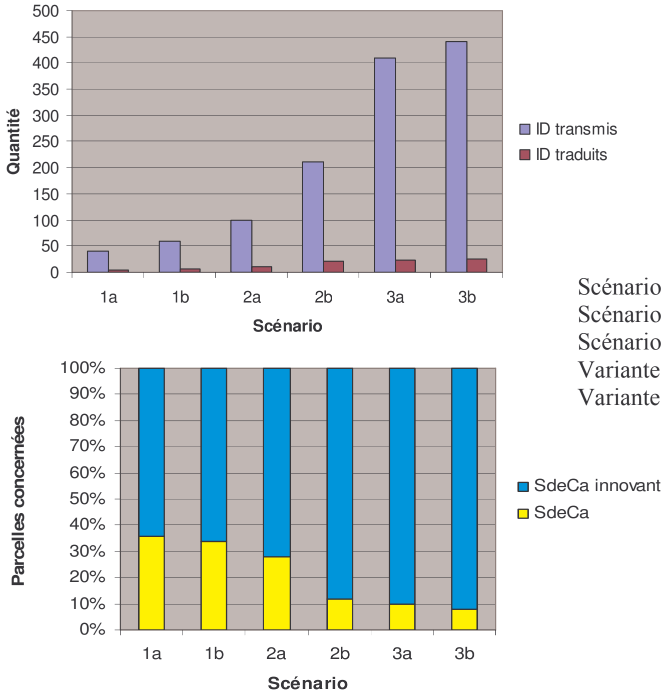


Fig. 4. Diagramme d'activité.

4.4. Résultats

D’une manière générale, plus les liens sont nombreux, plus la quantité d’ID est importante et ce quelle que soit la force des liens (Fig. 5).



Scénario 1 : groupe unique
Scénario 2 : groupes reliés
Scénario 3 : groupes non reliés
Variante a : liens faibles
Variante b : liens forts

Fig. 5. Synthèse des résultats en fin de simulations.

Concernant les IDemergents, leur émission dépendant en début de simulation de la part des agents "pionniers" par rapport aux agents non pionniers, l'évolution varie en fonction de ces paramètres statistiques.

Concernant les ID transmis, la quantité de liens joue un rôle majeur (Fig. 6). Si les quantités d'ID changent selon la force des liens, le rapport entre les différents cas de groupe est globalement toujours le même. Le nombre de liens introduit toujours une différence importante.

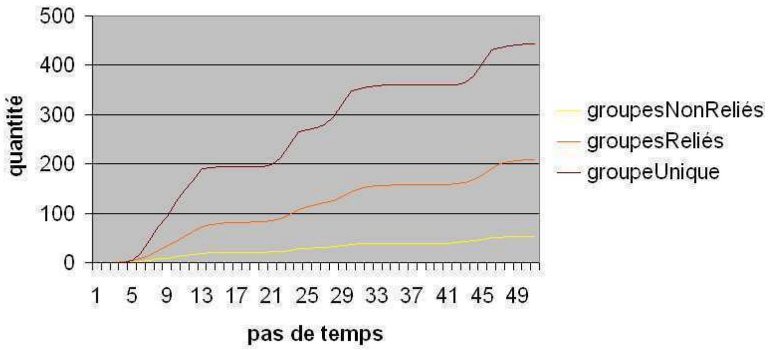


Fig. 6. Evolution comparée du nombre d'ID transmis selon la structure du réseau, dans le cas de liens forts.

Pour les ID provenant de l'expérience individuelle chez les non pionniers, le nombre de liens joue un rôle important, quelle que soit la force des liens (Fig. 7). Dans les trois cas de force de liens possible, il y a toujours un grand écart de production de ces ID selon le nombre de liens. En liaison avec une plus grande production d'ID transmis, plus il y a de liens entre exploitants, plus il y a d'ID provenant de l'expérience individuelle chez les non pionniers.

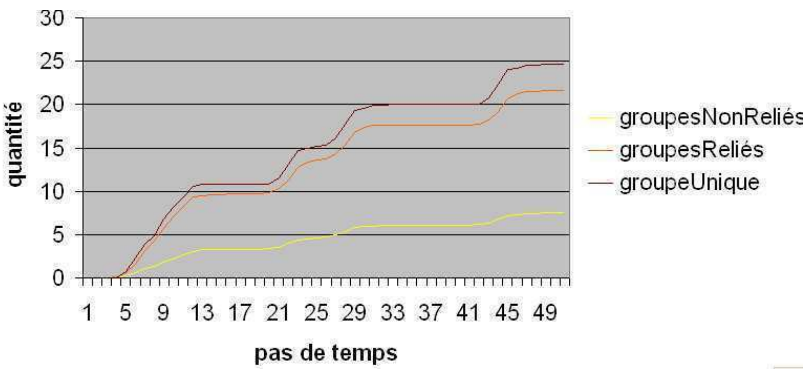


Fig. 7. Evolution comparée du nombre d'ID traduits selon la structure du réseau, dans le cas de liens forts.

Concernant la mise en œuvre de systèmes de culture innovants, il apparaît que dans le cas où les liens sont forts, la quantité de liens introduit une grande différence entre une structure à groupe unique et une structure en groupes reliés d'une part et une structure basée sur plusieurs groupes non reliés d'autre part (Fig.8). Dans le cas de liens faibles la quantité de liens introduit effectivement une différence très forte dans la dynamique des systèmes de culture.

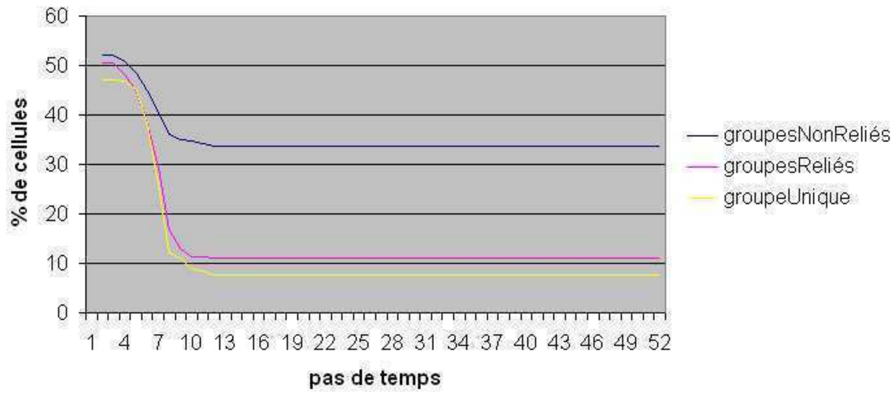


Fig. 8. Evolution de la distribution du SdeCa selon la structure du réseau (dans le cas de liens forts).

Quant à la force des liens, elle semble introduire, plus qu'une différence de quantités d'ID, une différence dans la vitesse de production de ces ID. Plus les liens sont forts, plus les ID transmis évoluent par palier marqué, alors que les liens faibles adoucissent les courbes (Fig. 9). En effet, les liens forts augmentent la vitesse de diffusion des ID : dans le cas de liens très forts, tout est envoyé d'un seul coup ce qui génère le palier, la rupture dans la dynamique. Plus la probabilité d'envoi des ID est faible, plus la courbe est lissée.

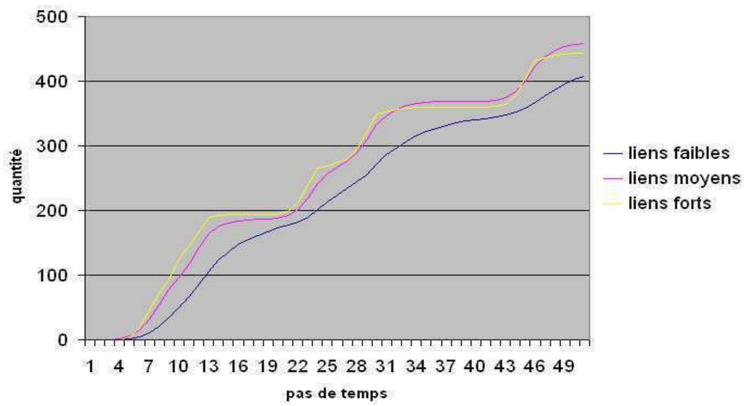


Fig. 9. Evolution comparée du nombre d'ID transmis selon la force des liens (dans le cas d'un groupe unique).

La force des liens introduit peu de différence dans la mise en œuvre des systèmes de culture innovants (Fig. 10). Dans le cas où les liens entre agents sont nombreux par exemple (scenario du groupe unique), la force des liens joue peu dans la dynamique des SdeC : que la probabilité soit de 0.2 ou de 1, les courbes d'évolution des différents SdeC sont très proches.

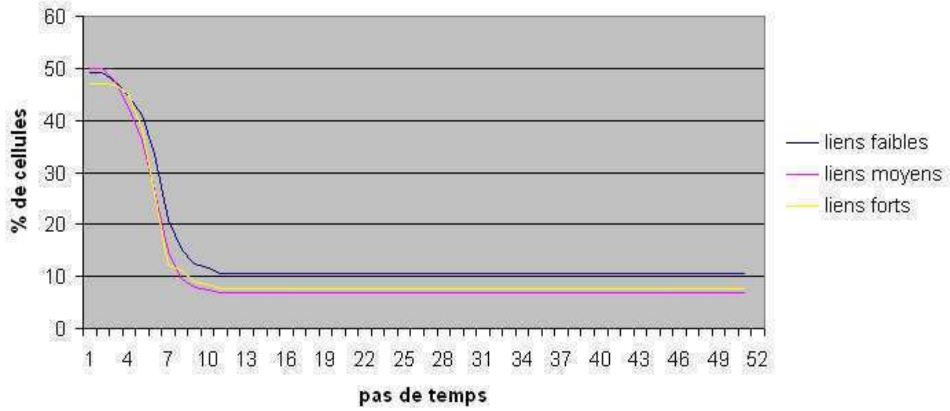


Fig. 10. Evolution de la distribution de SdeCa selon la force des liens (dans le cas d'un groupe unique).

5. Discussion/conclusion

A l'issue des simulations, il apparaît que le nombre de liens est plus important que la force des liens, à la fois pour des échanges autour des stratégies plus nombreux que pour la mise en œuvre plus fréquente de systèmes de culture innovants. Le modèle élaboré selon la théorie de la construction de l'innovation laisse penser que la force des liens ne joue pas dans la vitesse de diffusion ni dans la quantité d'innovation produite. Seul le nombre de liens revêt une importance majeure. Cela soulève de nombreuses questions théoriques. La simulation de la dynamique de l'ID permet d'analyser les processus qui sous-tendent la dynamique de l'innovation. Cependant, constater à quel point la force des liens semble être minimisée dans l'émergence et la transmission de l'ID et, par suite, dans la fréquence de mise en œuvre de l'innovation, amène à repenser la conceptualisation de cette théorie de construction de l'innovation, soit en revoyant les principes mêmes de formalisation, soit en confrontant avec d'autres scénarii, incluant notamment la dynamique même des réseaux.

La traduction de la théorie retenue mérite d'être approfondie. Selon la théorie des réseaux d'acteurs, tout réseau est toujours changeant : par exemple avec l'intégration d'un nouvel agent qui provoque de nouvelles alliances, ou la disparition d'un agent. C'est donc l'une des améliorations à donner à ce prototype : intégrer la dynamique des réseaux [28]. Selon l'approche structuraliste des réseaux de dialogues en liaison avec les processus de construction de l'innovation, la mise en œuvre d'une innovation par un acteur provoque un changement dans la structure du réseau. Dans le modèle, il faut intégrer la modification des réseaux de dialogue et/ou l'émergence de nouveaux réseaux suite à la mise en œuvre d'une innovation.

Concernant les questions de l'apprentissage, d'autres aspects méritent d'être améliorés. Pour ce qui est de la notion d'apprentissage individuel, on part du principe dans ce modèle que l'information est acquise avant toute prise de décision. Il n'y a pas de rôle actif imputé à l'agent pour aller chercher l'information lorsque se pose le problème. Quant à la notion d'apprentissage collectif, la notion de prise de risque n'est pas prise en compte (erreurs durant la phase d'apprentissage, risque que l'adoption ne soit pas intéressante si les conditions économiques/réglementaires changent). La possibilité d'observation de l'efficacité de l'innovation n'est pas intégrée.

Les résultats des simulations amènent également à réfléchir aux solutions éventuellement envisageables pour développer les capacités du groupe d'agriculteurs enquêtés en Guadeloupe à faire émerger des innovations. Ainsi, dans un cadre où les innovations sont "visibles" (ce qui est le cas d'un changement de système de culture) la mise en relations de tous les agents par observation serait aussi efficace que la constitution d'un groupe de travail fédérant l'ensemble des agriculteurs d'une zone pour la mise en œuvre généralisée de systèmes de culture innovants. Cela amène à réfléchir par exemple à la pertinence de la création de groupes de travail au regard des difficultés de mise en œuvre que cela peut susciter.

Outre ces aspects sur d'éventuels leviers d'actions, les résultats sur l'efficacité similaire d'une observation et d'un travail en commun sur des questions d'innovations stratégiques, renforce l'intérêt et l'importance d'une approche territoriale dans l'analyse des innovations agro-écologiques. Plus que l'élaboration d'hypothèses qu'il conviendrait de vérifier, le modèle aide alors à engager des démarches pluri-disciplinaires spécifiques : à l'analyse des réseaux dialogiques doit s'associer l'analyse géographique.

Enfin, le modèle peut être interfacé avec d'autres modèles, tels que des modèles biophysiques sur l'impact environnemental de certains systèmes de culture ou leur performance économique. Les spécificités spatiales des SMA constituent en ce sens un atout majeur. Pour cela, des règles de décision des exploitants propres à la zone doivent être développées et implémentées, règles à la fois économiques mais également spatiales. En ce sens, ce modèle d'aide à la construction de dispositif de terrain peut devenir un outil d'aide à la décision par la mise en regard de modèles biophysiques.

Références

1. Bousquet, F. & Le Page, C. (2004). Multi-agent simulations and ecosystem management: a review. Ecological Modelling 176(3-4): 313-332.
2. Gilbert, N. (2005). Agent-based social simulation: dealing with complexity. from http://www.complexityscience.org/NoE/ABSS-dealing%20with%20complexity-1-1.pdf
3. Janssen, M. A. & Ostrom, E. (2006). Empirically based, Agent-based models. Ecology and Society 11(2): 37.
4. Bonin, M., Cattan, P., Dorel, M., Malézieux, E. (2006). L'émergence d'innovations techniques face aux risques environnementaux. Le cas de la culture bananière en

Guadeloupe : entre solutions explorées par la recherche et évolution des pratiques. Agronomes et Innovations. C. J. (ed.). 123-135.
5. Schumpeter, J. A. (1934). The Theory of Economic Development. Cambride: Harvard University Press.
6. Hägerstrand, T. (1952). The propagation of innovations waves. Lund Studies in Geography série B: 4.
7. Rogers, E. M. (1995). Diffusion of innovations. New York, Macmillan Publishing Co.
8. Callon, M., Courtial, J. P., Turner, W.A. (1983). From translations to problematic networks: an introduction to co-word analysis. Social Science Information 22(2): 191-235.
9. Latour, B. (1989). La science en action. Paris, La Découverte
10. Bourdieu, P. (1980). Le sens pratique. Paris, Ed. de Minuit.
11. Darré, J.P., Guen (le), R., Lemery, R., (1989). Changement technique et structure professionnelle locale en agriculture. Economie rurale, n°192-193, pp.115-122.
12. Crozier, M. & Friedberg, E. (1977). L'acteur et le système : les contraintes de l'action collective. Paris, Seuil.
13. Compagnone, C. (2004). Agriculture raisonnée et dynamique de changement en viticulture bourguignone. Recherches sociologiques 3: 103-21.
14. Bousquet, F., Barreteau, O., Le Page, C., Mullon, C., Weber, J. (1999). An environnemental modelling approach: the use of multi-agent simulations. Advances in environmental modelling. Elsevier: 113-122.
15. Benoît, M., Chicoisne, G., Deffontaines, J.-P., Hervé, D., Lardon, S., Le Ber, F., Mullon, C., Papy, F., Souchère, V., Thinon, P., Tichit, M., Treuil, J.-P. (1999). Coordonner des choix de cultures sous contraintes environnementales : des jeux de rôles aux modèles multi-agents. Modèles et systèmes multi-agents pour la gestion de l'environnement et des territoires, Clermont-Ferrand.
16. Bonnefoy, J.-L., Bousquet, F., Rouchier, J. (2001). Modélisation d'une interaction individus, espace et société par les systèmes multi-agents : pâture en forêt virtuelle. L'espace géographique 1: 13-25.
17. Barreteau, O., Cernesson, F., Garin, P., Dumontier, A., Abrami, G. (2003). Agent-based facilitation of water allocation: case study in the Drome river valley. Group Decision and Negotiation 12: 441-461.
18. Steyer, A. & Zimmermann, J.-B. (2004). Influence sociale et diffusion de l'innovation. Mathematics and Social Sciences 168(4): 43-57.
19. Daudé, E. (2004). Apports de la simulation multi-agents à l'étude des processus de diffusion. Cybergeo : Revue européenne de géographie n° 255: 15p.
20. Valente, T. W. (1996). Social network thresholds in the diffusion of innovations. Social Networks 18(1): 69-89.
21. Amblard, F. & Ferrand, N. (1999). Modélisation multi-agents de l'évolution de réseaux sociaux. Modèles et Systèmes Multi-Agents pour la Gestion de l'Environnement et des Territoires. Cemagref. Clermont-Ferrand: 153-168.
22. Barreteau, O. & Bousquet, F. (1999). Jeux de rôle et validation de systèmes multi-agents. Actes des 7ème journées francophones d'Intelligence Artificielle et systèmes multi-agents (JFIADSMA'99), Paris.
23. Rouchier, J., F. Bousquet, Requier-Desjardins, M., Antona, M. (2001). A multi-agent model for describing transhumance in North Cameroon: Comparison of different rationality to develop a routine. Journal of Economic Dynamics and Control 25: 527-559.
24. Abrami, G. (2004). Niveaux d'organisation dans la modélisation multi-agents pour la gestion des ressources renouvelables. Application à la mise en œuvre de gestion collective de l'eau d'irrigation dans la basse-vallée de la Drôme. Cemagref/ENGREF. Montpellier: 305 p. + annexes.
25. Dupouët, O., M. Yildizoglu, Cohendet, P. (2003). Morphogenèse de communautés de pratique. Revue d'économie industrielle n°103 2eme et 3eme trimestre 2003: 91-110

26. Nunes, L. and E. Oliveira (2004). Learning from multiple sources. AAMAS'04, New York.
27. Darré, J.P. (1986). Le rôle des réseaux de dialogue entre agriculteurs. Comment les façons de faire et de penser se transforment : l'étude des réseaux de dialogue. Agriscope 7: 143-151.
28. Bala, V. & Goyal, G. (2000). A noncooperative model of network formation. Econometrica 68(5): 1181-1229.

Modelling bilingualism in language competition: the effects of complex social structure

Xavier Castelló, Riitta Toivonen, Víctor M. Eguíluz, and Maxi San Miguel

IFISC (CSIC-UIB) Universitat Illes Balears, E-07122 Palma de Mallorca, Spain

Abstract. In the general context of dynamics of social consensus, we study an agent based model for the competition between two socially equivalent languages, addressing the role of bilingualism and social structure. In a regular network, we study the formation of linguistic domains and their interaction across the boundaries. We analyze also the dynamics on a small world network and on a network with community structure. In all cases, a final scenario of dominance of one language and extinction of the other is obtained. In comparison with the regular network, smaller times for extinction are found in the small world network. In the network with communities instead, the average time for extinction is not representative of the dynamics and metastable states are observed at all time scales.

1 Introduction

Language competition belongs to the general class of processes that can be modelled by the interaction of heterogeneous agents as an example of collective phenomena in problems of social consensus [1]. In this respect, a specific feature of language dynamics is that agents can share two of the social options that are chosen by the agents in the consensus dynamics. In the present work, these are the bilingual agents, that is, agents that use both language A and B, who have been claimed to play a relevant role in the evolution of multilingual societies [2].

In this work we are interested in the emergent phenomena appearing as a result of a self-organized dynamics in the case of two equally prestigious competing languages. With the aim of elucidating possible mechanisms that could stabilize the coexistence of these languages, we wish to discuss the role of bilingual individuals and social structure in the process of language competition. To this end, and along the lines of the original proposal by Minett and Wang [2], we study an agent based model that incorporates bilingual agents on different networks: a regular lattice, a small world network, and a social type network with community structure [3]. We compare the results obtained with the agent-based version [4] of Abrams-Strogatz two-state model [5], where bilingualism was not taken into account.

2 The Bilinguals Model

We consider a model of two socially equivalent (i.e. equally prestigious) competing languages in which an agent i sits in a node within a network of N

individuals and has k_i neighbours. It can be in three possible states: A, agent using [1] language A; B, agent using language B; and AB, bilingual agent using both languages, A and B.

The state of an agent evolves according to the following rules: at each iteration we first choose one agent i at random, and, then, we compute the local densities of language users of each linguistic community in the neighbourhood of agent i: σ_i^l (l=A, B, AB; i=$1, N$; $\sigma_i^A + \sigma_i^B + \sigma_i^{AB} = 1$). The agent i changes its state of language use according to the following transition probabilities [2]:

$$p_{i,A\to AB} = \frac{1}{2}\sigma_i^B \ , \qquad p_{i,B\to AB} = \frac{1}{2}\sigma_i^A \tag{1}$$

$$p_{i,AB\to B} = \frac{1}{2}(1-\sigma_i^A) \ , \qquad p_{i,AB\to A} = \frac{1}{2}(1-\sigma_i^B). \tag{2}$$

Equation (1) gives the probabilities for an agent to move away from a monolingual community to the bilingual community AB. They are proportional to the density of monolingual speakers of the other language in its neighbourhood. On the other hand, equation (2) gives the probabilities for an agent to move from the bilingual community towards one of the monolingual communities. Such probabilities are proportional to the density of speakers of the adopting language including bilinguals ($1-\sigma_i^l = \sigma_i^j + \sigma_i^{AB}$, l, j=A, B; $l \neq j$). It is important to note that a change from being monolingual A to monolingual B or vice versa always implies an intermediate step through the bilingual community. The transition probabilities (1) and (2) are fully symmetric under the exchange of A and B, which is consistent with the fact that both languages are socially equivalent in terms of prestige.

We recover the agent-based version of Abrams-Strogatz two-state model when bilinguals are not present [4]. In this model, an agent essentially imitates language use of a randomly chosen neighbour.

For a quantitative description of the emergence and dynamics of linguistic spatial domains we use the ensemble average interface density $\langle\rho\rangle$ as an order parameter. This is defined as the density of links joining nodes in the network which are in different states [1]. The ensemble average, indicated as $\langle\cdot\rangle$, denotes average over realizations of the stochastic dynamics starting from different random distributions of initial conditions. During the time evolution, the decrease of ρ from its initial value describes the ordering dynamics, where linguistic spatial domains, in which agents are in the same state, grow in time. The minimum value $\rho = 0$ corresponds to a stationary configuration in which all the agents belong to the same linguistic community.

3 Results

a)*Regular and small world networks*

[1] Note that we always refer to language use rather than competence.

[2] Non-equivalent languages were considered in the original version of the model [2]. The prefactor 1/2 corresponds to the special case of equivalence between A and B.

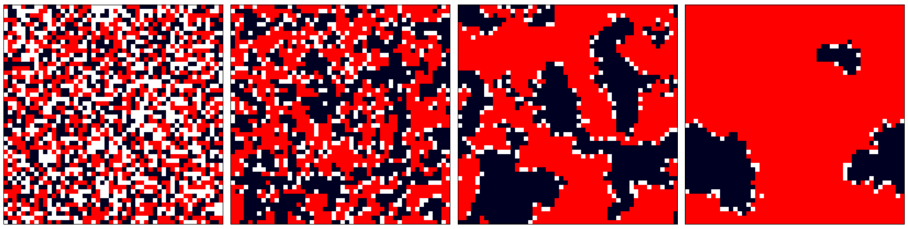


Fig. 1. Random initial conditions: snapshots of a typical simulation of the dynamics in a regular lattice of 2500 individuals. t=0, 2, 20, 200 from left to right. Red: monolinguals A, black: monolinguals B, white: bilinguals.

The bilinguals model has been extensively studied in two-dimensional lattices, and small world networks [6]. In two-dimensional lattices, and starting from a randomly distributed state of the agents, spatial domains of each monolingual community are formed and grow in size (Fig 3). This is known in the physics literature as *coarsening*. Meanwhile, domains of bilingual agents are never formed. Instead, bilingual agents place themselves in a narrow band between monolingual domains (Fig 3). Finally a finite size fluctuation drives the system to an absorbing state, where all the agents become monolingual, while the other monolingual community together with the bilingual agents face extinction. Average interface density $\langle\rho\rangle$ decays as a power law $\langle\rho\rangle \sim t^{-\gamma}$, $\gamma \simeq 0.45$ [6]. This indicates that the growth law found for the bilinguals model is compatible with the well known exponent 0.5 associated with domain growth driven by mean curvature and surface tension reduction observed in SFKI (spin flip kinetic Ising model) [7]. The characteristic time to reach an absorbing state τ scales with system size as $\tau \sim N^{1.8}$. A very different behaviour is found for the agent based Abrams-Strogatz model, where bilingual agents are not present: coarsening is slower ($\langle\rho\rangle \sim \ln t$) and driven by interfacial noise.

In Watts-Strogatz small world networks $\tau \sim \ln N$ [6]. While for the agent based Abrams-Strogatz model the long range connections inhibit coarsening by producing long-lived metastable states, in the bilinguals model adding long range connections to the two-dimensional lattice slows down coarsening, but domains keep growing in size. In addition, they speed up the decay to an absorbing state due to finite size fluctuations [6].

b) *Social type network with community structure*

A combination of random attachment with search for new contacts in the neighbourhood has proved fruitful in generating cohesive structures (algorithm in [3]). We choose this model, which produces well-known features of social networks, such as assortativity, broad degree distributions, and community structure.

The most important result regarding this topology, is the behaviour of the characteristic time to reach an absorbing state. To this end, we analyze the fraction $f(t)$ of runs still alive at any time t, i.e. the fraction of runs which have not reached the absorbing state. We average over different realizations of the network, and several runs in each. For the agent based Abrams-Strogatz model, the fraction of alive runs decreases exponentially. Results are more interesting for the bilinguals model: $f(t)$ appears to have power law behaviour $f(t) \sim t^{-\alpha}$, $\alpha \approx 1.3$. Since the exponent $\alpha < 2$, the average decay time for the bilinguals model does not give a characteristic time scale, but alive realizations which have not reached the absorbing state are found at any time scale. Analyzing these simulations, we observe configurations of language domains in which one of the monolingual communities has nearly taken over the whole system, while the domain of the minority language eventually only resides in small communities which are loosely connected to the rest of the network. Bilingual agents remain at the interfaces between domains. These facts indicate correlation between community structure and linguistic domains. As the bilinguals model is effectively a majority rule (SFKI), these configurations lead to metastable states.

4 Conclusion

We have analyzed the bilinguals model (in comparison to the agent-based version of Abrams-Strogatz model) in different topologies. Although the final state of the system is always an homogeneous state where one of the languages faces extinction, the transient towards this final state depends crucially on the network structure. Within the limitations and assumptions of the model, the study of the dynamics in the social type network with communities shows that there exist metastable states at all time scales; indicating that in presence of bilingual individuals, minority languages might survive for very long periods when the social network displays community structure.

References

1. San Miguel, M., Eguíluz, V., Toral, R., Klemm, K.: Binary and multivariate stochastic models of consensus formation. Computer in Science and Engineering **7** (2005) 67–73
2. Wang, W.S.Y., Minett, J.W.: The invasion of language: emergence, change and death. Trends in Ecology and Evolution **20** (2005) 263–269
3. Toivonen, R., Onnela, J., Saramäki, J., Hyvönen, J., Kertész, J., Kaski, K. Physica A **371(2)** (2006)
4. Stauffer, D., Castello, X., Eguiluz, V.M., San Miguel, M.: Microscopic abrams-strogatz model of language competition. Physica A **374** (2007) 835–842
5. Abrams, D.M., Strogatz, S.H.: Modelling the dynamics of language death. Nature **424** (2003) 900
6. Castelló, X., Eguíluz, V.M., San Miguel, M.: Ordering dynamics with two non-excluding options: Bilingualism in language competition. New Journal of Physics **8** (2006) 308–322
7. Gunton, J.D., San Miguel, M., Sahni, P.: The dynamics of first order phase transitions. In: Phase Transitions and Critical Phenomena. Volume 8. Academic Press, London (1983) 269–446

Vulnerability of Reputation Management System due to Tolerant Evaluation

Hitoshi Yamamoto[1], Isamu Okada[2] and Toshizumi Ohta[3]

[1] Faculty of Business Administration, Rissho University, Shinagawa-ku, Tokyo 141-8602, Japan
hitoshi@ris.ac.jp
[2] Faculty of Business Administration, Soka University, 1-236 Tangi, Hachioji City, Tokyo 192-8577, Japan.
okada@soka.ac.jp
[3] The Graduate School of Information Systems, University of Electro-Communications, Choufu-shi, Tokyo 182-8585, Japan
ohta@is.uec.ac.jp

Abstract. Reputation management system is effective for promotion of cooperative behaviors in online transaction. However, it may be hard for a trader to evaluate his/her partner accurately on the transaction because of his/her fear for the retaliatory evaluation or of the expectation to positive evaluation from his/her partner. Traders have not only the strategy for behaviors but also that for evaluations. In order to analyze the influences of inaccurate evaluation in the reputation management system, we model an online consumer-to-consumer market. The results show that the dominant strategy among traders for evaluation is cooperative while that for behavior is non-cooperative. The domination of non-cooperative behaviors is derived from that the tolerant evaluation cannot eliminate those behaviors. We conclude that reciprocal evaluations by unmalicious participants cause an inflation of reputation and may prevent the reputation management system from functioning properly.

1 Introduction

The importance of the development of online transactions and the reputation management systems is widely insisted (Kollock, 1999), (Dellarocas, 2000). Especially, an online market among consumers in which many consumers participate uncertainly as sellers and buyers usually demands a reputation management system due to choose trustable transaction opponents essentially. A reputation management system (RMS) is a system of which participants in an online market can evaluate transaction opponents and share their information each other. One of the goods examples of successive markets using such a RMS is eBay. The eBay has a RMS in which participants evaluate opponents qualities.

Evaluations in the RMS are almost positive. The case study of the eBay online trading by Resnick and Zeckhauser (2001) report that a ratio of positive opinions in the system exceeds 99%. Baron (2001), by the way, points that a dissatisfied considering giving negative feedback might fear that the other trader would retaliate negative feedback. Does the 99% of traders satisfy their transactions really? Or do the evaluations have some kinds of biases? Miyata and Ikeda (2001) survey for users of the internet auctions in Japan. They found the following opinions in the survey. “When I write an opponent negative evaluation, I fear that the opponent may write a negative evaluation about me later”. “When a user evaluates her partner positively, she expects reciprocal evaluation from the partner”. These opinions suggest a possibility of that even an unmalicious participant evaluates inaccurately because of a fear of a retaliative evaluation for one’s own evaluation and an expectation of a high evaluation from one’s partner. In a summary, the RMS using interactive evaluations essentially distinguishes non-cooperative participants, however in practice, over-positive evaluations emerge and the participants may not judge good partners. Why does the RMS in online markets bring the over-positive evaluations? What kind of influences does the bias have for the soundness of the market?

Generally, it is natural that a person returns a positive or negative evaluation for the other person who evaluates one positively or negatively. Such attitude is called as reciprocal. Thompson (1967) investigates that people behave for a partner’s action reciprocally when they interact. Does the RMS in online markets functionalize effectively if person’s attitude is reciprocal when one wants to evaluate the other? The purpose of the RMS is to promote cooperative behaviors among participants due to evaluate their behaviors and chare the information. However, this discussion suggests that the participants have not only strategies of behaviors but also those of evaluations because they have the fears and the expectations for the reactions from their opponents when they evaluate them.

We develop a model of reciprocal evaluations. The model describes transactions and interact evaluations in an online market based on a hypothesis of which the evaluations are reciprocal. The goal of this research is obtain how evaluating participants’ partners reciprocally has an influence on the soundness of the online markets?

In Section 2, we explain a relationship between some studies treating evolution of cooperation using the prisoners’ dilemma games and a RMS of the online markets. We construct a simulation model in Section 3. Section 4 shows experiments and results of the simulation. Section 5 discusses mechanisms and their background of which the inflations of reputation emerge and the RMS cannot work effectively by which participants in the markets evaluate their opponents reciprocally. In Section 6, we conclude.

2 Prisoners’ Dilemma Game and Reputation Management System

In this Section, we discuss that an online transaction market is able to be described a model with prisoners’ dilemma games and show basic ideas for the model of the market with the reputation management system.

2.1. Online Transactions and Prisoners' Dilemma Games

A player who participates in an online transaction always has an incentive to cheat others (i.e., to defect) due to anonymity and the ease of entry and exit from transactions. For example, a buyer may demands to discount by irrational complaints though a seller provides an item with high quality. In contrary, a seller can sell an item off at a higher price with low quality and denies complaints from the buyer. Non-cooperative actions, in addition, are exemplified the delay for responses, the cancellation of transactions, the imposition of shipping and handling, and so on. These actions can be described as a prisoners' dilemma game. In the game, there are two players, whom we can refer to as player-1 and player-2, and they cannot communicate with each other directly because they are in solitary confinement in a prison. Each player has two strategies, namely cooperation (C) and defection (D). We can consider these strategies within a payoff matrix, as shown in Table 1.

Table 1: Payoff matrix for prisoner's dilemma

		Action of player-2	
		C	D
Action of player-1	C	R_1, R_2	S_1, T_2
	D	T_1, S_2	P_1, P_2

The necessary conditions for prisoner's dilemma are the following two inequalities (1):

$$\begin{cases} T_i > R_i > P_i > S_i & , i = 1,2 \\ 2R_i > T_i + S_i \end{cases} \tag{1}$$

In the prisoner's dilemma of an online transaction, a seller can take two possible actions: cooperation, i.e., sending goods with high quality, and defection, that with low quality. Likewise, a buyer can also cooperate or defect, i.e. no complaints about goods or claims to discount by irrational complaints. Under the situation, if a trader acts cooperative, one keeps being exploited with non-cooperative traders. To keep the online market safely, a manager of the market needs a system which protects cooperative actors and exploit non-cooperative actors.

2.2. Model of Reputation Management System and Evolution of Cooperation

Evolutionary dynamics of cooperative behaviors is being discussed by many papers with various interests and approaches e.g. Axelrod (1984). Cohen et.al. (2001) is a systematical and exhaustive paper about the evolution of cooperation in the iterated prisoners' dilemma game. The main result of Cohen is that "context preservation" is an important trigger for promotion of cooperative behaviors. The context preservation is derived from a situation that individuals located in a two-dimensional lattice torus space have fixed von Neumann neighborhood as opponents of games.

The Cohen's context preservation model restricts that the information of which the individuals can refer to for their next actions is just opponents' actions before one period. In other words, they focus on the evolution of cooperation under a situation without reputation management systems because the individuals cannot use indirect information concerning opponents' actions. Besides, in the Cohen model, the individual imitates a strategy of whoever has the highest payoff of her neighbors when her strategy evolves. This rule is also premised on that the individuals' actions and strategies are visible from the others.

Yamamoto et.al. (2004a,b) develop a model of what a new entry person enters an online market by imitating a strategy of whoever has the highest payoff. The model assumes that the participant strategies do not change and a new participant enters the market as alternated. The model follows a situation of which winners of the market invite their friends. The individual who has higher payoff invite her acquaintances to the market and they can imitate her strategy.

Generally speaking, the higher the participant gains payoff, the more she can survive in markets. The evolution of individual strategy is studied with two approaches: one using the genetic algorithm (GA) such as Yao and Darwen (1999), Ashlock et.al. (1995) and Nowak and Sigmund (1998), and the other imitating the neighbor's strategy such as the Cohen model or the Yamamoto model. In this paper, we develop a model as not imaging a particular situation such as Yamamoto et. al. (2004a,b) but describing a general adaptive process of which the fittest strategy for the market is dominant. Our purpose is to discuss on the essential traits of the reputation systems.

It is natural that the other participants' payoffs and strategies are invisible on the markets because they are inner status of persons. In turn, we cannot watch the others' penny in purse and inside their heads. Therefore, we assume the neighbors' payoffs and strategies are hard to see and the adaptation process of individual strategies follows the GA as a method of which the individuals who fit in the market can survive evolutionarily.

We discuss on the influence of which tolerant reputation gives on a RMS by comparing three models with the model of reciprocal evaluations introduced above. They are a model of correct evaluations, that of invisible payoffs, and that of visible payoffs with basic prisoners’ dilemma.

3 Model of Reciprocal Evaluations

In order to design a simulation model of an online marketplace, we discuss on transactions within the framework of the prisoner’s dilemma, in which the players are represented by buyers and sellers. In this paper, we treat of a market such that sellers and buyers deal with each others in the bidding. The sellers and buyers are actors who have their strategies and deal with the others automatically. We describe the sellers and buyers as agents in our model.

3.1Agents

Agents are two types: Sellers and Buyers. An agent makes two types of decisions. One decision making is about one’s behavior of the trading, whether one deals with one's partner cooperatively or non-cooperatively (defectively). Another one is about one's evaluation of the partner, a positive evaluation or a negative one.

An agent, x, has an action strategy $Act^x = (a_i^x, a_p^x, a_q^x) \in [0,1]^3$ and an evaluation strategy $Eva^x = (e_i^x, e_p^x, e_q^x) \in [0,1]^3$. The probability of cooperating on the first dealing is a_i^x. The probability of cooperating if the other agent cooperated on the previous dealing is a_p^x, and a_q^x is the probability of cooperating if the other agent defected on the previous dealing. On the other hand, the probability of positively evaluating (appreciation) on the first dealing is e_i^x. The probability of appreciation if the other agent cooperated on the dealing this time is e_p^x, and e_q^x is the probability of appreciation if the other agent defected on the dealing this time. For a simplicity a la Cohen, the agent has one step memory for strategies, and let $a_i^x = a_p^x$ and $e_i^x = e_p^x$. Using this notation, we can define the basic strategies of the prisoner's dilemma. If an agent has (a_p^x, a_q^x)=(1,1), she is a perfect cooperator, so we call her strategy All-C. When (a_p^x, a_q^x)=(1,0), the strategy is called tit-for-tat (TFT), and (0,0) is the strategy of a perfect defector, or All-D.

Each agent can have one single role, either buyer or seller. All the buyers put in position on a two-dimensional lattice torus Buyers' space and all the sellers are also on the same sellers' space. Shown in Figure 1, a buyer chooses one seller from her neighbors in the sellers' space. Same as a seller's decision. So, possible trading partners of a seller (a buyer) are defined as the fixed four buyers (sellers) located on von Neumann neighborhood in the buyer's (seller's) space.

An agent evaluates a trading partner on her evaluation strategy for every dealing action. A reputation of an agent is defined as summing up evaluations for one. Using this reputation, agents choose their partners to deal it.

A dealing process on a market consists of four phases: bidding, trading, evaluation, and learning of strategies.

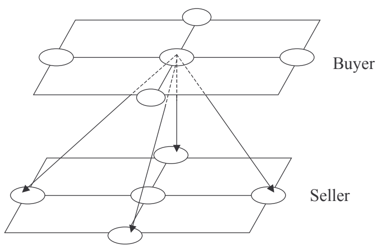


Fig. 1. Neighborhoods of tradable sellers for a buyer

Bidding: A buyer bids for a seller who has the highest reputation in one's neighbor. After the biddings of all the buyers, the sellers contract trading actions with particular buyers who have the highest reputations in the bidders for them.

Trading: For all contracts, a buyer, *b*, and a seller, *s*, deal with each other following their action strategies, Act^b and Act^s. Based on the strategies, *b* (*s*) selects one's action, of two alternatives, cooperation (C) and defect (D).

Evaluation: After trading, b (s) evaluates s (b) on one's evaluation strategy, Eva^b (Eva^s). The order of evaluation is random for every trading, so the probabilities of what b and s are the first evaluator is 0.5, respectively. When the first evaluator, x, evaluates the second evaluator, y, one's evaluation, E(y<-x) in {P,N}, is decided by which if Act^y=C then P (positive) with the probability e_p^x, so N (negative) with the probability 1-e_p^x. Likewise, if Act^y=D then P (positive) with the probability e_q^x, and N (negative) with the probability 1-e_q^x. The evaluation of the second evaluator is the same value of that of the first evaluator, so E(x<-y) = E(y<-x). The reason is explained later.

Learning strategy: After a set period of simulation, agents learn their strategies. Using the GA with trading payoffs as a fitness function, an agent's action strategy Act^x and evaluation strategy Eva^x evolve.

In the evaluation phase, why is the evaluation of the second evaluator the same of that of the first? Thompson (1967) shows that, people tend to act reciprocally to their partners when they interact. If the partner acts favorably, the person evaluates it positively, and if the partner acts unfavorably, the person did it negatively. From the view of Heider's balance theory (1958), an impression of an interaction between x and y depends on the opponent's impression. When the evaluation of x is positive, that of y is also positive, and vice versa. From these insights, it is appropriate to assume that an evaluated person evaluates one's partner reciprocally based on the partner's evaluation. In the evaluation phase of our model, the second evaluator decides a reciprocal (tit-for-tat) evaluation (reciprocal) of the partner. Therefore, we assume E(x<-y) = E(y<-x).

3.2 Reputation Management System

An agent x's evaluations, E^x, are recorded on a reputation management system (RMS). The RMS is managed intensively, and the agents can directly refer to the reputations of all agents. We define a reputation of an agent. A set of a history of evaluations to an agent, x, at a period of t is formulated as the Equation (2).

$$T_t^x = \left\{ E_k^x \middle| k \in \{0,1,\cdots,t\} \right\} \tag{2}$$

A number of what the value of the evaluation is P (positive) in T_t^x is defined as $\left|T_{P,t}^x\right|$, and a number of what the value of the evaluation is N (negative) in T_t^x is as $\left|T_{N,t}^x\right|$. A reputation of an agent, x, at a period of t, R_t^x, can be calculated as difference of the sum of positive evaluations and that of negative evaluations. See Equation (3).

$$R_t^x = \left|T_{P,t}^x\right| - \left|T_{N,t}^x\right| \tag{3}$$

A model of reciprocal evaluations is constructed as the extension of a model of correct evaluations, which is also constructed as the extension of a model of invisible payoffs, and which is as the extension of a model of visible payoffs with basic prisoners' dilemma. So, we explain four models in a methodical way.

In the model of visible payoffs based on the Cohen model, agents deal with the neighbor agents and do not evaluate their actions interactively. The information the agent can refer is only one's partner's action on the previous trading. The buyers and sellers have no difference in the model. The model of invisible payoffs is different from the model of visible payoffs in the point of the cooperation of agents, from imitating to GA. By introducing a RMS reflected with agents' actions correctly to the model of invisible payoffs, we make a model of correct evaluations. So, it is $E_t^x = A_t^x$ in the model of correct evaluations. Finally, when agents have strategies of evaluations in the model of correct evaluations, the model becomes a model of reciprocal evaluations.

The model of reciprocal evaluations and the three models are compared in Table 2.

Table 2: Comparison of the models

	Reciprocal Eva.	**Correct Eva.**	**Invisible Payoffs**	**Visible Payoffs**
RMS	with reciprocal evaluation	with actions correctly	None	None
Invisibility of a partners' payoff	Invisible	Invisible	Invisible	Visible
Bidding	for whoever has the highest reputation	for whoever has the highest reputation	None	None
Trading	a trading partner selected by bidding	a trading partner selected by bidding	all agents located on von Neumann neighborhood	all agents located on von Neumann neighborhood
Evaluation	evaluations based on an agent's evaluation strategy	trading actions with accuracy	None	None
Learning strategy	GA	GA	GA	imitating a neighbor agent's strategy

4 Simulation Experiments

We simulate our model developed in Section 3 with the agent-based approach. The simulations are carried out using C language. The parameter settings are as follows, the number of agents is 100, the number of periods for a generation is 100, and the number of generations is 200.

4.1 Model of visible payoffs based on the Cohen model

Cohen et.al. (2001) developed a basic model of the evolution of cooperation without a RMS. So, we examine the model as a pilot study.

Results of the simulations of the model are shown in Figure 2, with the vertical axis indicating the average (Ap, Aq) of all the agents and the horizontal axis indicating the simulation generation. Figure 2 shows that the All-D (perfect defector) strategy is dominant at first, and in time the tit-for-tat (TFT) is dominant. As is explained by Cohen, this mechanism is explained as follows: in the initial random state, the All-D strategy is the most advantageous strategy because it can perfectly exploit agents whose strategies are the All-C. When the All-D strategy is dominant, two TFT agents, who are located in each other's neighborhoods due to mutation, can overcome neighbor All-D agents because two TFT agents can gain higher payoffs than the neighbor All-D agents. The TFT agents then can form their colonies. As a result, the TFT strategy is dominant, and cooperation emerges.

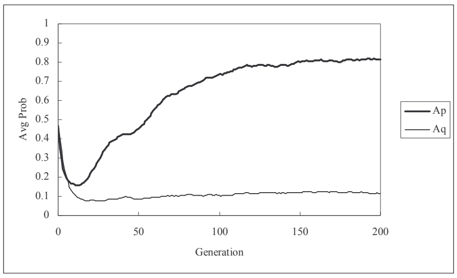


Fig. 2. The evolution of strategy in Cohen model

4.2 Model of invisible payoffs using GA as learning strategies

In the basic model explained above, we assume that the other agents' actions an agent dealt with and neighbor agents' payoffs are visible from the agent. But that is unnatural. On the other hand, this model premises that neighbor agents' payoffs are invisible from the agent. Results of this model are shown in Figure 3.

As shown in Figure 3, the All-D strategy is dominant in the model, while the TFT strategy is dominant in basic Cohen model. Why? Invisibility of neighbors' payoffs prevents two TFT agents who are located in each other's neighborhoods from surviving because a difference between payoffs of the All-D agents and those of the

TFT agents is a slight. According to the GA, the surviving probability of agents (a fitness function) is proportional to the payoffs of agents. Therefore, the All-D strategy continues dominant. The result indicates that the evolution of cooperation is hard when agents can use just dyadic information (which agents gain through their own experiences).

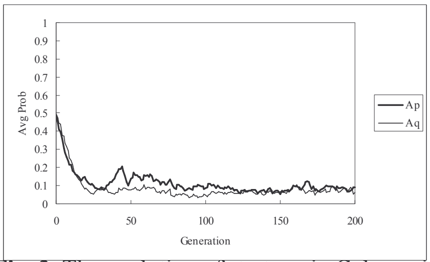


Fig. 3. The evolution of strategy in Cohen with GA model

4.3 Model of correct evaluations

A RMS is a system of which visible information is not only dyadic actions of whom the agent deals with directly, but also histories of innocent bystanders. We develop a model of a RMS which can record agents' actions correctly as a basic model for our research purpose. We refer to it as a model of correct evaluations. In this model (and a model of reciprocal evaluations explained later), it is able to choose trading partners because of using agents' reputations. We then introduce a bidding process to the model to choose trading partners.

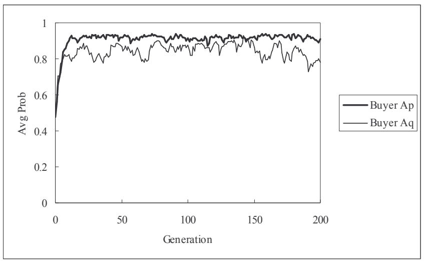


Fig. 4. The evolution of the buyers' strategies in the model of correct evaluations

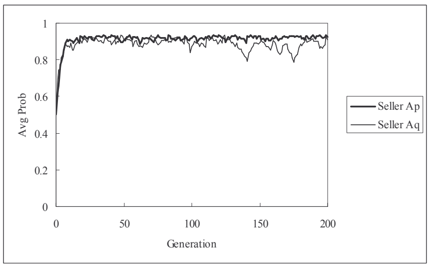


Fig. 5. The evolution of the sellers' strategies in the model of correct evaluations

Figure 4 and 5 show the average (Ap, Aq) of the buyers and sellers in this model. All the agents can see the information of all the agents' actions. If an agent acts non-cooperatively and therefore one can gain a high payoff at that time, one is not chosen by the others for the bidding process because all the agents know one’s irrelevant action. One cannot exploit the All-C agents when one does not take part in any trading. Therefore, the All-D agents cannot survive in the model. Besides, the TFT agents are more disadvantage than the All-C agents. When the TFT agent deals with the All-D agent, one acts defectively as a punishment for the All-D agent. This is why a history of the TFT agents has irrelevant records. As a result of these processes, the All-C strategies become dominant strategies.

4.4. Model of reciprocal evaluations

People's actions cannot be recorded on the real RMSs correctly. Because they tend to have psychological pressures that they are afraid of others' revenges for themselves when they evaluate the others, and to have expectations of which they want others to evaluate him positively. It may be general that individuals behave strategically in the evaluation phase. From the point of view, we introduce the evaluation function, Eva, discussed in Section 3. The evolutions of the buyers' and sellers' actions are shown in Figure 6 and 7. Figure 8 and 9 show, on the other hand, the evolutions of the buyers' and sellers' evaluations.

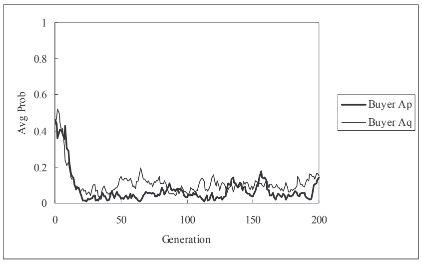


Fig. 6. The evolution of the buyers' action strategies in the model of reciprocal evaluations

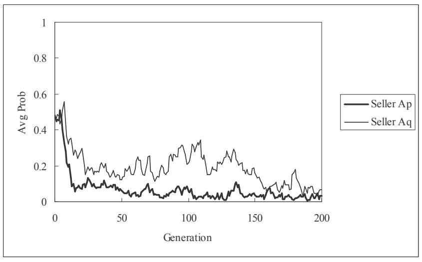


Fig. 7. The evolution of the sellers' action strategies in the model of reciprocal evaluations

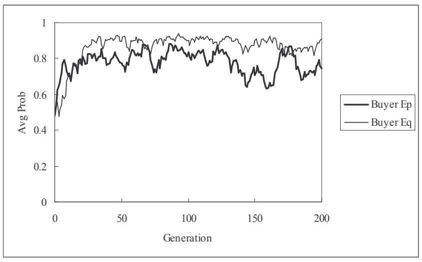


Fig. 8. The evolution of the buyers' evaluation strategies in the model of reciprocal evaluations

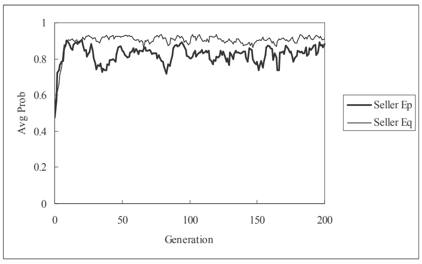


Fig. 9. The evolution of the sellers' evaluation strategies in the model of reciprocal evaluations

As shown in Figure 6 and 7, dominant action strategies of both the buyers and sellers are the All-D strategies; while Figure 8 and 9 show that dominant evaluation strategies of them are the All-C strategies. Because the agent who evaluates trading partners tolerantly tends to be chosen as a trading partner. So the All-C strategy has the advantage in the evaluation strategy. The agent whose evaluation strategy is the TFT punishes non-cooperative agents, so one has a history of one's revenge for non-cooperative ones. On the other hand, the All-C agents have no history of the revenges. Therefore, the All-C agents are more advantageous than the TFT agents because the All-C agents are free riders for the punishment of non-cooperative partners. However, it is the mechanism that brings that the function for the punishment of non-cooperative agents cannot work and the agent whose action strategy is the All-D has the advantage. There are no agents, then, whose evaluations are the TFT strategies when the other agent acts non-cooperatively. Therefore, the dominant action strategies are the All-D.

5 Conclusion

As shown in the simulation results, individuals who evaluate tolerantly have an advantage because evaluating tolerantly is being chosen as a trading partner. As is well known that the wise man keeps away from danger, we can often experience a situation of what people are hesitative about punishing non-cooperative actions. However, tolerant evaluations are regarded as a free riding for costs of punishing non-cooperative actions. Besides, we think it is hard to eliminate the tolerant evaluations because they are not based on malicious attitude. The tit-for-tat strategy which punishes non-cooperation is an evolutionary stable strategy in the game theory, however, this action costs for punishing in real life. The free riding for costs for punishing non-cooperators is known as a second order free riding problem (Axelrod, 1986). Our results show that the RMS in online markets has such a problem intrinsically due to reciprocal evaluations. We have an important theme for study on designing a RMS in how participants share the cost to develop the online trading actions in the future.

In the RMS, inaccurate evaluations are permitted due to share the information introduced on participants' subjective evaluations. Many papers pointed the necessity for exploiting malicious evaluations and arbitrary evaluations in collusion. For example, Dellarocas (2000) shows that the combination of "Controlled Anonymity" and "Cluster Filtering" makes the online trading robust for inequitable evaluations. Kamber et.al. (2003) studied a distributed hast table can prevent malicious groups in a peer-to peer environment. Does the RMS work effectively and the participants share the accurate evaluations if it exploits malicious evaluations? We conclude that a trap of the RMS is never exploiting the non-cooperative actions because the participants evaluate reciprocally and they prefer to tolerant participants for their evaluations. That is to say, the reputations cannot work as signals for the quality.

References

1. Ashlock, D., M. Smucker, E. Stanley, and L. Tesfatsion, "Preferential Partner Selection in an Evolutionary Study of Prisoner's Dilemma", BioSystems 37, No. 1-2, pp. 99-125. 1996.

2. Axelrod, R., "The Evolution of Cooperation", New York, Basic Books, 1984.
3. Axelrod, R., "An evolutionary approach to norms", American Political Science Review, 80, pp.1095-1111, 1986.
4. Axelrod, R.,"The Complexity of Cooperation", Princeton University Press, 1997.
5. Baron,D.,"Private Ordering on the Internet: The eBay Community of Traders",Business and Politics, vol.4,issue 3, pp.245-274,2002.
6. Cohen, M., Riolo, R. and Axelrod, R., "The Role of Social Structure in the Maintenance of Cooperative Regimes", Rationality and Society, Vol.13, pp.5-32, 2001.
7. Dellarocas, C., "Immunizing Online Reputation Reporting Systems Against Unfair Ratings and Discriminatory Behavior", Proceedings of the 2nd ACM Conference on Electronic Commerce, pp.17-20, 2000.
8. Dellarocas C. ,"The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms",Management Science, Vol. 49, No. 10, pp.1407-1424, 2003.
9. Heider, F., "The Psychology of Interpersonal Relations", Wiley,1958.
10. Kamvar,S., M., Schlosser, and H. Garcia-Molina, "The EigenTrust Algorithm for Reputation Management in P2P Network,",Proceedings of the Twelfth International World Wide Web Conference,2003.
11. Kollock, P., "The Production of Trust in Online Markets", Advances in Group Processes, Vol.16, pp.99-123, 1999.
12. Miyata,K. and Ikeda, K.,"The effect of reputations on virtual communities and consumer behavior: How Can Participants trust each other?", Annual Research Report, No.34, pp.166-176. Yoshida Hideo Memorial Foundation, 2001.
13. Nowak, A., and K., Sigmund, "Evolution of indirect reciprocity by image scoring", Nature, Vol.393, pp.573-577, 11,June,1998.
14. Resnick, P. and Zeckhauser, R., "Trust among strangers in internet transactions: Empirical analysis of eBay's reputation system", Technical report, University of Michigan, 2001.
15. Resnick, P., Zeckhauser, R., Friedman, E. and Kuwabara K., "Reputation Systems", Communications of the ACM, Vol.43, No.12, pp.45-48, 2000.
16. Thompson, J.D., Organizations in Action, New York: McGraw-Hill, 1967.
17. Yamamoto, H., K. Ishida and T. Ohta, "Modeling Reputation Management System on Online C2C Market", Computational & Mathematical Organization Theory, Vol. 10, No. 2, pp.165-178, 2004a.
18. Yamamoto, H., K., Ishida and T., Ohta,"Promotion of Cooperative Behavior in C2C market: Effect of Reputation Management System",Proc. of The Third International Workshop on Agent-based Approaches in Economic and Social Complex Systems pp.97-104,2004b
19. Yao, X., and P., Darwen, "How Important Is Your Reputation in a Multi-Agent Environment", Proc. of the 1999 IEEE Conference on Systems, Man, and Cybernetics, IEEE Press, pp.II-575 - II-580, 1999.

Enforcing Prosocial Behaviour*

Gennaro Di Tosto, Francesca Giardini, and Rosaria Conte

Institute of Cognitive Science and Technology, CNR, Italy
{gennaro.ditosto|francesca.giardini|rosaria.conte}@istc.cnr.it

Abstract. In this paper we present an explorative work framed within a research project aimed at the study of the emergence and evolution of prosocial behaviour (including altruism, cooperation, and compliance with norms) among autonomous intelligent agents. With the use of agent-based social simulation, we analysed different mechanisms for the evolution of cooperation: partner selection, punishment and communication. Results from a population of agents playing a modified version of the prisoner's dilemma are shown and compared statistically to point out the impact of the mechanisms proposed on the performance of the population.

1 Introduction

This paper presents an explorative work framed within a research project aimed at the study of the emergence and evolution of prosocial behaviour (including altruism, cooperation, and compliance with norms) among autonomous intelligent agents. As shown within a huge literature on the iterated prisoners dilemma (IPD), cooperation among non-kin needs to be sustained by enforcing mechanisms, the most frequent of which is punishment, i.e. a propensity to shift to defection with defectors (i.e. tit-for-tat). However, results obtained by this means are found to be sensitive to errors in strategy execution as well as invasions by free riders. Furthermore, some authors [1] convincingly argued that reactive strategies are not so frequently followed in human societies and, together with other authors[2], insist on the importance of social networks in the emergence of cooperation. Hence, variants of reactive strategies that fit with long-term relationships have been proposed, like cliquishness, i.e. a propensity to defect with strangers [2], or commitment [1]. The latter, in particular was found to 'benefit more from being unconditionally cooperative' although unconditional cooperation makes strategies vulnerable to exploitation.

An enforcing mechanism that rapidly gained popularity in the literature of reference concerns partners' assortation. Generally speaking, in evolutionary game theory, play is forced and attention is preferably given to the use of partner selection for retaliation. Some game theory studies have allowed players to avoid

* This work was partially supported by the European Community under the FP6 programme (eRep project, contract number CIT5-028575 and EMIL project, contract number IST-FP6-33841)

unwanted interactions, or more precisely to affect the probability of interaction with other players through their own actions (see [3–12] for more detailed discussions of related game theory work). Thanks to partner choice and refusal, payoff scores are found to increase since players can protect themselves from defections without having to defect themselves, and defectors get ostracized. On the other hand, choice and refusal also permit opportunistic players to home in quickly on exploitable players and form parasitic relationships. In particular, cooperators seem to take advantage of choice and refusal over nonreciprocators.

Finally, a fundamental mechanism supporting both punishment and partner choice is communication. As it requires agent-based rather than equation-based simulation, communication has none or poor tradition in the study of the emergence of altruism and cooperation. Nonetheless, its role in promoting informational cooperation as a means for material cooperation, at least in human societies, can hardly be denied.

Enforcing mechanisms have usually been observed in a fragmentary, non-systematic way, often starting from unclear concepts, not uniquely defined. Apparently, neither the interplay of punishment and partner selection and refusal nor the role of specific modalities of communication in supporting them has been addressed explicitly in the study of prosocial behaviour.

Last but not least, the variety of prosocial behaviour has not been dutifully considered. Again, altruism and cooperation are often treated as interchangeable notions, or at least as if potential differences among them did not affect the conditions under which they emerge. On the contrary, we believe that the peculiar features of these various forms of prosocial behaviour ought to be more carefully analysed and distinguished.

In this paper, we will turn the reader's attention on the following issues:

- Which is the most efficient mechanism of enforcement of prosocial action? Usually, this question is raised in the context of IPD. What about altruism?
- To what extent do agents contribute with their social intelligence to the efficiency of these mechanisms? In particular,
  - to what extent does communication contribute, and how does it interact with partner selection and punishment?
  - which modalities of communication can be envisaged, and which one is more beneficial?

2 Previous Notions

Prosocial action varies on many dimensions and in many ways. One main source of variability is mental: agents pursue different goals and are guided by different beliefs while executing different types of prosocial action (for a discussion of these aspects, which exceed the scope of the present paper, the reader is turned to [13, 14]). However, types of prosocial action, like altruism and cooperation, differ also in directly observable features:

– Symmetry *vs* asymmetry: this consists of two specific components, role complementarity and direction of benefit. Altruism is asymmetric in both senses, as for each episode the roles of donor and recipient are played by different agents and the direction of benefit is from donor to recipient; cooperation is symmetric as it allows for role identity and bidirectional benefit. It is of some interest to notice that in other types of prosocial action, for example social exchange, roles are complementary but benefit is bidirectional. Conversely, role identity and unidirectional benefit occurs only in antisocial action, i.e. exploitation.
– Individual *vs* shared benefit: this consists of the recipient of benefit, which in altruism can only be the beneficiary, while cooperation, at least in principle, allows for a shared benefit. Again, for the sake of analysis it may be interesting to observe that in social exchange by definition no shared benefit is allowed.

The interplay between these features leads to a third fundamental distinction:

– One-shot exploitation: in altruism, the donor is exploited only if she is not reciprocated later on. In cooperation, instead—think of the classic one-shot Prisoner's Dilemma—a cooperator may be exploited online by a partner playing defection.

Hence, in our terms, altruism is an asymmetric form of prosocial interaction characterized by individual benefit and no immediate exploitation. On the contrary, cooperation is a symmetric interaction, where benefit may be shared but immediate exploitation is also possible. The question is which enforcing mechanisms are needed to promote prosocial action? We explored the impact of the following mechanisms:

– Punishment (P), defined as a change of strategy (namely from cooperation to defection in presence of a known defector). To note that punishment here is something different from what is usually intended in the literature about altruistic punishment, where it denotes a costly behaviour usefull to prevent free-riding of public goods [15, 16]. In the present setting, however, benefits are not shared among the population, and we call punishment *every defection towards a known noncooperator* (based on personal experience, if repeted interaction are allowed; or on reported social evaluations, if communication is allowed—see sections 4, 5 for further details).
– Partner selection (PS) and refusal (PR): this is a non-random, rule-based assortation, such that each agent tries to associate with the partner from which they expect the highest payoff *given their own prosocial attitude*, and reject any other. Both cooperators and defectors expect to obtain the highest payoff by playing with a cooperator.

To these, we add the communication mechanism in two distinct modalities:

– Narrowcast (NC): one-to-one or one-to-few exchange of messages. Cooperators spread social evaluations they have acquired both directly through own

experience, or by means of communication, to known cooperators. Social evaluations concern the prosocial attitudes of potential partners.
– Broadcast (BC): one-to-many or blackboard-like exchange of messages about social evaluations from cooperators to the whole population.

The intuition behind is that broadcast is faster and therefore more efficient, but more dangerous as it exposes cooperators to belief-based exploitation from defectors.

3 Simulation Model

Starting from the assumption that prosocial behaviours among autonomous agents require special conditions to emerge and proliferate, we draw a simulation model both to test the effects, alone and in combination, of three enforcing mechanisms and to evaluate which of them is more suitable for making cooperation and altruism emerge. Traditionally, the main mechanisms to empower cooperators to protect themselves against cheaters are: punishment, partner selection, and communication. These three tools work differently and produce diverse effects depending upon the kind of prosocial situation considered. In what follows we will describe how each of these mechanisms works in the two games, and what effects they are expected to determine.

3.1 Games of Cooperation and Games of Altruism

The PD game has been extensively used by game theorists to address the issue of cooperation. In this game, both players have the choice to cooperate (C) or to defect (D), and the equilibrium outcome is defection for both players. This outcome is deficient, whereas cooperation is the Pareto-optimal outcome for both players. The PD game structure is well-known and it does not deserves further explanations. The simplicity of the PD game has led many scholars to use it to model several social and biological phenomena [17].

With regard to altruism, we used an asymmetric game, modeled after the *image scoring game* [18]. In each period, players are partnered and one is given the chance to take a costly action that helps the other. Cooperating in this manner is socially efficient, but the only way to monitor free riding is through the image score which, in this context comes to an accounting of a player's past altruistic actions. The image score is a property of agents immediately and universally accessible to everybody, with no actual communication process involved. Modelling reputation as an explicit feature of the agent is a solution that has been proved to be a good answer to the problem of indirect reciprocity. Image score is not the only one; tag-based systems [19] are based on a similar idea and generalize on the notion of image and image score to develop a system in which interactions take place between agents with similar tags. Tags are empty labels and, unlike image score, do not have to possess a semantic. But, like image score, they are public information which are compared by the agents before performing cooperative interactions. Tag-based systems, thanks to their level of abstraction, are now applied to different problems in different domains [20].

These techniques seems to share a common intuition which can be found elsewhere in the literature about the evolution of cooperation. Dawkins [21] discusses the possible evolutionary effects of a phenotypic trait capable of signaling the presence of its gene(s) to other organisms. Dawkins named this phenotypic trait as green beard, and the following effect consists in an altruistic behaviour towards the organism with the green beard, independent to the degree of relatedness of the other organisms. Despite the fact that an example of the green beard effect was actually found in nature [22], this theory, with the relative mechanisms, paves the way to a series of consideration. In nature, phenotypic traits are common indicators of status and identities and are involved, e.g., in defensive strategies and conflicts; in some cases they are used to bluff. But, in all these cases, they are used to avoid dangerous interactions or to find solutions to the problems of the social life that minimizes costs for all the organisms.

Altruists and cooperators in both games are doomed to extinction in a population where half of the agents are neither reciprocators nor cooperators, i.e. cheaters, as it is in our model. In fact, in the basic version, cooperators always play C in the PD game, and they always donate in the altruistic game, without differentiating between cooperative and non cooperative partners. This behaviour easily exposed themselves to exploitation to death by cheaters. This extreme situation allows us to put to the test 4 mechanisms that are supposed to enforce prosocial behaviours such as altruism and cooperation. Given the two games, we explored the effects of these variables, once per time or in combination, in terms of the average payoffs of agents.

4 How to help altruists and cooperators to survive: Punishment, Active and Passive Partner Selection, Communication

In our terms, punishment is the possibility to react to a defection playing Defect in the PD game, or playing keep in the game of altruism. When punishment is not active, agents simply play their built-in strategy, without the possibility of shifting it when facing dishonest partners. Partner Selection can be twofold: active and passive Partner selection. The former means that agents can select their partners in the interaction. In the PD game, this should lead cooperators to home in, avoiding defectors. Anyway, this same mechanism also permits cheaters to choose cooperators and exploit them. In the altruistic game, altruists should choose altruist to play give, expecting a reciprocal donation in the following of the game. On the other hand, we call Refusal the possibility to avoid an interaction, i.e. to escape from a cheater. Every time an agent refuses to interact, both agents of the couple loose an opportunity to interact. When Partner Selection was not available, agents were randomly paired. Finally, we explored three communication conditions:

1. Silent: no communication allowed.
2. Narrowcast: the access to this kind of communication is limited to cooperators. Cooperative agents send messages about their partners to other cooperative agents previously met. Cheaters can neither send nor receive messages in this condition.
3. Broadcast: this modality works as the image score does. Once an interaction is over, the receiver sends a message about the nature of the mover, i.e. a cheater or a cooperator. The message is posted on a blackboard accessible to both cooperative and non cooperative agents.

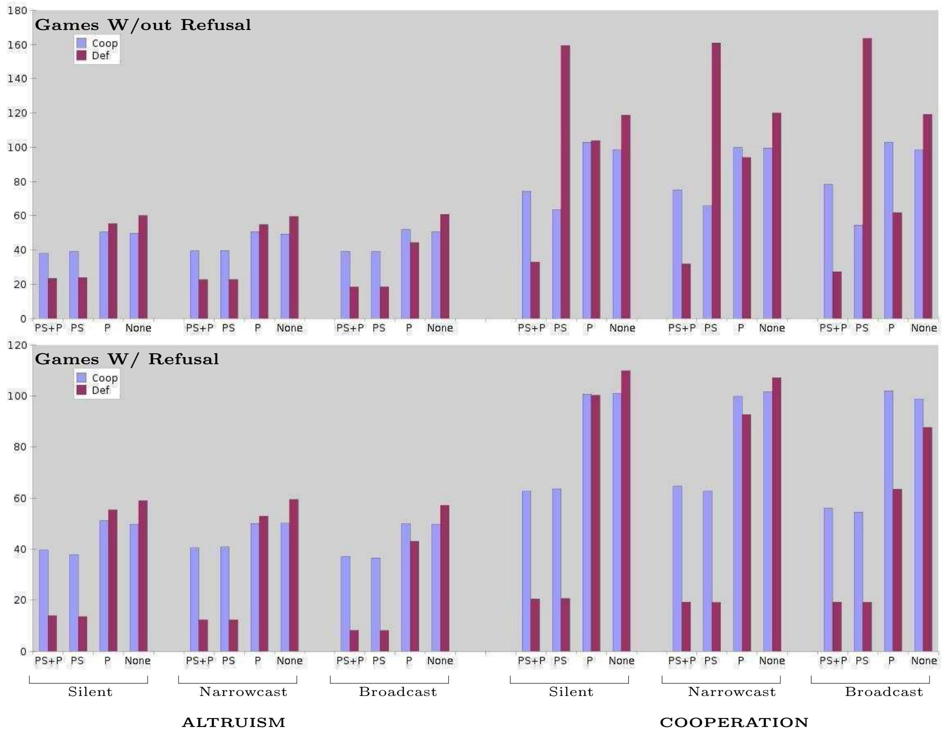


Fig. 1. Average payoffs after 200 simulation ticks of the two subgroups of agents (cooperators and defectors) as effect of PSR and P. The results for all the experimental conditions are shown (150 simulation runs for each experimental condition): altruism (image-score-like game) and cooperation (PD-like game), in all three forms of communication.

5 Experimental Settings and Results

A population of 500 agents, randomly assigned to one of two groups, cooperators and defectors, is created to play the game of altruism, and the game of cooperation. At each turn of the simulation, agents are coupled and face the option to perform or not a prosocial action, which will confer a benefit $b = 1.0$ to another agent, at a cost $c = 0.1$ to himself. After 200 interactions the gain of each agent are collected and analysed.

Performances of the two groups of agents are tested in several experimental conditions, under which we observed the effects of the mechanisms proposed. In figure 1 the average payoffs for the groups of cooperators and defectors are reported. Results are divided by game type and communication type. Furthermore the effects of PS and P are presented, with and without the possibility of refusal.

Table 1 shows the results of a regression analysis conducted with agents' payoffs as independent variable in order to asses the relative importance of PS and P in the interpretation of simulations data. Generally, both PS and P have the effect to lower the average payoffs, in both the groups of agents. PS is found to be the most important factor for the explanation of the variation of simulation outcomes, with few exceptions: in the cooperative game, without the possibility of Partner Refusal, Punishment is the most important factor to interpret the outcome of the defector strategy, independently of the communication mechanism. The cooperative game without PR is the condition where defectors obtain their higher average payoffs (see Fig. 1), and the punishment algorithm is the only mechanism through which cooperators can compete and avoid exploitation. In all the other settings PS is found to be responsible the results and appear to be the most effective tool for the enforcement of prosocial behaviour.

6 Discussion and Concluding Remarks

In this paper, we presented a simulation study of the effects of different enforcement mechanisms, such as punishment (P), partner selection (PS), in interaction with each other and with one-to-many and one-to-few communication, on symmetric prosocial behaviour, namely cooperation, and asymmetric one, namely altruism. P consists of a shift in strategy (from prosocial to antisocial) in presence of cheaters, while PSR reduces the potential number of interactions to those occurring between known prosocial partners (partner refusal is a variant of PS in which chosen partners can escape interaction).

In particular, P and PS are found to be

- almost perfectly complementary, i.e. PS favours altruists rather than cooperators whereas P at the opposite enforces cooperators but has no effect on altruists.
- P in general is found to produce average payoffs higher than PS.

Table 1. OLS regression model with dependent variable *gain*; regressors (PS and P) are dummy variables.

	Asymmetric game without refusal											
Communication	Silent				Narrowcast				Broadcast			
Agent Type	Cooperators		Defectors		Cooperators		Defectors		Cooperators		Defectors	
	B	**Std. Error**	**B**	**Std. Error**	**B**	**Std. Error**	**B**	**Std. Error**	**B**	**Std. Error**	**B**	**Std. Error**
(Intercept)	61.8171***	0.5299	95.61725***	0.20895	59.6090***	0.5369	95.18848***	0.20811	62.5209***	0.4647	98.8703***	0.2614
punish	-0.1048	0.2425	-2.57903***	0.09613	0.6980**	0.2457	-2.43151***	0.09573	0.6623**	0.2136	-8.2626***	0.1197
partner	-11.4568***	0.2425	-34.05266***	0.09613	-10.4758***	0.2457	-34.44588***	0.09573	-12.1315***	0.2136	-34.0189***	0.1197
R-Squared	0.1824		0.9268		0.155		0.9284		0.2431		0.8959	
Adj. R-squared	0.1822		0.9267		0.1548		0.9284		0.2429		0.8959	

	Symmetric game without refusal											
Communication	Silent				Narrowcast				Broadcast			
Agent Type	Cooperators		Defectors		Cooperators		Defectors		Cooperators		Defectors	
	B	**Std. Error**	**B**	**Std. Error**	**B**	**Std. Error**	**B**	**Std. Error**	**B**	**Std. Error**	**B**	**Std. Error**
(Intercept)	121.3757***	0.6174	231.260	1.257	121.9467***	0.6177	233.8512	1.1616	114.0966***	0.4852	230.427	0.910
punish	7.4729***	0.2817	-70.247	0.580	4.7609***	0.2832	-76.7944	0.5334	14.1121***	0.2229	-96.307	0.417
partner	-31.7445***	0.2817	-15.082	0.580	-29.2487***	0.2832	-11.4454	0.5334	-34.3346***	0.2229	4.613	0.417
R-Squared	0.573		0.6071		0.5238		0.6789		0.7374		0.8406	
Adj. R-squared	0.573		0.607		0.5237		0.6788		0.7374		0.8406	

	Asymmetric game with refusal											
Communication	Silent				Narrowcast				Broadcast			
Agent Type	Cooperators		Defectors		Cooperators		Defectors		Cooperators		Defectors	
	B	**Std. Error**	**B**	**Std. Error**	**B**	**Std. Error**	**B**	**Std. Error**	**B**	**Std. Error**	**B**	**Std. Error**
(Intercept)	59.9390***	0.5303	102.91870***	0.21474	59.9047***	0.5350	104.86628***	0.20758	62.4438***	0.4332	102.5312***	0.2473
punish	1.6025***	0.2431	-1.61888***	0.09863	-0.2126	0.2460	-3.23283***	0.09506	0.4736*	0.1993	-7.0825***	0.1132
partner	-11.7385***	0.2431	-43.37943***	0.09863	-9.4042***	0.2460	-43.85994***	0.09505	-13.1087***	0.1993	-41.8904***	0.1132
R-Squared	0.1904		0.9514		0.1278		0.9553		0.3026		0.9335	
Adj. R-squared	0.1902		0.9514		0.1276		0.9553		0.3025		0.9334	

	Symmetric game with refusal											
Communication	Silent				Narrowcast				Broadcast			
Agent Type	Cooperators		Defectors		Cooperators		Defectors		Cooperators		Defectors	
	B	**Std. Error**	**B**	**Std. Error**	**B**	**Std. Error**	**B**	**Std. Error**	**B**	**Std. Error**	**B**	**Std. Error**
(Intercept)	139.7175***	0.6398	196.6988***	0.2692	138.24267***	0.60848	190.7708***	0.2821	141.9865***	0.4203	150.0152***	0.3093
punish	-0.5302 .	0.2939	-4.9999***	0.1233	-0.02781	0.28113	-6.9913***	0.1285	2.4354***	0.1927	-12.0476***	0.1420
partner	-37.7808***	0.2939	-84.3273***	0.1233	-37.20654***	0.28113	-80.5589***	0.1285	-45.0963***	0.1927	-56.3656***	0.1420
R-Squared	0.621		0.9793		0.6372		0.9753		0.8466		0.9425	
Adj. R-squared	0.621		0.9793		0.6371		0.9753		0.8466		0.9425	

However, communication mitigates these results. In particular, although the narrowcast (NC) modality allows higher payoffs for cooperators in partner selection than the broadcast (BC) modality (as it does not expose them to exploiting defectors), the latter modality

- favours altruists even without partner selection;
- favours cooperators even without punishment.

To be noted, in cooperation the coupling of punishment and partner selection yields worse results than punishment alone.

Furthermore, our findings allow for PS and partner refusal (PR) to be compared. In particular, PR is extremely competitive with punishment even in cooperation. In such a condition, this is the only experimental condition beside punishment in which cooperators are better-off than noncooperators: this is rather obvious, since it is the only condition in which interactions between a cooperator and a known defector is impossible.

Finally, communication helps both altruists and cooperators. With the former, it works even in absence of PS, at least in the BC modality, and gives better outcomes to altruists and cooperators in the NC modality. As expected, the BC modality is faster and can be decisive, but the NC one is less dangerous and can give a stronger advantage to the good guys.

Not surprisingly, all the mechanisms examined contribute, although in different ways and degrees, to enforce prosocial behaviour. However, unlike expected, they are not hierarchically ordered in terms of efficiency. Rather, their efficiency depends on the type of interaction in which they are observed: in particular PS is required to promote altruism and is irrelevant in cooperation, whereas P is needed for enforcing cooperation but is useless in altruism. Why is this the case?

The rationale of these results is incorporated in the very nature of the two forms of prosocial action considered. Thanks to agent-based computational modelling, requiring these two forms of prosocial action to be formally modelled before and in order to be compared, it was possible to observe that features such as symmetry/asymmetry not only allow altruism and cooperation to be kept distinct but also call for different enforcing mechanisms. In particular, punishment reduces exploitation when immediate exploitation is possible, i.e. in cooperation but not in altruism. Conversely, partner selection promotes altruism thanks to an increased proportion of donations to the benefit of altruists on the total number of donations. But since this is obtained by reducing the total number of donations, the outcomes obtained by altruists are lower than those they would obtained by means of punishment. Still, in the latter condition, altruists are worse-off than nonaltruists. Otherwise stated, partner selection is more efficient when the promoted benefit is individual rather than shared, whereas punishment performs in the opposite way: as it excludes no-one from interaction, neither the good nor the bad guys, it allows higher payoffs to be obtained, and these benefit cooperators more than defectors when the game is symmetric and each has something to gain from interaction, i.e. the benefit is shared.

Interestingly, partner refusal sometimes makes it of its own: this happens in cooperation. Whereas with PS only, cooperators that are not known and therefore not chosen by their fellows are fully exposed to exploitation, with PR they can find an escape, and end up with being even better-off than defectors, even without the help of punishment.

In future works other forms of prosocial action will be investigated and compared with the ones examined in this work, also including other modalities of communication. Furthermore, we will compare the results obtained by cheaters and honest agents in non-homogeneous populations with varying percentages of the two behaviours in both conditions, in order to test how robust are altruism and cooperation to cheaters' invasion. Finally, the present findings will be re-analyzed for different values of the individual attitude to cheating and with different payoffs structures.

References

1. Back, I., Flache, A.: The viability of cooperation based on interpersonal commitment. Journal of Artificial Societies and Social Simulation **9**(1) (2006) http://jasss.soc.surrey.ac.uk/9/1/12.html.
2. Hruschka, D., Henrich, J.: Friendship, cliqueness, and the emergence of cooperation. Journal of Theoretical Biology **239** (2006) 1–15
3. de Vos, H., Smaniotto, R., Elsas, D.: Reciprocal altruism under conditions of partner selection. Rationality and Society **13**(2) (2001) 139–183
4. Fogel, D.B.: On the relationship between the duration of an encounter and the evolution of cooperation in the iterated prisoner's dilemma. Evolutionary Computation **3** (1995) 349–363
5. Guriev, S., Shakhova, M.: Self-organization of trade networks in an economy with imperfect infrastructure. In Schweitzer, F., ed.: Self-Organization of Complex Structures: From Individual to Collective Dynamics. Gordon and Breach Scientific Publishers, London (1996)
6. Hirshleifer, D., Rasmusen, E.: Cooperation in a repeated prisoners' dilemma with ostracism. Journal of Economic Behavior and Organization **12** (1989) 87–106
7. Kitcher, P.: The evolution of altruism. The Journal of Philosophy **90** (1993) 497–516
8. Mailath, G., Samuelson, L., Shaked, A.: Evolution and endogenous interactions. SSRI working paper, UW-Madison (June 1994)
9. Orbell, J.M., Dawes, R.M.: Social welfare, cooperators' advantage, and the option of not playing the game. American Sociological Review **58** (1993) 787–800
10. Stanley, E.A., Ashlock, D., Tesfatsion, L.: Iterated prisoner's dilemma with choice and refusal of partners. In: Artificial Life III. Volume 17 of Santa Fe Institute Studies in the Sciences of Complexity., Reading, MA (1994) 131–175
11. Ashlock, D., Smucker, M., Stanley, E.A., Tesfatsion, L.: Preferential partner selection in an evolutionary study of prisoner's dilemma. BioSystems **37**(1-2) (1996) 99–125
12. Hauk, E.: Leaving the prison: Permitting partner choice and refusal in prisoner's dilemma games. Computational Economics **18**(1) (2001) 65–87
13. Conte, R., Castelfranchi, C.: Cognitive and Social Acition. UCL Press, London (1995)
14. Conte, R.: Social intelligence among autonomous agents. Computational and Mathematical Organization Theory **5** (1999) 202–228

15. Boyd, R., Gintis, H., Bowles, S., Richerson, P.J.: The evolution of altruistic punishment. PNAS **100**(6) (March 2003) 3531–3535
16. Fehr, E., Gachter, S.: Altruistic punishment in humans. Nature **415**(6868) (January 2002) 137–140
17. Doebeli, M., Hauert, C.: Models of cooperation based on the prisoner's dilemma and the snowdrift game. Ecology Letters **8** (2005) 748–766
18. Nowak, M.A., Sigmund, K.: Evolution of indirect reciprocity by image scoring. Nature **393**(6685) (June 1998) 573–577
19. Riolo, R.L., Cohen, M.D., Axelrod, R.: Evolution of cooperation without reciprocity. Nature **414** (2001) 441–443
20. Hales, D., Edmonds, B.: Applying a socially-inspired technique (tags) to improve cooperation in p2p networks. IEEE Transactions in Systems, Man and Cybernetics – Part A: Systems and Humans **35**(3) (2005) 385–395
21. Dawkins, R.: The Selfish Gene. Oxford University Press (1976)
22. Grafen, A.: Evolutionary biology: Green beard as death warrant. Nature **394**(6693) (1998) 521–522

Session on Spatial Dynamics

Simulating pedestrian behavior in subway stations with agents

Arnaud Banos[1], Angèle Charpentier[2]
[1] Image et Ville (UMR 7011 CNRS), ULP, France, arnaud.banos@lorraine.u-strasbg.fr
[2] RATP, UMR Géographie-Cité, France, angele.charpentier@ratp.fr

Abstract. Despite the central and fundamental role pedestrian walking plays within the urban transport system, it still remains a badly known transportation mode. Generally speaking, while most of the developed countries have been developing, for the last 40 years, a wide variety of sophisticated methods and tools aimed at studying urban mobility, only a few of them were really designed to deal with pedestrian movement, especially in interaction with the other transportation modes. Pedestrian motion indeed occurs in an ever changing environment, defined by constraints and opportunities. In such a perspective, we propose an agent-based prototype, MAGE, allowing simulating pedestrian movements in very constrained and limited spaces, like subway stations.

Keywords: Multi-agents system, Pedestrian motion, Simulation, Transportation Systems

1 Introduction

Despite the central and fundamental role pedestrian walking plays within the urban transport system, it still remains a badly known transportation mode. Generally speaking, while most of the developed countries have been developing, for the last 40 years, a wide variety of sophisticated methods and tools aimed at studying urban mobility, only a few of them were really designed to deal with pedestrian movement, especially in interaction with the other transportation modes. Noticing the tremendous asymmetry that has been existing for long in the scientific literature between traffic flow and pedestrian movement, Weifeng and his colleagues [22] propose several explanations:
"The reason may be that the pedestrian movement is more complex than vehicular flow. First, pedestrians are more intelligent than vehicles and they can choose an optimum route according to the environment around. Secondly, pedestrians are more flexible in changing directions and not limited to the "lanes" as in vehicular flow. Thirdly, the slight bumping is acceptable and need not be absolutely avoided as in traffic flow models. So the model developed for pedestrian movement should fully consider these differences in order to study the special phenomena in pedestrian movement. It is pity that till now, most pedestrian movement models are established based on the rules used for traffic flow and consider little of the special characteristics of pedestrian movement itself"[pp. 634].

The key point we will try to defend here, is that pedestrian movement needs not only to be considered as a specific phenomenon. It also needs to be included in a much more global and complex perspective, the urban system as a whole. Pedestrian motion indeed occurs in an ever changing environment, defined by constraints and opportunities. The MAGE prototype has been precisely designed to explore the behavior of pedestrians in motion and in interaction, in a virtual space where most of the phenomenon can be mastered and studied.

This idea of designing “virtual laboratories” [3], within which “artificial societies” can be grown for example [2], [8], [20] has become very popular in the recent years and is largely related to two other fields, science of complex systems and agent based modeling. More, it is firmly embedded in a microscopic approach of urban mobility, where the world is represented as closely as possible in a one-to-one way, which means that “people should be represented as people, cars should be represented as cars and traffic lights should be represented as traffic lights and not as, say, departure rates, traffic streams and capacities respectively” [18]. Reaching such a modeling level, without being flooded with microscopic details, requires then an ad-hoc procedure for designing agent-based computer simulations. The two crucial principles of reductionism and parsimony may therefore constitute main guidelines, in our quest for the identification of the micro-specifications sufficient to generate macrostructures of interest [8].

2 Pedestrian movement modeling

Since the first quantitative approaches to pedestrian movement in the 1950s [20], a considerable amount of work has been done in the field, from queuing models [16] to traffic-likes models, with some incursions towards more behavioral ones and more recently large investments in micro-simulation approaches [1], [4], [12], [13], [10], [15].

The last family is of special interest for our purpose, as it specifically deals with interactions at the pedestrian level, a key point in our project. For example, focusing on the emergence of collective pedestrian behavior from very simple specifications made at the level of individuals, Blue and Adler [4] managed to reproduce, with a cellular automata, collective characteristics such as formation of pedestrian lanes or speed-flow diagrams. Then, the social force model proposed by Helbing [12], [13] focuses more directly on the influence of the environment and other pedestrians on individual behavior , the whole being formalized as a fluid-dynamic pedestrian model.

While being fundamental steps, these works have anyway the drawback of relying on an excessive simplification of the urban environment (a corridor, a place or a room). Motivations and goals of pedestrians are also particularly simplified, reduced to the couple destination to reach / obstacles to avoid. Theses limitations encouraged other researchers to explore agent-based models of pedestrian movement [1], [10], [15], [20]. In this last family, each pedestrian/agent is defined by a set of capacities tries to achieve a set of goals, interacting locally with its environment and with other agents. MAGE directly relies to that specific field, its originality being defined by its focus on local interactions.

2.1 Design of the model

Every subway station may be seen as an open complex system, part of a wider system dedicated to mobility. Simulating pedestrian's behavior in such a peculiar system faces different issues, related to:

- Equipments and facilities locations (ticket dispensers, shops, services) ;
- Traffic management (speed/density/flow of users) ;
- Emergency procedures (evacuation).

2.1.1 Design of the model

A map of the station has been rasterized and divided into a large number (nearly 7000) of small cells (0.4 m^2). Then, a distinction has been made between what we called "physical space" and "direction graph". The physical space is defined as the set of cells $E_p = \{C_1, C_2, ...C_i, ...C_n\}$, each cell $C_{i,[1,n]}$ being defined by a set of attributes $V = \{V_1, V_2, ...V_i, ...V_n\}$. Secondly, we sought to enrich the space, by distinguishing specific landmarks and directions. To this end, we took inspiration from classic research demonstrating the important role played by certain nodes and landmarks on individuals' mental representations of their environments (Lynch, 1998). We drew a direction graph $G\theta = (S, \theta)$, with S the summits, defined as localised decision points[1] , and θ the curves defined as headings between adjacent summits. A decision point can be seen as a pivot, a space where a choice of heading must be made. This point may correspond in a given physical space E_p, to a landmark (signalling), a particular service (store, ticket office or machine, etc.), or simply to a bifurcation. Therefore, the direction graph may be seen as an underlying invisible skeleton, approximating long distance vision field of individuals. Figure 1 shows the kind of graph drawn for Montparnasse station in Paris, France. Note that in this case, for simplicity, the summits of these graphs are solely restricted to bifurcation points. In other words, we assume that agents are unfamiliar with the availability and location of services (ticket offices, ticket machines, shops), they shall identify locally when passing nearby through a vision sensor.

[1] This idea of decision points comes from research by Timmermans and Djikstra (1999), although our interpretation is a bit different.

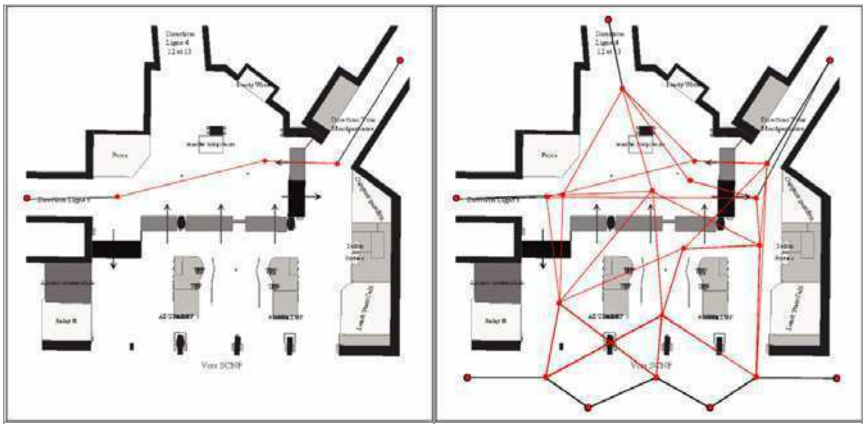


Fig.1a	Fig.1b

Fig. 1. Montparnasse station and the direction graph built for a given O/D (fig.1a) and for every O/D (fig.1b)

The complexity of this graph depends largely on the assumptions we make on pedestrians. If we accept the idea of an average pedestrian with complete information and boundless rationality (Homo Oeconomicus), then $G\theta^{i}_{(OD)} = G\theta_{(OD)}$ with $G\theta_{(OD)}$ the minimum energy path between a given origin and destination. Each pedestrian may then have the same optimized orientation guide for a given O/D. On the contrary, if we move towards more heterogeneous agents, with limited information and bounded rationality, then $G\theta^{i}_{(OD)} = (S', \theta') \ni S' \subset E_p$ and $\theta' \subset E_p$. Each agent may then possess his own graph for a given O/D.

2.1.2 Designing pedestrians

In MAGE, each pedestrian is represented by an agent A_i, described by a vector of attributes $X(A_i) = \{G\theta^{i}_{(OD)}; \delta_i; \phi_i; l_i^t; \theta_i^t; v_i^t\}$ with: $G\theta^{i}_{(OD)}$ his orientation graph for a given origin/destination ; δ_i his vision field ; ϕ_i his action field ; l_i^t his location *(x, y)* at time *t* ; θ_i^t his orientation at time *t*, $0° \leq \theta_i^t \leq 360°$ and v_i^t his speed at time t (figure 2).

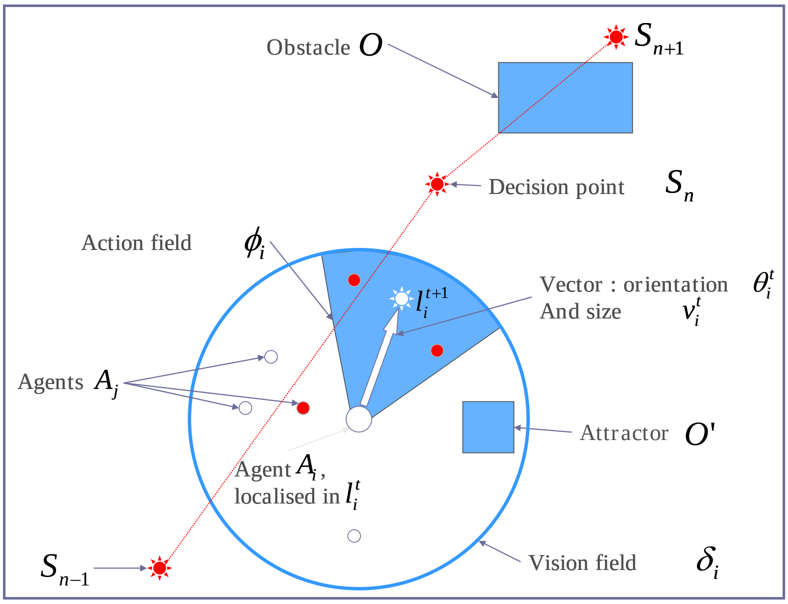


Fig. 2. Pedestrian as a Brownian agent

Each agent then updates at each time *t* his location $l_i^t \rightarrow l_i^{t+1}$, by modifying his orientation and/or his speed, depending on the various constraints and opportunities encountered locally :

$$\theta_i^t = f\left(S_n, O \subset \delta_i, O' \subset \delta_i, A_{i,j \neq i} \subset \delta_i, \varepsilon\right) \tag{1}$$

and

$$v_i^t = f\left(\# A_{i,j \neq i} \subset \phi_i\right) \tag{2}$$

In the absence of interaction with environment and other agents, each agent computes at each time *t* his next location l_i^{t+1} :

$$x_i^{t+1} = x_i^t + \Delta x \, , \Delta x = v_i^t \times \sin\left(\theta_i^t + \varepsilon\right) \tag{3}$$

$$y_i^{t+1} = y_i^t + \Delta y \, , \Delta y = v_i^t \times \cos\left(\theta_i^t + \varepsilon\right)$$

with: θ_i^t the angle location l_i^t of agent A_i and his next decision point S_n ; v_i^t speed of A_i (m/s) ; ε an angular value to be chosen by user in the interval [0 , 360].

Then, following the set of decision points $S_n \to S_{n+1}$ constituting his orientation graph $G\theta^i_{(OD)}$, each agent may face obstacles and then try to bypass them:

if $l^{t+1}_{i,\theta_i^t,v_i^t} \neq O$ (4)

then $l_i^t \to l^{t+1}_{i,\theta_i^t,v_i^t}$,

else $\theta_i^t = \min\left|(\theta_i^t \pm \alpha) - \theta_i^t\right| \ni l^{t+1}_{i,\theta_i^t,v_i^t} \neq O,\ 0 \leq \alpha \leq 180°$

On the contrary, various equipments (ticket dispensers, shops…) may attract an agent with a given probability:

if $O' \subset \delta_i$

and if $\beta_i \leq \mathrm{B}$, with B a constant belonging to interval (0 ; 1) and β_i a random (uniform) value chosen in the same interval (0,1),

then $O' \to S_n, S_n \to S_{n+1}$ and more generally $n \to n+1$.

Then, the presence of agents in the vicinity of agent A_i may cause interferences. Indeed, if every agent is supposed to avoid any agent on his way, he may change his behavior when the level of pedestrian density increases. In this case, we implemented specific crowd behaviors, inspired from Reynold's Boids [19]. Each agent then adopts a crowd behavior when the density of agents in his neighbor reaches a given value and when some of these neighbors share the same decision point than agent A_i:

$$\left(\# A_{j,j\neq i} \subset \delta_i\right) \geq \lambda \text{ and } \left(\# A_{j,j\neq i} \subset \delta_i \ni S_{nj} = S_{ni}\right) \neq 0 \quad (5)$$

In this situation, agent A_i tries to align himself to neighboring agents sharing the same decision point:

$$\theta_i^t = \frac{1}{\# A_{j,S_{nj}=S_{ni}}} \sum_{j=1}^{n} \theta_j^t \quad (6)$$

In the same time, he also tries to move nearer the barycenter $\overline{A}_{j,S_{nj}=S_{ni}}$ of this group:

$$x = \frac{1}{\# A_{j,S_{nj}=S_{ni}}} \sum_{j=1}^{n} x_j \text{ and } y = \frac{1}{\# A_{j,S_{nj}=S_{ni}}} \sum_{j=1}^{n} y_j \quad (7)$$

Finally, each agent modifies his speed according to the number of agents in his action field: $v_i^t = f\left(\# A_{i,j\neq i} \subset \phi_i\right)$. A non linear model based on Fruin's data [9] has been adjusted:

$$v_i^t = \begin{cases} v_{\max} \text{ if } \# A_{i,j\neq i} \subset \phi_i = 0 \\ -0{,}376 * S^{-1{,}5} + 1{,}372 \text{ if } S > 0{,}4\text{m}^2 \\ 0 \text{ if } S \leq 0{,}4\text{m}^2 \end{cases} \quad (8)$$

with :

– v_i^t the walking speed (m/s) of agent A_i at time t ;

– S the free surface (m^2/pedestrian) available, i.e. $\dfrac{S(\phi_i)}{\# A_{i,j\neq i} \subset \phi_i}$

3 Implementation and first results

3.1 An overview of MAGE

A first version of MAGE has been implemented within the simulation platform Netlogo (http://ccl.northwestern.edu/netlogo/). Figures 3a and 3b show snapshots of the kind of environment created.

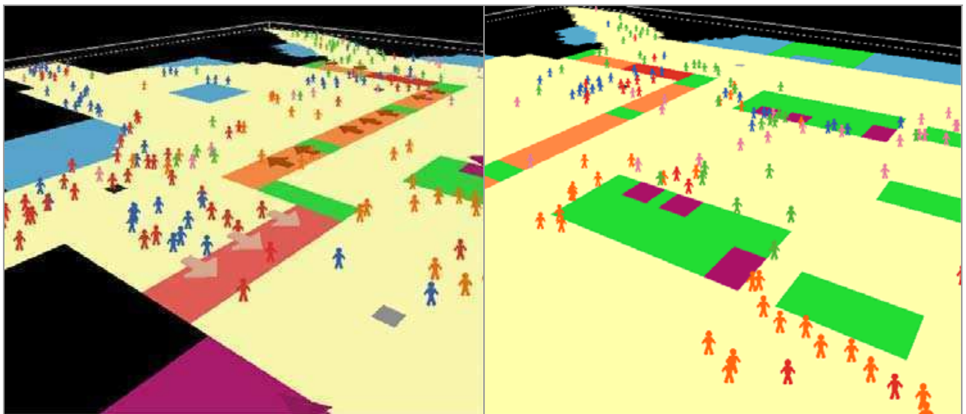


Fig.3a Virtual Montparnasse station	Fig.3b Agents queuing at turnstiles, exit doors and tickets dispensers

Fig. 3. An overview of MAGE

The focus may also be put on global structures, through real time display of density maps (figure 4). It may be noticed that map 4a (no interactions with local environment and few noise) is very close to the orientation graph in figure 1b. The more interactions and noise we add, the more deviations we obtain from this initial skeleton (map 4c).

	
Fig. 4a $\varepsilon = 0°$, $B = 0$	Fig. 4b $\varepsilon = 30°$, $B = 0$
	Density maps obtained for different values of parameters ε (random deviation from straight line) and B (probability of using a given equipment). Each cell $C_{i,[1,n]}$ records at each time t the number of pedestrians passing. This cumulated number is then displayed using a gradation of reds (the darker the higher).
Fig. 4c $\varepsilon = 30°$, $B = 0.5$	

Fig. 4. Revealing macrostructures : density of agents

3.2 Exploring the role of interaction through the fundamental diagram

One of the objectives of the MAGE model is to allow, further ahead, the introduction of an experimental approach to pedestrian behaviour in the confines of a subway station. Using individual behaviour, the aim is to achieve a clearer understanding of the emergence of observable collective behaviour, which is fundamental to the design of public transport areas.

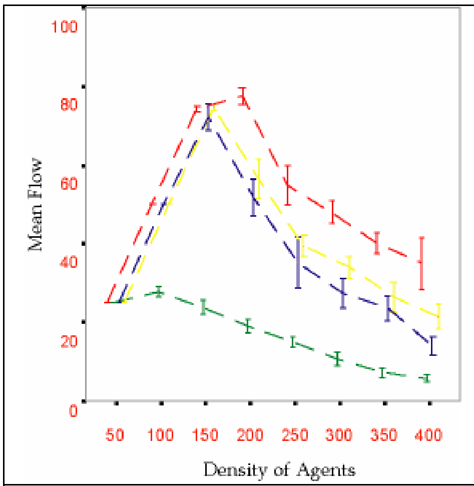


Fig. 5a Flow-Density diagrams in the absence of constraints (no ticket gates and no exit doors)

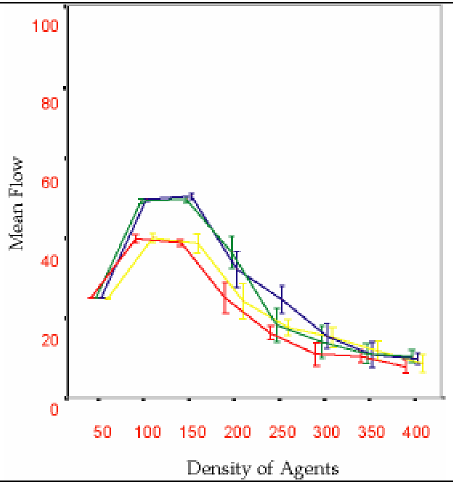


Fig. 5b Flow-Density diagrams in the presence of constraints (ticket gates and exit doors)

Scenario	Curve color	Avoid other agents	Crowd behaviour
1		No	No
2		Yes	No
3		No	Yes
4		Yes	Yes

Fig. 5c Scenarios tested and curve colors

Fig. 5. Flow-Density diagram showing the impact of local interactions on bi-directional flows ("Density" is the number of agents created, half at each entrance of the path highlighted Fig.1a ; "Mean Flow" is the number of agents reaching the opposite exit in two minutes, averaged over 10 simulations ; vertical bars show the variability of this last output variable over the ten simulations)

We defined an experimental protocol in order to illustrate the relevance of an individual-centred approach centred on computer simulation. Within Montparnasse station, a space characterised by a multitude of flows leading in every direction, a benchmark path was selected (see Fig. 1.a), to explore the influence of individual interactions on the flow-density diagram in the case of bi-directional flows. The diagram was constructed by way of an exit variable (flow) where values change depending on the values of the entry variable: density. This last variable is defined by the number of pedestrians following the itinerary, while flow corresponds to the number of passengers reaching their destination in two minutes or less (Figures 5). These two variables are defined (density) and measured (flow) in a systematic way, for several combinations of behaviour and spatial configurations. Agents' ability to avoid collision (equation 4) and adopt crowd behaviour (equations 5-7) was favoured. Note too the presence or absence of facilities constricting pedestrian movement (ticket gates and exit doors). Figure 5 highlights the variation of bi-directional flow relative to the density parameter, based on different scenarios.

In both diagrams, curves are characterized by two distinct variation ranges, on both sides by a so-called critical density value, after which pedestrian flow decreases as the number of agents increases. When there are no doors or gates (Fig. 5a), the capacity of agents to avoid collision and adopt crowd behaviour after certain density values increases pedestrian flow significantly (no overlapping of confidence intervals). Under this configuration, avoidance without crowd behaviour has a penalising effect, as each agent wastes time avoiding all other agents he comes across. Adding doors and gates alters the hierarchy of the lines in a remarkable fashion (Fig. 5b).

Under this configuration, the critical density value is systematically replaced by an interval (between 100 and 150 agents), while the maximum flow values obtained are relatively lower than in a situation where there are no facilities to constrain agents. The change in hierarchy of the lines highlights the regulating role that these facilities play in one direction, which automatically encourages a certain level of flow separation.

4 Conclusion

This first implementation of MAGE, aimed at simulating pedestrian behavior in subway stations, already allows to test "what if scenarios" and to explore some fundamental aspects of complex spatial systems. Anyway, reaching such a modeling level - even still simplified - without being flooded with microscopic details, requires an ad-hoc procedure. Crucial principles like reductionism and parsimony proved to be relevant guidelines, in our quest for the identification of the micro-specifications sufficient to generate macrostructures of interest. The systematic exploration of global structures seeming to emerge from local interactions is a further step we wish to address, focusing particularly on the influence of local interactions on flows. Our initial investigations demonstrate the major role played by individual interactions and their influence on the overall behaviour of the system. However, a certain amount of fine-tuning is needed before this type of application can really become operational, above all with regard to: multi-directional flow; the inclusion of a wide variety of individual behaviour (heterogeneous agents) and spaces and finally, using larger populations in the simulations.

References

1. Batty M.: Agent-based pedestrian modelling, CASA paper n° 61 (2003), 26 p.
2. Batty M.: Cities and complexity, MIT Press, Cambridge (2005), 565 p.
3. Batty M., Torrens P.; Modeling and prediction in a complex world, Futures, N°37 (2005), pp. 745-766
4. Blue V., Adler J.: Emergent fundamental pedestrian flows from cellular automata microsimulation, Transportation Research Record, vol. 1644 (1998), pp. 29-36
5. Cunningham, P., Cullen D.: Pedestrian flow data collection and analysis, Proceedings of the Institution of Civil Engineers, Transport Vol. 100 ((1993), pp.59-69
6. Djikstra J, Timmermans H.: Towards a multi-agent model for visualizing simulated user behavior to support the assessment of design performance, In: Ataman O, Bermúdez (eds) ACADIA99 – Media and design process, Acadia (1999), pp. 226-237
7. Epstein J.: Agent-based computational models and generative social science, Complexity, Vol. 4, N°5 (1999)
8. Epstein J., Axtell R.: Growing artificial societies: social science from the bottom up, Brookings Institution Press, MIT Press, Washington DC (1996)
9. Fruin J.: Pedestrian planning and design, New York, Elevator world's publication (1987), 206 p.
10. Hacklay M., Thurstain-Goodwin M., O'Sullivan D., Schelhorn T.: So go downtown: simulating pedestrian movement in town centers, Environment and Planning B, n° 28 (2001), pp. 343-359
11. Hankin, B-D., Wright R-A.: Passenger flow in subways, Operational research quaterly, Vol. 9, n°2 (1958), pp. 81-88.
12. Helbing D., Molnar P., Farkas I., Bolay K.: Self-organizing pedestrian movement, Environment and Planning B, n° 28 (2001), pp. 361-383
13. Helbing, D., Buzna, L., Werner, T.: Self-organised pedestrian crowd dynamics and design solutions, Transportation science (2003), 39 p.
14. Jiang, B.: SimPed: Simulating pedestrian flows in a virtual urban environment », Journal of geographic information and decision analysis, Vol. 3, n°1 (1999), pp.21-30
15. Kerridge J., Hine J., Wigan M.: Agent-based modeling of pedestrian movements: the questions that need to be asked and answered, Environment and Planning B, vol. 28 (2001), pp. 327-341
16. Lovas G.: Modeling and simulation of pedestrian traffic flow, Transportation Research B, n° 28-6 (1994), pp. 429-443
17. Nagel K., Schreckenberg M.: Cellular automaton models for freeway traffic, Physics. I-2 (1992), pp. 2221-2229
18. Nagel K., Esser J., Rickert M.: Large-scale traffic simulations for transportation planning, Annual Reviews of Computational Physics VII (2000), pp. 151-202
19. Reynolds C.: Flocks, Herds, and Schools: A Distributed Behavioral Model, Computer Graphics, vol. 21, n°4 (1987), pages 25-34.
20. Schweitzer F.: Brownian Agents and active particles, Springer-Verlag, Berlin (2003)
21. T.R.B.: Highway capacity manual USA, Transportation research board (2000), 1134 p.
22. Weifeng F., Lizhong Y., Weicheng F.: Simulation of bi-direction pedestrian movement using a cellular automata model, Physica A 321 (2003), p. 633 – 640

Quantitative Agent-Based Modeling of Human Interactions in Space and Time

Dirk Helbing[1,2,a] and Anders Johansson[1]

[1] Dresden University of Technology, Andreas-Schubert-Str. 23, 01062 Dresden, Germany

[2] Collegium Budapest – Institute for Advanced Study, Szentháromság utca 2, 1014 Budapest, Hungary

Abstract. The study of spatio-temporal pedestrian interactions has led to a deeper understanding of how collective behavior on a macroscopic scale emerges from individual human interactions. The results have been related to game theory and transferred to the description of other social interactions such as collective opinion formation. Here, we explain how individual interactions lead to self-organization phenomena, social order, and "collective intelligence". We also present recent efforts to quantatively determine pedestrian interactions from empirical observations. The findings complement game theory by additional mechanisms leading to the coordination and cooperation of human beings.

1 Introduction

The emergence of new, functional or complex collective behaviors in social systems has fascinated many scientists. One of the primary questions in this field is how cooperation or coordination patterns originate based on elementary individual interactions. In the social sciences, there are different approaches to this problem. This is partly a consequence of the fact that social systems tend to be so complex that the relevant variables and parameters involved are often hard to identify and to measure. In view of this situation, some social scientists try to give an accurate as possible description of a social phenomenon under investigation. This is, for example, done within the framework of case studies, and the result may be a narrative with a large degree of detail and, therefore, a great descriptive power. Recently, however, there is a growing tendency to develop simplified, abstracted descriptions by means of mathematical or computer models [1–13]. For example, agent-based models are developed by a growing community [14–20]. However, this approach is not yet standardized. Some scientists try to model each single effect with separate mathematical terms and parameters, thereby aiming at a large descriptive power. In many cases, however, there are not enough data available to fit the parameters accurately. This can lead to a small predictive power, i.e. a model that fits well to the data to which it was

[a] Dirk Helbing is currently changing to the ETH Zurich in order to work as Professor for Sociology, particularly Modeling and Simulation in the future.

calibrated may describe another data set or situation rather poorly, if the same parameter values are used. Therefore, in many cases, the predictive power of simple models with a few parameters can be higher than that of more detailed models. This observation has established the KISS principle: "Keep It Simple, Stupid." In other words, some scientists propose to develop simplified models to describe the "stylized facts" [21, 22]. This tends to apply not only to researchers from the sociophysics community. It essentially also underlies the approach of game theory [23, 24, 5], which may be considered to be one of the most successful and powerful mathematical theories in the social sciences. From the point of view of the authors, all the above mentioned approaches are needed, and they complement each other.

In this contribution, we will introduce a simplified model of pedestrian interactions, the "social force model", and discuss the surprisingly many empirical findings it can describe. Originally, this model aimed only at understanding qualitative discoveries such as the observed kinds of collective motion (see Sec. 5). This goal could be successfully reached by phenomenological specifications of the interaction forces, as the simulated kinds of collective motion did not depend on the details of the interactions (i.e. the social force model was robust with respect to modifications in the force functions and parameters). For example, the qualitative findings were not sensitive to whether we chose an exponential [25] or power-law decay [26] of repulsive interaction forces with distance. Circular and elliptical, isotropic and anisotropic specifications [25–28] could all reproduce similar collective patterns of motion.

In this, contribution, however, we want to go beyond this stage of research by trying to quantify pedestrian interactions based on video-extracted data. That is, we will not be satisfied with a *qualitative* description anymore, but we will describe methods to obtain a model, which is also in *quantitative* agreement with empirical data. While the exponential decay of repulsive interaction forces turns out to be well compatible with previous specifications of the social forces [28], other aspects of the interaction forces must be improved. Finally, we will discuss some of the challenges on the way of developing quantitative and verifiable agent-based models.

2 Short History of Pedestrian Modelling

Pedestrian crowds have been empirically studied for more than four decades [29–31]. The evaluation methods applied were based on direct observation, photographs, and time-lapse films. Apart from behavioral investigations [32, 33], the main goal of these studies was to develop a *level-of-service concept* [34], *design elements* of pedestrian facilities [35–38], or *planning guidelines* [39, 40]. The latter have usually the form of *regression relations*, which are, however, not very well suited for the prediction of pedestrian flows in pedestrian zones and buildings with an exceptional architecture, or in extreme conditions such as evacuation. Therefore, a number of simulation models have been proposed, e.g. *queueing models* [41], *transition matrix models* [42], and *stochastic models* [43], which are partly related to each other. In addition, there are models for the *route choice behavior* of pedestrians [44, 45] and experimental studies [38, 46–50].

None of these concepts adequately takes into account the self-organization effects occuring in pedestrian crowds. These may, however, lead to unexpected obstructions due to mutual disturbances of pedestrian flows. More promising with regard to this is the approach by Henderson. He conjectured that pedestrian crowds behave similar to gases or fluids ([51], see also [52]). This could be partially confirmed, but a realistic gas-kinetic or fluid-dynamic theory for pedestrians must contain corrections due to their particular interactions (i.e. avoidance and deceleration maneuvers) which, of course, do not obey momentum and energy conservation. Although such a theory can be actually formulated [53, 54], for practical applications a direct simulation of *individual* pedestrian motion is favourable, since this is more flexible. As a consequence, current pedestrian research focusses on *agent-based models* of pedestrian crowds, which also allow one to consider local coordination problems. The "social force model" is maybe the most well-known of these models, but we also like to mention *cellular automata* of pedestrian dynamics [55–61] and *AI-based models* [62, 63].

3 The Social Force Concept

In the following, we shall shortly introduce the social force concept. Human behavior often seems to be "chaotic", irregular, and unpredictable. So, why and under what conditions can we model it by means of forces? First of all, we need to be confronted with a phenomenon of motion in some (quasi-)continuous space, which may be also an abstract behavioral space such as an opinion scale [64, 2]. Moreover, it is favourable to have a system where the fluctuations due to unknown influences are not large compared to the systematic, deterministic part of motion. This is usually the case in pedestrian traffic, where people are confronted with standard situations and react "automatically" rather than taking complicated decisions, e.g. if they have to evade others.

This "automatic" behavior can be interpreted as the result of a *learning process* based on trial and error [65], which can be simulated with *evolutionary algorithms* [66]. For example, pedestrians have a preferred side of walking [30, 31], since an asymmetrical avoidance behavior turns out to be profitable [14, 56, 65]. The related *formation of a behavioral convention* can be described by means of *evolutionary game theory* [14, 67, 2].

Another requirement is the vectorial additivity of the separate force terms reflecting different environmental influences. This is probably an approximation, but there is some experimental evidence for it. Based on quantitative measurements for animals and test persons subject to separately or simultaneously applied stimuli of different nature and strength, one could show that the behavior in conflict situations can be described by a superposition of forces [68, 69]. This fits well into a concept by Lewin [70], according to which behavioral changes are guided by so-called *social fields* or *social forces*, which has later on been put into mathematical terms [14, 71, 2]. In some cases, social forces, which determine the amount and direction of systematic behavioral changes, can be expressed as gradients of dynamically varying potentials, which reflect the social or behavioral fields resulting from the interactions of individuals. Such a social force concept was applied to opinion formation and migration [71, 2], and it was particularly successful in the description of collective pedestrian behavior [14, 25, 38, 65].

For reliable simulations of pedestrian crowds we do not need to know whether a certain pedestrian, say, turns to the right at the next intersection. It is sufficient to have a good estimate what percentage of pedestrians turns to the right. This can be either empirically measured or calculated by means of route choice models [44]. In some sense, the uncertainty about the individual behaviors is averaged out at the macroscopic level of description. Nevertheless, we will use the more flexible microscopic simulation approach based on the social force concept. According to this, the temporal change of the location $\boldsymbol{r}_\alpha(t)$ of pedestrian α obeys the equation of motion

$$\frac{d\boldsymbol{r}_\alpha(t)}{dt} = \boldsymbol{v}_\alpha(t)\,. \tag{1}$$

Moreover, if $\boldsymbol{f}_\alpha(t)$ denotes the sum of social forces influencing pedestrian i and if $\boldsymbol{\xi}_\alpha(t)$ are individual fluctuations reflecting unsystematic behavioral variations, the velocity changes are given by the *acceleration equation*

$$\frac{d\boldsymbol{v}_\alpha}{dt} = \boldsymbol{f}_\alpha(t) + \boldsymbol{\xi}_\alpha(t)\,. \tag{2}$$

A particular advantage of this approach is that we can take into account the flexible usage of space by pedestrians, requiring a continous treatment of motion. It turns out that this point is essential to reproduce the observations in a natural and robust way, i.e. without having to adjust the model to each single situation and measurement site. Furthermore, it is interesting to note that, if the fluctuation term is neglected, the social force model can be interpreted as a particular *differential game*, i.e. its dynamics can be derived from the minimization of a special utility function [72].

4 Specification of the Social Force Model

The social force model for pedestrians assumes that each individual α is trying to move in a desired direction $\boldsymbol{e}_\alpha^0$ with a desired speed v_α^0, and that it adapts the actual velocity $\boldsymbol{v}_\alpha$ to the desired one, $\boldsymbol{v}_\alpha^0 = v_\alpha^0 \boldsymbol{e}_\alpha^0$, within a certain relaxation time τ_α. The systematic part $\boldsymbol{f}_\alpha(t)$ of the acceleration force of pedestrian α is then given by

$$\boldsymbol{f}_\alpha(t) = \frac{1}{\tau_\alpha}(v_\alpha^0 \boldsymbol{e}_\alpha^0 - \boldsymbol{v}_\alpha) + \sum_{\beta(\neq\alpha)} \boldsymbol{f}_{\alpha\beta}(t) + \sum_i \boldsymbol{f}_{\alpha i}(t)\,, \tag{3}$$

where the terms $\boldsymbol{f}_{\alpha\beta}(t)$ and $\boldsymbol{f}_{\alpha i}(t)$ denote the repulsive forces describing attempts to keep a certain safety distance to other pedestrians β and obstacles

i. In very crowded situations, additional physical contact forces come into play [73]. Further forces may be added to reflect attraction effects between members of a group or other influences. For details see Refs. [37, 65].

In this contribution, we will assume a simplified interaction force of the form

$$\boldsymbol{f}_{\alpha\beta}(t) = \boldsymbol{f}\big(d_{\alpha\beta}(t)\big)\,, \tag{4}$$

where $\boldsymbol{d}_{\alpha\beta} = \boldsymbol{r}_\alpha - \boldsymbol{r}_\beta$ is the distance vector pointing from pedestrian β to α and $d_{\alpha\beta} = \|\boldsymbol{d}_{\alpha\beta}\|$. (Angular-dependent shielding effects may be taken into account by a prefactor describing the anisotropic reaction to situations in front of as compared to behind a pedestrian [25, 28]).

The distance-dependent force $f(d_{\alpha\beta})$ has been specified in different ways. We will start with the **circular specification** of the interaction force,

$$\boldsymbol{f}(d_{\alpha\beta}) = A_\alpha \mathrm{e}^{-d_{\alpha\beta}/B_\alpha} \frac{\boldsymbol{d}_{\alpha\beta}}{\|\boldsymbol{d}_{\alpha\beta}\|}\,, \tag{5}$$

where A_α and B_α are parameters. A_α reflects the strength of interaction, while B_α corresponds to the interaction range. While the dependence on α explicitly allows for a dependence of these parameters on the single individual, we will assume $A_\alpha = A$ and $B_\alpha = B$ in the following. Otherwise, it would be hard to collect enough data for parameter calibration.

Elliptical specification I: In Ref. [25], a generalization of Eq. (5) was formulated, which assumed that the repulsive potential

$$V_{\alpha\beta}(b_{\alpha\beta}) = AB\,\mathrm{e}^{-b_{\alpha\beta}/B} \tag{6}$$

is an exponentially decreasing function of $b_{\alpha\beta}$ with equipotential lines having the form of an ellipse directed into the direction of motion. The semi-minor axis $b_{\alpha\beta}$ was determined by

$$2b_{\alpha\beta} = \sqrt{(\|\boldsymbol{d}_{\alpha\beta}\| + \|\boldsymbol{d}_{\alpha\beta} - v_\beta\,\Delta t\,\boldsymbol{e}_\beta\|)^2 - (v_\beta\,\Delta t)^2} \tag{7}$$

in order to take into account the length $v_\beta\,\Delta t$ of the stride (step size) of pedestrian β, where $v_\beta = \|\boldsymbol{v}_\beta\|$. The reason for this specification was that pedestrians require space for movement, which is taken into account by other pedestrians.

The repulsive force is related to the repulsive potential via

$$\boldsymbol{f}_{\alpha\beta}(\boldsymbol{d}_{\alpha\beta}) = -\boldsymbol{\nabla}_{\boldsymbol{d}_{\alpha\beta}} V_{\alpha\beta}(b_{\alpha\beta}) = -\frac{dV_{\alpha\beta}(b_{\alpha\beta})}{db_{\alpha\beta}} \boldsymbol{\nabla}_{\boldsymbol{d}_{\alpha\beta}} b_{\alpha\beta}(\boldsymbol{d}_{\alpha\beta})\,, \tag{8}$$

where $\boldsymbol{\nabla}_{\boldsymbol{d}_{\alpha\beta}}$ represents the gradient with respect to $\boldsymbol{d}_{\alpha\beta}$. Considering the chain rule, $\|\boldsymbol{z}\| = \sqrt{\boldsymbol{z}^2}$, and $\boldsymbol{\nabla}_{\boldsymbol{z}}\|\boldsymbol{z}\| = \boldsymbol{z}/\sqrt{\boldsymbol{z}^2} = \boldsymbol{z}/\|\boldsymbol{z}\|$, this leads to the explicit formula

$$\boldsymbol{f}_{\alpha\beta}(\boldsymbol{d}_{\alpha\beta}) = A\mathrm{e}^{-b_{\alpha\beta}/B} \cdot \frac{\|\boldsymbol{d}_{\alpha\beta}\| + \|\boldsymbol{d}_{\alpha\beta} - \boldsymbol{y}_{\alpha\beta}\|}{2b_{\alpha\beta}} \cdot \frac{1}{2}\left(\frac{\boldsymbol{d}_{\alpha\beta}}{\|\boldsymbol{d}_{\alpha\beta}\|} + \frac{\boldsymbol{d}_{\alpha\beta} - \boldsymbol{y}_{\alpha\beta}}{\|\boldsymbol{d}_{\alpha\beta} - \boldsymbol{y}_{\alpha\beta}\|}\right) \tag{9}$$

with $\boldsymbol{y}_{\alpha\beta} = v_\beta\,\Delta t\,\boldsymbol{e}_\beta$. For $\Delta t = 0$, we regain the expression of Eq. (5).

Elliptical specification II: Recently, a variant of this approach has been proposed [28], assuming

$$2b_{\alpha\beta} = \sqrt{(\|\boldsymbol{d}_{\alpha\beta}\| + \|\boldsymbol{d}_{\alpha\beta} - (\boldsymbol{v}_\beta - \boldsymbol{v}_\alpha)\Delta t\|)^2 - [(\boldsymbol{v}_\beta - \boldsymbol{v}_\alpha)\Delta t]^2}\,. \qquad (10)$$

The special feature of this approach is its symmetrical treatment of both pedestrians α and β. Note that further velocity-dependent specifications of pedestrian interaction forces were proposed in the past, but we will restrict to the above specifications, as these are sufficient to demonstrate our evolutionary calibration method.

4.1 Evolutionary Calibration with Video Tracking Data

We have made several video recordings of pedestrian crowds in different natural environments in Budapest (Hungary) and Stuttgart (Germany) from the top. The dimensions of the recorded areas were known, and the floor tiling or environment provided something like a "coordinate system". The heads were automatically determined by seaching for round moving structures, and the accuracy of tracking was improved by comparing actual with linearly extrapolated positions (so it would not happen so easily that the algorithm interchanged or "lost" closeby pedestrians). The trajectories of the heads were then projected on two-dimensional space in a way correcting for distortion by the camera perspective. A representative plot of the resulting trajectories is shown in Fig. 1. It should be noted that extracting trajectory data has been already carried out in the past with infra-red sensors [74] or video recordings [75, 76]. However, our algorithm is less sensitive to confusing different pedestrians, and it can handle more than one thousand pedestrians simultaneously.

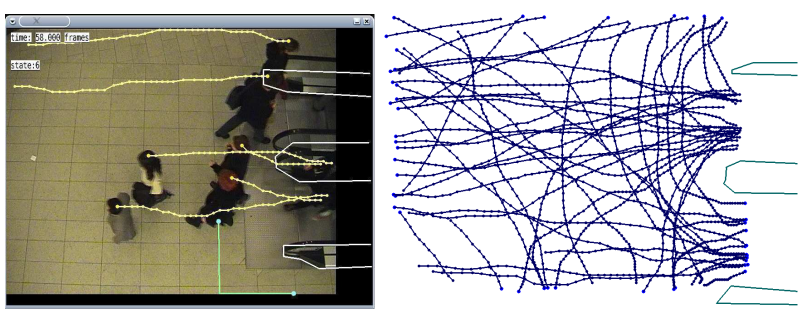


Fig. 1. Video tracking used to extract the trajectories of pedestrians from video recordings close to two escalators (after [28]). Left: Illustration of the tracking of pedestrian heads. Right: Resulting trajectories after being transformed onto the two-dimensional plane.

For the subsequent model calibration, we have used a hybrid method fusing empirical trajectory data and microscopic simulation data of pedestrian movement in space. To each tracked pedestrian, we have assigned a virtual pedestrian in the simulation domain. We have then started a simulation for $T = 1.5$ seconds, in which one pedestrian α was moved according to a simulation of the social force model, while the others were moved exactly according to the trajectories extracted from the videos. This procedure was performed for all pedestrians α and for several different starting times t, using a fixed parameter set for the social force model.

Each simulation run was performed according to the following scheme:

1. Define a starting point and calculate the state (position $\boldsymbol{r}_\alpha$, velocity $\boldsymbol{v}_\alpha$, and acceleration $\boldsymbol{a}_\alpha = d\boldsymbol{v}_\alpha/dt$) for each pedestrian α.
2. Assign a desired speed v_α^0 to each pedestrian. In our simulations, we have specified it by the maximum speed during the pedestrian's tracking time, which is sufficiently accurate, if the overall pedestrian density is not too high and the desired speed is constant in time.
3. Assign a desired goal point for each pedestrian. We have assumed it would correspond to the point at the end of the trajectory.
4. Given the tracked motion of the surrounding pedestrians β, simulate the trajectory of pedestrian α over a time period T based on the social force model, starting at the actual location $\boldsymbol{r}_\alpha(t)$.

After each simulation run, we determined the relative distance error

$$\frac{\|\boldsymbol{r}_\alpha^{\text{simulated}}(t+T) - \boldsymbol{r}_\alpha^{\text{tracked}}(t+T)\|}{\|\boldsymbol{r}_\alpha^{\text{tracked}}(t+T) - \boldsymbol{r}_\alpha^{\text{tracked}}(t)\|} . \tag{11}$$

After averaging the relative distance errors over the pedestrians α and starting times t, 1 minus the result was taken as measure of the goodness of fit (the "fitness") of the parameter set used in the pedestrian simulation. Hence, the best possible value of the "fitness" was 1, but any deviation from the real pedestrian trajectories would imply lower values.

One result of our parameter optimization was that, for each video, there was a broad range of parameter combinations of A and B which performed almost equally well [28]. This allowed us to apply additional goal functions in our optimization. We used this fact to determine among the best performing parameter values such parameter combinations, which performed well for *all three* video recordings, using a fitness function which weighted the fitness reached in each single video equally (i.e. with a factor of 1/3). This is how we determined the common parameter values listed in Table 1. It turns out that, in order to reach a good model performance, the pedestrian interaction force must be specified velocity dependent.

5 Self-Organization of Pedestrian Crowds

Despite its simplifications, the behavioral force model of pedestrian dynamics describes a lot of observed phenomena quite realistically. Especially, it allows one

Model	A	B	"Fitness"
Extrapolation	0	–	0.34
Circular	0.11 ± 0.06	0.84 ± 0.63	0.35
Elliptical I	1.52 ± 1.65	0.21 ± 0.08	0.33
Elliptical II	4.30 ± 3.91	1.07 ± 1.35	0.53

Table 1. Optimal parameter values resulting from our evolutionary parameter optimization for 4 specifications of the interaction forces between pedestrians (see main text). The calibration was based on 3 different video recordings. The values are specified as the mean value ± the standard deviation. Note that, with angular-dependent specifications of the interaction forces, "fitness" values upto 0.61 have been reached [28].

to explain various self-organized spatio-temporal patterns that are not externally planned, prescribed, or organized, e.g. by traffic signs, laws, or behavioral conventions. Instead, the spatio-temporal patterns discussed below emerge due to the non-linear interactions of pedestrians even without assuming strategical considerations, communication, or imitative behavior of pedestrians. Despite this, we may still interpret the forming patterns as phenomena that establish coordination and cooperation, i.e. social order on short time scales. It is actually surprising that strangers coordinate each other within seconds, if they have grown up in a similar enviroment. People from different countries, however, are sometime irritated about local walking habits, which indicates that learning effects and cultural backgrounds still play a role in social interactions as simple as random pedestrian encounters. Rather than on particular features, however, we will focus on the common, internationally reproducible observations in the following.

5.1 Lane Formation

In pedestrian flows one can often observe that oppositely moving pedestrians are forming lanes of uniform walking direction (see Fig. 2). This phenomenon even occurs when there is not a large distance to separate each other, e.g. on zebra crossings. However, the width of lanes increases (and their number decreases), if the interaction continues over longer distances (and if perturbations, e.g. by flows entering or leaving on the sides, are low; otherwise the phenomenon of lane formation may break down [26]).

Lane formation may be viewed as *segregation phenomenon* [77, 78]. Although there is a weak preference for one side (with the corresponding behavioral convention depending on the country), the observations can only be well reproduced when repulsive pedestrian interactions are taken into account. The most relevant factor for the lane formation phenomenon is the higher relative velocity of pedestrians walking in opposite directions. Compared to people following each other, oppositely moving pedestrians have more frequent interactions until they have segregated into separate lanes by stepping aside whenever another pedestrian is

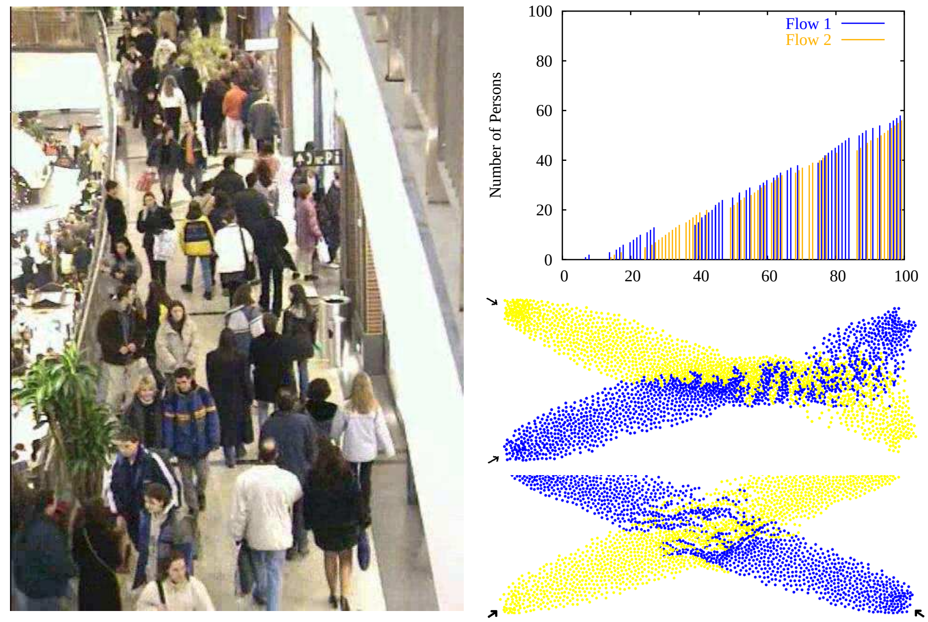


Fig. 2. Self-organization of pedestrian crowds. Left: Photograph of lanes formed in a shopping center. Computer simulations reproduce the self-organization of such lanes very well. Top right: Evaluation of the cumulative number of pedestrians passing a bottleneck from different sides. One can clearly see that the narrowing is often passed by groups of people in an oscillatory way rather than one by one. Bottom right: Multi-agent simulation of two crossing pedestrian streams, showing the phenomenon of stripe formation. This self-organized pattern allows pedestrian to pass the other stream without having to stop, namely by moving sidewards in a forwardly moving stripe.

encountered. The most long-lived patterns of motion are the ones which change the least. It is obvious that such patterns correspond to lanes, as they minimize the frequency and strength of avoidance maneuvers. Interestingly enough, as computer simulations show, lane formation occurs as well when there is no preference for any side.

Since lanes minimize frictional effects and delays in oppositely moving crowds, one could say that they are a pattern reflecting "collective intelligence". In fact, it is not possible for a single pedestrian to reach such a collective pattern of motion. Lane formation is a self-organized collaborative pattern of motion originating from simple pedestrian interactions. Particularly in cases of no side preference, the system behavior cannot be understood by adding up the behavior of the single individuals. This is a typical feature of complex, self-organizing systems and, in fact, a wide-spread characteristics of social systems. It is worth noting, however, that it does not require a conscious behavior to reach forms of social organization like the segregation of oppositely moving pedestrians into lanes. This organization occurs automatically, and most people are not even aware of the existence of this phenomenon.

5.2 Oscillatory Flows at Bottlenecks

At bottlenecks, bidirectional flows of moderate density are often characterized by oscillatory changes in the flow direction (see Fig. 2). The authors have observed this, for example, at entrances of museums during crowded arts exhibitions and at entrances of mensas during lunch time. While these oscillatory flows may be interpreted as an effect of friendly behavior ("you go first, please"), computer simulations of the social force model indicate that the collective behavior may again be understood by simple pedestrian interactions. That is, oscillatory flows occur even in the absence of communication. Therefore, they may be viewed as another self-organization phenomenon, which again reduces frictional effects and delays. That is, oscillatory flows have features of "collective intelligence". While this may be interpreted as result of a learning effect in a large number of similar situations (a "repeated game"), our simulations suggest an even simpler, "many-particle" interpretation: Once a pedestrian is able to pass the narrowing, pedestrians with the same walking direction can easily follow. Hence, the number and "pressure" of waiting, "pushy" pedestrians on one side of the bottleneck becomes less than on the other side. This eventually increases their chance to occupy the passage. Finally, the "pressure difference" is large enough to stop the flow and turn the passing direction at the bottleneck. This reverses the situation, and eventually the flow direction changes again, giving rise to oscillatory flows.

5.3 Stripe Formation in Intersecting Flows

In intersection areas, the flow of people often appears to be irregular or "chaotic". In fact, it can be shown that there are several possible collective patterns of motion, among them rotary and oscillating flows. However, these patterns continuously compete with each other, and a dominating patterns is destroyed by another one after a short time. Obviously, there has not evolved any social convention that would establish and stabilize an ordered and efficient flow.

Self-organized patterns of motion, however, are observed in situations where pedestrian flows intersect each other only in two directions. In such situations, the phenomenon of stripe formation is observed [79]. Stripe formation allows two flows to penetrate each other without requiring the pedestrians to stop. For an illustration see Fig. 2. Like lanes, stripes are a segregation phenomenon, but not a stationary one. Instead, the stripes are density waves moving into the direction of the sum of the directional vectors of both intersecting flows. Naturally, the stripes extend sidewards into the direction which is perpendicular to their direction of motion. Therefore, the pedestrians move forward with the stripes and sidewards within the stripes. Lane formation corresponds to the particular case of stripe formation where both directions are exactly opposite. In this case, no intersection takes place, and the stripes do not move systematically. As in lane formation, stripe formation allows to minimize obstructing interactions and to maximize the average pedestrian speeds, i.e. simple, repulsive pedestrian interactions again lead to an "intelligent" collective behavior.

5.4 Self-Organization in Dense Crowds and Learning

At high densities, one can observe additional self-organization phenomena. This includes intermittent clogging phenomena at bottlenecks ("faster-is-slower effect") [73, 80], the formation of stop-and-go waves [80, 81], and the occurence of "turbulent" flows [81]. These collective behaviors can also be understood as results of pedestrian interactions, but more complicated ones. Some of these phenomena show features of a breakdown of social order, particularly in situations of "crowd panics" [73, 81]. For most people, these situations are so special that they do not have previous experience with them. Hence, there are normally no opportunities to learn an optimal behavior. This is in marked contrast to the situations a pedestrian faces every day. For these, it can be assumed that pedestrians (in their childhood) have learned a behavior that minimizes collisions and delays. This learning process can be simulated [56, 65]. The result is a surprisingly predictable behavior, which can be reflected by a mathematical model such as the social force model with suitably specified forces and parameters.

6 Summary, Discussion, and Outlook

In this contribution, we have presented a multi-agent approach to pedestrian and crowd dynamics. Pedestrian interactions are simple, but easy to measure, and they show a great variety of self-organized patterns and short-lived social phenomena, where coordination or cooperation emerges spontaneously. For this reason, they are interesting to study, particularly as we expect new insights into coordination mechanisms of social beings beyond the scope of classical game theory. Examples for observed and simulated self-organization phenomena are lane formation, stripe formation, oscillations and intermittent clogging effects at bottlenecks, and the evolution of behavioral conventions (such as the preference of the right-hand side in continental Europe). As similar observations are made in other systems, we expect that our findings may also promote the understanding of opinion formation and other kinds of collective behaviors. Our hope is that, based on the discovered elementary mechanisms of emergence and self-organization, one can eventually also obtain a better understanding of the constituting principles of more complex social systems. At least the same underlying factors are found in many social systems: non-linear interactions of individuals, time-dependence, heterogeneity, stochasticity, competition for scarce resources (here: space and time), decision-making, and learning. Our future work will also address issues of perception, anticipation, and communication.

Here, we have specifically discussed how to calibrate a multi-agent model of pedestrian behavior based on empirical data determined from video tracking. While the elliptical model II performed well, one surprising conclusion from Table 1 is that the circular model and the elliptical model I do not provide better predictions of pedestrian trajectories than a simple extrapolation. The latter assumes no interaction effects at all, but a continuation of pedestrian motion with the previous velocity ("null model"). Nevertheless, all three specifications of pedestrian interactions reproduce lane formation, oscillatory flows, and stripe formation in pedestrian crowds, in contrast to the null model. This suggests several conclusions:

1. A model which, based on some error measure, has been assessed to be not significantly better than another model may still perform better when judged with other criteria. Therefore, it is essential to find not only a good model, but also a suitable error measure.
2. One problem of calibrating multi-agent simulation models [82–84] is that there is currently no established procedure how to determine a suitable error measure. Conventional error measures can be insufficient to differentiate between models which are capable of describing self-organization phenomena and those which are not.
3. Self-organized, collective behavior is possible despite of weak interactions, since the non-linear interactions between pedestrians persist over a certain time period, which leads to feedback and re-inforcement effects.

These aspects will be addressed in more detail in future publications. The same applies to *controlled experiments* with pedestrians that have been recently performed to check the social force model itself [85]. Evaluations of these experiments have particularly shown that a model with an angular dependence of pedestrian interactions performs better than the models discussed above. The improved model reproduces the observed self-organization effects in pedestrian crowds well, but it is also in good *quantitative* agreement with individual pedestrian behavior. In conclusion, human interactions in space and time are currently measured in detail, and they can be formulated in terms of mathematical models. Multi-agent simulations of such models do not only reproduce the original data well, but they also provide a deeper understanding of observed self-organization phenomena and features of "collective intelligence". Further studies in this field appear to be promising.

Acknowledgments

The authors are grateful for partial financial support by the German Research Foundation (research projects He 2789/7-1, 8-1) and by the "Cooperative Center for Communication Networks Data Analysis", a NAP project sponsored by the Hungarian National Office of Research and Technology under grant No. KCKHA005.

References

1. W. Weidlich and G. Haag, Concepts and Models of a Quantitative Sociology. The Dynamics of Interacting Populations (Springer, Berlin, 1983).
2. D. Helbing, Quantitative Sociodynamics (Kluwer Academic, Dordrecht, 1995).
3. K. G. Troitzsch, U. Mueller, G. N. Gilbert, and J. E. Doran (eds.), Social Science Microsimulation (Springer, 1996).
4. J. M. Epstein and R. L. Axtell: Growing Artificial Societies: Social Science from the Bottom Up (Brookings Institution Press, 1996).

5. R. M. Axelrod: The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration (Princeton University, 1997).
6. W. B. G. Liebrand, A. Nowak, and R. Hegselmann (eds.) Computer Modeling of Social Processes (Sage, 1998).
7. Richard J. Gaylord and Louis J. D'Andria: Simulating Society: A Mathematica Toolkit for Modeling Socioeconomic Behavior (Springer, 1998).
8. G. Ballot and G. Weisbuch (eds.) Applications of Simulation to Social Sciences (Hermes Science, 2000).
9. N. Gilbert and K. G. Troitzsch: Simulation for the Social Scientist (Open University Press, 1999).
10. N. J. Saam and A. Harrer, JASSS 2(1), 2 (1999).
11. W. Thorngate, JASSS 3(1), 1 (2000).
12. W. Weidlich, Sociodynamics: A Systemic Approach to Mathematical Modelling in the Social Sciences (Harwood Academic, Amsterdam, 2000).
13. A. Lomi and E. R. Larsen (eds.), Dynamics of Organizations: Computational Modeling and Organizational Theories (AAAI Press, 2001).
14. D. Helbing, Behavioral Science 36, 298–310 (1991).
15. M. W. Macy and A. Flache, Proceedings of the National Academy of Sciences USA **99**, Suppl. 3, 7229–7236 (2002).
16. F. Schweitzer, Brownian Agents and Active Particles: Collective Dynamics in the Natural and Social Sciences (Springer, 2003).
17. F. C. Billari, T. Fent, A. Prskawetz, and J. Scheffran (eds.) Agent-Based Computational Modelling (Physica, Heidelberg, 2006).
18. P. Davidsson, JASSS 5(1), 7 (2002).
19. R. Hegselmann and A. Flache, JASSS 1(3), 1 (1998).
20. F. S. Beltran, L. Salas, and V. Quera, JASSS 9(3), 5 (2006).
21. A. C. Harvey and A. Jaeger, J. Applied Econometrics 8(3), 231–247 (1993).
22. D. Challet, M. Marsili, and Y.-C. Zhang, Physica A 294(3/4), 514–524 (2001).
23. D. Fudenberg and J. Tirole, Game Theory (MIT Press, 1991).
24. M. J. Osborne and A. Rubinstein, A Course in Game Theory (MIT Press, 1994).
25. D. Helbing and P. Molnár, Physical Review E 51, 4282–4286 (1995).
26. D. Helbing, I. Farkas, and T. Vicsek, Physical Review Letters 84, 1240–1243 (2000).
27. P. Molnár, *Modellierung und Simulation der Dynamik von Fußgängerströmen*, (Shaker, Aachen, 1996).
28. A. Johansson, D. Helbing, and P. K. Shukla, Advances in Complex Systems, in print (2007).
29. B. D. Hankin and R. A. Wright, Operational Research Quarterly 9, 81–88 (1958).
30. S. J. Older, Traffic Engineering & Control 10, 160–163 (1968).
31. U. Weidmann, Transporttechnik der Fußgänger, (Institut für Verkehrsplanung, Transporttechnik, Straßen- und Eisenbahnbau, ETH Zürich, 1993).
32. M. R. Hill, *Walking, Crossing Streets, and Choosing Pedestrian Routes* (University of Nebraska, Lincoln, 1984).
33. M. Batty, Nature 388, 19–20 (1997).
34. J. J. Fruin, *Designing for pedestrians: A level-of-service concept*, in *Highway Research Record, Number 355: Pedestrians*, pp. 1–15 (Highway Research Board, Washington, D.C., 1971).
35. J. Pauls, Fire Technology 20, 27–47 (1984).
36. W. H. Whyte, City. Rediscovering the Center (Doubleday, New York, 1988).
37. D. Helbing, Verkehrsdynamik (Springer, Berlin, 1997).
38. D. Helbing, L. Buzna, A. Johansson, and T. Werner, Transportation Science 39(1), 1-24 (2005).

39. W. M. Predtetschenski and A. I. Milinski, *Personenströme in Gebäuden – Berechnungsmethoden für die Projektierung* – (Rudolf Müller, Köln-Braunsfeld, 1971).
40. Transportation Research Board, *Highway Capacity Manual*, Special Report 209 (Transportation Research Board, Washington, D.C., 1985).
41. S. J. Yuhaski Jr., J. M. Macgregor Smith, Queueing Systems 4, 319–338 (1989).
42. D. Garbrecht, Traffic Quarterly 27, 89–109 (1973).
43. N. Ashford, M. O'Leary, and P. D. McGinity, Traffic Engineering & Control 17, 207–210 (1976).
44. A. Borgers and H. Timmermans, Socio-Economic Planning Science 20, 25–31 (1986).
45. D. Helbing, *Stochastische Methoden, nichtlineare Dynamik und quantitative Modelle sozialer Prozesse*, Ph.D. thesis (University of Stuttgart, 1992, published by Shaker, Aachen, 1993).
46. D. Helbing, M. Isobe, T. Nagatani, and K. Takimoto, *Physical Review E* 67, 067101 (2003).
47. W. Daamen and S. P. Hoogendoorn, in *Proceedings of the 82nd Annual Meeting at the Transportation Research Board* (CDROM, Washington D.C., 2003).
48. M. Isobe, D. Helbing, and T. Nagatani, *Physical Review E* 69, 066132 (2004).
49. A. Seyfried, B. Steffen, W. Klingsch, and M. Boltes, *J. Stat. Mech.* P10002 (2005).
50. T. Kretz, M. Wölki, and M. Schreckenberg, *J. Stat. Mech.* P02005 (2006).
51. L. F. Henderson, Transportation Research 8, 509–515 (1974).
52. R. L. Hughes, Transportation Research B, in print (2001).
53. D. Helbing, Complex Systems 6, 391–415 (1992).
54. S. P. Hoogendoorn and P. H. L. Bovy, Transportation Research Records 1710, 28–36 (2000).
55. P. G. Gipps and B. Marksjö, Math. Comp. Simul. 27, 95–105 (1985).
56. K. Bolay, *Nichtlineare Phänomene in einem fluid-dynamischen Verkehrsmodell* (Master's thesis, University of Stuttgart, 1998).
57. V. J. Blue and J. L. Adler, Transportation Research Records 1644, 29–36 (1998).
58. M. Fukui and Y. Ishibashi, Journal of the Physical Society of Japan 68, 2861–2863 (1999).
59. M. Muramatsu, T. Irie, and T. Nagatani, Physica A 267, 487–498 (1999).
60. H. Klüpfel, M. Meyer-König, J. Wahle, and M. Schreckenberg, in S. Bandini and T. Worsch (eds.) *Theory and Practical Issues on Cellular Automata* (Springer, London, 2000).
61. C. Burstedde, K. Klauck, A. Schadschneider, and J. Zittartz, Physica A 295, 507–525 (2001).
62. S. Gopal and T. R. Smith, in M. M. Fischer, P. Nijkamp, and Y. Y. Papageorgiou (eds.) *Spatial Choices and Processes* (North-Holland, Amsterdam, 1990), pp. 169–200.
63. C. W. Reynolds, in D. Cliff, P. Husbands, J.-A. Meyer, and S. Wilson (eds.) From Animals to Animats 3: Proceedings of the Third International Conference on Simulation of Adaptive Behavior (MIT Press, Cambridge, Massachusetts, 1994), pp. 402–410.
64. D. Helbing, Behavioral Science 37, 190–214 (1992).
65. D. Helbing, P. Molnár, I. Farkas, and K. Bolay, Environment and Planning B 28, 361-383 (2001).
66. J. Klockgether and H.-P. Schwefel, in D. G. Elliott (ed.) Proceedings of the Eleventh Symposium on Engineering Aspects of Magnetohydrodynamics (California Institute of Technology, Pasadena, CA, 1970), pp. 141–148.

67. D. Helbing, in G. Haag, U. Mueller, and K. G. Troitzsch (eds.) Economic Evolution and Demographic Change. Formal Models in Social Sciences (Springer, Berlin, 1992), pp. 330–348.
68. N. E. Miller, in J. McV. Hunt (ed.) Personality and the behavior disorders, Vol. 1, (Ronald, New York, 1944).
69. N. E. Miller, in S. Koch (ed.) *Psychology: A Study of Science*, Vol. 2 (McGraw Hill, New York, 1959).
70. K. Lewin, Field Theory in Social Science (Harper & Brothers, New York, 1951).
71. D. Helbing, Journal of Mathematical Sociology 19(3), 189–219 (1994).
72. S. Hoogendoorn and P. H. L. Bovy, Optimal Control Applications and Methods 24(3), 153–172 (2003).
73. D. Helbing, I. Farkas, and T. Vicsek, Nature 407, 487–490 (2000).
74. J. Kerridge and T. Chamberlain, in N. Waldau, P. Gattermann, H. Knoflacher, and M. Schreckenberg (eds.) Pedestrian and Evacuation Dynamics '05 (Springer, Berlin, 2005).
75. S. P. Hoogendoorn, W. Daamen, and P. H. L. Bovy, in Proceedings of the 82nd Annual Meeting at the Transportation Research Board (CDROM, Mira Digital Publishing, Washington D.C., 2003).
76. K. Teknomo, Microscopic pedestrian flow characteristics: Development of an image processing data collection and simulation model (PhD thesis, Tohoku University Japan, Sendai, 2002).
77. T. Schelling, Journal of Mathematical Sociology 1, 143–186 (1971).
78. D. Helbing and T. Platkowski, International Journal of Chaos Theory and Applications 5(4), 47–62 (2000).
79. K. Ando, H. Oto and T. Aoki, Railway Research Review 45 (8), 8-13 (1988).
80. D. Helbing, A. Johansson, J. Mathiesen, M.H. Jensen and A. Hansen, Physical Review Letters 97, 168001 (2006).
81. D. Helbing, A. Johansson and H. Z. Al-Abideen, Physical Review E, in print (2007).
82. P. Windrum, G. Fagiolo, and A. Moneta, JASSS 10(2), 8 (2007).
83. R. Boero and F. Squazzoni, Does empirical embeddedness matter? Methodological issues on agent-based models for analytical social science, JASSS 8(4), 6 (2005).
84. R. Leombruni, M. Richiardi, N. J. Saam, and M. Sonnessa, TTA common protocol for agent-based social simulation. in K. G. Troitzsch (ed.) Representing Social Reality: Approaches and Results (Proceedings of the 3rd Annual Conference of the European Social Simulation Association, Koblenz, September 5–9, 2005).
85. M. Moussaid, S. Garnier, D. Helbing, and G. Theraulaz, Internal Working Paper (2007).

Emergence in Social Networks: Modeling the Intentional Properties of Multi-Agent Systems

Jorge Louçã[1], John Symons[2], David Rodrigues[1], André Morais[1]

[1] ISCTE / LabMAg, Av.das Forças Armadas, 1649-026 Lisboa, Portugal
Jorge.L@iscte.pt, {david.rodrigues, anoeee}@gmail.com

[2] University of Texas at El Paso, 500 W.University Ave., El Paso, Texas 79968, USA
jsymons@utep.edu

Abstract. This paper examines some methodological and conceptual challenges concerning the study of emergent patterns of behavior in communication networks. We discuss the notion of intentionality in collectives in the light of two case studies. The first one regards the mechanisms of communication in a community of insects. The second case study concerns the "smart mobs" phenomena in human societies. The focus of our discussion is the detection of what we call *emergent fingerprints*, which signal the presence of a system of agents behaving in a manner identical to a social structure whose intentional behavior under some known set of constraints we have modeled. These structures are used to attribute some intentionality to social systems, even in the presence of noisy data sets.

Keywords: agent-based model, pattern-oriented modelling, social simulation, communication patterns.

1 Introduction: Intentionality in Collectives

Social networking technologies have given rise to new social dynamics and distinctive kinds of collective behavior. This paper describes how collective behaviors can be modeled and examines some of the associated methodological and conceptual challenges.

We assume an anti-reductionist conception of collective behavior, recognizing, for instance, the effect of the social context on the behavior of its constituents. For instance, as Bibb Latané (8) argues, when people engage in group activity, their sense of individual responsibility attenuates and their behavior changes accordingly. Cases where we observe a downward causal effect of the social context on its constituents support a distinction between collective and individual intentionality (14). We therefore assume that socially networked behavior is not merely the composition of constituent individuals.

Understanding the emergent features of group behavior has obvious practical relevance. Take for instance Howard Rheingold's (10) notion of a smart mob. Rheingold introduced the expression "smart mob" to describe the concept of a "mobile ad hoc social network". Smart mobs are social networks where people communicate using mobile and wireless internet technologies. As a first introduction to the intentional character of smart mobs consider the relationship between a network of friends and some scarce and desirable resource. One important reason to have friends is that it allows one to share in a larger store of relevant information with respect to valuable resources. Technologically extended social networks amplify and alter the effect of familiar friendship relations in dramatic ways. For example, the press has reported several cases of smart mobs aiming to exploit scarce resources. The June 29th 2006 edition of *The Economist* described the example of Guangzhou, where 500 shoppers assembled in front of a popular electronics superstore. They arrived *en masse*, at a given moment previously agreed to online and by cellphone messages. Shoppers left the store having secured 10-30% discounts on cameras, DVD players and flat-screen televisions. Smart mobs of this kind play are an obvious and easily explained example of collective action in pursuit of bargaining power.

Smart mobs are becoming increasingly familiar for their role in social and political expression. Street demonstrators at the 1999 anti-WTO protests in Seattle used dynamically updated websites, cellphones, and swarming tactics (13). Demonstrators using SMS were able to quickly converge on a single place from highly dispersed positions. They overwhelmed the local police presence with brief but intense protests, quickly dispersing and blending back into the crowds before the police could reallocate their forces. More recently, SMS communication was used to organize mass protests all over Spain in the aftermath of the Madrid train bombings of March 11th 2004 (6). Again, viral communication strongly spread through social networks mainly composed of friends, where trust between members of the network is extremely high.

The increasing importance of networked forms of political, cultural and economic expression were noted by Tony Negri and Michael Hardt in their study of globalization *Empire* (7). While Negri and Hardt's work is highly impressionistic and imprecise, it correctly points to the possibility that networked activity (what they call 'the multitude') can give rise to radically new political and economic changes. Given the increasingly pervasive role of networking technologies, it is highly probable that collective activity like smart mobs will play an increasingly important role.

In our analysis, we will point towards the possibility that collective behavior exhibits properties which are not as evident as those exhibited by the shoppers in Guanghzhou or the demonstrators in Seattle. In those cases, both the emergent feature and the explanation of those features are relatively obvious. Investigating emergent properties can be more than simply attempting to explain obvious emergent features of groups. Instead, our approach opens the possibility that previously unnoticed emergent features can be discovered in complex systems and that these features can play a role in explanations at various levels of analysis.

2 Outline and Assumptions

Our principal goals involve identifying emergent patterns of behavior in a social environment. To this end, we explore the practicality of adopting what Daniel Dennett has called 'the intentional stance' towards collective behaviors (1). We test several related research hypotheses via the modeling and simulation of case studies in a variety of social contexts, arguing that modeling such systems both reveals and explains emergent patterns in communication. The focus of our discussion is the detection of what we are calling *emergent fingerprints*, which signal the presence of a system of agents behaving in a manner identical to a social structure whose behavior under some known set of constraints we have modeled. We begin with relatively simple simulated social environments. From these simulations we extract emergent properties of the behavior of some group of agents which we then treat as a template for pattern-matching within the massive and noisy data sets which are derived from the study of natural systems.

To study this hypothesis, we propose to distinguish patterns or forms that arise by virtue of constraints on the flow of information. In our modeling projects, we identify sequences and combinations of those forms. This approach will lead us to interpret and predict emergent properties of various forms of social systems, even in the presence of massive and noisy data sets.

We build our models on the assumption that knowing only the links between the nodes of the communications network is enough to characterize the flow of information in a social environment. One advantage of the assumption of this hypothesis is its independence from the semantic properties of communication between agents. While most current treatments of emergent features of group knowers involve propositional or semantic features of the systems under consideration, our strategy is to examine patterns that appear at the level of behavioral patterns. We anticipate that social structures in which we can understand the characteristics and constraints on communication will exhibit behavioral features when modeled. We predict that some of these patterns will be robust enough to detect in real world settings.

Aiming to test these hypotheses, our research derives from two main scientific domains: the well known field of multi-agent based simulation (4) and the research that has recently been labelled pattern-oriented modeling (5).

This paper presents the broad trajectory of our research, starting with the case study *Cicada barbara lusitanica*, detailed in a previous text (9). From this initial study concerning the mechanisms of communication in a community of insects, we have developed another case study, concerning the emergence of self-organized social structures in human communication.

We provide a brief characterization of the theoretical framework supporting our research before discussing its relevance to the understanding of social phenomena. Next, our initial case study from zoology is discussed as the basis for the implementation of a multi-agent simulation environment illustrating some of the general features of our approach. In the simulations, we compare the input communication patterns and resulting emergent movement patterns from different species of cicadas. The second case study concerns a multi-agent model of communication and movement, illustrating the social dynamics of the "smart mobs" phenomena. This case study allows the identifying of patterns, both at the micro and macro levels of observation. Finally we discuss open questions and directions for further research.

3 Grasping Macro Intentionality

In our work, we take a commonsense attitude towards the relationship between individuals and groups. Contrary to a rigidly reductionist approach, we assume that the constituents of social structures and the social structures themselves mutually influence and transform one another. In this tangled and complex circumstance, some features will strike observers as stable and significant. Such features of human life or of complex systems are not necessarily (or generally) amenable to reductive explanation. Such features are sometimes called emergent. Just because patterns are not easily reduced to the behavior of their constituents, does not mean that they are somehow mysterious or that they cannot be explained scientifically. They can often be understood in terms of the kinds of mechanisms that computational modeling reveals (15). In fact, our approach supports the possibility that through modeling, not only are we in a position to explain previously puzzling emergent properties, but perhaps more interestingly, our models allow us to discover previously unknown emergent patterns in complex systems of various kinds.

In ordinary human affairs, we assume that other people have beliefs and desires and that their actions can be explained in terms of those beliefs and desires. Daniel Dennett (1) called this pattern of belief-desire-action explanation folk psychology and argued that when we engage in folk psychological reasoning, we are adopting the intentional stance towards our fellow human beings (14).

For Dennett, when we take the intentional stance towards something or someone, we project the virtual world of beliefs and desires onto the other person or animal in somewhat the same way a geographer might project lines of latitude and longitude onto the Earth's surface. In both cases, the projections permit us a means of manipulating the objects in questions and in both cases the question of whether these virtual objects *really exist* is misguided. The intentional stance is a strategy that begins with the assumption that other animals believe what they should believe given their perceptions and desire what they should desire given their needs. This is what Dennett calls the assumption of optimal design. We assume that other animals (including people) tend to pursue outcomes that serve their interests and that they have been equipped, by natural selection with suitable perceptual and cognitive capacities to manipulate their environments appropriately. Ascriptions of beliefs and desire are often objectively true, he grants, but not by virtue of describing inner mechanisms, any more than references to centers of gravity, vectors, equators and other useful virtual notions.

The most distinctive aspect of Dennett's approach to the mind is his choice of starting point. Rather than diving immediately into traditional questions about the nature of mental entities, Dennett begins by investigating what we ordinarily say or write about other minds. Conversation or language, Dennett writes, is the royal road to the knowledge of other minds (2, p13). So, rather than immediately focusing on the *terra incognita* of what the mind actually is, he focuses instead on the practical role played by the terms we use to talk about the mind, terms like 'belief', 'desire', 'the will', 'consciousness' etc. Dennett's interest lies in understanding the ordinary commonsense descriptions of other minds that we find so useful in our daily lives. What are we doing, for instance, when we say that another person believes, desires or imagines something? And why does our talk of what people think, want, or hope, seem to work so well in ordinary life?

Traditionally, philosophers have been interested in understanding the nature of "belief" rather than the process of interpretation by which we come to know the beliefs of another person or animal. Generally, they began with the assumption that words like 'belief', 'desire' and the like refer to mental entities of some kind. They were not overly concerned with the ways we come to know those mental entities, since presumably, even if we encounter some difficulty when it comes to other minds, our own minds are easy enough to examine. Dennett does not assume that these mental entities exist. Instead, his approach is to remain agnostic (at least initially) towards the meaning of psychological terms like 'belief', 'desire', etc. Rather than assuming that the word 'belief' refers to some kind of entity in the mind or brain, he focuses instead on the role such terms play in practice. In effect, this means that he begins by considering the reasons we use words like 'belief' and 'desire' rather than the nature of belief and desire *per se*. By focusing on what we do when we interpret or describe the mental life of another person or animal, Dennett redirects our attention from traditional philosophical puzzles and opens fertile new lines of inquiry for both scientists and philosophers.

This subtle shift in perspective changes the contour of philosophical problems considerably. Instead of worrying about the mysterious inner workings of other minds, or the problem of how the mind and body interact, we look instead to the publicly observable utterances we use to describe one another. These utterances, along with the contexts in which they appear, form an objective starting point for a more scientific approach to the problems posed by mental life. In principle, it should be possible to study the relationship between these patterns of utterances and the underlying processes that take place in the brain and central nervous system.

According to Dennett, our use of psychological terms arises primarily out of our interest in predicting the behavior of people and other animals. For obvious reasons, animals need a way to anticipate the behavior of possible mates, predators, prey and competition. This is accomplished in a variety of ways in different species. Naturally, we humans are most familiar with the techniques employed by mammals like ourselves. While mammals have an array of techniques for predicting behavior, humans tend to be especially biased towards visual cues of various kinds. By watching the eyes, posture and motion of other mammals we can generally predict their behavior into the very near future with some accuracy. So, how do we manage these remarkable prophecies? Is it a matter of knowing the mind of the other person or animal?

For Dennett, Darwin's theory of evolution through natural selection provides the key to understanding our amazing ability to predict the behavior of other animals. Over the course of natural history, those who fail to reliably anticipate the future behavior of relevant others soon perish. As a result, a significant number of species have evolved to become masters in the art of predicting the futures of massively complex biological systems. For example, without knowing anything about the physiological processes taking place under the animal's skin or in its mind, most of us can predict, with relative certainty, what a hungry dog is likely to do when we offer it a bowl of food. Our ability to predict the behavior of other biological systems is not the result of an ability to somehow see inside the minds of other animals. Instead, according to Dennett, our interpretations are the product of a skill that has been sculpted by a long process of natural selection. Our ability to accurately predict the future behavior of other creatures based on the evidence of their past behavior is a skill that we, along with many other animals, simply inherited.

By contrast, we humans are not suited by natural selection to predict the behavior of collectives in complex, technologically-mediated social contexts. The simple fact that the new social context results from communications technologies which postdate the selective pressures on our species means that technological enhancement is needed in order for us to adopt something analogous to the intentional stance in such cases. It is also worth noting that language, Dennett's royal road, drops out of the picture in the cases that are of interest to us.

The case studies presented in the following sections exemplify two very different examples of social interaction. Nevertheless, they both illustrate that the attribution of intentionality to groups in a social network allow the explanation of social dynamics via the only detection of emergent behavioral fingerprints. Computational modeling allows us to understand how the collective behavior of the system is shaped by constraints on communication i.e., without any access to the semantic content of that information. As such, this perspective allows deep insight into the study of social behavior. In addition to offering insight which was previously unavailable to us, this approach creates many new questions and opens up many new lines of research.

4 *Cicada barbara lusitanica* case study

Our first case study rests on experimental results obtained at the Faculty of Sciences at the University of Lisbon, concerning live experiments on the stereotyped singing response behaviors of cicadas (3, 9). We used their results to design a multi-agent simulation platform which allowed us to represent the behavior of two cicada species, *Cicada barbara lusitanica* and *Cicada orni*.

4.1 Previous experimental results

Previous experimental results concerning the study of cicada communication (2) allowed us to conclude that insect songs encode specific information about the identity of species, which are used by individuals to discriminate conspecific from heterospecific sympatric species. Figures 1 and 2 depict examples of temporal calling song configurations. The *Cicada barbara* song, characterized by an uninterrupted pulse, illustrates a continuous pattern (figure 1). By contrast, the *Cicada orni* song shows a clear discontinuous pattern (figure 2).

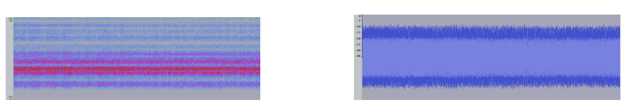


Fig. 1. *Cicada barbara* calling song analysis: spectrum and waveform db.

Figure 1 shows that the *Cicada barbara* calling song has a broad spectrum (image on the left), although some frequencies are more intense than others. The waveform analysis shows that the calling song is continuous, with no pauses.

Fig. 2. *Cicada orni* calling song analysis: spectrum and waveform db.

The *Cicada orni* spectrum depicts a clearly discontinuous calling song. This particular species of cicada uses a calling song characterized by pulses and pauses. The spectrum occupies a narrower band than that of *Cicada barbara*. Nevertheless, it is very intense in some frequencies. The waveform represents the discontinuity in the calling and regular pauses (e.g., silence) are evident.

Concerning frequency, the analysis of a pulse period in *Cicada orni* calling song reveals that very few frequencies are significant. By contrast, *Cicada barbara* might have to produce a broader spectrum of frequencies as it will not use temporal patterns (the song is continuous). Fonseca and Revez conclude that such configurations of calls are generally used by cicadas to identify members of their own species.

Fonseca and Revez stated that both frequency spectrum of the signal and temporal pattern carries information about the species-identity of a calling male, but the use of only one parameter might not be sufficient (3). The pre-copulatory isolating mechanism based on song analysis, used to maintain species integrity, uses one or/and another parameter according to the species environment.

4.2 Multi-agent simulation of cicada communication behaviors

Fonseca and Revez experiments and conclusions were used to implement a multi-agent simulation of cicada communication behaviors. The model considers the existence of two patterns used by *Cicada barbara* and *Cicada orni* to recognize conspecifics: temporal (pulse and pause duration) and frequency patterns. *Cicada barbara* and *Cicada orni* individuals are randomly placed in a finite environment. Males are static and females move, attracted by songs from their conspecific males. Each female cicada has a given initial energy, which she expends by traveling through the environment. When a female meets a male, two things can occur: either they belong to the same species and in this case her energy is set to a maximum level, or they belong to different species and then her energy is set to a minimum level. The key issue under consideration in the model is the pattern-matching mechanism used by females to recognize conspecific song. Several simulation experiments were modeled and executed using temporal patterns, frequency patterns separately and simultaneously.

Results of the simulations are illustrated by the images in figure 3. The image on the left compares the number of contacts achieved by *Cicada barbara* individuals in three situations: using the temporal pattern uniquely, using the frequency pattern and using both temporal and frequency patterns. The image on the right presents the same comparison concerning the number of contacts achieved by *Cicada orni* individuals, using its species specific pattern values.

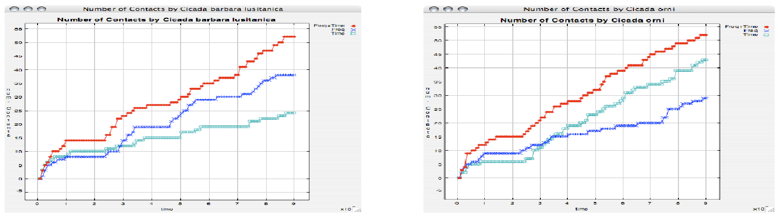


Fig. 3. Running the simulation with different singing recognition patterns used by *Cicada barbara* (left) and *Cicada orni* (right).

From the observation of graphs, we can conclude that, in *Cicada barbara* species the frequency pattern leads to a better performance, in *Cicada orni* is the contrary, and in both species the combination of the two patterns result in a more accurate perception about the origin of some calling song.

Let's consider now the coexistence, in a cicada's field, of the following behavior types:

Type I – cicadas do not use a recognition pattern and follow a song randomly;
Type II – cicadas use a temporal recognition pattern;
Type III – cicadas use a frequency recognition pattern;
Type IV – cicadas use both recognition patterns.

Cicada's paths are identified by gray marks in cells. Initially all cells are white, and they became darker when crossed by females traveling in search of conspecific males.

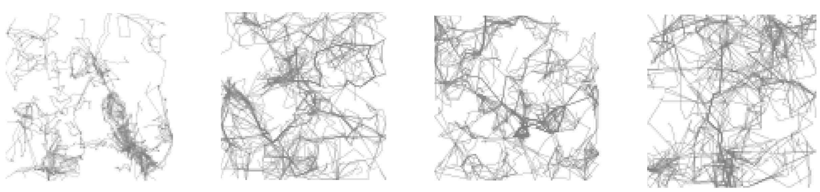


Fig. 4. From left to right: results of simulations showing Type I, II, III and IV paths.

Figure 4 represents the result of simulations, characterized by marks from Type I, II, III and IV, e.g. cicadas not using a detection pattern (on the left), using temporal or recognition pattern or using both (on the right). The extreme left image (Type I) shows that each female cicada is not moving far from a small radius, and her direction changes randomly. The result is a set of small and diffuse zones of dark cells, with no clear tracks. Males act like nodes of a network. Dark cells surround males. The center left image (Type II), concerning cicadas using the frequency pattern, shows a longer dislocation of females and the existence of clear tracks. In what concerns Type III (center right image), where cicadas use the temporal pattern, tracks are also well defined. Type IV paths (left image), regarding cicadas using both patterns, represent a thinner network, composed by a diversity of tracks. Some kind of star-like pattern concerning males positions is visible in the networks.

These experiments support the conclusion that, when cicadas use song recognition patterns, females dislocation is structured, covering the field with clear tracks. The particular case of simultaneously using both recognition patterns results in a thin network, with more tracks and more rapid access to males of the corresponding species. On another hand, when there is no use of recognition patterns, dislocations are short and direction random, with no clear existence of tracks. On another hand, the results of field and simulation experiences, support correlating the behavior of cicadas with their need of differentiating species trough communication patterns. This seems to be a clear intentional mark of cicadas social behaviors, aiming to preserve their species.

The assumption of emergent intentionality characterizing the actions observed in a group can be very useful in a diversity of social contexts, as illustrated by the following case study.

5 Smart mob case study

In Section One we introduced the notion of "smart mobs" and discussed a number of putative cases. These cases of smart mobs were characterized by viral propagation of messages through the social network of each individual. The objective was generally to achieve the gathering of a significant number of individuals at a given moment and in a given place. The result of a smart mob, whatever its goal, is on the one hand the grouping of a large number of individuals and, on the other hand, the surprise of some unintended collective action which was not previously broadcast by the media.

These basic characteristics allow the design of a generic model of smart mob dynamics, where the viral propagation of communication through the social networks of individuals coexists with the coordinated dislocation of individuals to some meeting point. The technologically-mediated cases are, in some respects, more easily modeled than the biological cases because the constraints on communication can be specified or described quite precisely.

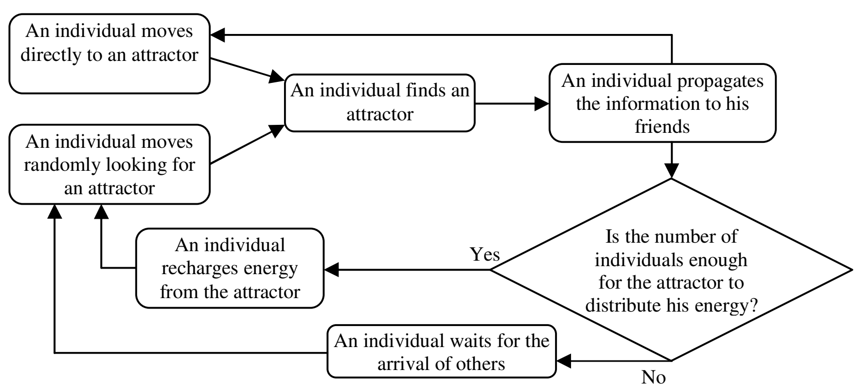


Fig. 5. Flux diagram of the behavior of individuals

Our model of smart mobs comprises two types of agents, individuals and attractors, initially placed in a bi-dimensional space. An individual circulates in the space while he has energy left; when he runs out of energy, he disappears. His goal is to recharge his energy. To achieve his goal, the individual searches for an attractor, moving randomly throughout the space. When an individual finds an attractor, he propagates this information to all his friends; consequently, they will then move towards the attractor, aiming to recharge their own energy. Also, when an individual is notified of the existence of an attractor, he will propagate this information to his friends. When the individuals surrounding an attractor exceed a given number, the attractor distributes his energy to the attending individuals; afterwards, each one reinitiates his random movement, until finding or being notified of the existence of a new attractor.

Besides the viral propagation mechanism through his social network, an individual can decide to move as a result of observing the dislocation of his neighbors. If he verifies that more than a given number of individuals is moving in the same direction, he concludes the existence of an attractor in that region of the space. On their side, each attractor accumulates energy all through the time of the simulation, and distributes his energy when a given number of individuals surrounds him.

The results of the simulation can be observed in the following images:

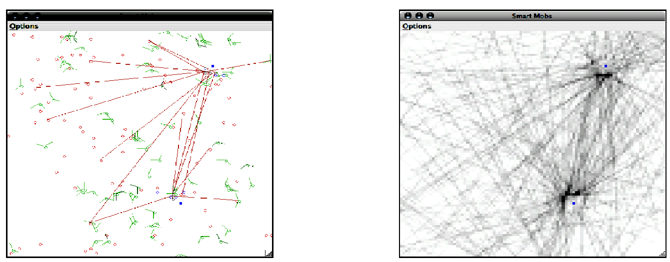


Fig. 6. The Smart Mobs model: macro observation of patterns

The image on the left of Figure 6 represents two fixed attractors and several individuals moving randomly. When an individual finds an attractor, he informs the members of his social network. These messages are represented by the traces connecting the individuals surrounding the attractor to the others dispersed throughout the space. The image on the right of Figure 10 represents a trace of the dislocation of all individuals. It is patent that straight-line movement is followed by those informed of the place of the attractors. The pattern characterizing the joining movements made during a smart mob is clearly visible. Meanwhile, individuals also move as a consequence of the dislocation they observe among their neighbors. This aspect is depicted by the following images.

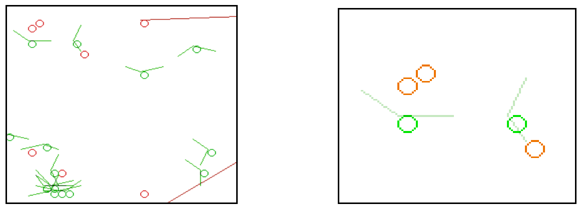


Fig. 7. The Smart Mobs model: micro observation of patterns

The image on the left of Figure 7 represents several moving individuals. Some of them are specifically directed to attain an attractor – these individuals are graphically represented by simple circles. Others move randomly – these have a cone-shaped vision, graphically represented by two traits connected to an individual (a circle), such as can be more closely observed in the image on the right.

Figures 6 and 7 illustrate the existence of patterns at different abstraction levels: the trace of the movement of the individuals indicates specific macro communication and dislocation patterns. The observer can assign the intentionality of a collective movement of the smart mob kind to these patterns. From another side, the observation by an individual of a coherent movement of his neighbors will trigger the detection of a dislocation pattern. Also, the observing individual will attribute some intentionality to this pattern, i.e., dislocation towards some attractor.

6 Discussion

The case studies described above can be discussed in the light of the research hypothesis initially announced. In both cases, we have identified a clear fingerprint that can serve as the basis for predicting and explaining the behavior of groups in natural settings. In addition, since fingerprints are identified and group intentions assigned to the structures, we can conclude that knowing only the links between the nodes of the communications network is enough to characterize the flow of information in these specific social situations.

Another research hypothesis concerns the characterization of social structures through the combination of several elementary patterns. Both case studies suggest that the combination of patterns, identified at different abstraction levels, is a way of identifying social structures determined by communication processes. In the case of *Cicada barbara*, different input (pulse and frequency) and output (structured star-like dislocation) patterns are considered. In the case of the smart mob, communication and dislocation macro patterns are also considered, simultaneously with micro patterns of dislocation of the neighbors of an individual.

On the other hand, the assignment of a given intentionality to collective behavior patterns uniquely from identifying these patterns at a macro level and without knowing the semantic content of the communication allows the attribution of specific intentions to a group of individuals. For instance, observing star-like communication patterns centered on an attractor suggests a social dynamic with the intentionality of a movement with the configuration of a smart mob.

7 Further research

The work presented in this paper incorporates a larger research program, based on the study of the mechanisms and patterns of communication in different social contexts. We consider that our approach can be advantageously applied to several application domains. This way, the research will continue through modeling and simulating other case studies, namely concerning communication domains at a micro-cellular level, as well as new domains relating to sensor networks.

The tasks of designing models and programming multi-agent simulations will be supported by the methodology and programming library developed to the case studies described in this paper. These tools, as well as other resources such as texts, images and videos, are freely available to the community on our web page: <http://www.listaweb.com.pt/projects/cells>.

Acknowledgments. We express our gratitude to Professor Paulo Fonseca, from the Faculty of Sciences of the University of Lisbon, for allowing the use of the results of his experiences in the development of the *Cicada barbara lusitanica* case study.

References

1. Dennett, D.: The Intentional Stance. MIT Press, ISBN 0262540533 (1989)
2. Dennett, D.: Kinds of Minds: Towards an Understanding of Consciousness. Science Masters Series, ISBN 0465073514 (1997)
3. Fonseca, P. J., Revez, M. A.: Song discrimination by male Cicada barbara lusitanica (Homoptera, Cicadidae). The Journal of Experimental Biology No.205, 1285-1292, Ed. The Company of Biologists Limited, U.K. (2002)
4. Goldstone, R., Janssen, M.: Computational models of collective behaviour. TRENDS in Cognitive Sciences, Vol.9, No.9 (2005)
5. Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W. M., Railsback, S. F., Thulke, H.-H., Weiner, J., Wiegand T., DeAngelis, D. L.: Pattern-oriented modeling of agent-based complex systems: lessons from ecology. Science, No.310, 987-991 (2005)
6. The Guardian. April 7th edition (2004)
7. Hardt, M. and Negri, A.: Empire. Cambridge MA.: Harvard University Press (2000)
8. Latané, B.: The Psychology of Social Impact. American Psychologist 36: 343–56 (1981)
9. Louçã, J., Symons, J., Rodrigues, D., Morais, A.: Pattern-oriented analysis of communication flow: the case study of *Cicada barbara lusitanica*. 21st European Conference on Modelling and Simulation - ECMS 2007, Prague, Czech Republic (2007)
10. Rheingold, H.: Smart Mobs: the next social revolution. Basic Books, ISBN 0-7382-0608-3 (2002)
11. Sallach, D., Macal, C., North, M.: Proceedings of the Agent 2006 Conference on Social Agents: Results and Prospects. University of Chicago (2006)
12. Sawyer, K.: Social Emergence: Societies as Complex Systems. Cambridge University Press (2005)
13. Seattle Post-Intelligencer. December 3rd on-line edition (1999)
14. Symons, J.: *On Dennett.* Belmont, CA.: Wadsworth (2002)
15. Symons, J.: Computational Models and Emergence. *Minds and Machines* (forthcoming)

Session on Organization

Group Diversity Dynamics and Decision Quality

J. Richard Harrison and Orlando C. Richard
University of Texas at Dallas
P.O. Box 831327
Richardson, TX 75083, USA
harrison@utdllas.edu

Abstract. Group diversity—including both visible (e.g., race and gender) and non-visible (e.g., values and attitudes) characteristics—is linked to the quality of group decisions. Group diversity affects decision quality by influencing the number of alternatives and the number of decision attributes considered by the group. Group diversity also changes over time as a function of group dynamics, including group entry and exit processes and group socialization. In this paper, we develop a computational model of these processes and present our simulation findings. Implications for future research and for managers are also discussed.

Keywords: Group diversity, group decision making, decision quality, group demography.

1 Introduction

Group diversity consists of visible traits such as race and gender and non-visible traits such as differences in values and attitudes. Social interaction among diverse individuals offers conceptual restructuring and more insights [3], leading diversity proponents to maintain that diversity positively impacts work group outcomes. In fact, both visible and non-visible diversity have been found to relate to decision quality [4], [6]. The literature, however, offers no clear analytical foundation or framework for linking diversity to organizational processes and group outcomes in general, and to decision quality in particular.

In this paper, we develop a computer simulation model linking group diversity to decision quality. The model includes three components: (1) how group diversity changes over time through entry, exit, and socialization processes; (2) how diversity at any point in time affects decision quality through influencing the number of alternatives considered in the decision making process; and (3) how the number of alternatives considered influences decision quality. The first model component is based on Harrison and Carroll's [1] model of team dynamics, adapted so that socialization operates only on non-visible traits since visible diversity traits cannot be changed by socialization. The second model component draws on work by Milliken and Martin [5] and Williams and O'Reilly [7], among others, which shows that more diverse groups consider more decision alternatives. The final component is taken from Harrison and March [2], which shows how the number of alternatives affects the value of the chosen alternative. Essentially, each alternative i is evaluated by observing $z_i = x_i + y_i$, where x_i is the signal (true value) and y_i is a noise term with zero mean, and the highest z_i chosen; because this process tends to select for high values of both signal and noise, the true value of the choice is typically less than the observed value, and the chosen alternative may not even have the highest true value among the set of alternatives. Figure 1 shows the basic model.

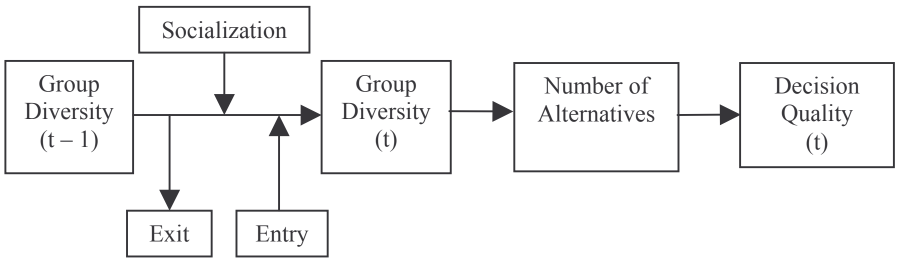


Fig. 1. Model for group diversity and decision quality

2 Experimental Design

We consider two basic situations: one for visible diversity, which we simplify as members of either the majority or a minority and measure as the proportion of the largest group; and one for non-visible diversity, which we will call culture and measure as the standard deviation of individual enculturation scores (see [1]).

For each situation, we examine two processes each for entry and exit. Entry is based either on replication (selection based on replicating the current diversity score of the group) or on diversity (selection attempting to restore the diversity target for the group); the group's diversity target is set experimentally to either 10% or 30% minorities. Attempts to select new minority members slow down the replacement process, and the final choices are not always minorities. Exit can be either random or fit-based, where group members with diversity scores further from the group's diversity measure are more likely to leave due to alienation. The base exit rate is set experimentally to either .01 or .03, and the alienation factor for fit-based exits is set to either a low level or a high level (four times the low level factor). The peer socialization process within the group operates only in the case of non-visible (cultural) diversity. All of these processes are stochastic. The number of alternatives considered increases with diversity as a monotonic step function for each type of diversity. These structural elements of the decision process affect decision outcomes in a stochastic manner described by Harrison and March [2], moderated by the quality of information designated by the signal/noise ratio w, the ratio of the variance of the signal to the variance of the noise. We vary w experimentally, using values of 1 (signal and noise of equal strength), 5, and 9 (high quality information), where the true values have standard normal distributions. Decision quality is measured as the true value of the chosen alternative.

Finally, we consider initial group sizes of 5, 10, and 15, although the actual group sizes may be lower due to unfilled positions due to exit. Overall, our experimental design examines two entry modes, two diversity targets, two exit modes, two base turnover rates, two alienation levels, three signal/noise ratios, and three group sizes, for a total of 288 experimental conditions for each type of diversity, or 576 conditions in all. Each condition is simulated for 120 time periods (months) and repeated for 100 trials; the averages across trial are reported. The computational models are taken from Harrison and Carroll [1] and from Harrison and March [2]; because of space limitations, they cannot be fully specified in a short paper.

3 Findings

Because the model is stochastic, group diversity varies substantially over time. Figure 2 shows this variation for typical runs for each type of diversity.

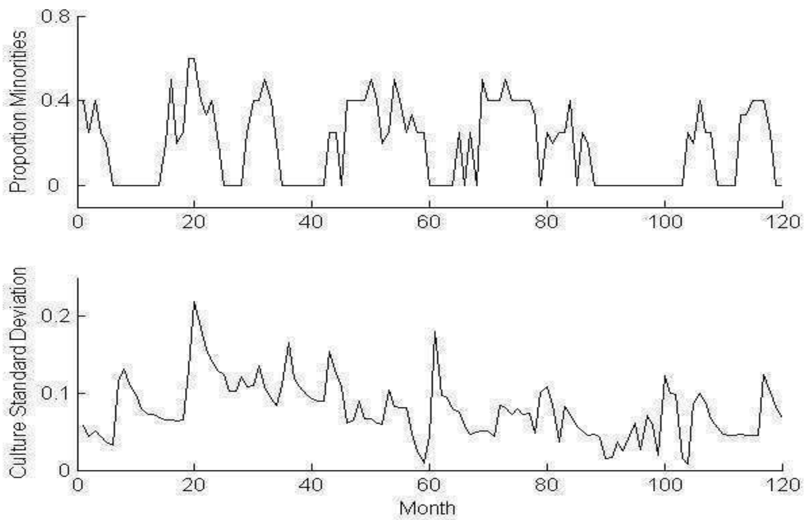


Fig. 2. Variation in group diversity over time

Since group diversity determines decision quality to within stochastic variations, we now explore how group dynamics influences diversity. Table 1 shows regression results for effects of group size, month, and the experimental manipulations on group diversity (the signal/noise ratio is omitted because it has no role in this part of the model). Group size and month are actual values; other variables are coded 0 for low and 1 for high unless otherwise noted.

As Table 1 indicates, the effects are consistent for both types of diversity. Diversity tends to be higher for larger groups but declines slightly over time. As expected, it is higher when selection is diversity-based and when targets are higher, and when turnover is not driven by fit and not influenced by alienation. Interestingly, higher base turnover rates promote diversity.

Table 1. Structural influences on group diversity

Explanatory Variable	**Visible Diversity**	**Non-Visible Diversity**
Group Size	.0025	.0005
Month	-.0003	-.0000
Entry Mode (0 = replication, 1 = diversity)	.0762	.0142
Diversity Target	.0883	
Exit Mode (0 = random, 1 = fit-based)	-.1221	-.0124
Base Exit Rate	.0013	.0131
Alienation	-.0099	-.0017
Constant	.0786	.0499
Adjusted R^2	.788	.875
Number of Observations	34560	34560
*All estimates significant at $p< .001$		

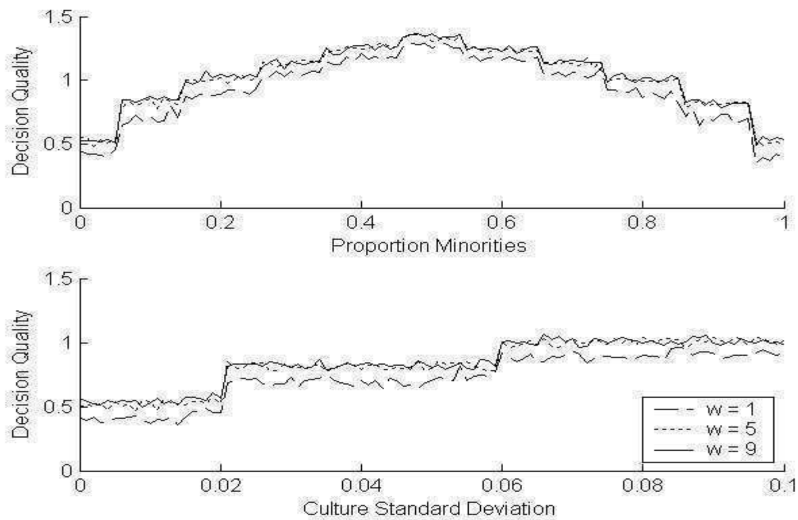


Fig. 3. Findings for group diversity and decision quality

Figure 3 shows how diversity influences decision quality in our model, for each type of diversity. The “steps” in the figure are due to the quantized nature of alternatives: diversity has to increase by some amount before the number of alternatives considered jumps from two to three, for example. Even though these figures are averages across 100 runs, the stochastic nature of the process is still apparent. Note that in the case of visible diversity, maximum diversity is associated with proportion minorities = .5, and that decision quality is somewhat higher when w is higher.

4 Discussion

Our simulation data were subjected to additional analysis using statistical methods. The findings highlight the influence of group dynamics on diversity and decision quality. The basic findings are that processes of group dynamics influence diversity, and that diversity promotes decision quality. While most of these findings are obvious, a formal model linking the underlying processes has not been previously presented. Future research can focus on further exploration and refinement of the model and on empirical work to test its predictions. Note also that the effects of diversity on decision quality arise through the simple mechanism of influencing the number of alternatives considered by the group. There are undoubtedly other mechanisms linking diversity with decision quality, and that diversity is expected to affect other group behaviors; these are also fruitful topics for further research.

A somewhat unexpected finding concerns the positive effects of group size and base turnover rates on diversity and decision quality. A single new member in a small group strongly impacts the group’s composition and can cause it to overshoot its diversity target; in larger groups, a more incremental approach to promoting diversity may be easier. And there are more vacancies available to promote diversity when turnover is higher and groups are larger.

Our findings have implications for managers interested in increasing diversity and consequently decision quality. They include expanding group size, promoting movement of personnel in and out of groups, encouraging selection based on diversity with explicit diversity targets, and providing incentives for members who are “different” to stay in groups.

References

1. Harrison, J.R., Carroll, G.R.: Culture and Demography in Organizations. Princeton University Press, Princeton, NJ (2006)
2. Harrison, J.R., March, J.G.: Decision Making and Postdecision Surprises. Administrative Science Quarterly 29 (1984) 26-42
3. Jehn, K.: Diversity, Conflict, and Team Performance: Summary of Program of Research. Performance Impact Quarterly 12 (1999) 6-19
4. McLeod, P., Lobel, S.A., Cox, T.H.: Ethnic Diversity and Creativity in Small Groups. Small Group Research, 27 (1996) 827-847

5. Milliken, F.J., Martin, L.L.: Searching for Common Threads: Understanding the Multiple Effects of Diversity in Organizational Groups. Academy of Management Review 21 (1996) 402-433
6. Watson, W.E., Kumar, K., Michaelsen, L.K.: Cultural Diversity's Impact on Interaction Process and Performance: Comparing Homogeneous and Diverse Task Groups. Academy of Management Journal 36 (1993) 590-602
7. Williams, K.Y., O'Reilly, C.A. III: Demography and Diversity in Organizations. Research in Organizational Behavior 20 (1998) 77-140

Agent-based Organizational Cybernetics for Organizational Learning

Yusuke Goto[1] and Shingo Takahashi[2]

[1,2]Department of Industrial and Management Systems Engineering, Waseda University
[1]ysk5@toki.waseda.jp, [2]shingo@waseda.jp

Abstract. This paper proposes guidelines to evaluate agent's internal model effectively in organizational learning. We developed an approach named *Agent-based Organizational Cybernetics* (AOC) to describe problems in organizational learning with an agent-based model and to design some prescriptions for them. This paper focuses on four typical evaluation problems of internal model in organizational learning: (1) environmental scanning, (2) policy making, (3) coordination, and (4) task resolution. We design and test some guidelines for each problem. The results provide some effective guidelines in organizational learning. *Organizational-performance(OP)-based* evaluation realizes organization's high performance in all four problem situations. However, the OP-based evaluation cannot guarantee that an agent gets a right internal model all the time. So, it is effective to also consider some other complementary evaluations.

Keywords: organizational cybernetics, organizational learning, computational organization theory, agent-based organizational cybernetics

1 Introduction

Agent-based approach for organizational learning sounds promising. Agent-based modeling can describe the behavior of an organization as a result of individual agents' behaviors. So we can discuss the micro-macro problems of complex organizational systems in an operational manner.

Computational organization theory has attacked some problems in organizational learning. March has described a simple double-loop learning mechanism with a concept of exploration and exploitation (March 1991). Takadama et al. have developed an Organizational-learning oriented Classifier System to model four-loop learning in organizational learning operationally (Takadama et al. 1999). Some models have considered organizational structure in organizational learning (Carley and Svoboda 1996; Takahashi and Goto 2005).

However, computational organization theoretic approach describes either only a "flat" organization that has no hierarchical relationship between subsystems, or higher subsystems as super-agent activities (Takahashi 2006; Chang and Harrington 2006).

The model in this paper represents a "hierarchical" organization that has four functional layers defined in organizational cybernetics: intelligence, policy, coordination and resolution. All functional layers are described as an agent-based system. We call this approach *Agent-based Organizational Cybernetics* (AOC).

This paper focuses on typical evaluation problems of internal model in each functional layer: environment scanning, policy making, coordination, and task resolution. Effective evaluation will reduce the equivocality (Daft and Lengel 1986) and lead successful organizational double-loop learning (Argyris and Schon 1978). We design and test some guidelines for each problem.

In the following sections we briefly introduce the main features of AOC, and describe the typical problems in organizational learning, simulation model, and guidelines for effective organizational learning. In the last part, we report initial results and discussion.

2 Agent-based Organizational Cybernetics

AOC is a hybrid model that combines organizational cybernetic framework and computational organization theoretic approach. The main target of AOC is to describe problems in organizational learning with an agent-based model and to design some prescriptions for them (Takahashi 2006). The basic features of AOC can be listed: (1) interaction between environment and agents, (2) agent's decision making according to his decision principle, (3) multi-layer hierarchical organization with some functional subsystems, (4) agent-group in each layer of the hierarchy, (5) situatedness of an agent, (6) organizational learning of revising and sharing agents' internal models.

3 Typical problems in organizational learning

It is natural to be thought that problems in organizational learning are different within organizational functions. We define four functional layers in the AOC manner, and describe typical problems for each layer. An organization considered has the following four functional layers.

1. *Intelligence*. An organization should recognize the environmental structure properly to realize an adaptive organizational policy. The environmental structure is scanned.
2. *Policy*. An organization defines own structure and input to be viable in an environment. A task resolution structure and a resource allocation are defined as an organizational policy.
3. *Coordination*. An organization has to coordinate inferior subsystems to realize better organizational performance. The values of coordination variables are decided.
4. *Resolution*. An organization resolves tasks in an environment. A plan of task resolution is selected.

We can specify a typical problem in each layer.

1. Scanning environment. In intelligence layer the right recognition of the environmental structure is offered.
2. Policy making. In policy layer it is difficult to be viable or adaptive without an adequate decision principle to evaluate organizational policies.
3. Coordination. In coordination layer the proper anticipation of a coordination variable value to a subsystems performance level is essential for better coordination.
4. Task resolution. In resolution layer the right recognition of the task is required for effective task resolution.

Learning is essential for an organization to tackle these four problems. We especially focus on internal model evaluation in learning. If an agent cannot evaluate his internal model properly, he will fail to revise it successfully. So we, in this paper, seek guidelines for the evaluation of an agent's internal model in each typical problem in organizational learning.

4 Model

We here consider a simple maker-type organization that has some processes. The organization anticipates characteristics of an environment, defines a task resolution structure and a resource allocation, distributes resources to each process, and resolves a task in each process by using the resources (see Figure 1).

Task

A task means a series of demand for a service or a product that a process provides. For effective task resolution, supply need to fit demand. Much supply requires a lot of resources. There exist n tasks $\mathbf{u}_{e1},\cdots,\mathbf{u}_{ei},\cdots,\mathbf{u}_{en}$ in an environment. An organization selects tasks to resolve, and resolves them.

We assume that an organization use the common resource that can resolve every task. The i th task $\mathbf{u}_{ei}$ that has a q -long string is expressed as $\left(u_{ei1},\cdots,u_{eil},\cdots,u_{eiq}\right)$. The l th element of the u_{eil} means the amount of the demand in the l th term.

Environment

In the AOC framework an agent is confronting her own environment, and an organization is also confronting its own environment. We build an environment model of the organization as a task generator and a response function.

The task generator provides the organization with n tasks every step. We define the task resolution plan $\mathbf{m}_i$ as $\left(m_{i1},\cdots,m_{il},\cdots,m_{iq}\right)$. Task resolution performance with the $\mathbf{m}_i$ is evaluated by the fitness of the demand and the supply (Eq. 1).

$$y_i=\sum_{l=1}^{q}\hat{m}_{il}\Big/u_{eil}\times q,\ \hat{m}_{il}=\begin{cases}u_{eil}, & \text{if } m_{il}\geq u_{eil}\\ m_{il}, & \text{otherwise}\end{cases} \quad (1)$$

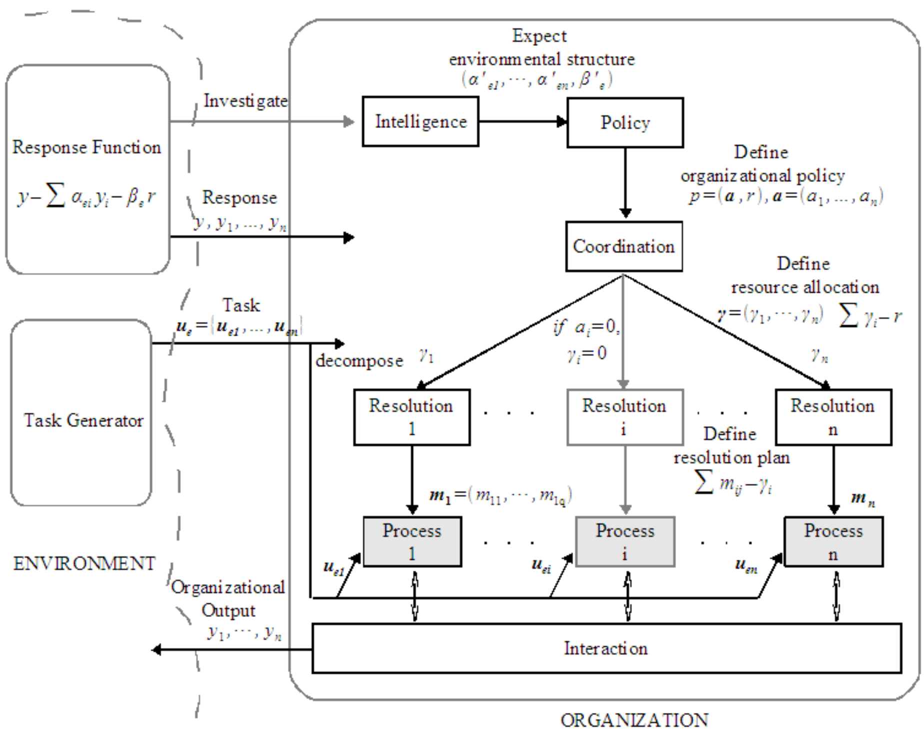


Fig. 1. Hierarchical organization model. Based on a model appeared in WCSS06 (Goto and Takahashi 2006), the model in this paper is newly developed in the AOC manner (Takahashi 2006).

The response function evaluates the output of the task resolution activities of an organization. An organization's performance is evaluated by the response function

$$y = \sum_{i=1}^{n} \alpha_{ei} y_i - \beta_e r \,, \tag{2}$$

where α_{ei} and β_e respectively show a real profit coefficient and a cost coefficient, y_i is a task resolution performance in the i th process.

Functional hierarchy

An organization is a three levels hierarchical system: adaptive, coordination and operational levels defined in organizational cybernetics (Beer 1981; Takahara and Mesarovic 2003). Every agent belongs to one of the subsystems of the hierarchy, and makes a decision of the subsystem.

Adaptive level consists of intelligence and policy functions. In an intelligence function the environmental structure *st* that considers the profit coefficients $\alpha_{e1},\cdots,\alpha_{ei},\cdots,\alpha_{en}$ and the cost coefficient β_e of Eq (2) is observed. In a policy function a task resolution structure $a_1,\cdots,a_i,\cdots,a_n$ and a resource allocation r are defined as an organizational policy *p*. The task resolution structure defines processes in which the tasks are resolved; if task resolution is taken in the ith process, then we set $a_i = 1$, otherwise $a_i = 0$. We assume that the tasks and the processes are one-to-one correspondence: the task $\mathbf{u}_{ei}$ in the environment is resolved in the ith process in an organization.

In coordination level a coordination function sets the values of coordination variables to manage the operational level under the organizational policy. The values of coordination variables $\boldsymbol{\gamma} = (\gamma_1, \cdots, \gamma_i, \cdots, \gamma_n)$ represent how to allocate resources to the processes; γ_i represents that agents in the ith process use γ_i resources to resolve the task $\mathbf{u}_{ei}$; if $a_i = 0$, then $\gamma_i = 0$.

Operational level consists of a resolution function and n processes. The resolution function has n agent-groups. Based on the coordination variables every agent makes a plan of task resolution $\mathbf{m}_i = (m_{i1}, \cdots, m_{in}), \sum_{l=1}^{q} m_{il} = \gamma_i$ in the ith process. The plan $\mathbf{m}_i$ is a series of the supply for the service of the product.

Agent and Internal Model

An organization has some agents. Every agent belongs to one of the four functions. We call an agent in an intelligence function a *scanner*, in a policy function a *policy maker*, in a coordination function a *coordinator*, and in a resolution function a *resolver*, respectively. Every agent refers his internal model and makes a decision as in the cybernetic control concept (Takahashi 2006).

A scanner observes an environment structure st, has the recognition as his internal model, and reports the aggregated one to the policy function.

A policy maker gets the anticipated environmental structure from the intelligence function, applies his decision principle of his internal model, and selects the most preferable organization policy. The decision principle consists of three types of preference: coverage pursuit (CP), profitability pursuit (PP) and cost aversion (CA). A policy maker has an attitude for each type: agree (1), disagree (-1) and not considered (0). The decision principle is represented as $DP = (CP, PP, CA)$, and every preference takes a value of $\{1, 0, -1\}$. The CP is a preference for resolving more tasks. The PP is a preference for resolving more profitable tasks. The CA is a preference for paying more costs. We here skip detailed description of them. A policy maker selects the alternative that maximizes the sum of the preference values.

A coordinator has an anticipation of a task resolution performance $y'_{1,} \cdots, y'_i, \cdots y'_n$ realized by a coordination $\boldsymbol{\gamma}$.

A resolver in the ith process recognizes the task $\mathbf{u}_{ei}$, and has the recognition $\mathbf{u}'_{ei}$ as his internal model, and selects the plan of task resolution $\mathbf{m}_i$ that seems to realize the highest task resolution performance.

Organizational Learning

We reinterpret the four-loop learning (Espejo et al. 1996) in an operational way as an agent-based organizational learning model. We represents double-loop learning as three processes: evaluation of agent's internal model, revision of the internal model, and sharing of agents' internal models in an agent-group. The described organizational learning cycle is taken in each organizational layer in an organization.

1. *Individual single-loop learning*. Every agent refers his internal model that describes confronting environment, applies his decision principle to optimize his decision alternatives, and selects a decision alternative. This process has no influence with the progress in the ability of organization's decision making.
2. *Organizational single-loop learning*. Agent-group's decision is taken by unifying agents' decisions. The most selected decision alternative is adopted as the group's one.
3. *Individual double-loop learning*. An agent evaluates his own internal model from the result of the decision made just before. Then he revises his internal model. These processes are implemented by using genetic algorithm (GA): a fitness function corresponds with agent's internal model evaluation, and genetic operation corresponds with revision of the internal models. The implementation by GA represents the evolutionary aspect of organizational learning. Design of the fitness function is the key to success of effective organizational learning.
4. *Organizational double-loop learning*. As a result of effective individual double-loop learning, agents in the group share their helpful internal models. The group (or an organization) progresses its ability of decision making, and can be more viable or adaptive.

Guidelines in Organizational Learning

We described four typical problems in organizational learning: scanning environment, policy making, coordination, and task resolution. In these problems proper evaluation of agent's internal model is essential. So we, in this paper, seek effective guidelines to define how helpful agent's internal model is to a situated environment.

A critical constraint in these problems is that agent does not know the "right" internal model in organizational learning. Then we assume that an agent can get data of organizational performance and task resolution performance as a response from an environment.

For scanners' scanning environment

Scanners' problem is to recognize the environmental structure $st = (\alpha_{e1}, \cdots, \alpha_{en}, \beta_e)$ correctly. We design the following three guidelines to evaluate scanners' internal model properly.

1. Organizational-performance-based evaluation (OP). If a scanner has a right recognition of the st, then the organizational performance will be improved.
2. Task-resolution-performance-based evaluation (TRP). If the organizational performance will be improved when a scanner has a right recognition of the st, then the task resolution performance will be improved.
3. Organizational-performance-anticipation-based evaluation (OPA). A scanner can get information of an organizational policy: the task resolution structure and the

resource allocation. If a scanner has a right recognition st , then the difference between an expected performance y and a realized performance y' will gets smaller. We set an expected environmental structure $st' = (\alpha'_{e1}, \cdots, \alpha'_{en}, \beta'_e)$ and define the expected performance as $y' = \sum \alpha'_{ei} a_i - \beta'_e r$.

There exists a problem. Each of these three guidelines cannot apply to scanners directly, because these correspond with the function's (group's) decision making. So we take an indirect approach. A scanner uses retrospective data that the organization got, and he evaluates his internal model by the retrospective decision that is most similar with his decision. This approach is common in the following problems. In the retrospective data we set most similar organizational performance y^u , the expected performance of the decision $y^{u\prime}$ and the realized task resolution performance $y^u{}_1, \cdots, y^u{}_n$. Then OP, TRP and OPA are defined as

$$\mathrm{OP} = y^u \big/ \max \text{of } y \,, \tag{3}$$

$$\mathrm{TRP} = \sum a_i y_i^u \big/ \sum a_i \,, \tag{4}$$

$$\mathrm{OPA} = 1 \Big/ 1 + \left| y^u - y^{u\prime} \right| . \tag{5}$$

If the value of evaluation is high, then he has a helpful internal model in terms of the guideline.

For Policy makers' policy making

Policy makers' problem is to get an adequate decision principle. We design the following two guidelines to evaluate policy maker's internal model effectively.

1. OP. If a policy maker's decision principle is desirable, then the organization will realize high-level performance naturally (see Eq. 3).
2. TRP. When an organization realizes high-level performance, the task resolution performance will be also high-level (see Eq. 4).

For Coordinators' coordination

Coordinators aim at having a sure anticipation of a task resolution performance realized by a resource allocation.

We design the following three guidelines.

1. OP. If a coordinator's internal model is helpful and realizes effective coordination, then the organization will finally realize high-level performance (see Eq. 3).
2. TRP. Effective coordination realizes high task resolution performance naturally (see Eq. 4).
3. Fitness for retrospective data (FRD). If the anticipated task resolution performance with a coordinator's internal model fits to the retrospective fact data, the internal model seems to be helpful. We set the j th retrospective data of the task resolution

performance $y^j{}_1,\cdots,y^j{}_n$ by the resource allocation $\boldsymbol{\gamma}$, the coordinator's anticipation for the $\boldsymbol{\gamma}$ as $y^{j'}{}_1,\cdots,y^{j'}{}_n$.

$$\mathrm{FRD} = 1\Big/1 + \sum_j \sum\nolimits_{i=1}^{n} \left| a_i y_i^j - a_i y_i^{j'} \right| \tag{6}$$

For Resolvers' task resolution
Resolvers in the ith process should have a right recognition of the task $\mathbf{u}_{ei}$. We design the following three guidelines.

1. OP. The helpful internal model realizes effective task resolution. As a result, the organization will achieve high-level performance (see Eq. 4).
2. TRP-2. If a resolver's internal model is helpful, then task resolution will be taken effectively in his process.

$$\text{TRP-2} = y^u{}_i \tag{7}$$

3. Difference between task and plan (DTP). If a resolver gets the magnitude correlation between task and plan, he will be able to evaluate his decision. We set the resolver's decision $\bar{\mathbf{m}}_i = \left(\bar{m}_{i1},\cdots,\bar{m}_{in}\right)$.

$$\mathrm{DTP} = \sum\nolimits_{l=1}^{n} e_l \Big/ q,\ e_l = \begin{cases} 1, & \text{if } \left(u_{eil} \ge m_{il}, m_{il} \le \bar{m}_{il}\right) \text{or} \left(u_{eil} < m_{il}, m_{il} > \bar{m}_{il}\right) \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

Experimental design and results

First we confirm two classic propositions in organizational learning for model validation. Next we review which guideline is more effective in each problem, and analyze the characteristics of organizational learning in each layer.

Model validation
We here test whether the built model can generate the result that is compatible with the propositions of foregoing empirical studies. We choose the famous propositions by Argyris (Argyris and Schon 1978).

- P 1: *"If the organization has a right recognition of the environment, the optimization of organizational actions by the organizational single-loop learning (OSL) will lead to be adaptive to the environment."*
- P 2: *"If the organization doesn't have a right recognition of the environment, organizational learning is not significant. The organization requires organizational double-loop learning (ODL) that also revises the recognition of the environment."*

In this model the Argyris's OSL can be interpreted as the learning in resolution layer. The ODL can be interpreted as the learning in all four layers. In the proposition 1 situation, scanners, policy makers and coordinators have a proper internal model, though resolvers can have an inadequate internal model that is generated randomly. In the proposition 2 situation all agents can have an inadequate internal model that is randomly generated.

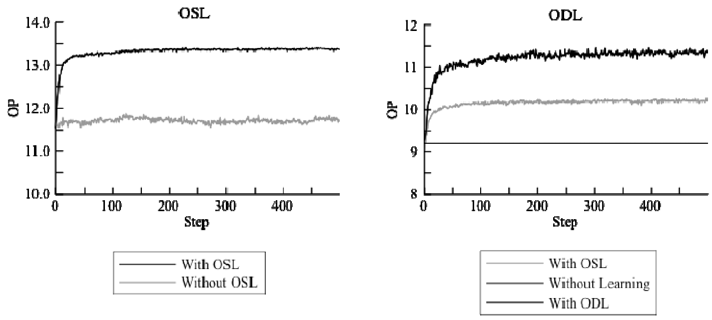


Fig. 2. Process of adaptation with OSL or ODL. OP represents the average organizational evaluation value of 1000 runs in a setting.

The left side of the Fig. 2 shows the result for the P. 1. An organization with OSL improves OP rapidly. On the other hand, an organization without OSL seems to stagnate. Then the P. 1 is confirmed.

The right side of the Fig. 2 shows the result for the P. 2. Both of organizations with OSL and with ODL improve OP. The organization with ODL improves OP twice as much as the organization with OSL does. An organization without learning seems to stagnate. Then the P. 2 is also confirmed.

Effective guidelines in organizational learning

We here review which guideline is more effective for each problem. We have got results that the OPA (Eq. 5) in intelligence layer and the DTP (Eq. 8) in resolution layer are effective (Goto and Takahashi 2005; Goto and Takahashi 2006). However we haven't compared them with other guidelines.

We set the number of tasks $n = 3$, the length of tasks $q = 5$, the response function $y = 2y_1 + y_2 + 4y_3 - 2r$, and the number of agents in each layer 30. The task $\mathbf{u}_{ei}$ is randomly set its values. A run of simulation consists of 500 steps of task resolving and learning.

Fig. 3 shows the review of proposed guidelines. Every guideline is effective itself, because they improve organizational performance. In the case of single guideline, the OP-based evaluation is most desirable in every layer, because the evaluation improves organizational performance most.

However, if we consider combined guidelines, the OP-based evaluation is not always the best one. In coordination layer a combined evaluation of the OP and the FRD is most desirable. In resolution layer a combined evaluation of the OP and the TRP is as desirable as the OP only. We think that the effectiveness of the combined evaluation comes from a complementary relation between the OP evaluation and other evaluations.

We show a typical case in intelligence layer. We define Intelligence Fitness (IF) to measure the correctness of scanners' internal model. Higher the IF is, more correct the scanner's recognition of the environmental structure st is.

$$IF = 1/1+d \qquad (\mathbf{9})$$

d : hamming distance between the actual structure and the recognized one

When we review the guidelines in terms of the IF, the OPA is outstanding. On the other hand, the OP is very bad. This data suggest us that the OP evaluation learns for the organizational performance, not for the "right" recognition of the environment. So we need to consider some other complementary evaluation for simultaneous pursuit of better organizational performance and right internal model.

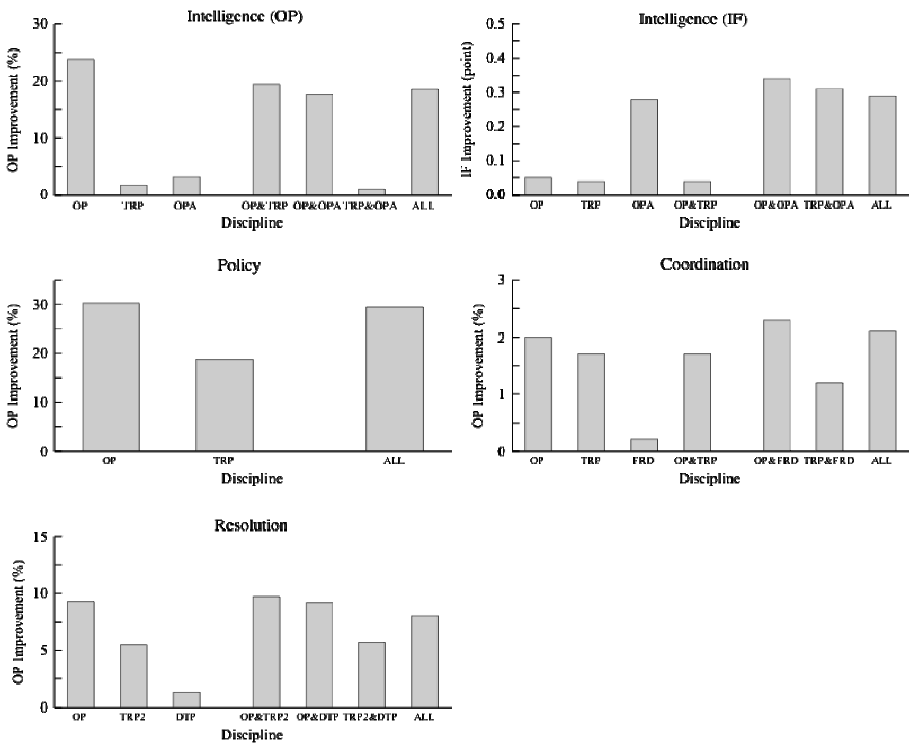


Fig. 3. Review of proposed guidelines. OP (IF) improve means the percent of organizational performance (IF) improvement from the first step to the last 100 step.

Adaptive behavior in each layer

We compare the behaviors of organizational learning. Most of studies discuss the behavior of organizational learning in a single layer. So it is impossible to compare the behaviors between different layers. An AOC-based model can discuss the characteristics of adaptive behavior in multi-layers and co-affection of organizational learning among the layers.

Figure 4 shows the process of adaptation in each layer. In intelligence, policy, and resolution layer organizational leaning is realized rapidly in the first 100 steps. On the other hand, in coordination layer organizational learning is taken gradually in about 400 steps. We can say that organizational learning in each layer has the characteristics described above.

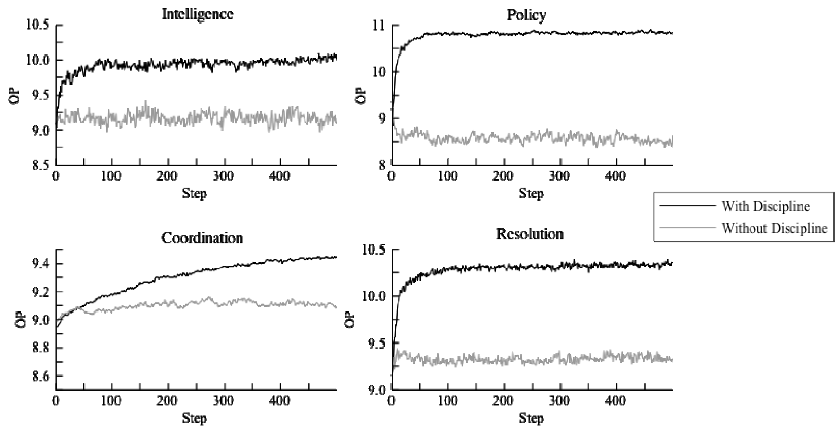


Fig. 4. Process of adaptation in each layer.

Discussion

We, in the preceding paper, have presented an effective guideline for each layer (Goto and Takahashi 2006). We, in this paper, proposed some guidelines for each layer, and reviewed which guideline is more effective. The results in this paper provide some new implications for effective organizational learning.

An agent must consider the organizational performance that is realized by his decision when he evaluates his internal model. The OP-based evaluation generally leads the agent in each layer to achieve high organizational performance. However, achieving better organizational performance doesn't always assure that an agent has a "right" internal model. In intelligence layer a scanner should also consider the difference between the realized organizational performance and the anticipated one. We found some complementary relations between two evaluations. For example, in coordination layer a combined evaluation of the OP and the FRD is more effective than the OP, though the FRD itself is the worst among the three evaluations.

Guidelines in organizational learning should be effective in any situations universally, because it is hard for agents in an organization to appreciate the situation in which they are. We presented the initial results that are examined in particular settings. Since we need to test the robustness of the results, we are now testing some key parameters: the task number n, the task length q, the environmental structure st, and the parameters of GA operations.

Conclusion

We proposed some guidelines to evaluate agent's internal model properly in each layer, and reviewed which guideline is more effective. The guidelines tested as more effective are as follows. In intelligence layer it is desirable to consider both of the OP and OPA. In policy layer a policy maker should evaluate his internal model in terms of the OP. In coordination layer a coordinator had better consider both of the OP and FRD. In resolution layer it is necessary for a resolver to consider either the OP or the OP and the TRP-2. The guidelines suggest some effective organizational learning strategies: which information should be gained by an organization and how to utilize the information in the organization.

References

Argyris C, Schon DA (1978) Organizational Learning. Addison-Wesley

Beer S (1981) Brain of The Firm (2nd edition). John Wiley & Sons

Carley KM, Svoboda DM (1996) Modeling Organizational Adaptation as a Simulated Annealing Process. Sociological Methods & Research 25: 138-168

Chang MH and Harrington Jr. JEH (2006) Agent-based Models of Organizations. In: Tesfatsion L, Judd KL (eds) Handbook of Computational Economics. North-Holland, 1273-1337

Daft RH, Lengel RH (1986) Organizational Information Requirements, Media Richness and Structural Design. Management Science 32:554-571

Espejo R, Schuhmann W, Schwaninger M, Bilello U (1996) Organizational Transformation and Learning. John Wiley & Sons

Goto Y, Takahashi S (2005) Organizational Learning Oriented Model of Organizational Adaptation. Proceedings of the First World Congress of the International Federation for Systems Research

Goto Y, Takahashi S (2006) Effective Guidelines for Organizational Learning in the Organizational Cybernetics Framework. Proceedings of The First World Congress on Social Simulation

March, JG (1991) Exploration and Exploitation in Organizational Learning. Organization Science 2: 71-87

Takadama K, Terano T, Shimohara K, Hori K, Nakasuka S (1999) Making Organizational Learning Operational: Implications from Learning Classifier System. Computational & Mathematical Organization Theory 5 (3): 229-252

Takahashi S (2006) Agent-based Organizational Cybernetic Approach to Organizational Learning. Proceedings of SICE-ICASE International Joint Conference

Takahashi S, Goto Y (2005) Agent-based Simulation of Adaptive Organizational Structures to Environmental Change. In: Terano T, Kita H, Kaneda T, Arai K, Deguchi H (eds) Agent-based Simulation (vol. 1). Springer, 99-110

Takahara Y, Mesarovic M (2003) Organization Structure. Kluwer Academic/Plenum Publishers

Agent-based simulation to analyze business office activities using reinforcement learning

Yukinao Kenjo*, Takashi Yamada, and Takao Terano

Department of Computational Intelligence and System Science,
Tokyo Institute of Technology,
4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8502, Japan
yukinao.kenjo@gmail.com tyamada@trn.dis.titech.ac.jp
terano@dis.titech.ac.jp

Abstract. This paper attempts to clarify team behavior in cooperative organizations by agent-based simulations. We focus attention on both the roles of manages and the initiatives of staffs, and then model them using agent-based model concepts. This enables us to investigate phenomena in organizations at micro-level and macro-level. Besides, we formulate the task processing of each member in real organizations as learning for maze problem. The advantages of applying maze problem for our simulation model are as follows: It is possible to describe agents who acquire skills by reinforcement learning and to represent environmental uncertainty by changing block placements dynamically. Several computational experiments show what the whole organization behaves from microscopic points of view. At the same time, the authors confirm that the ability to adapt environments under uncertainty is different from the characters of organization.

1 Introduction

Agent-based simulation has been researched in recent years by a number of fields beginning with social science, and its usefulness is recognized [2, 7]. This paper attempts to clarify organizational behavior in corporate bodies through an agent-based simulation approach.

Attempts to analyze organizational phenomena computationally began in the 1960s [6] and have evolved into the computational organization theory in recent years [3, 4]. For example, the garbage can model provides an operational explanation of "not necessarily rational decisions" made in group decision making situations [5]. It can be said that the garbage can model is a good work of explaining not necessarily rational decision making that may occur in group decision situations. However, the garbage can model is highly abstract and only touching the surface of phenomena occurring due to the complex group interactions within actual organizations.

In response, this research proposes a model based on the bottom-up technique [9]. This allows for simulation of organizational phenomena from an ABS stance.

* Now at Sony Corporation

Which means, we focus on two local organizational elements, the role of the superior and the independence of the subordinate, in modeling this under the agent concept. We believe that this makes it possible to carry out both micro and macro analyses of complex organizational phenomena.

The relationship between superior and subordinate is especially important amongst the elements that compose organizations. "Indifference index" and "identification index" are proposed from organizational theory in order to detail the relations above in the proposed model. On the one hand, indifference index represents conformity with orders given by a superior. A high indifference index represents a passive attitude within the organization. On the other hand, identification index indicates agreement between purpose and values of an individual, and those of the organization. A high identification index becomes the motivation to actively work within the organization. Introduction of these indexes makes it possible to implement a simulation that handles a variety of organizational forms observed in actual organizations.

Furthermore, the simulation is conducted in the proposed model assuming that dealing with maze problems by reinforcement learning is a task process. In this study, we postulate that to reach a goal in a maze is consistent with grappling with a task. In other words, the task processing of every member within the actual organization is formulated as a maze problem learning. In this way, the following two points are seen as advantages of introducing maze problem learning: Firstly, reinforcement learning allows expression of agents acquiring skills and abilities, namely reinforcement learning lets agents solve a maze problem more easily. This is based on observations of actual society such as actively working members of the organization becoming able to process that work in a short period of time. Secondly, dynamically changing maze blocks allows expression of environmental uncertainty. This corresponds to, for example, situations in which the need arises to learn new work processing methods due to such happenings as technical innovation.

The rest of this paper is organized as follows: The next section explains the details of our simulation model. Section 3 presents some of the highlighted results taken from our proposed model. Finally, this paper ends with some concluding remarks in section 4.

2 Setup

2.1 Model overview

The model proposed deals with hierarchical team observed in general industrial organizational forms and aims to analyze what group form is applicable in response to environmental change. In this part of the section we will present an overview the model.

In the proposed model, each agent carries out an appropriate action from the following three decision making rules at a certain time, namely interaction with others through 1) the decision making rule to manage the work of subordinates as a manager, 2) the one to follow the directions of the superior, or 3) the non-interactive one to process your own work. Although the agent, who is placed in mid-level (middle management), is caught between the management of subordinates and directions from upper-level agents, behavioral rules as a staff member take precedence in this case. Interaction with others is given priority with regard to prioritization of each behavioral rule, and placed in staff-manager-individual order.

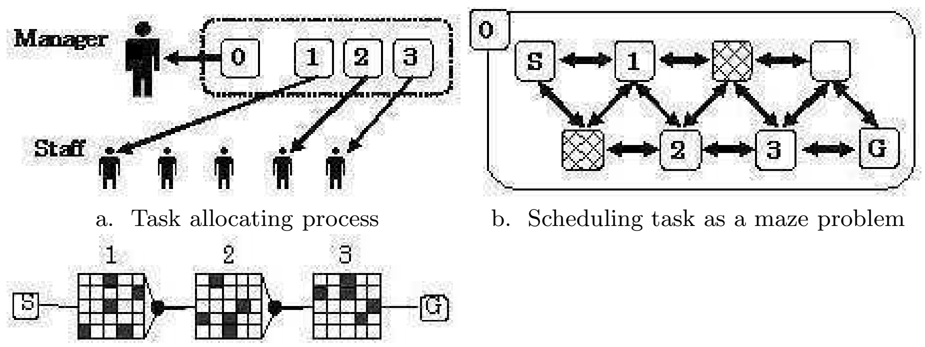
2.2 Object definition

The main constituents of the proposed model and their attributes are defined in Table 1. Organizations are represented as Organization Object, and agents as Member Object. Task Object represents all work content processed by members of the organization, and is referred to as a unit task in this study. Desk Object is individually owned by members and correspond to the place of task processing.

- Attributes defined by organization objects are related to the organization as a whole.
  Structure (organizational structure) manages authority structure of the organization.
  Finished (finished task group) retains finished tasks.
- Attributes defined by member objects play an important role in distinguishing members.
  Role stands for top, middle, or bottom level.
  State is a present condition.
  Intention means intent in regard to working in the affiliated organization.
  Interest is an attentiveness in the work content of the unit task handled.
  Identification (identification index) represents identification with superior. This is derived from intention attributes of the two.
  Indifference (indifference index) means how eagerly an agent is going to deal with the presently held tasks. That is derived from interest attributes and attraction attributes of Task.
  Knowledge (knowledge of unit task) retains work content knowledge of unit task. Includes Q-table of reinforcement learning.
- Attributes defined by desk objects are inherent to each member.
  Buffer (task group waiting for processing) retains a list of tasks presently held by members.
  Phone (calling flag) is used when called by superior.
- Attributes defined by task object distinguish each task.
  MazeId means the id number labeled in each maze problem corresponding to this unit task.
  MazeListId is an ID of generated task (maze string).
  Attraction is used to derive indifference attributes of a member. Represents the attraction of the work content of this unit task, set as a factor to induce members' interest.

Table 1. Definition of objects

Object	Attribute
Organization	structure (int[])
	finished (Set<Task>)
Member	role (String)
	state (String)
	intention (Character[])
	interest (Character[])
	identification (double)
	indiffereence (double)
	knowledge (Map<mazeId, Strategy>)
Desk	buffer (List<Task>)
	phone (boolean)
Task	mazeID (integer)
	mazeListID (integer)
	attraction (Character[])



a. Task allocating process

b. Scheduling task as a maze problem

c. A series of maze problems

Fig. 1. Processing flow of a maze list

2.3 Task processing and reinforcement learning

This section describes the flow of task generation to partial task/unit task division in detail, and goes on to explain maze problem learning, reporting to superiors, and task completion.

1. Task Generation
   A task is generated within the organization at fixed periods. The generated task is piled on to a given managers desk. This task includes several unit tasks to be handled by the manager's subordinates and is determined by the following algorithm.
   (a) A manager is chosen arbitrarily at fixed periods.

(b) Only agents without subordinates are taken from the agent group corresponding to the subordinates under the manager chosen in the previous step.
(c) Each agent taken in the previous step is drawn according to predetermined selective probability.
(d) The unit task, handled by agents chosen in the previous step along with tasks from him handled by agents right up to managers chosen in the first step, is included in the generated task.

Since unit tasks are set as mazes, a generated task consists of two or more arranged maze strings. Unit tasks handled here are work content held by each agent, and signify unit tasks that must be processed by that agent.

2. Division and distribution of tasks
Tasks generated are broken down into partial tasks by the manager and allocated to the appropriate staff member who is called up to the desk. This situation is represented in Figure 1a. The manager begins to process the single unit task left to him from the original task once all partial tasks have been allocated. This unit task, like the others, is a maze problem that must be processed by the manager, but is treated here as the manager solving the management scheduling problems of his subordinates. This is represented in Figure 1b and is a maze problem that includes the partial tasks allocated to subordinates in the task that must be processed by the manager.
3. Learning
Their exists a unit task to be handled by each agent which is processed by learning maze problems. Maze problems are used extensively as the benchmark for reinforcement learning. The proposed model employees the particularly basic Q learning. The renewal formula for the Q value in Q learning is shown as follows:

$$Q(S_t,\ a_t) \leftarrow Q(S_t,\ a_t) + \alpha(r_t + \gamma \max_{a \in A} Q(S_{t+1},\ a) - Q(S_t,\ a)), \quad (1)$$

where S_t, a_t is the agent's condition and behavior, and $Q(S_t,\ a_t)$ is the value of selecting a_t in regard to condition S_t. Moreover, $\alpha,\ \gamma,\ r_t$ are learning rate, discount rate, and compensation respectively. While agents processing unit tasks are to solve the maze step by step, the behavior selection technique employed is the ϵ-greedy method. This ϵ value makes use of the indifference index which will be explained in the next section and is set with the following formula:

$$\epsilon = 1 - indifference \quad (2)$$

where $indifference$ takes the real value in $[0,\ 1]$.
Therefore, an agent with a low indifference index becomes an exploratory agent. Unit task completion signifies arrival at the maze's goal, acquisition of compensation and a return to the starting point. In other words, processing one unit task is equivalent to moving through one level of maze problem learning.

4. Reports
   Agents who have completed unit tasks or partial tasks given by their superior go to the superior's desk to report. They join the end on the line if other agents are waiting at the desk to report. Staffs called up during task distribution are to join the front of the line.
5. Task completion
   Generated tasks and partial ones received from superiors are completed when unit tasks included have been processed. Therefore, tasks passed on to subordinates are explained as part of the critical path for the scheduling problem maze to be processed, as shown by the situation in Figure 1b. In other words, the given maze string as seen in Figure 1c of the same figure is completed when a single path has been laid to the goal. Partial tasks are reported to the superior upon completion, and generated tasks are registered on the processed task list when completed.

Incidentally, maze strings are produced by complex processes during generation of the above tasks. This results from modeling that dictates that the higher the level in the organization the more prolonged the job handled is to be, and the lower the level the shorter. The manager cannot reach process completion even if they have completed their own unit tasks until all partial task process reports are received from lower levels. Therefore, the higher the level of the manager who is given a task the greater the number of agents involved and hence the longer the period. Moreover, choosing a manager ad lib induces the election of lower level managers in a pyramidal structure. As a result, lower level managers frequently handle short-term jobs, while high level managers incidentally handle long-term jobs.

2.4 Indifference index and identification index

This section introduces organizational theory knowledge embedded in the proposed model as a distinguishing index of members composing the organization. In other words, the introduction of the independence of members due to indifference index and identification index is explained. Indifference index is an indication of compliance with orders from superiors. And identification index is an indication of the extent of identification of the member's purpose and values against that of the organization. Members are classified into the following four types in accordance with the values of two indexes [8]:

type 1 passive instrument type
The passive instrument type which has both a high indifference index and identification index is seen as an agent that is loyal to organization orders while also sharing common purpose and values. Although they carry out given tasks, there is little self-initiated action. Additionally, motivation does not become a problem as they share common purpose and values with the organization.

type 2 alienated worker type

The alienated worker type has a high indifference index and low identification index. Although drawing the same line with the organization in terms of purpose and values, it is referred to as a public official or bureaucrat type as orders are followed in behavioral terms.

type3 problem solving type

The problem solving type which has a low indifference index and high identification index is not simply a loyal follower of orders, but rather makes decisions through problem identification from an organization point of view based on shared purposes and values.

type4 non-contributor type

The non-contributor type which has both a low indifference index and identification index not only shares no common purpose or values, but is also in-compliant to orders and can not be expected to act organizationally. Not very common in stably function actual organizations.

In this way, even though each constituent of the organization has several factors governing their environment, they are broken down by these two indicators.

The identification attributes and indifference attributes shown in Table 1 correspond to identification index and indifference index and are derived from interest/attraction attributes. In the proposed model, identification index is the consistency of intention attributes with superiors, and indifference index is the inconsistency of the attraction attributes of presently held tasks with one's interest attributes. The length of these intention attributes and interest/attraction attributes match respectively, and are expressed by an ad lib character string containing defined quantity characteristics. This is a tag that can be seen in Axelrod's model "The Dissemination of Culture" [1], where a tag has number quantity properties for each length quantity characteristic.

Although identification index was conventionally an indicator of each member against the organization, it is used in the proposed model in terms of identification with the direct superior for greater micro-analysis.

In this model, identification attributes and indifference attributes are represented as real numbers in the domain [0, 1], and are derived from the following formula:

$$identification = \frac{1 - distance(staff.intention,\ manager.intention)}{length} \tag{3}$$

$$indifference = \frac{distance(staff.interest,\ task.attraction)}{length} \tag{4}$$

where $distance(tag1,\ tag2)$ is a function that seeks the hamming distance of tag1 and tag2 which have the same characteristics.

2.5 Environmental uncertainty

In addition, the proposed model handles an environmental uncertainty as the variability of maze problems. A task can be said to be an element with strong

Table 2. Parameters

Parameter		Value
Structure		1/3/9
Intention	Features	30
	Traits	2
Interest	Features	10
	Traits	3
Maze size		5
Block rate		0.2
Generate span		5

Table 3. Identification and indifference values

Agents	Identification	Indifference
Member 1	1.00	0.60
Member 2	0.53	0.80
Member 3	0.50	0.60
Member 4	0.30	0.90
Member 5	0.27	0.50
Member 6	0.43	0.60
Member 7	0.73	0.70
Member 8	0.53	0.70
Member 9	0.57	0.50
Member 10	0.50	0.60
Member 11	0.50	0.70
Member 12	0.50	0.80
Member 13	0.60	0.60

links outside of the organization, and it is thought that the extent of environmental uncertainty can be adjusted by giving it a rate-of-change. In simulation experiments it is expressed by the transfer of a block within the maze to a section without blocks at regular intervals.

3 Results

3.1 Basic experiment: under low uncertainty conditions

First, we implemented a simulation assuming a low level of environmental uncertainty. This is shown by an experiment that does not vary maze blocks. Parameters for this experiment are shown in Table 2. An organization where one top-level, three mid-level, and nine bottom-level members are active on three levels from the top-down is used as the organizational structure. In other words, this is a simple structure in which each agent has three subordinates. The agent name index is prioritized in term of width from the top of the organization down. With this parameter condition *Member 1*'s subordinates are *Member 2, 3*, and

4 and *Member i*'s subordinates are *Member* $3i-1$, $3i$, and $3i+1$ ($i = 2, \cdots, 4$). The length of the intention attribute tag associated with the identification index is 30, and the length of the interest/attraction attribute tag associated with the indifference index is 10. Maze problems of 5×5 in size with a block ratio of 0.2 are treated, and tasks (maze strings) are generated every five steps.

Simulation results obtained due from parameter conditions are shown in the left panel of Figure 2. The vertical axis is set as the number of steps involved from when the task is given by the superior until the report is completed, and follows variations during each step. The left panel of Figure 2a is a plot of the nine bottom level agents, while that of Figure 2b is a plot of the three mid-level agents. The superior and subordinate's id are given in the two legend brackets. The identification and indifference of each agent are shown in Table 3. Please note that all the parameters are fixed in the whole simulation for simplicity or tractability. In other words, all the agents in this model are not evaluated by their bosses.

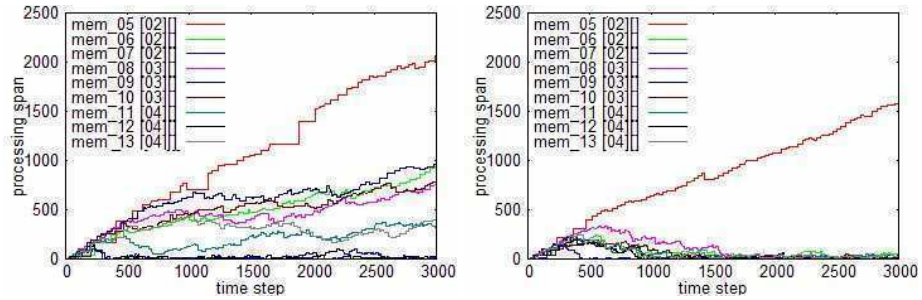
Firstly, an apparent fact from Figure 2a is that *Member 5*'s task processing efficiency is strikingly poor due to a low identification index. Although unprocessed tasks are piling up on *Member 5*'s desk, he/she does not actively process tasks since he has a low identification with his superior, *Member 2*. Other agents seem to be carrying out routine work because their task processing cycle is broadly flat. Since *Member 7* and *Member 12* require few steps for each task process, we can consider them as high performers in this environment.

Upon examination of mid-level results in Figure 2b, one's attention is drawn to the behavior of *Member 2*. This is due to the coexistence of the high performing *Member 7* and poorly motivated *Member 5* in the same department. If tasks related to *Member 5* are contained within the maze string given by *Member 2* who acts as manager, then process completion is kept waiting. Conversely, tasks handed to *Member 7* return quickly as *Member 7* has a high task processing ability. This misfit is represented as this kind of rectangular wave.

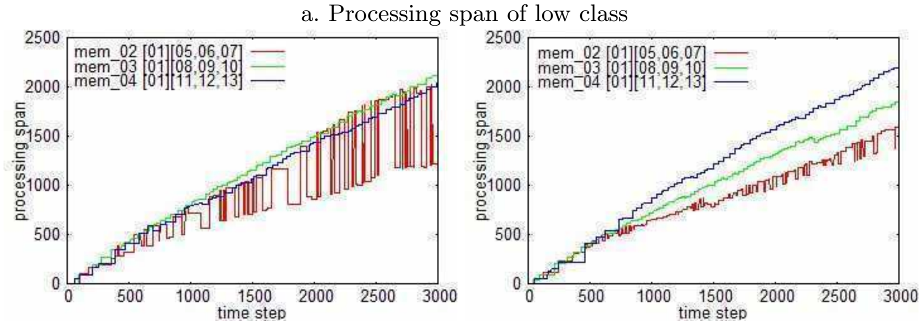
Also, Figure 2c represents the state ratio of each member after 1000 steps, allowing intuitive understanding of their various predicaments. The distinguishably low motivation of *Member 5* is also represented by this result. More than half of all work time is spent shirking, or avoiding responsibility. *Member 5* has a strong 'non-contributor' tendency. Furthermore, this chart shows that time allocated to the management of subordinates by middle management *Member 2, 3, and 4* is more distinguishable than expected, and it can be seen that they have almost no free time.

When reconsidering Figure 2b with this in mind, the task processing period is ever-increasing on a management level even though each agent is task learning. The cause is a short task generation cycle, and it can be said that this result reflects the high-load of middle management in the organization.

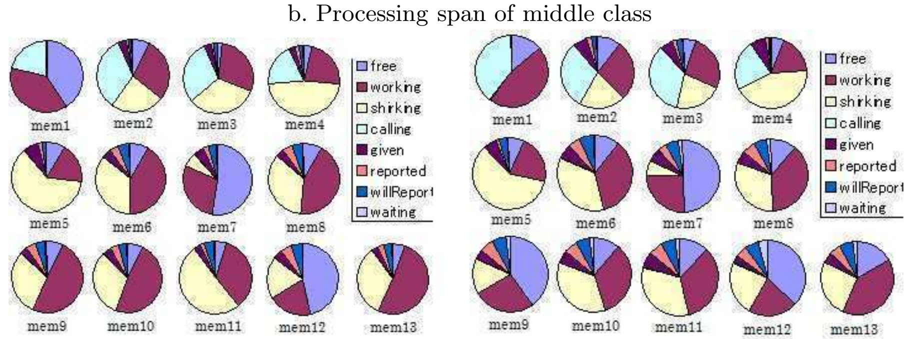
The above result allows a glimpse of one facade of organizational activity. However, it is difficult to tell from a member's indifference index distribution whether or not maze problem learning is functional as the ϵ value is relatively large. Consequently, the value of ϵ was set to zero, and the ϵ-greedy method



a. Processing span of low class

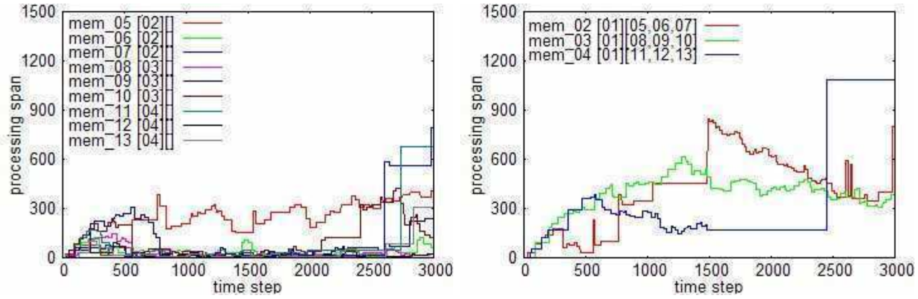


b. Processing span of middle class



c. Task processing situation of each member after 1000 steps

Fig. 2. Comparative performances (left panel: basic, right panel: in case of ϵ-greedy method)



a. Performance of the low class b. Performance of middle class

Fig. 3. Comparative performances after adding maze change parameter

utilized in an experiment assuming only maze solving agents. The result is shown in the right panel of Figure 2.

Member 5 is again the bottom level agent as was the case with the right panel of Figure 2a. However, an improvement due to learning can be seen in the task processing efficiency of agents excluding *Member 5*. This processing efficiency arises solely from the identification index as only knowledge utilization is relied on without any exploration. It can be said that *Member 5* with his low identification index has few learning opportunities, as processing improvement cannot keep pace with task generation. Additionally, agents were more exploratory when attacking maze problems in the first experiment than this one as the ϵ value is set by the method described in the proposed model. This is represented by the fact that the work time of most agents is longer in the left panel of Figure 2c results than those in the right panel of the same figure. The above comparative experiment shows that exploratory attributes work negatively when there is no maze variation and low uncertainty, resulting in more time tied down by work.

3.2 Advanced experiment: increases in uncertainty

Next, we will describe an experiment aimed at analyzing the effects of environmental uncertainty. In other words, an experiment was conducted where maze blocks varied in a given cycle.

The parameters used in this experiment are the same as those of the previously mentioned basic experiment, only that the task generation cycle is doubled, assuming an organization of reduced load. Parameters distinguishing each member are the same in Table 3 as intention, interest/attraction attributes are unchanged.

Furthermore, an element is added in which a maze block held by each member changes position every 500 steps during the simulation period. The simulation result is shown in Figure 3. The result Figure 3a for bottom level agents and Figure 3b for mid-level agents are both of increased complexity. Of particular note here is the behavior of *Member 7* who was seen as a high performer in the previous experiment. Processing time rapidly increases around the 2700 step mark. These phenomena can also be observed for other agents in the later half of the simulation. It seems that repeated learning has taken place, and the path to the goal has been reinforced by over-learning. The higher the performer, the stronger this tendency is within stationary organizations. Therefore, he is more likely influenced when there is variation in the maze.

4 Conclusion

This research modeled and conducted various types of experiments regarding agent simulation that expresses the task processing of active members of the organization using the maze problem solving of reinforcement learning agents. While taking the bottom-up method as the fundamental rule when modeling, it became possible to analyze organizational phenomena with environmental uncertainty taken into account through sufficient utilization of organizational theory knowledge. Simulation results have clarified what happens in organizations from a micro point of view, and also helped confirm that appropriate organizational form differs per environment for organizations effected by environmental uncertainty. In other words, a high performer in the organization in a given environment is not necessarily so when placed in a different environment. That is to say, this phenomena which occurred in a section of the organization expands bottom-up to a macro phenomena encompassing the entire organization. As a consequence, it can be clearly stated from an agent-based simulation stance that an individual's drop in performance effects the entire organization, and that appropriate organizational form differs per environment.

We will study the relations between the results and actual activities of organizations, and the influences when introduced evaluation of agents by others.

References

1. Axelrod, R.: The complexity of cooperation: agent-based models of competition and collaboration. Princeton University Press (1997)
2. Axelrod, R., Cohen, M.D.: Harnessing complexity: organizational implications of a scientific frontier. Free Press (2000)
3. Carley, K. M.: Computational and mathematical organization theory: perspective and directions. Compu. and Math. Org. Th. **1** (1995) 39–56
4. Carley, K. M., Gasser, L.: Computational organization theory.in Weiss, G. (ed.): Multiagent systems - a modern approach to distributed artificial intelligence. MIT Press (1999) 299–330
5. Cohen, M. D., and J. G. March: A garbage can model of organizational choice. Administrative Sci. Quart. **17** (1972) 1–25
6. Cyert, R. M., March, J. G.: A behavioral theory of the firm. Prentice-Hall (1963)
7. Epstein, J. M., Axtell, R. L.: Growing artificial societies: social science from the bottom up. MIT Press (1996)
8. Takahashi, N.: An evaluation of organizational activation. Omega: Int. J. Manag. Sci. **20** (1992) 149–159
9. Terano, T.: Beyond the KISS principle: from agent-based social simulation. Proc. of 2nd AESCS (2005) 69–78

Session on Methodological Issues

Capturing heterogeneity in empirical agent-based models: A guideline

Alex Smajgl[1], Erin Bohensky[2], and Iris Bohnet[3]

[1] CSIRO Sustainable Ecosystems, University Drive,
Townsville QLD 4810, Australia, alex.smajgl@csiro.au
[2] CSIRO Sustainable Ecosystems, University Drive,
Townsville QLD 4810, Australia, erin.bohensky@csiro.au
[3] CSIRO Sustainable Ecosystems, University Drive,
Atherton QLD 4883, Australia, iris.bohnet@csiro.au

Abstract. Empirical applications of agent-based modeling to research questions related to social groups face the challenge to capture heterogeneity of the group population without compromising the practicability of such applied project work. Clustering is often a techniques used for implement behavioural aspects of cognitive agents. This paper is based on experiences from two case studies and develops a first step towards a guideline for parameterising behavioural variables and how different methodologies can be effectively sequenced.

1 The Challenge of Modelling Heterogeneity

Agent-based modelling captures interactions between heterogeneous agents which lead to emerging macro patterns [6]. However, empirical agent-based models that use real-world data cannot represent the heterogeneity of hundreds or thousands of individuals or households. Capturing heterogeneity at a level appropriate for the given research question therefore becomes a crucial modelling challenge. Developing a typology to group individuals or households according to similar traits that are identified as critical is a potential solution [9]. Unfortunately, the majority of agent-based models remain in the theoretical domain and are based on hypothetical assumptions. A few examples of empirical agent-based models exist, but the methodological discussion of parameterisation or calibration of cognitive agents is not well documented [2].

This paper discusses a methodology for the calibration of behavioural aspects of agent based models which involves five steps:

1. Gathering of data for a typology of agents
2. Definition of types of agents
3. Collation of data on behavioural responses
4. Definition of behavioural agent rules
5. Assigning of agent rules and spatial locations to real world population

We first present two case studies that demonstrate the range of variation in agent-based modelling approaches and why empirical calibration guidelines are needed. This guideline is based on two major model characteristics that define method selection: (1) the degree of behavioural complexity among agents and (2) the population size. From these, six modelling domains are identified and recommendations made for effective sequences of methods to calibrate behavioural rules in different types of empirical agent-based models.

2 Experiences in the Bowen Broken, Australia

The first case study simulates cattle grazing activities in the Bowen Broken sub-catchment in the Great Barrier Reef (GBR) region, which comprises 68 households at low population densities. The research objective is to model management practices – specifically the effects of incentive structures on graziers' practices regarding fencing and stocking rates of cattle – on sediment delivery into the GBR lagoon. The small number of agents and the small number of dimensions for defining the relevant aspects for grouping the population into heterogeneous groups led to a decision to use a qualitative approach [11]. The study commenced with a literature review that was carried out to provide background and context to the research. Subsequently, meetings were held with researchers and natural resource managers who are familiar with the Bowen-Broken catchment and the graziers. Relevant maps were then developed before the qualitative interview schedules and landscape assessments were designed [3, 12]. From a range of possible qualitative interview techniques, the 'semi-structured interview' was chosen as the best suited for capturing the same information in all interviews. The interview questions were open ended and were focussed on management strategies, history of the grazing property and the grazier's values, economics of their business and possible responses to new incentives, and future aims and aspirations related to the property. This database was analysed using a general inductive approach, whereby transcripts were coded for emerging themes and categories as well as for predefined issues. Data was coded with the qualitative software package ATLAS.ti (Muhr, 1997).

In order to capture all possible types, the interviewees were sampled so that those selected included graziers from different locations within the catchment, different age groups, female and male graziers, and graziers who manage their properties using different approaches. In addition, a snowballing technique, where an interviewee suggests potential other interviewees, was used to gain access to as many different graziers as possible. This sampling procedure was repeated until no new information could be gained from the interviews. Nine in-depth interviews and landscape assessments with 13 graziers were carried out at their stations and six further interviews were carried out over the phone. Four types of agents were identified based on: management strategy (rotational grazing, cell razing, continuous grazing, and additional mining interests); average paddock size; stocking rate; property improvement plans; supplementary feeding; spelling and rotation. The agent types were mapped into the model with satellite imagery and average paddock size was used as an identifier for the typology.

3 Experiences in East Kalimantan, Indonesia

The East Kalimantan example reflects very different conditions for the development and implementation of an agent-based model. The model, currently in development, aims to simulate household responses to energy policy changes such as kerosene subsidy removal or fuel mix change, which drive household livelihood strategies involving forests and other natural resources. This is a multidimensional problem that has implications for a broad range of livelihood aspects that must be brought into the model. To simplify the problem somewhat for the purposes of the model, households with particular characteristics that are expected to respond to policy changes similarly were grouped.

Additionally, the number of discriminators (19 livelihood options, 10 wellbeing variables, 16 demographic attributes), and the size of the population (2 million people) makes the process that was used for the Bowen Broken study unrealistic in the East Kalimantan case. Instead, a two-step approach was used. First, a survey instrument was developed in order to collect household data on livelihoods and wellbeing. Based on this data a cluster analysis was employed to develop a typology of agents. Semi-structured interviews were then conducted in order to develop (in a series of "what-if" questions) the behavioural responses to policy changes. At the beginning of the interviews, questions were asked that enabled interviewees to be categorised according to the typology defined previously. Following this, in workshop situations with focus group discussions, agent rules were identified from the qualitative and quantitative data collected during the interviews. Finally, in order to apply the typology to all households in the East Kalimantan study area, clear discriminators were isolated that linked to one agent type only and were able to be obtained from government census data. Census information was also used to assign the spatial location of agents.

4 Lessons learnt from case studies

The development of empirical agent-based models requires an effective design of agents and their behavioural responses to scenario assumptions. While data sets for passive (non-cognitive) agents are often available, behavioural data tends to be more problematic. Developing such behavioural data based on social science techniques often follows an ad-hoc approach based on the availability of resources. Other agent-based models have employed ethnographic methods [8], semi-structured interviews [10], census data [7], surveys [2, 4], and role-playing games [1].

Empirical agent-based modelling experiences such as those offered by the two case studies above have helped to identify two main drivers of the methodological selection: the behavioural (or system) complexity and the population size (Figure 1). Behavioural complexity is clearly linked to the research question. This aspect ranges from one-dimensional marketing-oriented demand responses to price changes to multi-dimensional research questions (i.e. policy impact on land management in Bowen Broken) to very complex research questions (i.e. cross-sectoral policy impacts on livelihood and well-being in East Kalimantan). Sample size determines the methodological selection as well. For instance, a survey may not be appropriate for a population of 20, while in-depth interviews are unlikely to be realistic for a sample size of 1,000.

Given these two drivers, six domains of agent-based model applications can be distinguished and methodological guidelines can be identified for the five steps of calibration listed above.

Domain A: Models in this domain are developed for a highly focused one-dimensional research question such as sales responses to a price change of a specific brand. If the number of consumers is very small, existing data can be analysed with statistical methods or new data can be gathered by a structured interview. This step can often provide data for a typology (step 1, above) and behavioural responses (step 3) simultaneously. Often typologies are determined by the one-dimensional research focus (i.e. such as whether a consumer will change brands). Otherwise a workshop situation can effectively allow the development of a typology (step 2) and agent rules (step 4). Step 5 is often possible based on existing customer databases.

Domain B: Semi-structured interviews have proven to be very effective in this domain to gather typology data (step 1), as in the Bowen Broken case study described above [3, 12]. Simultaneously, step 3 can be considered in the interview instrument. In a workshop situation typologies can be identified (step 2) and rules can be developed (step 4). Applying the typology-specific agent rules to the whole agent population (and their spatial distribution) can be based on census data or satellite imagery (step 5). A clear identifier is required for each agent type that is contained in either the database (all agents that are students are type X) or the imagery (all fenced properties are managed by type Y).

Domain C: The complexity of the research question does not allow steps 1 and 3 to be combined. Instead, data for the typology has to be gathered separately and semi-structured or structured interviews have proven to be very effective. Statistical grouping methods such as cluster analysis or vector analysis can be applied to identify agent typology (step 2). Given the typology, semi-structured or in-depth interviews can be conducted for collating what-if responses (step 3). Focus group discussions that combine experts and modellers are best suited for translating the data into agent rules (step 4). A critical component here is often the inclusion of expert or local knowledge of the problem context. Depending on the research focus, clear identifiers for agent types can be linked to satellite imagery or census data to assign the typology to the entire agent population.

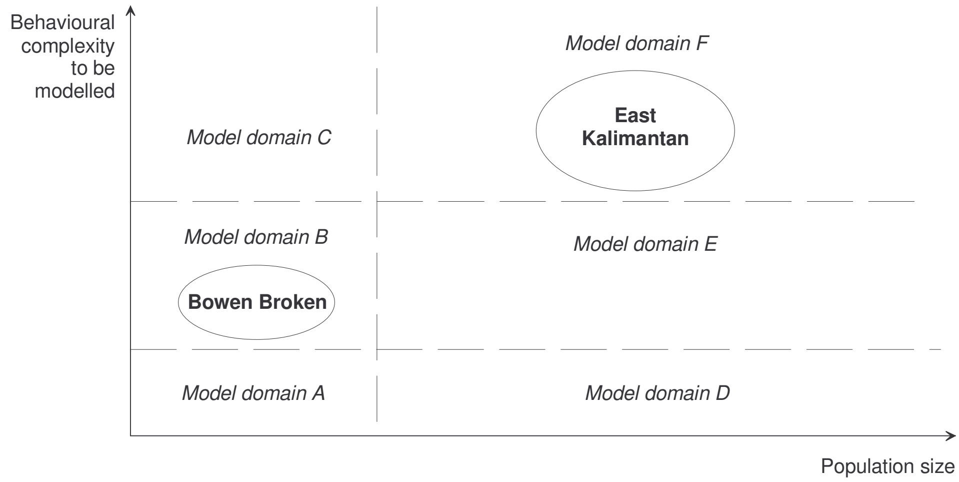


Figure 1: Empirical model domains depending on behavioural complexity and population size

Domain D: Survey instruments are most likely to be effective when modelling a large population with a low level of complexity, and steps 1 and 3 can be combined in the same instrument. Alternatively, the reduced dimensionality can allow a clustering based on existing data bases such as census data [5, 7]. Similar to *Domain A*, step 2s and 4 can be pre-defined by the research focus. In step 5 census data can be used based on clearly discriminating attributes.

Domain E: With growing complexity a census based clustering becomes less effective as attributes of interest are not typically captured in such datasets. Surveys are effective for step 1 while step 3 requires a methodology that allows the modeller to explore "what-if" contexts, such as semi-structured interviews. Statistical methods (i.e. cluster analysis) can determine agent types (step 2). Translating interview data into behavioural agent rules (step 4) can be accomplished in a workshop that links local experts and modellers. Step 5 requires distributing the agent types across a given landscape and/or the entire agent population. Depending on the identifier for each typology, census data or satellite imagery can be effective.

Domain F: In this domain of greatest population size and complexity, as illustrated by the East Kalimantan case study, surveys or semi-structured interviews can be used to collate data for an agent typology (step 1) and will generally rely on statistical grouping methods (step 2). Step 3 can sometimes be based on semi-structured interviews but can also require in-depth interviews depending on required sample size and the complexity of scenario assumptions that define the "what if" context for the behavioural agent rules. Agent rules can be developed in a focus group discussion (step 4) and step 5 can be based on satellite imagery or census data.

The two case studies provide examples of two different approaches to empirical agent-based modelling, each of which requires the development of a typology of agents and uses a tested sequence of methods. The six domains illustrated above represent the majority of agent-based modelling cases, and collectively provide a guideline for calibrating behavioural aspects of agent-based models in a way that captures heterogeneity in the system of interest but deals realistically with logistical or modelling constraints.

References

1. Barreteau O, LePage C, D'Aquino P. Role-playing games, models and negotiation processes. JASSS - The Journal of Artificial Societies and Social Simulation 2003; 6 (2)
2. Berger T, Schreinemachers P. Creating Agents and Landscapes for Multiagent Systems from Random
   Samples. Ecology and Society 2006; 11 (2)
3. Bohnet, I. Sustainable Futures for graziers and grazing communities in the Bowen Broken catchment. 2007. Atherton, CSIRO. Water for a Healthy Country working paper. (GENERIC)
   Ref Type: Report
4. Brown DG, Robinson DT. Effects of Heterogeneity in Residential Preferences on an Agent-Based Model of Urban Sprawl. Ecology and Society 2006; 11 (2)
5. Dewi, S., Puntodewo, A., Tarigan, J., Sitorus, S., Achdiawan, R., and Belcher, B. Modelling Deforestation in East Kalimantan at the pixel level and village level. 2007. BOGOR, CIFOR. (GENERIC)
   Ref Type: Report
6. Epstein JM, Axtell R. Growing Artificial Societies: Social science from the bottom up, Washington D.C: The Brookings Institution. 1996
7. Happe K, Kellermann K, Balmann A. Agent-based Analysis of Agricultural Policies: an Illustration of the Agricultural Policy Simulator AgriPoliS, its Adaptation and Behavior. Ecology and Society 2006; 11 (1)
8. Huigen MGA, Overmars KP, de Groot WT. Multiactor Modeling of Settling Decisions and Behavior in the San
   Mariano Watershed, the Philippines: a First Application with the MameLuke Framework. Ecology and Society 2006; 11 (2)
9. Janssen MA, Ahn TK. Learning, signaling and social preferences in public good games. Ecology and Society 2006; 11 (2)
10. Purnomo H, Yasmi Y, Prabhu R, Yuliana L, Priyadi H, Vanclay JK. Multi-agent Simulation of Alternative Scenarios of Collaborative Forest Management. Small-scale Forest Economics, Management, and Policy 2003; 2 (2):277-292
11. Silverman D. Doing Qualitataive Research - A Practical Handbook, London: Sage. 2000
12. Smajgl, A., Heckbert, S., and Bohnet, I. Simulating the grazing system in the Bowen Broken catchment: An agent-based modelling approach. 2007. Townsville, CSIRO. Water for a Healthy Country working paper. (GENERIC)
    Ref Type: Report

Brief Note on the Logic of Replicating Implementations Before and After Publishing a Model

Nuno David

Department of Information Science and Technology
Instituto Superior de Ciências do Trabalho e da Empresa (ISCTE)
Lisboa, Portugal
Nuno.David@iscte.pt

Abstract. This short paper introduces a hypothesis complementary to the current logic of replicating computerized models in social simulations. It is submitted that the logic of cumulative knowledge provided by the process of peer reviewing may become more resourceful once the effort of replication becomes focused on producing multiple computerized versions of the same conceptual model *before* actually submitting it for peer reviewing, in contrast to the usual practice of having other teams aligning and replicating models after they have already been reviewed, accepted and published. Note that while this is a similar approach to Edmonds and Hales' proposal, it differs from a methodological point of view, insofar as they double replicated an already published model. While this approach would require more manpower in simulation projects, it is argued here that the resulting benefits may override the costs, particularly in participative-based simulations, where potential under-verified simulations may lead, nevertheless, to the actual implementation of policies with unpredictable effects in the economic and social life of stakeholders.

1 The Logic of Replication and Challenges in Simulation

The goal of this short article is to contribute a few reflections for the problem of replicating social science simulations. A hypotheses is introduced which may be both complementary and alternative to the current logic of replicating computerized models in social simulations, namely: that the logic of cumulative knowledge provided by the process of peer reviewing may become more resourceful once the effort of replication becomes focused on producing multiple computerized versions of the conceptual model before submitting the latter for peer reviewing. This contrasts the usual practice of having other teams replicating models after they have already been reviewed, accepted and published.

This proposal draws on methodological and epistemological arguments, as well as one technological argument:

a) Rather than the computerized model, it is the conceptual model that is susceptible to being confirmed or infirmed by the scientific community. It is the conceptual model that is internalized, transformed, generalized and connected to other disciplines and methods, and eventually incorporated into a community consciousness,[1] in contrast to the inability of computer programs to become scrutinized by the scientific community.

b) Only very simple and highly abstract models – in the sense of the Axelrod's KISS motto – are susceptible to being re-implemented; however, if one adopts a KIDS philosophy for constructing simulations,[2] which is a typical approach in participative-based simulations, computerized models are hardly susceptible to becoming re-implemented by other teams, given the context dependency of individual and institutional stakeholders' opinions; however, these models can – and, it is argued, should – be subject to replication by the team proposing the model, *before* publishing or applying its results in the relevant field. This seems particularly important in policy making, where un-replicated simulations in complex domains run the risk of being under verified, while leading to the implementation of specific policies which influence the economic and social context of individual and institutional stakeholders.

c) Often, the correspondence among conceptual and computerized models is not only established formally and empirically, but also intentionally by researchers and stakeholders. Insofar as the semantic richness of social processes surpasses the empirical expressiveness of computer programs, the intended meanings ascribed to computer programs are evaluated experimentally in an intentional way. This means that the conceptual model is actually the one being internalized by the researchers and stakeholders along the executions of the computerized model, whereas the syntactic structure and operational semantics of the computerized model plays a secondary role in interpreting experiments.[3]

d) Articles on replication seem to have a low academic return (Galan et al. 2002). Most replications seem to be presented in forums especially dedicated to replication[4], rather than resulting from spontaneous confirmation or infirmation by other scientists according to the logic of peer reviewing and cumulative science, except when the goal is to extend the original conceptual model.

e) The increasing portability of programming languages for the Web, such as Java and XML, provides technologies for executing and visualizing models through Applets or Servlets on the Internet. These technologies may be able to facilitate the comparative analysis of different implementations of the

---

[1] C.f. De Millo et al. (1979).
[2] See Edmonds and Moss (2005).
[3] See David et al. (2005;2007).
[4] For instance, the well known Model-to-Model workshops.

same model, and are susceptible to being submitted to peer-review together with the description of the conceptual model.

2. Replication through N-Version in the Development Process

Given all these reasons, it is suggested that the following topics be introduced into the methodological debate:

1) A distinction between replicating a model before and after it has been submitted for peer reviewing and published for the first time;

2) Two or more replicas of the computerized model should be submitted to the peer-review process along with the conceptual model, in order to confirm the verifiability of the results obtained, under the penalty of having the article rejected for reviewing and publication.

Like Edmonds and Hales (2003), the author believes that "an unreplicated simulation is an untrustworthy simulation – one should not rely on their results, they are almost certainly wrong," in the sense that the computerized model differs from what was intended or assumed in the conceptual model. This is the well-known problem of verification.[5]

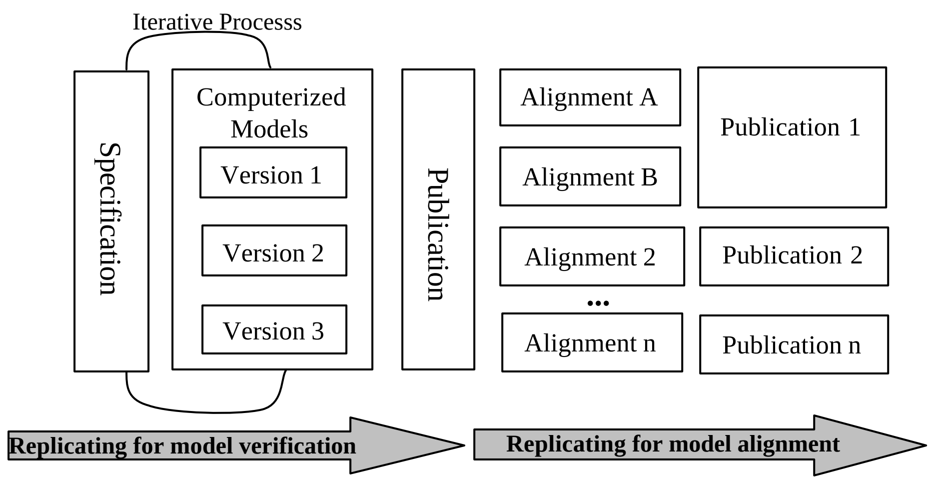


Fig. 1. Replication through multiple versions *vs*. replication for model alignment.

The report of Edmonds and Hales is indeed particularly informative in this context. In order to align with a published conceptual model, they had to rely on a double

[5] For a definition of the term "verification" in the context of simulation see David (2007).

replication process, insofar as the first replication did not seem to result in alignment. It was only after implementing a second replication of the original model that the authors were sure that the results reported in the original model did not align with their own results, and hence concluded that the original implementation was leading to possibly misleading results.

The essence of the present argument is that the burden of verifying computerized models through multiple implementations, on different platforms and by different people in the same team, should be carried out before the model is published. Consider Figure 1 above. The left part illustrates a methodology where a common specification should produce a number of different implementations carried out by different persons on the same team. The right part of the figure illustrates the usual approach to replication. Both kinds of replications have a useful role, and may be complementary.

As for the left part of Figure 1, the implementation of multiple versions before publication should be attained by an iterative process, whereby inconsistencies in the different versions would result in new corrected versions or a re-specification of the model. If all versions result in the same outputs, then there are reasonable grounds for confidence in the results obtained, and the model is ready to be published. A similar approach can be traced back to the so called N-Version Programming technique in software engineering.[6]

All versions should be made accessible to the reviewers, along with the description of the conceptual model, and reports on the equivalence of versions should be provided. The fact that the model is only published after going through a process of program verification via replication should increase overall confidence in the results obtained.

3. Prospects on N-version Programming in Simulation

It is submitted that the proposed methodology would have the following benefits:

a) It would foster the replication of simulations, regardless of the types of models and targets. Whereas the replication of participative-based simulations is rarely reported in the literature, this methodology would foster the provision of multiple implementations of the model before peer-reviewing even in the case of participative-based simulation, thereby increasing overall confidence in the model proposed, and the results obtained, whereas an un-replicated simulation in complex domains would run the risk of being under-verified but lead, nevertheless, to the actual implementation of policies

b) The burden of showing that the simulation is verified would be on the team proposing the simulation, rather than on the recipient researchers. Conversely, it would be the role of the recipient researchers to confirm or refute the validation and usefulness of the conceptual model, rather than to verify the computerized model.

[6] See Avizienis (1995).

It is also believed that the effort in replicating computerized models before submitting the corresponding conceptual model to publication comes with a significant overhead. In particular, it requires that the team enrols in multiple implementations, requiring more manpower in projects.

It is submitted, however, that the benefits obtained may override the costs, particularly in the context of participative-based simulations. If the dissemination of knowledge obtained via simulation is to acquire a reliable status of cumulative science, it is necessary to make sure that the verification stage of the simulation development process is rigorously carried out before publishing the results, and that this process can be adequately scrutinized by the scientific community after the model is published.

References

1. Avizienis, A. (1985) The N-version approach to fault-tolerant software. *IEEE Transactions on Software Engineering*, 11(12), 1491-1501.
2. David, N., Sichman, J. and Coelho, H. (2005). "The Logic of the Method of Agent-Based Simulation in the Social Sciences: Empirical and Intentional Adequacy of Computer Programs". *Journal of Artificial Societies and Social Simulation* 8(4).
3. David, N. (2007). "Verification and Validation in Social Simulation: Patterns and Clarification of Terminology Use". Article from the workshop Epistemological Perspectives on Simulation, Italy, submitted to Synthese Book Series. Draft version available at: http://iscte.pt/~nmcd/papers/drafts/PatternsNDavid.pdf.
4. David, N., Sichman J. and Coelho H. (2007. "Simulation as Formal and Generative Social Science: The Very Idea" In C. Gershenson, Aerts, D. and B. Edmonds (editors) *Worldviews, Science and Us: Philosophy and Complexity*, World Scientific, 266-284.
5. DeMillo R., Lipton R. and Perlis A. (1979). Social Processes and proofs of theorems and programs, *Communications of the ACM* 22, 5 (May), 271-280.
6. Edmonds, B. and Hales, D. (2003). Replication, replication and replication: Some hard lessons from model alignment. *Journal of Artificial Societies and Social Simulation* 6(4).
7. Edmonds, B. and Moss, S. (2005). From KISS to KIDS – an 'anti-simplistic' modelling approach. *Multi Agent Based Simulation*. Springer, Lecture Notes in Artificial Intelligence, **3415**:130–144.

6\. J.S. Galán, J.M. Tom Downing, A. López-Paredes (2002) "A standardization framework for replicating agent-based models. A case study SDML vs Repast.". First conference of the European Social Simulation Association, September 18-21, Groningen, The Netherlands.

Agent Based Simulation Framework for Quantitative and Qualitative Social Research: Statistics and Natural Language Generation

Samer Hassan[1], Juan Pavón[1], Millán Arroyo[2], Carlos León[1]

[1] Dep. Ingeniería del Software e Inteligencia Artificial. Universidad Complutense de Madrid
samer@fdi.ucm.es, jpavon@sip.ucm.es, cleon@fdi.ucm.es
[2] Dep. Sociología IV, Universidad Complutense de Madrid.
millan@cps.ucm.es

Abstract. Even though Agent Based Social Simulation is beginning to be spread out as a powerful quantitative method for sociologists, it is still far from attracting qualitative ones. We propose to broaden ABSS horizons with a system that returns outputs useful for both paradigms. The case used as example is the study of the evolution of religiosity in the Spanish post-modern society. From a "macro" perspective, it analyses social trends, using quantitative data from the European Values Survey and giving evolution statistics. From the "micro" one, it can generate a narrative personal story of a representative agent, much closer to the Interpretativism tools. Several methods for choosing the most representative agent are commented.

Keywords: Social simulation, abss, mas, qualitative, quantitative, social research, natural language generation

1 Introduction

Sociologists and other social scientists use different research methods in order to study human societies. Social methods can generally be subdivided into two broad categories, even though both involve a systematic interaction between theories and data [4][9]. Quantitative methods, based on natural science and positivism, are concerned with attempts to quantify social phenomena and collect and analyse numerical data, with focus on the links among a small number of variables across many cases. For them, the social world is something that is *out there*, external to the social scientist, and waiting to be researched (objective vision). Qualitative methods, on the other hand, are based on comprehension, which emphasizes personal experiences and interpretation (subjectivism) over quantification, and is more concerned with understanding the meaning of social phenomena, with focus on links among a larger number of attributes across relatively few cases. Interpretivists believe that the social world is constructed by social agency and therefore any intervention by a researcher will affect social reality (interaction builds social reality).

However, it is increasingly recognized that the significance of these differences should not be exaggerated and that quantitative and qualitative approaches can be and must be complementary [6]. Quantitative methods are useful for describing social phenomena, especially on a larger scale. Qualitative methods allow social scientists to provide richer explanations (and descriptions) of social phenomena, frequently on a smaller scale. By using two or more approaches researchers may be able to 'triangulate' their findings and provide a more valid representation of the social world [10]. In fact, we can see more and more how sociologists of "one side" use tools from "the other side" in an assistant way.

There are already several works that show how social simulation can be a useful tool for quantitative researches [12]. From this point of view, agent based modelling even could be thought as an experimental tool for theoretical quantitative approach, a platform for empirical studies of social systems. With the aim of supporting this statement, we are working on a social simulation system for a specific sociological problem, trying to use it as a social research lab.

Qualitative researches have been used in this field just in an assistant way (in a second level), working from the quantitative perspective and letting it to guide the analysis and results view, as it can be seen in [16]. On the other hand, we want to propose, even knowing that the use of computers is mostly linked to quantitative researchers, the use of the results of agent based social simulation (ABSS) also for qualitative social scientists, at least in an assistant way. The approach is to consider ABSS not only for getting results that may be useful for a quantitative analysis of a social model, but for providing some assistance from a qualitative perspective. This is made by having advantage of the ability of agents to evolve autonomously and their interactions with the environment and other agents. By recording "stories" of agents, it is possible to get knowledge about their particular evolution and build a narrative personal story, which may provide insight on their motivation along time.

This work extends some experiments that have been performed by our research group with some sociologists of our University concerning the study of the evolution of religiosity in Spain [14]. In principle, the ABSS used for this case was applied to get quantitative data around a set of features that characterize the individuals. Here we are adding the generation of stories on some particular agents of the society, in order to get some qualitative argumentation. The next section presents roughly the multi-agent system that has been used for modelling the society. Then, section 3 presents a macro perspective of the analysis of the system, which would correspond to the quantitative approach. And next section goes to the individuals, the micro perspective, and explains how stories are generated and their use for qualitative research. At the end we provide some discussion on the potential of using ABSS for gaining knowledge on a society from the combination of qualitative and quantitative approaches.

2 Modelling of Social Systems with Multi-Agent Systems

Social phenomena are extremely complicated and unpredictable, since they involve complex interaction and mutual interdependence networks. Quantitative sociological explanations deal with large complex models, involving many dynamic factors, not subject to laws, but to trends, which can affect individuals in a probabilistic way. A social system consists of a collection of individuals that interact among them, evolving autonomously and motivated by their own beliefs and personal goals, and the circumstances of their social environment.

The idea beneath ABSS is that we may be able to understand this huge complexity not by trying to model them at the global level but instead as emergent properties of local interaction among adaptive autonomous agents who influence one another in response to the influences they receive [12]. Because of that, the specification of characteristics and behaviour of each agent is critical, so it can manage the dimensions of the studied problem.

In the Multi-Agent System (MAS) designed, as explained in [14], agents have been developed with several attributes: from simple ones such as sex or age, to complex ones like for example ideology or educational level. The attributes have been chosen with respect to the context of the problem to solve, as we will see. The population in the agents' society (as in real societies) also experiments demographic changes: individuals are subject to some lifecycle patterns: they are born, get married, reproduce and die, going through several stages where they follow some intentional and behavioural patterns.

Moreover, agents/individuals can build and be part of relational groups with other agents: they can communicate with other close agents, leading to friendship relationships determined by a rate of similarity. Or, on the other hand, they can build family nuclei as children are born close to their parents. We can see a snapshot of the agents' space in Figure 1. The friendship relationships are represented by yellow links and families by green ones. Agents have different colour depending on their age and sex.

Taking the underlying sociological model, the parameters of the social simulation system fit all together logically. In this way, the system may be configured to reflect the parameters (such as average number of children per couple, or mean of male average age of death) from a specific country or even import data from surveys that specify the attributes of the agents, reflecting the behaviour of the given population.

Besides, due to the relative simplicity of the agents, the system can manage hundreds (and even thousands) of them, reaching the necessary amount for observing an emergent behaviour that results from the interactions of individuals, leading to the appearance of social patterns than can be studied [3]. And for this study, during and after the execution of the simulation tool several graphs may be plotted that reflect the evolution of the main attributes of the social system.

The system robustness has been tested enough to demonstrate the stability of the results, needed for the macroscopic analysis. Besides, the system has a deep diversity, attending to the differences between individuals. This fact will be useful for the micro analysis.

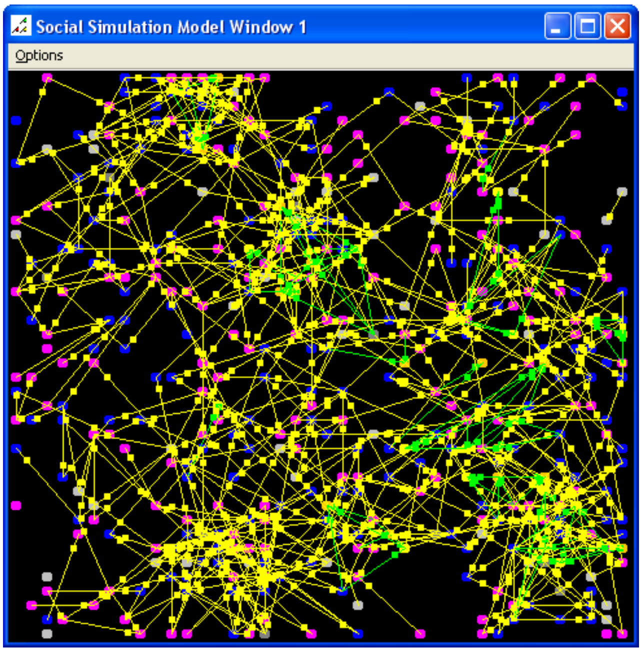


Fig. 1. Snapshot of the social simulation, after a certain period of time. We can see complex networks of friends and family nuclei all over the space.

3 The "Macro" Approach: Experimentation for Quantitative Research

The case under study, which is used to illustrate the approach in this paper, makes an analysis of the evolution of religiosity in Spain between 1980 and 2000. Initially, the aim was to assess the usability of ABSS tools for sociologists, who are not skilled in computer programming. Therefore, we looked for a real sociological issue, as the one of religiosity in society. The problem was how to model this social problem with agent concepts.

Initial data for the model and the simulation has been taken from results of the European Values Survey (EVS), which periodically make surveys in all European countries. EVS provides a source of quantitative information and periodical results offer data for validation of the simulation model [1]. The problem chosen and its simulation is widely explained in [14], so here we only focus on the main points. We just want to notice that religious people were divided into four types [2], as we can see in Table 1: ecclesiastical are clearly the most conservative, followed by low-intensity. Then, alternatives have a more modernized mentality, while the most modern and left-winged are those non-religious.

Taking the EVS-1980 of Spain, the sociologist was able to build an Excel spreadsheet for the characterization of 500 individuals, which statistically represent the Spanish population. These data were taken as input to generate a population of agents in the model, which was simulated for evolution for a period of 20 years, till year 2000. The similarity of results of the simulation with those of EVS-2000, as shown in Table 1 and Figure 2, allow to validate MAS model for this case, on the attributes under consideration (more specifically, those defining religious values). Note that as the system is non-deterministic, the graphical results have some variations at each execution. Then, Figure 2 should not be taken as static output. Anyway, as it was previously commented, the trends are always very similar, even though the exact data have some small comparison error. Therefore, the system executions have structural similarity, as defined in [7].

By using this kind of simulations it is possible to experiment on assumptions about the influence of attributes, relationships and interactions of agents in the evolution of societal values, from a quantitative point of view.

Table 1. Evolution of the religious patterns of the Spanish society, according to the European Values Survey.

	1980	*1990*	*1999*
Ecclesiastical	33	25	22
Low-intensity	22	26	23
Alternatives	14	17	19
Non-religious	31	32	35

4 The “Micro” Approach: Tool for Qualitative Research

Given the results that can be obtained by analysis of the emergent behaviour of the MAS under simulation, some explanations and theories could be also envisaged at the “micro” level, on the individuals. If we make a U-turn of perspective, we will try to overcome the limitations of the pure statistical analysis of characteristics from individuals (variables). To achieve this aim we will simulate their biographic behaviour (that follows some rules), taking each individual as a whole, with the holistic perspective of qualitative researchers. From this point of view, the evolution of each individual has much more importance, and instead of considering the global emergent trend, we focus on people lives. Each person is unique and unrepeatable, and therefore extremely important. All this statements are consistent with qualitative methods.

We have pursued this discursive line till the point of giving name and surname for each agent: now each one represents a person instead of a number or an ID (it’s quite different “i214 died” than “Pablo Martínez died”). Also, we have given the agents the possibility of “living” events across their lives, events that could change their future decisions.

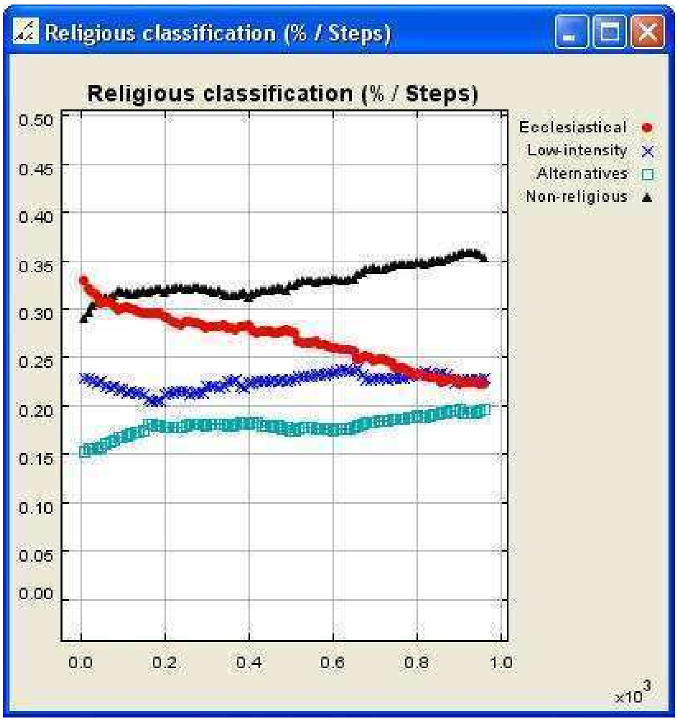


Fig. 2. Evolution of the religious patterns according to the social simulation system (1000 steps are equivalent to 20 years)

All these changes converge in the new output of the system: a narrated life-story in natural language. This is achieved by adding a new module to the system: a natural language processing (NLP) tool. This tool is a simple automatic narrator designed with most usual techniques on this area [15]. The main purpose of this system is to add more content to the analysis of the multi-agent simulation, and in this way completing the graphical output.

This Natural Language Generation System has been designed following simple and usual approaches based on rules. The system is heavily oriented on Content Determination (filtering the facts, telling only those facts that are considered to be really interesting for the reader), and Discourse Planning (ordering that filtered facts in such a way that the reader perceives a coherent story). Also very simple Surface Realization based on templates for creating the final text is addressed, to show a human readable form of the final content.

The most natural way of choosing one (or a few) individual stories to tell is to find the qualitative “ideal types”. We want the life-stories of the most representative individuals, the ones that show what has happened during the simulation. These individuals reflect the “macro” changes of the complex society. For example, a hyper-inflation process can be shown through the fall of living quality of a representative person. This person story would capture how he/she is forced to buy cheaper products and stop having some luxuries.

To achieve Content Determination, the NLP tool takes every individual as a whole being, and analyses the set of agents depending on its configured rules. Those rules, clearly context-dependant, have to be defined in order to choose correctly the representative individual(s). They should measure the interest (weight) of different life-events of the agents, and their relationships with others. In this way, only the most representative individual will be selected (the closest agent to the qualitative ideal type) from a set of many possible agents (500), the tool selects only one of them, that who is considered to be the most interesting one.

This filter is applied by computing a numerical value over every agent based on his life-events. This value, called *interest*, is based on a lookup table that assigns an heuristic value to each fact. Each of these values represent how much importance we give to each fact. Then the agents receives the *interest* value corresponding to the sum of the *interests* of their facts.

Then, with the *interest*, we apply some set of rules like filtering those agent whose interest falls below a given threshold, or removing redundant facts or relationships between agents.

Discourse Planning is performed by the use of templates that gives logical and coherent order to the text that has been filtered in the previous stage of Content Determination. The templates have simple sequences like time-ordering of the facts and describing the most important relationships.

To apply the rules of the original system to the social simulation domain it has been necessary to adapt the rules to this context. To perform this task we built a vocabulary with the aid of an expert sociologist in the domain. In this way we have built the vocabulary of the possible events and a logical path tree which defines the possible events that each agent can follow on each stage based on its own story (for example, a very religious woman will have a low possibility of deciding to abort). More details about how this tool works can be found on [11].

Biographies of agents that have been selected to build a story are not as expressive as real textual narrations, but can be directly compared with them (because of its natural language format). Due to the huge amount of textual and narrative material that qualitative research deals with, a narrative story is the best way to help them. This could help to understand the real individuals, compared with other possible life-stories, and it could be very useful due to the lack of individuals that usually qualitative researches have. If we use clustering or statistics for building the ideal types of the simulation, the real individuals of the qualitative research can be better understood compared with the simulated stories of these ideal types. Even we could use it to compare different qualitative researches over the same field.

Other applications of this new output could be to give some information very simple to understand for any human, not necessarily familiarized with the social simulation environment: for example an expertise on social behaviour. This expertise could read the text, that resumes with an example a life of some character, and extract information about the collective of people that have been simulated.

Moreover, this narration can be a good explanatory complement of quantitative researches, as other qualitative materials can be presented for supporting the final conclusions. This point will be deeply explained in next section.

For the example under study, as it is in a prototype phase, we have not followed a formal path for building the NLP rules that decided the most representative agent. We have decided the conditioned events and the implemented rules just using advises from an expert of the field. For computing this degree of representativeness, some simple formulas based on lookup tables has been applied, as can be seen in [11]. Even so, the generated text, in Figure 4, can give an idea.

The execution of the simulation of the MAS model, which resulted in quantitative macro results, generates also an XML file. This XML is the result of logging every event of every agent along their simulated lives. Agents that die and are born, matching or friendship relationships, life events: all is recorded. Next, the NLP module will process this XML, as explained in [11] and futher, to give the narrative output. We can compare Figure 3 and Figure 4 to check the text generation, although the big effort is to choose "Rosa Pérez" between the hundreds of individuals.


```xml
<Log Id="i49">
  <Description />
  <Attribute Id="name" Value="rosa" />
  <Attribute Id="last_name" Value="pérez" />
  <Attribute Id="sex" Value="female" />
  <Attribute Id="ideology" Value="left" />
  <Attribute Id="education" Value="high" />
        ...
  <Events>
  <Event Id="e1" Time="1955" Action="birth" Param="" />
  <Event Id="e2" Time="1960" Action="friend" Param="i344" />
  <Event Id="e3" Time="1960" Action="friend" Param="i439" />
  <Event Id="e4" Time="1961" Action="friend" Param="i151" />
  <Event Id="e5" Time="1962" Action="horrible" Param="childhood" />
  <Event Id="e6" Time="1963" Action="best friend" Param="i151" />
  <Event Id="e7" Time="1964" Action="believe" Param="god" />
  <Event Id="e8" Time="1964" Action="every week go" Param="church" />
        ...
  <Event Id="e16" Time="1968" Action="problems" Param="drugs" />
  <Event Id="e17" Time="1971" Action="grow" Param="adult" />
  <Event Id="e18" Time="1971" Action="friend" Param="i98" />
  <Event Id="e19" Time="1972" Action="involved" Param="labour union" />
  <Event Id="e20" Time="1972" Action="friend" Param="i156" />
  <Event Id="e21" Time="1973" Action="get" Param="arrested" />
  <Event Id="e22" Time="1973" Action="learn" Param="play guitar" />
  <Event Id="e23" Time="1975" Action="became" Param="hippy" />
  <Event Id="e24" Time="1976" Action="involved" Param="NGO" />
        ...
  <Event Id="e29" Time="1980" Action="impossible love" Param="i469" />
  <Event Id="e30" Time="1984" Action="couple" Param="i439" />
  <Event Id="e31" Time="1984" Action="live together" Param="i439" />
  <Event Id="e32" Time="1984" Action="child" Param="i679" />
  <Event Id="e33" Time="1984" Action="child" Param="i680" />
  <Event Id="e35" Time="1985" Action="friend" Param="i102" />
  <Event Id="e36" Time="1985" Action="divorce" Param="i439" />
  <Event Id="e37" Time="1987" Action="couple" Param="i102" />
  <Event Id="e38" Time="1987" Action="live together" Param="i102" />
  <Event Id="e39" Time="1987" Action="have" Param="abortion" />
        ...
  <Event Id="e48" Time="1995" Action="be unfaithful" Param="i102" />
  <Event Id="e49" Time="1995" Action="fired" Param="job" />
</Events>
</Log>
```


Fig. 3. XML example of life story

5 Conclusions and Further Work

We have seen the potential of the ABSS system for the quantitative analysis, reaching to the possibilities of social experimentation. We have experimented a way for expanding the possibilities of typical agent-based simulation, so they can build an output appropriate for the qualitative perspective.

```
Rosa Pérez was born in 1955, and she met Luis Martínez, and she met Miguel López, and she met María Valdés, and she suffered a horrible childhood, and she had a very good friend: María Valdés, and she believe in God, and she went to church every week, and she met David García, and she wanted to be a priest, and she suffered an incredible accident, and she met Marta Alonso.

When she was a teenager, she messed with a gang, and she met Claudia Sánchez, and she went to confession every week, and she had problems with drugs, and she became an adult, and she met Marci Boyle, and she was involved in a labour union, and she met Carla González, and she got arrested, and she learned how to play the guitar, and she became a hippy, and she was involved in a NGO.

She met Sara Hernández, and she stopped going to church, and she met Marcos Torres, and she fell in love, desperately, with Marcos Torres, but in the end she went out with Miguel López, and she lived together with no wedding with Miguel López, and she had a child: Melvin López.

She had a child: Andrea López, and she met Sergio Ruiz, and she separated from Miguel López, and she went out with Sergio Ruiz, and she lived together with no wedding with Sergio Ruiz.

She had a abortion, and she bought a house, and she had a depression, and she had a crisis of values, and she was involved in a NGO, and she had a huge debt, and she inherited a great fortune, and she met Daniel Lorenzo, and she bought a car, and she was unfaithful to Sergio Ruiz with another man, and she was fired from her job.

 Nowadays she is an atheist.
```

Fig. 4. The life story of a representative individual

With respect to the final appearance of the resulting text describing the life of one agent, it does not have yet good literary quality. It is necessary to generate more complex templates of text generation, as well as more sophisticated realization methods (those related with the final textual form of the narration). By defining better rules for choosing relevant agents, and for choosing most relevant facts of its life, the final text will be much clearer for a human reading it. To achieve this aim we must test the different ways of building those rules (as were defined previously) and compare their results. These rules could be:

- Rules built from the ideal types of a previous qualitative research, guiding the analysis
- Rules built from statistics or by an expert in the field
- Rules built from clustering of the agent lives.

Furthermore, it is possible to create narrations about more than one character, by inserting parts of other agents' life. There is, of course, much work in progress about this, and, although natural language generation has many problems, it is possible to create much better descriptions of social agents.

To evaluate the results of the NLP system there are many possible approaches. In [8] an evaluation of the texts is presented. It could be interesting to compare this system with other representation systems based on qualitative approaches.

It is possible to use a conceptual ontology for the communication between the expert sociologist for representing the domain semantics. Thus, it would be possible to formalize the hierarchy of the facts and their meanings.

Also, we are considering enriching the kind of information that is extracted from agents' behaviour, about their mental state. This is specially interesting with agents that follow BDI model [5]. In this way, the qualitative aspects of the agents will be empowered, as well as their stories, completed with mental states and motivations (instead of only describing facts). If agents' interactions base on this model, the quantitative macro results obtained in this way would be much more directly connected to the micro behaviour. With this aimed achieved, we could a) use qualitative researches as input of the system, in the form of BDI data (attitudes, perceptions, behavioural patterns), as is done in [16]; or b) contrast our simulated biographies with qualitative empirical results, and validate the system model: if there is convergence between data, the guarantees of accuracy of the system will be bigger [13].

The application of the obtained stories as a complement for quantitative results can be quite useful. It would explain the "why" of certain phenomena in the social system: why they evolve in a way and not another. These questions usually arise in the social sciences when only quantitative data is analyzed, and are referred to phenomena derived from the complexity of individuals' interaction and system dynamics.

However, due to is relatively usual to 'triangulate' the methodological perspectives (as it was explained previously), we have to analyze which are the benefits of simulated biographies above real ones. They are faster, easier and cheaper to obtain (if we already have a simulated model), but the important issue of trust should be tackled. We have to test the simulation with known empiric data (mainly quantitative, but if it is possible, qualitative too) before we can consider its biographies trustworthy. Of course, if there are discrepancies, we should trust real data. But when both converge in the main aspects and there are some holes in real biographies that are covered with simulated ones, we could trust simulated results. This limitation can be assumed for the practice of social research, because it's usual to use interpretations and explanatory models that are not verified.

Acknowledgments. This work has been developed with support of Dirección General de Universidades e Investigación de la Consejería de Educación de la Comunidad de Madrid (Spain) and Universidad Complutense de Madrid (Grupo de investigación consolidado 910494) and the project TIN2005-08501-C03-01, funded by the Spanish Council for Science and Technology.

References

1. Andrés Orizo, F., Elzo, J.: España entre el localismo y la globalidad. La Encuesta Europea de Valores en su tercera aplicación, 1981-1999. SM, Madrid (2000)
2. Arroyo Menéndez, M.: Cambio cultural y cambio religioso, tendencias y formas de religiosidad en la España de fin de siglo. Servicio de Publicaciones de la UCM. Madrid (2004)
3. Axelrod R.: Advancing the Art of Simulation in the Social Sciences. Conte, Rosario., Hegselman, Rainer., and Terna, Pietro. (eds.). Simulating Social Phenomena, Berlin: Springer. (1997) 21-40.
4. Babbie, E.: The Practice of Social Research, 11th edition, Wadsworth, Thomson Learning Inc. (2007)
5. Bratman, M.E.: Intentions, Plans and Practical Reason. Harvard University Press. (1987)
6. Corbetta, P.: Social research: theory, methods and techniques. Sage, London, (2003)
7. Gilbert, N., K. G. Troitzsch: Simulation for the Social Scientist. Open University Press, Buckingham, U.K. (1999)
8. Hassan, S., Gervás, P., León, C., Hervás, R.: “A Computer Model that Generates Biography-like Narratives”, 4th International Joint Workshop on Computational Creativity. University of London (2007)
9. Ibáñez, J.: “Del algoritmo al sujeto. Perspectivas de la investigación social”. Madrid. Siglo XXI. (1985)
10. Jick, T.D.: Mixing qualitative and quantitative methods: triangulation in action. Administrative Science Quarterly, 24 (1979) 602-611
11. León, C., Hassan, S., Gervas, P.: From the Event Log of a Social Simulation to Narrative Discourse: Content Planning in Story Generation. AISB'07, Artificial and Ambient Intelligence, Proceedings of the AISB Annual Convention, Patrick Olivier and Christian Kray, Newcastle, UK (2007) 402--409
12. Macy, M.W., Willer, R.: From Factors to Actors: Computational Sociology and Agent-Based Modeling, Annual Review of Sociology, Vol. 28 (2002) 143-166
13. Moss, S., Edmonds, B.: Sociology and Simulation: Statistical and Qualitative Cross-Validation. American Journal of Sociology, volume 110 (2005), 1095–1131
14. Pavón, J., Arroyo, M., Hassan, S., Sansores, C.: Simulación de sistemas sociales con agentes software. Actas del Campus Multidisciplinar en Percepción e Inteligencia, Vol. I (2006) 389-400
15. Reiter, E., Dale, R.: Building Natural Language Generation Systems, Cambridge University Press, (2000)
16. Taylor, R.I.: Agent-Based Modelling Incorporating Qualitative and Quantitative Methods: A Case Study Investigating the Impact of E-commerce upon the Value Chain. Doctoral Thesis, Manchester Metropolitan University, Manchester, UK. (2003)

Session on Multi-modeling and Ontologies

Evaluating a prototype self-description feature in an agent-based model of land use change

J. Gary Polhill and Nicholas M. Gotts

Macaulay Institute, Craigiebuckler, Aberdeen. AB15 8QH. United Kingdom
{g.polhill, n.gotts}@macaulay.ac.uk

Abstract. We describe and demonstrate a prototype approach to automatically building an OWL ontology from an Obj-C Swarm implementation of an agent-based model of land use change. The lessons learnt from this exercise highlight the potential inadequacy of relying on object-oriented formalisms for describing models, such as UML, but also show that there are a number of challenges to be overcome in automatic ontology creation from agent-based models.

1 Introduction

The problem of describing complex agent-based models is well documented (Grimm et al., 2006). Ontologies (Gruber 1993) offer a formal description of a model that may offer a convenient half-way point between model code and text (Gotts and Polhill, 2006). Techniques are already in development for automatically developing ontologies from text (Gómez-Pérez and Manzano-Macho 2005). Here we consider an approach to developing OWL (Web Ontology Language) ontologies (Antoniou and van Harmelen 2004) from the source code of a model, specifically, the FEARLUS (Polhill, Gotts and Law, 2001) modelling system.

FEARLUS is coded in Objective-C, an object-oriented flavour of C with similarities to Smalltalk. Objective-C has rudimentary features for what is, in Java, termed reflection, including the ability to inspect the class hierarchy, instance variables of classes, and instance variable values in objects. Thus the object-oriented structure of the model, and the state of the system at any one time, can be recorded, but without requiring access to the source code at run time. The approach used was to loosely couple the ontology creation code with FEARLUS so that it could reasonably easily be used with other Obj-C Swarm models.

Clearly, one can expect object-oriented source code in a land use change model to describe ontologically significant entities in land use change: classes representing land use decision-makers, parcels of land, and the land uses themselves for example. Parker et al. (in press) have created a 'conceptual design pattern' of all the classes one might expect to see in an agent-based model of land use change. However, there are key differences between the semantics of object-oriented programs and the description logics on which OWL ontologies are based that mean the translation process is not trivial, as the following will illustrate.

2 A trivial approach to translation from source code to ontology

The prototype ontology creation facility added to FEARLUS (version 1-0) uses a trivial mapping process from object-oriented program to ontology. Classes in the program become concepts in the ontology, and the inheritance hierarchy is reflected in the ontology using OWL's `subClassOf` relation. Instance variables of classes in the program become OWL properties. However, since OWL properties enjoy a rather more independent existence from concepts than instance variables do from classes, the name of the class is appended to the name of the instance variable to create the property in the ontology. Objects in the program become individuals in the ontology. Methods of classes are ignored, as are protocols (interfaces).

For example, land manager agents in FEARLUS may approve or disapprove of each other, and they store this information in their instance variables. Using ontology diagramming tools, it is possible to use this information in the generated ontology to create diagrams such as that in Figure 1. Ontology diagramming tools can also be used to depict the structure of the model itself. The hierarchy of some of the classes appearing in the model is shown in Figure 2 using the OWLViz tool, and relationships between key classes in the model are shown using OntoViz in Figure 3. Protégé itself (as opposed to plug-in tools) can be used to explore the model. Figure 4 shows the user interface during exploration of instances in the model, allowing the settings of instance variables, and pointers to other objects visible.

3. Evaluation

Useful though the ontology is to allow exploitation of visualisation tools, there are issues with the trivial translation approach described above. Here we focus on just one, the mapping from the inheritance hierarchy to the `owl:subClassOf` relation. The key issue is the difference in semantics of the two relations. The inheritance of one class by another in object-oriented programming is not necessarily done for ontologically significant reasons, but to implement differential or incremental programming (Cook and Palsberg, 1994), whereby subclasses either change the functionality of superclasses, or add new functionality. For example, one class A may have a subclass B, which adds some extra functionality and instance variables to A. At the same time, it may be that there is no instance of A that cannot be replicated by an instance of B, by choosing the right settings for these variables. Ontologically speaking, therefore, instances of A are a *subtype* of instances of B: an A is simply a B with certain restrictions on its instance variable values, even though B is a subclass of A in the object-oriented inheritance hierarchy.

The `owl:subClassOf` relation thus captures some interesting information about the classes that is not captured by the object-oriented inheritance hierarchy. Since the former is founded on description logics, whilst the latter depends more on issues such as reuse of code, this is hardly surprising. However, UML (Rumbaugh et al., 1998), often argued as being a suitable formalism with which to represent model descriptions (e.g. Grimm et al. (2006) recommend UML class diagrams for part of their protocol), is closely tied to object-oriented representations, and not to ontological descriptions.

These issues are reflected in debate about distinctions between inheritance and subtyping in the object-oriented programming literature of the late 80s and early 90s. (See for example, Cook, Hill and Canning, 1990; LaLonde and Pugh, 1991.)

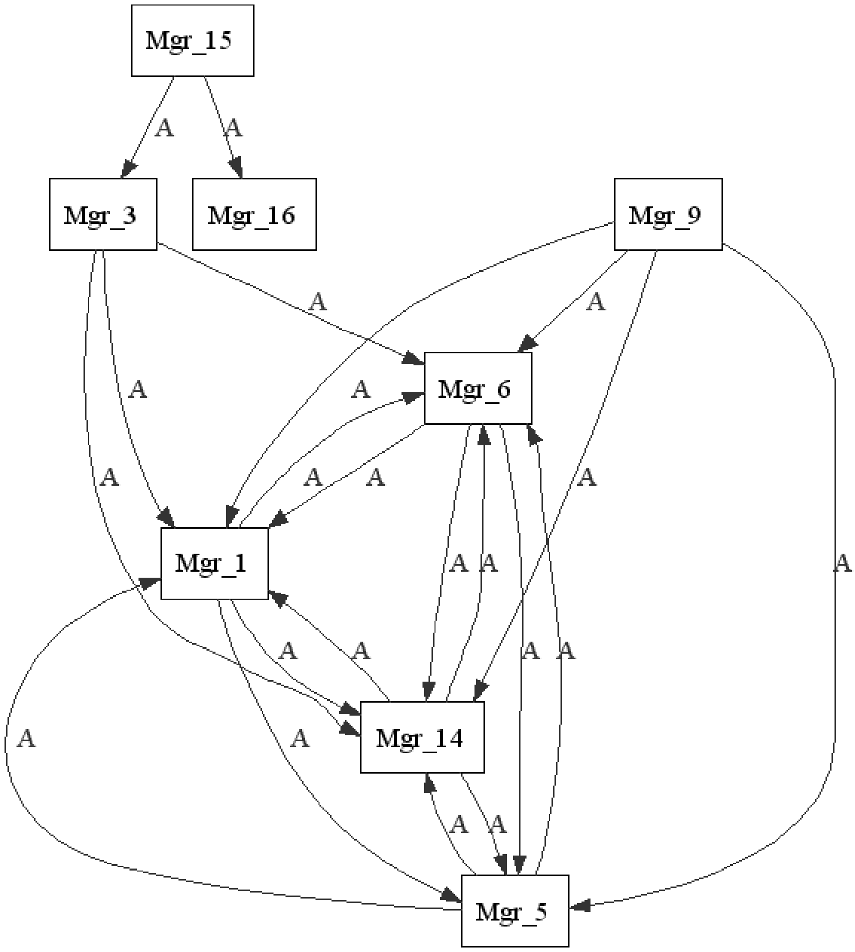


Fig. 1. Network of approval (A) among land manager agents (Mgr_*x*) created automatically from an edited ontology of FEARLUS using the OntoViz plug-in to Protégé. (The edits were done using sed to replace longer names for the approval relation and land manager identifiers with the shorter ones seen here.)

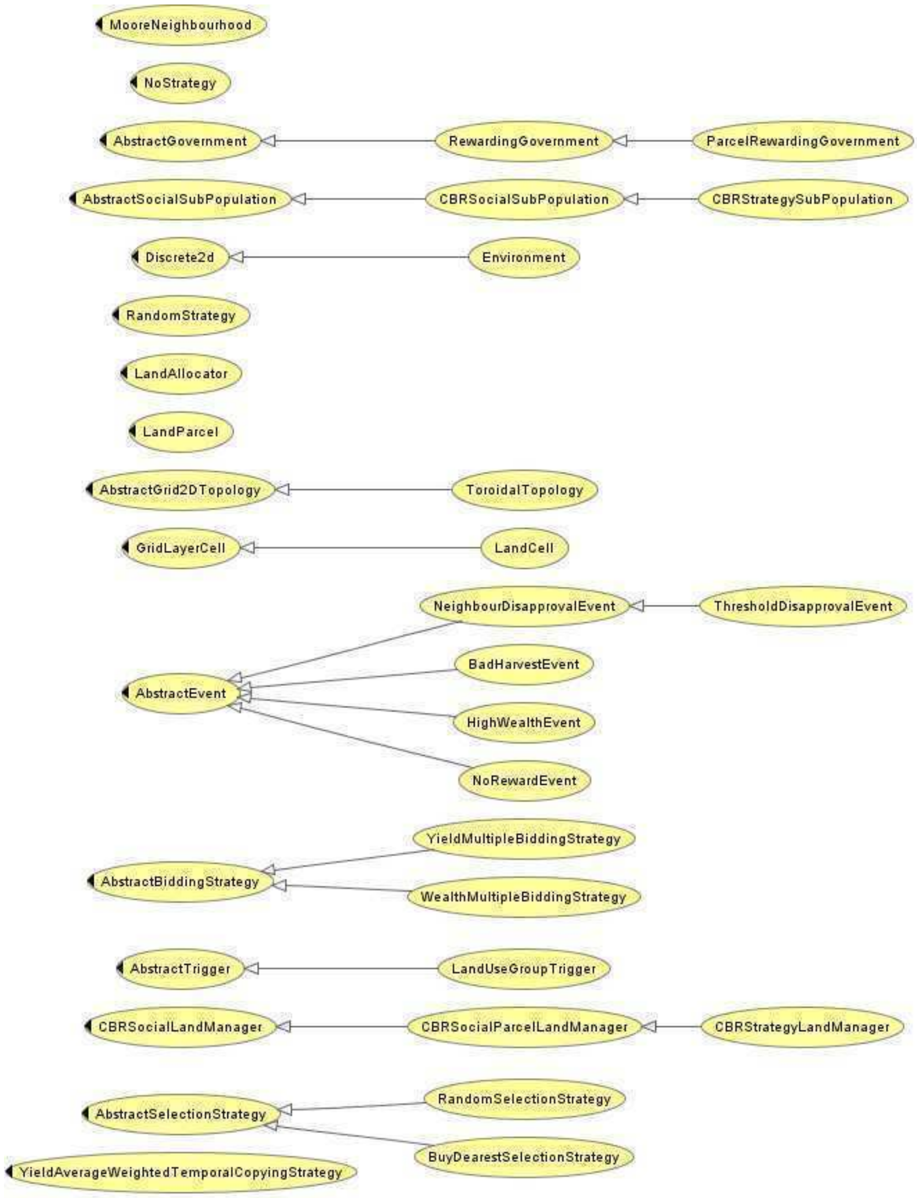


Fig. 2. Exploring the hierarchy of (part of) the model using the OWLViz tool.

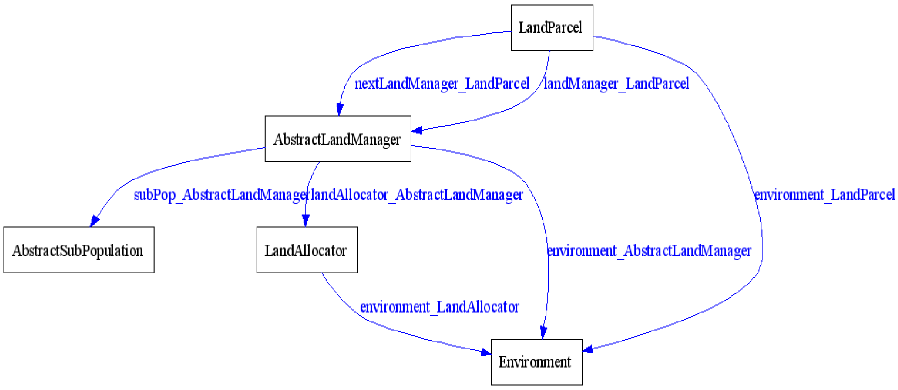


Fig. 3. Exploring the structure of the model with OntoViz. Links between classes, derived from instance variables, are shown as arrows.

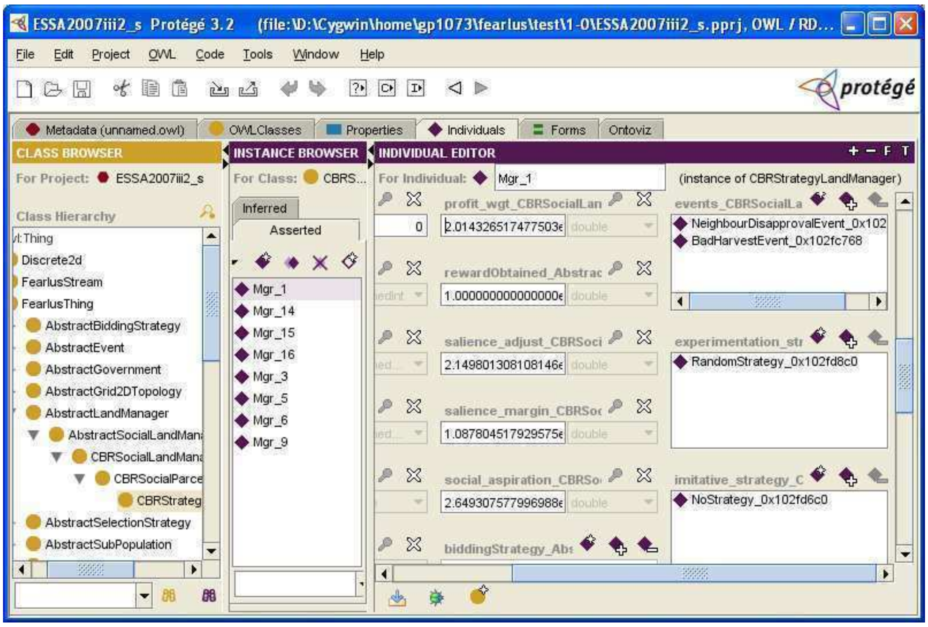


Fig. 4. Exploring instances in the Protégé user interface. The values of instance variables, translated to properties in the ontology, are shown according to their type, including the ability to cross-reference to other objects through representing the corresponding instance variables as object rather than datatype properties in the ontology.

For automated ontology generation, this poses a significant challenge, as it is equally possible for a particular subclass in the inheritance hierarchy to coincidentally have the same meaning as the `owl:subClassOf` relation. There may be heuristics that allow the various cases to be distinguished, or coding conventions could be used to declare the relationship explicitly. This example serves to show that differences in semantics between programming languages and ontologies mean that automatic ontology generation from modelling software is clearly a non-trivial task. If provision of such ontologies is deemed a worthwhile activity (e.g. in communicating ABMs and their structure), their automatic generation from source-code would facilitate their wider adoption. In doing so, we may have to change the way we program, adopting conventions that allow our modelling intentions in the code to be more easily derived ontologically. In the meantime, however, useful information can still be automatically recorded in an ontology, as the examples of figures 1-4 show.

4 Acknowledgements

This work was supported by funding from the Scottish Executive Environment and Rural Affairs Department, the Economic and Social Research Council, and the EU FP6 NEST Pathfinder Initiative on Tackling Complexity.

5 References

Antoniou, G. and van Harmelen, F. (2004) Web Ontology Language: OWL Ch.4 in Staab, S. and Studer, R. (eds) *Handbook on Ontologies* Springer

Cook, W. R., Hill, W. L. and Canning, P. S. (1990) Inheritance is not subtyping. *Proceedings of the 17th ACM Symposium on Principles of Programming Languages*. pp. 125-135.

Cook, W. and Palsberg, J. (1994) A denotational semantics of inheritance and its correctness. *Information and Computation* **114** (2), 329-350.

Grimm, V., Berger, U. Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S. K., Huse, G, Huth, A., Jepsen, J. U., Jørgensen, C., Mooij, W. M., Müller, B., Pe'er, G., Piou, C., Railsback, S. F., Robbins, A. M., Robbins, M. M., Rossmanith, E., Rüger, N. Strand, E., Souissi, S., Stillman, R. A., Vabø, R., Visser, U. and de Angelis, D. L. (2006) A standard protocol for describing individual-based and agent-based models. *Ecological Modelling* **198** (1-2), 115-126.

Gómez-Pérez, A. and Manzano-Macho, D. (2005) An overview of methods and tools for ontology learning from texts *The Knowledge Engineering Review* **19** (3), 187-212

Gotts, N. M. and Polhill, J. G. (2006) Narrative scenarios, mediating formalisms, and the agent-based simulation of land use change. *EPOS 2006 Epistemological Perspectives on Simulation, University of Brescia, Italy. October 5-6, 2006*.

Gruber, T. R. (1993) A translation approach to portable ontology specifications. *Knowledge Acquisition* **5**, 199-220

LaLonde, W. and Pugh, J. (1991) Subclassing ≠ subtyping ≠ is-a. *Journal of Object-Oriented Programming* January 1991, pp. 57-62.

Parker, D. C., Brown, D. G., Polhill, J. G., Deadman, P. J. and Manson, S. M, (in press). Illustrating a new 'conceptual design pattern' for agent-based models of land use via five case studies—the MR POTATOHEAD framework. In Paredes, L. and Iglesias C. H. (eds.) *Agent-Based Modelling in Natural Resource Management*.

Polhill, J. G., Gotts, N. M. and Law, A. N. R. (2001) Imitative versus nonimitative strategies in a land use simulation. *Cybernetics and Systems* **32** (1-2), 285-307.
Rumbaugh, J., Booch, G. and Jacobson, I. (1998) *The Unified Modeling Language Reference Manual*. Addison-Wesley.

Benefiting from the Other: Proposals on Incorporating Agent Based and System Dynamics Approaches

Gönenç Yücel, Catherine Chiong Meza

Faculty of Technology Policy and Management, Technology University of Delft,
Jaffalaan 5, 2628 BX Delft, The Netherlands
g.yucel@tudelft.nl, c.m.chiongmeza@tudelft.nl

Abstract. Two of the simulation-base approached broadly used in studying socio-economic systems are system dynamics (SD) approach and agent based (AB) approach. Despite their common point of using simulation in studying complex systems, their conceptualization of the system and the aggregation level used differ significantly. These two approaches are compared mainly based on these differences. Conversely, discussions on how one of them can benefit from the differing aspects and tools of the other approach are very limited. In this paper, we highlight two proposals about how these two approaches can be incorporated in order to strengthen them. These two proposals, which shall be considered as thought experiments, introduce feedback-loop-based conceptualization and behaviour analysis in AB approach and AB components in SD models in order to overcome some of the aggregation-related shortcomings of the latter.

Keywords. System Dynamics, Agent Base Modelling, Causal Loop Diagrams, Hybrid Modelling

1 Introduction

There is an evident increase in the popularity of the computational approaches, including computer-based simulations, when studying social phenomena. Simulation has been used for a long time as a very effective research approach in engineering and natural sciences, but its inclusion as a research tool in the portfolio of the social scientist is more recent. It is possible to identify various purposes in using simulation models, including prediction, training, experimentation and discovery [1]. Among those, it can be said that generating insights of the future and increasing the understanding about the temporal dynamics of a system either via testing a proposition or discovering important mechanisms or interactions in the system mainly constitute the relevance of these methods for the social sciences. In that respect, among various approaches involving simulation, like some econometric models, micro simulations, cellular automata, agent based models and system dynamics models [2, 3], the latter two seem to be the most relevant ones. The reasoning under this claim is the fact that both approaches can be seen as transparent-box models, where mechanisms yielding a phenomenon or dynamic behaviour are explicitly modelled. Hence, we will be focusing on these two approaches. In doing so, we aim to address the variety of modelling approaches from the perspective of how a system is visualized: from a holistic point of view (*top-down*) or addressing the heterogeneity (*bottom-up*) within the system. System Dynamics will represent top-down approaches and Agent-Based Modelling will do the same for bottom-up ones.

The paper starts with a brief discussion regarding the strong and weak points of both approaches with respect to our perspective. This discussion will highlight some potential points that may result in considerable added value when incorporating certain aspects of one approach to the other. In section four, we discuss two ways of incorporating these approaches. Being aware of the immature nature of these incorporation proposals, it is mainly aimed to initiate some discussion and possibly point out some direction for the research-to-follow.

2 System Dynamics (SD) Approach

The main characteristic of SD models is the fact that the dynamic behaviour of a system is defined in the form of a set of differential or difference equations that represent the dynamic relationships between state variables of the system. In this form of representation, the complexity of the dynamic system behaviour comes from the feedback relationships between state variables and the non-linear nature of these relationships. For a comprehensive review of the SD, the reader may refer to Forrester [4] and Sterman [5].

One of the important characteristics of SD models is the fact that the complex dynamic behaviour is caused by feedback relationships and delays in the system structure. Another important aspect is its aggregate-level focus, which is very significant with respect to the decision of using SD models in order to study a selected social system. This aspect of SD may best be clarified by a couple of simple examples. When modelling a national economy, the dynamic behaviour of the system will be defined in terms of relationships between aggregate level entities like gross domestic product (GDP), total industrial capital, total labour force, etc. Another example is an epidemic model. The dynamics of any epidemic will be represented by relationships between aggregate population subgroups of susceptible and infected populations. It is evident in both cases that individual behaviour of components (e.g. individual companies or consumers in the economy example, and infected and healthy individuals in the epidemic example), which is represented with the above mentioned aggregate level variables, is ignored. On the contrary, this can be justified under certain assumptions and in those cases an aggregate view of the system may be very fruitful in terms of analyzing the system behaviour.

The choice of the aggregation level is dependent on the degree of heterogeneity of individual system components being aggregated and the importance of this heterogeneity in terms of the overall dynamics of the system. If it can be assumed that the diffusion of the material and information is fast compared to the time frame of the process being studied, it can be argued that components can be treated as homogeneous and aggregated without any complications. In the epidemic case, every infected individual is the same with respect to the problem being studied, hence aggregation seems safe. Furthermore, even if heterogeneity prevails in the system it may not be crucial with respect to dynamics of the system and averaging out this heterogeneity does not yield any significant loss. Another example for this case can be found in the economy modelling example, in which different companies may own various amounts of industrial capitals that have different productivity values, but working with average capital productivity may be satisfactory to capture the dynamics of gross economic output.

On the contrary, consider again the case of epidemic example. This time individuals are distributed over a territory and contact is possible under certain physical proximity conditions. In this case, it is evident that spatial heterogeneity of individuals plays an important role in epidemic dynamics. Hence, sticking to a simple aggregation of susceptible and infected groups will not be sufficient to capture the actual behaviour of the system.

In some cases, an aggregate level phenomenon that emerges as a consequence of the interaction of heterogeneous individual components will be represented in the model as a deterministic relationship or a constant parameter. This issue stands as one of the points that attract numerous published or unpublished criticisms towards this approach. Even in the cases where aggregation can be justified, it may not be possible to construct these models due to lack of proper understanding of the relationships between aggregate entities. In such cases, the researcher may be forced to leave aside the aggregate approach and constrained to focus on individual components of the system, whose temporal behaviour could be formalized in a simple way.

Despite those aspects discussed above, SD models have important advantages that make them worth considering for studying socio-economic systems. The most crucial advantages are related to the model building and analysis stages. As an approach rooting back to the early 1960's, the know-how about model building process of SD type models is well established and homogeneous [6, 7]. Most crucial of all is the fact that formal validation procedures are available in SD type models, which cover both structural and behavioural validation of the models [8, 9]. Additionally, it is possible to trace back the causes of an observed complex dynamics in terms of model components and feedback relationships. Studying the cause of a behaviour is both easier in an informal sense and also achievable in a formal sense, using the recently developed analysis methods like eigenvalue elasticity analysis [10, 11], pathway participation method [12], and network theory based structure analysis [13], etc. In other words, in SD tradition linking and/or explaining the observed behaviour to the system structure is very important, and there are tools for establishing that link [7]. As it will be discussed under the corresponding section, this can be very complicated and problematic in the case of AB models.

Moreover, due to the focus on the aggregate system level, the model in general is composed of less elements and it is easy to grasp an idea about how a system works even just by observing the model structure. In that sense, building up some understanding about the dynamics of a system is evaluated to be much more efficient and easier in SD models.

3 Agent Based (AB) Modelling Approach

Compared to SD, the AB approach to modelling may be considered as a younger one that became especially popular in parallel to the popularity of the complexity theory. As it can be easily inferred from the name, the system of concern is represented in this type of models by a set of agents, whose behaviours and interactions constitute the system's dynamic behaviour. This perspective, focusing on individual components, also makes AB models an attractive method for methodological individualist social scientists. In those communities, the AB approach is characterized as the simulation approach that is most compatible with studying social phenomena, and referred to as "social simulation" [2, 14-16].

In contrast to SD models, heterogeneity at individual component level is not a problem for AB models. In the cases where individual level heterogeneity cannot be averaged out and plays an important role in the overall dynamics, an AB approach may be the logical option. This is one of the primarily appealing aspects of the AB models. Probably, the most important discussion point regarding the justification for using AB models is related to the nature and amount of knowledge that a researcher has about the system of concern. It may sometimes be a very challenging task to formalize a set of aggregate elements and relationships between those elements in order to explain the dynamic behaviour, especially in systems composed of interacting individual elements that are highly adaptive in their behaviour. This may be problematic in two ways. One the one hand, the researcher may have no knowledge about the aggregate description, but may be able to observe and imitate the behaviour of individual components. In this case, the only plausible option will be "to grow the phenomena being studied from bottom-up" [17] by simulating the individual components' interactions via AB models and allowing the dynamics of system to emerge. On the other hand, the researcher may have an initial aggregate representation, which represents the initial configuration of the system. However, due to the adaptive nature of the individual components and the co-evolution of their behaviours, initial aggregate configuration may lose its validity and needs to be updated. Unless the researcher is able to represent this co-evolution process in aggregate terms or foresees probable update requirements in the aggregate system representation and pre-builds them, an aggregate representation will fail to capture the dynamics of the system. These two cases (no aggregate knowledge and evolving aggregate configuration) seem to be the ones to which AB models fit best as models to study the dynamics of the system.

Despite the favourable aspects of AB models, this approach has also certain drawbacks and limitations. The first of these aspects is that AB models’ internal complexity may easily increase up to a level that makes it almost impossible for the researcher to deduct any conclusions from the simulations. The model itself becomes too complex to be understood. Although this risk exists for all modelling approaches, it is much higher in AB models from our perspective. This risk has already been pointed out by many other researchers indicating that in order to use AB models in an efficient way, the model should be very simple and should include only a minimal set of component that is capable of reproducing dynamics of the system under study [1, 17]. As long as it is applicable, the construction of modular models, i.e. detaching the complete model into manageable interactive sub-models, could offer a solution for such complications in extended AB models.

The ability to capture emergent system properties by focusing on individual components is discussed as the most important advantage of AB models. However, this aspect of the approach also causes serious and hard to avoid complications in terms of validating constructed models and gathering some insights based on the observed model behaviour. As mentioned before, the way these models are used aims to identify a minimal set of mechanisms that can reproduce the complex aggregate level phenomenon. The problem is that this minimal set is not unique. In other words, the mapping between mechanism sets and system behaviour is not one to one, but many to one [18]. Hence, it is a challenge for the researcher to validate his/her choice of individual components and mechanism when capturing the emergent phenomenon. In cases where the dynamics of the aggregate level phenomena (e.g. traffic, network density, segregation, etc.) are being studied, linking the changes in the temporal dynamics of the emergent phenomenon to micro-level agent behaviour constitutes another serious challenge.

4 Opportunities to Combine Strengths

The modelling approaches discussed above unfortunately are more commonly perceived as competitors. Meanwhile, discussions pointing out the opportunities of collaboration are relatively limited [7, 19, 20]. As discussed before, the SD approach has a reliable record on linking the system structure and the resultant dynamic behaviour while the fast diffusion of AB models in social sciences can be attributed to its orientation towards linking the behaviour of micro level components of the system to an aggregate level emergent phenomenon. In short, the emphasis of the SD approach is on linking the structure to behaviour, whereas the emphasis of the AB approach is on linking the micro level to the macro level [7]. This difference in the emphasis of the two presented approaches supports considering them as complementary rather than competing. In general, it is possible to identify similar attempts among modelling communities towards combining top-down and bottom-up modelling approaches. Researchers have started using hybrid models in order to compensate the limitations of top-down and bottom-up approaches: while top-down modellers are incorporating more explicitness in the dynamic behaviour of their models, bottom-up modellers are including macro-feedbacks at agents' level [21]. Similarly, we believe that modelling applications incorporating certain aspects of SD and AB in particular may allow us to study much more interesting systems and may provide richer insights than any of the approaches used alone.

The ways and degrees of incorporation of one approach into the other may vary. However, in this brief discussion paper we will mention two of them. The first one is using causal loop diagrams commonly used in SD as a conceptual analysis tool in AB modelling practices in both conceptualization and behaviour analysis stages. The second one aims to strengthen the aggregation related weak points of SD models via introducing AB components in those models.

4.1 Causal Loop Diagrams (CLD) in AB Modelling Practices

In almost all AB modelling practices the attention goes mainly on the totality of the discrete actions of agents in the system, which seems natural. However, what seems to be lacking appropriate attention is the fact that changes in the agents' behaviour are also influenced by feedback loops embedded in these systems.

CLD may be used to enhance the insights that can be gathered from an AB model and to communicate these insights. CLDs are a type of conceptualization tool for representing the feedback structure embedded in a system, and are widely used in the SD tradition [5, 6]. These diagrams include important variables about the system and arrows indicating "*causality*" between them. In order to avoid confusion about the concept of "*causality*", it can also be said that arrows – links between variables – indicate a sort of *influence* of one variable on another. Two types of links are identified. On the one hand, a positive link means that if the variable at the tail of a link (A) increases, the variable on the head of the link (B) increases above what it would otherwise have been, and if variable A decreases, variable B decreases below what it would otherwise have been. On the other hand, a negative link means that if the variable at the tail of a link (A) increases, the variable on the head of the link (B) decreases above what it would otherwise have been, and if A decreases, B increases below what it would otherwise have been [5].

Considering the conventional usage of CLDs, it could be argued that these types of diagrams are suitable for studying dynamics of continuous quantifiable variables, and do not work with systems including discrete events as in the case of AB models. We acknowledge that CLDs might not be appropriate and sufficient enough for conceptualizing the complete dynamic richness of an AB model. However, a similar conceptualization tool may be beneficial at some degree and it may support the modeller, in the course of a behaviour analysis or model conceptualization, in addition to other used tools or approaches.

In order to see what such a conceptualization may provide, we worked on a simple sketch of a feedback loop diagram related to individual agents' behaviours in a neighbourhood segregation model, which is mainly inspired by the famous work of Schelling [22]. In this model two types of inhabitants, yellow and blue, exist and they seek for neighbourhoods having their type as the dominant one. Once their neighbourhood is dominated by the inhabitants of the opposite colour, they move to a better spot, if possible. Figure 1 presents a CLD-like diagram that demonstrates how agents act and influence each other.

The variables represented by solid boxes are influenced by the actions of all agents; hence they belong to various loops like the ones given in this figure. Conversely, the variable represented by the circle is the basic action of the particular agent, whose behaviour is under focus in this diagram. The agent stays on a spot as long as certain fraction of her neighbours belongs to its type. In short, if that fraction is high, its likelihood of performing a *Move* action is low. Since agents move to spots that are better than the current one in terms of having more of the same coloured neighbours, after each *Move* action the percentage of the same colour neighbours will increase for the acting agent. These two relations constitute a loop (*Loop 1*) by which every agent tries to control the percentage of the same coloured neighbours.

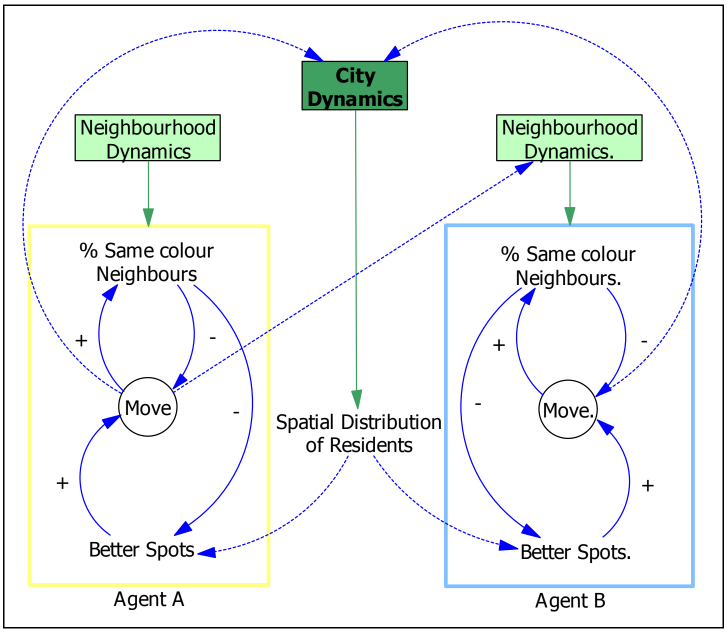


Figure 1: An experimental feedback loop diagram for a segregation model

Additionally, the *Move* action is also influenced by the availability of *Better Spots* in the city; the agent will not move if there is no better spot than the current one. Hence, the more *Better Spots* there are, the more likely the agent will move. Since being better, i.e. having a higher percentage of same colour neighbours, is a comparative evaluation, the number of *Better Spots* is dependent on the current spot, i.e. if the percentage of the same coloured neighbours is very high in the current spot, it will be less likely to identify better spots in the city. These three links constitute the second loop *(% Same Colour Neighbours – Better Spots – Move - % Same Colour Neighbours*). These two loops are related to how an individual agent acts towards a goal (i.e. live in a neighbourhood with more neighbours of its colour) and how an agent constrains its future moves by its own action. Moreover, actions of an agent also influence other agents. If we assume that agent A moves to the neighbourhood of agent B, agent A contributes to the Neighbourhood Dynamics of agent B, which in turn results in a change in percentage of same coloured neighbours for agent B. This change may activate the *Loop 1* for agent B (and probably for the other neighbours) and trigger further dynamics in the system.

This gives a rough picture of three loops via which an individual agent alters the general system and also its own behaviour. Looking at this experimental sketch, we can easily conclude that as long as Loop 1 is active for some agents, those agents will be dynamic and will be moving to neighbourhoods with more neighbours of their colour. This may have two different effects. Assume that a yellow agent moves to a new neighbourhood. Then this will make the neighbourhood more preferable for the existing yellow agents and they will be less likely to move, i.e. they will be locked-in to that neighbourhood in a sense (loop 2). For the blue agents opposite is true; neighbourhood will be less preferable and they will be more likely to move (loop 1). The whole system will reach to equilibrium only when all agents are satisfied with their current spot (loop 1 passive), or they cannot find a better spot even though they are not satisfied (loop 1 active, but suppressed by loop 2). The first occasion where all agents are satisfied is the situation where segregation is observed and this is a sort of stable dynamic equilibrium observed in almost all variants of this type of models. The case where agents want to move but cannot find a better spot, is a conceptually drawn conclusion from the diagram given above and we have not encountered any discussion about such a possibility. It is plausible that this option may be an unstable equilibrium that can be observed only when system is initiated at that state, so it is almost impossible to observe it with random model initiation. We are searching for a geometrical plausibility of an initial state on a two dimensional grid, which may support our conclusion from the presented diagram.

By itself, this diagram is not comprehensive enough to tell a complete story of the system behaviour as a whole. However, it provides some clues about why individual agents are almost never at rest and why they are inclined towards segregation. This type of diagram may help thinking on the agent's internal dynamics and communicating the conclusions. However, it will never be satisfactory in understanding why certain parameter values yield segregation and some do not. By introducing some new and AB-specific features in the representation conventions of CLDs, such diagrams may be used as promising conceptual tools in AB modelling studies.

4.2 Agent-Based Components in SD Models

As mentioned above, one of the aspects regarding the SD approach is its aggregate focus, or its more collectivist position in representing social systems and ignoring heterogeneity. Despite some credible discussion regarding the justification of this stance [23, 24], this point still constitutes a weak point of SD for many social scientists [25, 26].

In a typical SD model, an aggregate level phenomenon that emerges as a consequence of interacting heterogeneous individual components will be represented in the model as a deterministic relationship or a constant parameter. Combining a typical SD model with an AB component to represent that emergent phenomenon, instead of using a constant parameter is another possibility of combining the strengths of both approaches. This can be interpreted as a multi-resolution model, where a parameter or a relationship at the aggregate level is modelled in a finer grained way.

In order to demonstrate the proposed use, we reconciled the heterogeneity in AB approach and the aggregation level of SD by integrating them in a conceptual case. The model represents the dynamics of the welfare distribution in a hypothetical and very simple economy, which is regulated by an aggregate level regulator by imposing tax measures. The regulator is conceptualized as a ruler who reacts to the amount of population below a threshold welfare level. The economic interactions between individuals of the population are assumed to be random amounts exchanged between individuals in the system, i.e. an individual receives a random amount from the individual it encounters and gets it if the other individual has more than that amount, or gets all money the other individual has in the opposite case. If the amount of population under the threshold welfare is high, the regulator starts to alter the transactions by introducing an additional exchanged amount which is proportional to the individual wealth. More specifically, an agent pays a random amount plus a percentage (referred to as *tax percentage* in Figure 2) of his wealth to another randomly-selected agent.

However, considering the way this economic system is defined, it is not straightforward for clarifying the consequences of changes in the applied tax percentage. Hence, neither the polarity of the link between *Applied Tax Percentage* and *Percentage of Population under Threshold*, nor the delay between a change in *Applied Tax Percentage* and its consequences can be identified at this aggregate level system representation (See Figure 2).

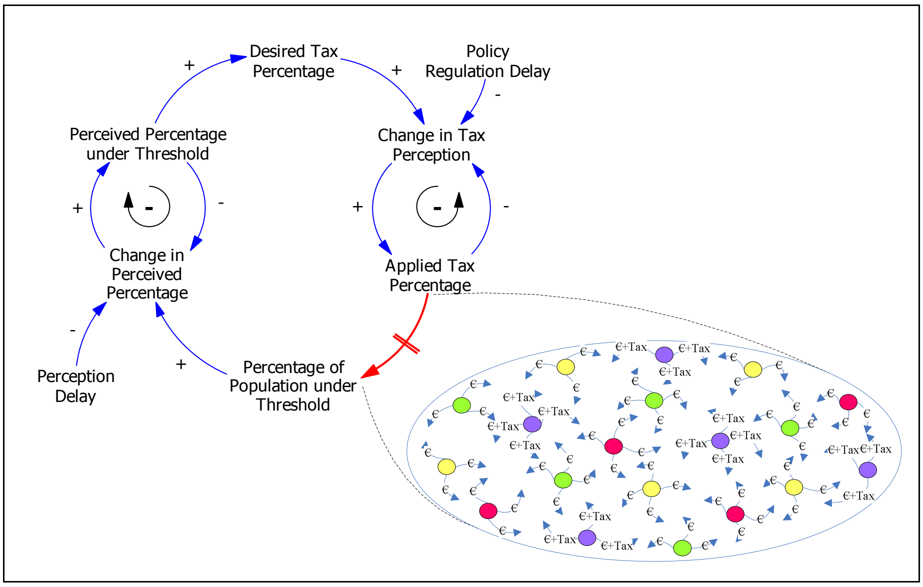


Figure 2: Welfare model as a combination of a SD model and an AB component

In order to overcome this complication, the linkage between these two aggregate variables is established via an AB component introduced into the model (See Figure 2). When the heterogeneity in the population is taken into account, the arrow between *Applied Tax Percentage* and *Percentage of Population under Threshold* is put under the microscope, leaving aside any theoretical conceptualization of the population as a whole. Instead, more attention is paid to the dynamics between individuals at this micro-level, where individuals hand in money to others from their own wealth, and receive money from other individuals, as long as they have some. In this specific case, the attention focuses on the agents' income and its variation due to the transactions – i.e. the handing-in and/or the reception of money – between agents. The agents' distribution under and above the threshold now varies depending on the income level of each agent. This income fluctuates depending on the transactions between agents which in turn are influenced by the introduction and the termination of the tax that the regulator would introduce in order to affect the amount of agents below the threshold welfare. In this way, *Percentage of Population under Threshold* becomes a statistic calculated based on the welfare level of the individuals in the population, and this is acknowledged by the regulator with a certain delay. The perception of the regulator alters its tax decision, and this tax will be later on introduced to the population's activities in order to transfer money from those having more income to those with less, i.e. interactions among individuals in the economy are altered with this dynamic decision.

Such applications with hybrid structure (hybrid in terms of having both SD and AB parts) are technically possible with the existing modelling tools. One available option regarding software is AnyLogic. This is a young simulation platform that appears with the promise of being capable of handling SD models incorporated to AB models [26]. Another option is integrating SD and AB models and software while preserving the respective components on their native platforms. For example Vensim, a widely used SD software, is capable of communicating with Java platform. This enables integration of models built in Vensim to an AB models built on RePast. Such sample implementations are tested before where some SD components are embedded into the agents in order to represent their perception and decision making processes [19]. The discussed example is another example of such an implementation from a technical point of view. However, conceptually they are different in the way how these two components (AB and SD) are used.

5 Conclusions

In studying complex socio-economical systems, simulation-based approaches are getting more popular. In this discussion paper, we focused on two of them; System Dynamics (SD) approach and Agent Based (AB) approach. Although both of them utilize simulation as the core of the approach, the way in which a system is conceptualized and modelled is significantly different. With an aggregate level focus, SD approach aims to identify feedback loops dominating the dynamics of the system. Conversely, AB approach focuses on individual level and works with agents that constitute the system. In this case, the aim is to identify agents' behaviour that yields system behaviour. This difference between the two approaches mainly induced to comparisons resulting in discussions about the superiority of each approach rather than proposals for collaboration. This paper attempted to highlight two of such proposals, which should be considered as thought experiments.

The first proposal is related to feedback-loop-focused conceptual analysis of agent behaviour and interactions. Although AB models have embedded feedback loops driving the agent actions and system behaviour, they seem to be lacking appropriate attention during the behaviour analysis and model conceptualization. In that sense, CLD tools customized for AB approach may make some contribution to the insights gathered about the system being studied and the behaviour being observed. Such tools may strengthen the toolset of AB modellers and help them to move more from a status of passive observer of the model output to a status of behaviour analyst. Future work aiming to come up with a proposal for a representation scheme that can be used in AB approach for this purpose will follow.

The second proposal is aiming to strengthen SD models regarding the aggregation-related shortcomings. The proposed approach aims to incorporate some AB modules in a SD model in order to replace deterministic relations or variable values, where appropriate. In this way it may be possible to incorporate certain aspects that are influenced by agent-level heterogeneity and interactions in a large SD model. We believe that such incorporation may make a SD model much stronger and a better representation of any socio-economic systems being studied. Differing from the former proposal, this proposal has some technical demands and, at the current state of the tools being used, this option is technically viable. Currently, we are working on such an implementation in order to see the added value as well as technical and conceptual difficulties.

References

1. Axelrod, R., *Advancing the art of simulation in the social sciences.* Complexity, 1997. **3**(2): p. 16-22.
2. Gilbert, N. and K. Troitzsch, *Simulation for the social scientist*. 2005, Berkshire: Open University Press.
3. Meadows, D.H. and J.M. Robinson, *The electronic oracle : computer models and social decisions*. 1985, Chichester ; New York: Wiley. xv, 445 p.
4. Forrester, J.W., *Industrial dynamics*. 1961, [Cambridge, Mass.]: M.I.T. Press. 464 p.
5. Sterman, J., *Business dynamics : systems thinking and modeling for a complex world*. 2000, Boston: Irwin/McGraw-Hill. xxvi, 982 p.
6. Richardson, G.P., *Feedback thought in social science and systems theory*. 1991, Philadelphia: University of Pennsylvania Press. x, 374 p.
7. Schieritz, N. and P.M. Milling. *Modeling the forest or modeling the trees: a comparison of system dynamics and agent-based simulation*. in *21st International Conference of the System Dynamics Society*. 2003. New York, USA.
8. Barlas, Y., *Formal aspects of model validity and validation in system dynamics.* System Dynamics Review, 1996. **12**(3): p. 183-210.
9. Forrester, J.W. and P.M. Senge, *Tests for building confidence in system dynamics models*, in *System Dynamics*, A.A. Legasto, J.W. Forrester, and J.M. Lyneis, Editors. 1980, North-Holland: Amsterdam.
10. Guneralp, B., *Towards coherent loop dominance analysis: progress in eigenvalue elasticity analysis.* System Dynamics Review, 2006. **22**(3): p. 263-289.
11. Kampmann, C.E. and R. Oliva, *Loop eigenvalue elasticity analysis: three case studies.* System Dynamics Review, 2006. **22**(2): p. 141-162.
12. Mojtahedzadeh, M., D. Andersen, and G.P. Richardson, *Using Digest to implement the pathway participation method for detecting influential system structure.* System Dynamics Review, 2004. **20**(1): p. 1-20.
13. Oliva, R., *Model structure analysis through graph theory: partition heuristics and feedback structure decomposition.* System Dynamics Review, 2004. **20**(4): p. 313-336.
14. Gilbert, N. and P. Terna, *How to build and use agent-based models in social science?* Mind and Society, 2000. **1**(1): p. 57-72.
15. Gilbert, N., *Holism, individualism and emergent properties*, in *Modelling and simulation in the social sciences from a philosophy of science point of view*, R.

Hegselmann, U. Mueller, and K. Troitzsch, Editors. 1996, Kluwer Academic Publishers: Dordrecht. p. 1-12.
16. Conte, R. and N. Gilbert, *Computer simulation for social theory*, in *Artificial societies: The computer simulation of social life* R. Conte and N. Gilbert, Editors. 1995, UCL Press: London. p. 1-18.
17. Epstein, J.M. and R. Axtell, *Growing artificial societies : social science from the bottom up*. Complex adaptive systems. 1996, Washington, D.C.: Brookings Institution Press. xv, 208 p.
18. Richardson, K.A. *On the limits of bottom-up computer simulation: towards a nonlinear modeling culture*. in *36th Hawaii International Conference on System Sciences*. 2003. Hawaii: IEEE.
19. Schieritz, N. and A. Grosler. *Emergent structures in supply chains - a study integrating agent-based and system dynamics modeling*. in *36th Hawaii International Conference on System Sciences*. 2003. Hawaii: IEEE.
20. Scholl, H.J. *Agent-based and system dynamics modeling: a call for cross study and joint research*. in *34th International Conference on System Sciences*. 2001. Hawaii.
21. Hourcade, J.-C., et al., *Hybrid Modeling: New Answers to Old Challenges. Introduction to the Special Issue of The ENergy journal.* The Energy Journal, 2006(Special Issue): p. 1 - 11.
22. Schelling, T.C., *Dynamic models of segregation.* Journal of Mathematical Sociology, 1971. **1**: p. 143-186.
23. Lane, D.C., *Rerum cognoscere causas: Part I - How do the ideas of system dynamics relate to traditional social theories and the voluntarism/determinism debate?* System Dynamics Review, 2001. **17**(2): p. 97-118.
24. Lane, D.C., *Rerum cognoscere causas: Part II - Oppourtunities generated by the agency/structure debate and suggestions for clarifying the social theoretic position of system dynamics.* System Dynamics Review, 2001. **17**(4): p. 293-309.
25. Parunak, H.V.D., R. Savit, and R.L. Riolo, *Agent-based modeling vs. equation-based modeling: a case study and users' manual*, in *Multi-agent systems and agent-based simulation*, J.S. Sichman, R. Conte, and N. Gilbert, Editors. 1998, Springer: Paris. p. 10-25.
26. Borshchev, A. and A. Filippov. *From system dynamics and discrete event to practical agent based modeling: reasons, techniques, tools*. in *22nd International Conference of the System Dynamics Society*. 2004. Oxford.

Linking CGE and Microsimulation Models: Different Approaches

Giulia Colombo[1]

[1]Research Fellow, Centre for European Economic Research (EW),
L7 1, 68161 Mannheim, Germany, colombo@zew.de
[1]Ph.D. in uantitative Models for Policy Analysis Catholic University of Piacenza, Italy

Abstract. The aim of this paper is to give an assessment of recent developments in the field that links Computable General Equilibrium (CGE) and microsimulation models, with a special concern for the different types of linking that are used in the literature on developing countries. In current literature we can observe two main trends: one that tries to fully integrate the two models (integrated model), while the other one develops separately the CGE and the microsimulation models and then links them through a set of equations (layered models). We build three models for the same economy (using fictitious data), according to the three main approaches that are used in literature to link CGE and Microsimulation models: the fully integrated approach, the Top-Down and the Top-Down Bottom-Up approaches. Implementingthe same simulation, we will obtain different results not only in terms of income distribution and poverty reduction, but also of changes in the main macroeconomic variables.

Keywords: CGE models, microsimulation, income distribution.

1 Introduction

Since the pioneering works by Adelman and Robinson [1] for South Korea and Lysy and Taylor [37] for Brazil, many Computable General Equilibrium (CGE) models for developing countries combine a highly disaggregated representation of the economy within a consistent macroeconomic framework and a description of the distribution of income through a small number of representative households (RHs). However, in order to account for heterogeneity among the main sources of the changes in household income, several “representative households” are necessary. Despite this need for variety, the number of RHs is generally small in these models (usually less than 10).

Typically, the level of disaggregation depends critically on the questions that the model is expected to answer: the household account is broken down into a number of relatively homogeneous household groups to reflect the socioeconomic characteristics of the country or region under consideration. The degree of homogeneity is essential in the design of classifications, and especially in the classification of household groups, where one would like to identify groups that are relatively homogeneous in terms of income sources and levels, and expenditure patterns, and that may be able to reproduce the socioeconomic and structural stratification observed within the society and the economy under study. It is noteworthy anyway that a household classification based on income or expenditure brackets does not satisfy any of these requirements – except perhaps the last one. Indeed, consider for instance the poorest segment of the society (say the bottom decile of the income pyramid): it may include very different household heads, such as a landless agricultural worker and a urban informal sector worker, and policies aimed at improving conditions in the two cases are likely to be very different.

The CGE RH framework sometimes also explicitly considers that households within a RH group are heterogeneous in a "constant" way. That is, in order to capture within-group inequality, it is assumed that the distribution of relative income within each RH follows an exogenous statistical law[1]. But the assumption that relative incomes are constant within household groups is not reflected in reality. Indeed, empirical analyses conducted on household surveys show that the within-group component of observed changes in income distribution is generally at least as important as the between-group component of these changes[2]. Thus, the RH approach based on this assumption may be misleading in several circumstances, and this is especially true when studying poverty. This argument may be better understood by presenting an example: consider a shock which increases the world price of a specific commodity, say maize, and reduces the world demand for this good. Under the small country assumption (that is, the country is price-taking on the world market), and assuming a demand elasticity with respect to price that is less than one in modulus, a country exporting this good will see a decrease in its exports and a domestic contraction of this sector. After the simulation of the shock with a CGE RH model, suppose that we find a little change in the mean income of a RH group, say workers in the agricultural sector. In this case, poverty might be increasing by much more than suggested by this drop in income: indeed, in some households there may be individuals that lost their job after the shock, or that encounter more difficulties to diversify their activity or their consumption than others. For these individuals or families, the relative fall in income is necessarily larger than for the whole group, and this fall in their income is not represented by the slight fall in the mean income of the whole group. Suppose moreover that the initial income of these individual was low. Then poverty may be increasing by much more than what predicted by a simple RH model, which is based

[1] For early applications of this type of models, see Adelman and Robinson [1], and Dervis *et al.* [29], who specified lognormal within-group distributions with exogenous variances. More recent examples of this kind of models can be found in Decaluwé *et al.* [27], Colatei and Round [21] and Agénor *et al.* [2].

[2] After Mookherjee and Shorrocks' study of UK [40], there are now other examples of "within between" decomposition analysis of changes in inequality that indicate that changes in overall inequality are usually due at least as much to changes in within-group inequality as to changes in the between-group component. Among the applications to developing countries, see Ahuja *et al.* [3], who applied this decomposition analysis to the case of Thailand, and Ferreira and Litchfield [31] for Brazil.

on the assumption of distribution neutral shocks. So, the RH approach does not capture the effects that a shock or a policy change may have on single individuals or households.
As it is well emphasized in Savard [46], another significant drawback in linking the intra-group distribution change to a statistical law that is completely exogenous is that no economic behaviour is considered behind this change in within-group distribution[3].
In order to overcome these problems, the recent literature has tried to develop new modelling tools which should be able at the same time to account for heterogeneity and for the possible general equilibrium effects of the policy reform (or the exogenous shock) under study. In view of the fact that most of the available economic models have either a microeconomic or a macroeconomic focus, and they do not address the question adequately, the recent literature has focused on the possibility of combining two different types of models. Most of the economic policies (structural adjustment programs or trade liberalizations, for example) and of the exogenous shocks commonly analyzed for developing countries (such as fluctuations in the world price of raw materials and agricultural exports) are often macroeconomic phenomena (or may have, at least, some structural effects on the economy), while poverty and inequality are mainly microeconomic issues. Thus, an approach that takes into account these important micro-macro linkages, seems to be the right answer to the problem. In particular, some authors have tried to link microsimulation models to CGE models[4], in order to account simultaneously for structural changes, for general equilibrium effects of the economic policies, and for their impacts on households' welfare, income distribution and poverty. The literature that follows this approach is quite flourishing in recent years: there are, among others, the important contributions by Decaluwé *et al.* [27] and [28], Cogneau and Robilliard [19] and [20], Cockburn [16], Cogneau [17], Bourguignon, Robilliard and Robinson [11], Boccanfuso *et al.* [9] and Savard [46].

The aim of the paper is to give an assessment of recent developments in this field, with a special concern for the different types of linking that are currently used in the literature.
In particular, we will link the microdata from a survey to a CGE model in three different ways: through a full integration of the survey data into a CGE framework, as

---

[3] The intra-group distribution change is usually linked to a theoretical statistical relationship between average and variance of the lognormal distribution. Savard [46] also underlines the fact that the average behaviour of a specific group is biased towards the richest in the group. Standard CGE models, indeed, use household groupings that take into account the total income and expenditure of each group and the behavioural parameters which are generally calibrated at the base year. In most of the models these parameters reflect the aggregate and not necessarily the average behaviour. Thus, as the richest of a group are endowed with most of the factors, their behaviour will be dominant in the group. Moreover, keeping in mind that when doing poverty analysis is very important to consider the behaviour around the poverty line, nothing really demonstrates that the average of aggregated behaviour will be representative of the households around the poverty line.
[4] More generally, this current of the literature develops the use of micro-data drawn from household surveys in the context of a general equilibrium setting, which is usually but not necessarily a CGE model.

it is done for instance in Cockburn [16]; by linking a behavioural microsimulation model to a CGE through a set of specific equations, which is the so called Top-Down method, as it is developed in Bourguignon *et al.* [11], and finally through a method which was developed by Savard [46], also known as Top-Down Bottom-Up model.
We will build all the three types of models using the same data from a fictitious economy. After this, by running an identical policy reform in the three models, we will analyse the different outcomes deriving from different types of linking. The choice for the use of fictitious data describing a simple economy is made with the aim of being able to understand better the differences that are observed in the results of the models, and to try to "go behind" these differences and look for the causes that generate them. Of course, this is of more difficult realization when using true data of a real and thus more complex economy, which naturally shows more a complex structure in its economic relationships.

2 The Integrated Approach

The main intuition behind this approach is to simply substitute the Representative Household Groups inside a standard CGE model with the real households that are found in the survey[5]. This way, one passes from a model with, for instance, ten representative agents to a model with thousands of agents, thus increasing the computational effort, but leaving substantially unchanged the modelling hypothesis of a standard CGE model. Basically, this approach does not include a true microsimulation module in the modelling framework, but it tries to incorporate the data from the household survey into the CGE model.
The first step to build such a model is to pass from the representative households' data of the survey to population values; to do this, one should weight each variable at the household level with the weights usually given in the survey, thus obtaining population values for each variable.
After this, we need a procedure to reconcile these population data coming from the survey (incomes and expenditures) with the accounts contained in the social accounting matrix (SAM). The literature on data reconciliation offers different alternatives. One may choose to keep fixed the structure of the SAM and adjust the household survey, or otherwise to adjust the SAM in order to meet the totals of the household survey. Another alternative would be that of using an intermediate approach. Whatever the method used, however, one necessarily loses the structure of the original data, which is one of the main drawbacks of the integrated approach. Our choice was for the first alternative, and we kept the original composition of households' incomes and expenditures unchanged.

[5] The first attempt in this direction was made by Decaluwé *et al.* [28]. Among the models following this approach there are the works by Cockburn [16] for Nepal, by Boccanfuso *et al.* [9] for Senegal, and by Cororaton and Cockburn [22], who studied the case of Philippine economy.

After these changes in the SAM, one encounters the problem of re-balancing it (row totals must be equal to column totals). To do this, we used an appropriate program that minimizes least squares[6].

The CGE model is the one described in section 3.2, except for the fact that we have added an index which refers to households.

A thing should be noted at this point: certain types of equations that are commonly included in a behavioural model, such as occupational choice equations, are not easily modelled within standard CGE modelling softwares[7], so that CGE-MS that follow the fully integrated approach are not always able to capture the behavioural responses of the agents to the policy reforms that are implemented. Instead, micro-econometric behavioural modelling provides much more flexibility in terms of the modelling structure used, and is more suitable to describe the complexity of household and individual behaviour, and the way this may be affected by the changes in the macroeconomic framework that are subsequent to a policy reform or an external shock.

The main point here is that with a CGE model like the one used for the integrated approach we are not able to predict which particular individual will enjoy the reduction (or will suffer from the rise) in the employment level on the basis of some characteristics of the individual or of the household that can be observed; this instead can be done through a behavioural microsimulation model.

Indeed, the main feature that differentiates a microsimulation model from a standard CGE framework (not only one with representative agents, but even one with thousands of households from a survey, as we have seen) is that it works at the individual level, selecting those individuals that show the highest probability of changing their labour market status, on the basis of their personal or family characteristics. This fact could bring above significant differences in the results between the two types of models, even after the same policy simulation, as we will see below.

3 The Top-Down Approach

We apply now the sequential or Top-Down approach as described in Bourguignon *et al.* [11].

The basic idea is to develop separately a MS model and then to run the simulation on the basis of changes in consumer producer prices, wages, and sectoral employment levels as predicted by the CGE model. This approach thus uses the two frameworks in a sequential way: first, the policy reform is simulated with the CGE model, and the second step consists of passing the simulated changes in some variables such as

---

[6] There exist different principles on which SAM-balancing programs can be based, such as the "Row and Sum" or RAS method (see Bacharach [7]), least squares minimization principles, known also as Stone-Byron methods (see Stone [49] and Byron [14]), or the more recent cross-entropy approach proposed by Robinson *et al.* [45] and Robilliard and Robinson [44].

[7] To this regard, see Savard's discussion about the limits and advantages of the various approaches of linking [46].

prices, wage rates, and employment levels[8] down to the MS module, as illustrated in Fig. 1.

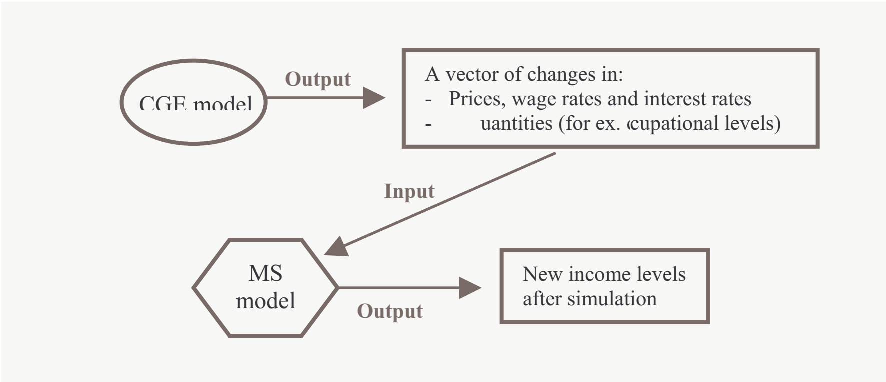


Fig. 1. The Top-Down Approach

3.1 The Microsimulation Module

The main role of the microsimulation module in the linked framework is to provide a detailed computation of net incomes at the household level, through a detailed description of the tax-benefit system of the economy, and to estimate individual behavioural responses to the policy change. For instance, through the use of microeconometric equations, we can model behaviours such as labour supply or consumption.

Behavioural Microsimulation (MS) models are developed to capture the possible reactions of the agents to the simulated policies, so that what happens after a reform can be very different from what is predicted by the simple arithmetical computations included in an accounting model.

In this section we will describe in detail a simple behavioural model, following quite closely the discrete labour supply choice model used in Bourguignon *et al.* [11]. Another description of a similar MS model for labour supply can be found in Bussolo and Lay [13] with their model for Colombia, and in Hérault [36], who built a model for the South African economy.

For the building of the model we will use fictitious data describing a very simple economy. In the household survey we have information about some individual

[8] When the assumption of imperfect labour market is adopted, or when the presence of a formal and an informal sector is predicted, the rationing in the labour market is usually carried out in the macro or CGE model, while the main use of the MS module is to select those households or individuals who will actually be barred out of, or let in, employment, or the formal sector. We will see this in more detail in the simulation section.

characteristics, such as age, sex, level of qualification, education, labour and capital income, the eventual receipt of public transfers, and the activity status. For the sake of simplicity, we have stated that each individual at working age (16-64) can choose between only two alternatives: being a full-time wage worker, or being unoccupied. There are other variables in the survey that are referred to households rather than to individuals, for example the area of residence, the number of household components, the number of adults (over 18 years old) and children (under 18), and so on.
All consumption goods of the economy are grouped in two main categories[9].
We derive income variables referring to households from initial individual data by summing up individual values for each household member; this way, we obtain households' labour and capital incomes, households' public transfers and households' total income:

Household *m*'s labour income: $YL_m = \sum_{i=1}^{NC_m} YL_{mi}$

Household *m*'s capital income: $YK_m = \sum_{i=1}^{NC_m} YK_{mi}$

Public transfers to household *m*: $TF_m = \sum_{i=1}^{NC_m} TF_{mi}$

Household *m*'s total income: $Y_m = YL_m + YK_m + TF_m$

where YL_{mi} is labour income of individual *i* member of household *m*, YK_{mi} his her capital income, and TF_{mi} are the public transfers he she receives from government. All these quantities are summed up for each family over all the individuals belonging to the family (NC_m is the number of components of household *m*); then, household *m*'s total income, Y_m, is the sum of all incomes received by the family: labour income, capital income, and public transfers.
For the benchmark situation, we assume all initial prices normalized at one.

The Model. The core of the behavioural model is represented by the following two equations:

Regression model for log-wage earnings: $$Log(YL_{mi}) = a + b \cdot x_{mi} + c \cdot \lambda_{mi} + v_{mi} \quad \textbf{(1)}$$

Choice of labour market status: $$W_{mi} = Ind[\alpha + \beta \cdot z_{mi} + \gamma \cdot rw_{mi} + \varepsilon_{mi} > 0] \quad \textbf{(2)}$$

The rest of the MS module is made up by simple arithmetical computations of price indices, incomes, savings and consumption levels. As the parameters entering the

[9] The focus of our distribution and poverty analysis will be on disposable income, even if an inequality and poverty analysis could also be conducted on expenditure rather than on income levels.

following equations (marginal propensity to save mps_m, income tax rates γ, and budget shares η_{mq}) are constant, this part of the model may be regarded as purely accounting, as it does not contain any possible behavioural response to policy simulations.

Household *m*'s income generation model:	$Y_m = \sum_{i=1}^{NC_m} YL_{mi} \cdot W_{mi} + YK_m + TF_m$	**(3)**
Household disposable (after tax) income:	$YD_m = (1-\gamma) \cdot Y_m$	**(4)**
Household specific consumer price index:	$CPI_m = \sum_{q=1}^{2} \eta_{mq} \cdot P_q$	**(5)**
Real disposable income:	$YDR_m = YD_m \quad CPI_m$	**(6)**
Savings:	$S_m = mps_m \cdot YD_m$	**(7)**
Household consumption budget:	$CEBUD_m = YD_m - S_m$	**(8)**
Consumption expenditure for commodity *q*:	$CE_{mq} = \eta_{mq} \cdot CEBUD_m$	**(9)**
Consumption level of commodity *q*:	$C_{mq} = \frac{CE_{mq}}{P_q}$	**(10)**
Household *m*'s capital income:	$YK_m = PK \cdot KS_m$	**(11)**

Description of the subscripts:

m	Households	$m = 1, 2, \ldots, 24$	
i	Individuals	$i = 1, \ldots, NC_m$	NC_m: number of components of household *m*
q	Goods	$q = 1,2$	

The *first equation* of the model, **(1)**, computes the logarithm of labour income (wage) of member *i* of household *m* as a linear function of his her personal characteristics (vector x_{mi} includes the logarithm of age, sex, skill level and educational attainment) and of λ_{mi}, which represents the inverse Mills ratio estimated for the selection model (for more details on the estimation process see below). The residual term v_{mi} describes the effects of unobserved components on wage earnings.

The *second equation* represents the choice of the labour status made by household members[10]. Each individual at working age has to choose between two alternatives:

[10] In the literature this kind of equation is known as occupational choice model, or selection model (and also discrete choice model of labour supply). However, it must be specified that this

being a wage worker, or being inactive. The variable W_{mi} is a dichotomic variable taking value one if individual *i* of household *m* decides to be a wage worker, and zero otherwise. The choice is made by each individual according to some criterion, the value of which is specific to the alternative, and the alternative with the highest criterion value is selected. A natural economic interpretation of this criterion value is utility: each individual chooses the alternative with the highest associated utility. Indeed, we will estimate the selection model using a binomial logit specification, which assigns each individual to the alternative with the highest associated probability. In our model we have arbitrarily set to zero the utility of being inactive. Function "*Ind*" is an indicator function taking value one if the condition is verified, and zero otherwise. Vector z_{mi} of explanatory variables includes some personal characteristics of individual *i* of household *m*, that is: age, sex, skill and educational level, the area of residence and the number of children under 6 living in the household. Variable rw_{mi} is the logarithm of real labour income. The equation is defined only for individuals at working age.

The *third equation* is an accounting identity that defines total household income, Y_m, as the sum of the wage income of its members YL_{mi}, of the exogenous household capital income YK_m, and of the total amount of public transfers received by household m, TF_m. In this equation, variable W_{mi} stands for a dummy variable that takes value one if member *i* is a wage worker and zero otherwise.

The *fourth equation* computes household disposable (after tax) income by applying income tax rates according to the rule reported in Table 1. In order to simplify computations, we have assumed that in this economy direct income taxes are imposed on households' total income Y_m, and not on individual incomes.

Table 1. Direct Income Tax Rates

Income brackets	Tax rate
Up to 10,000	0%
Up to 15,000	15%
Up to 26,000	24%
Up to 70,000	32%
Over 70,000	39%

equation is not really intended to explain the individual *choice* between being occupied or unemployed, but rather it tries to find out which characteristics strengthen the *probability* of being in one condition rather than in the other one for each individual, as it is described in more detail in the estimation section below.

Equation **(5)** computes an household specific consumer price index through the consumption shares η_{mq}. Real disposable income is then obtained by dividing households' disposable income by this index (equation **(6)**).
Then, to find out household *m*'s savings level, equation **(7)** multiplies the disposable income by the marginal propensity to save of each household, mps_m. The assumption underlying this equation is that household savings behaviour is unvarying, as the savings level is a fixed fraction of household disposable income. Then, subtracting savings from disposable income one obtains the budget that each household spends for consumption (equation **(8)**), which is spent on the two goods of the model according to the budget shares η_{mq} by equation **(9)**. Again, the assumption in this equation is that consumption behaviour is not flexible, that is, households spend a constant fraction of their consumption budget for each of the two goods.
To get the values of these exogenous parameters (marginal propensity to save mps_m and budget shares η_{mq}), we use the initial data from the survey in the following way:

Household *m*'s marginal propensity to save: $mps_m = \dfrac{S_m}{YD_m}$

Household *m*'s consumption budget shares: $\eta_{mq} = \dfrac{CE_{mq}}{CEBUD_m}$

Equation **(10)** derives then the consumption levels for each household by dividing the expenditure for each good by its price.
Finally, income from capital is obtained by multiplying capital endowment of each family, KS_m, by the return to capital, *PK* (equation **(11)**).
The initial values of the variables C_{mq} and KS_m (consumption levels and capital endowments, respectively) are derived from the initial data of the survey by making use of the assumption that in the benchmark situation all prices and returns are equal to one:

Household *m*'s consumption level of commodity *q*:	$C_{mq} = CE_{mq}$	**(12)**
Household *m*'s capital endowment:	$KS_m = YK_m$	**(13)**

Moreover, we assume that public transfers paid to households and household capital endowments are exogenously given. They are fixed at the level reported in the survey, for public transfers, and at the level as computed in equation **(13)**, for capital endowment, respectively.

Estimation of the Model. The only two equations in the MS module that need to be estimated are equations **(1)** and **(2)**.
The former, which expresses the logarithm of wage earnings as a linear function of some individual characteristics and of λ_{mi}, the inverse Mills ratio, was estimated using

an Heckman two-step model (see Heckman [34] and [35]). We follow this approach to correct for the selection bias which is implicit in a wage regression, that is, the fact that we observe a positive wage only for those individuals that are actually employed at the moment of the survey.
The results of the estimation are reported in Table 2 below. The estimation was conducted on the sub-sample of individuals at working age (16-64).
The interpretation of the coefficients for the wage equation thus follows that of a simple linear regression. As we can observe in Table 2, age, schooling and skill level have a positive effect on the wage, while being a woman shows a negative effect.
It is important to say that the aim of the wage equation within the model is that of obtaining an efficient estimate for an eventual wage income only for those individuals that are observed to be inactive in the survey, in the case that, after a policy reform, one or more of them will change their labour market status and become wage workers. In this case, through these estimates, we will be able to assign an estimated wage to the individual that has changed his her labour market status after the simulation run.
For all the other individuals that are observed to receive a wage in the survey, we use instead the observed wage level and not the estimated one.

Table 2. Heckman selection model, two-step estimates

Dependent variable: logarithm of Wage				
	Coefficient	Std. Error	z	P>|z|
constant	7.032117	0.3145104	22.36	0.000
ln(age)	0.697818	0.0833084	8.38	0.000
sex	-0.466210	0.1018222	-4.58	0.000
qualification	0.396613	0.0771516	5.14	0.000
education	0.525011	0.0871646	6.02	0.000
Mills ratio	0.216005	0.1473164	1.47	0.143
Selection				
ln(age)	0.338583	0.0807227	4.19	0.000
sex	-1.549158	0.2802896	-5.53	0.000
qualification	1.020388	0.2728658	3.74	0.000
children under 6	0.168214	0.2368365	0.71	0.478
region	-0.751549	0.2980307	-2.52	0.012
rho	0.762760			
sigma	0.283187			

Parameters of equation **(2)** were obtained through the estimation of a binomial logit model, assuming that the residual terms ε_i are distributed according to the Extreme

Value Distribution - Type I[11]. The estimation was conducted on the sub-sample of individuals at working age (16-64).
Our explanatory variables include individual characteristics such as the logarithm of predicted real wage, sex, skill and education level, the region of residence and a variable accounting for the presence or not of children under 6 years old in the household. The model is estimated by Maximum Likelihood. Results are presented in Table 3.
A binomial model states that the probability of observing the dependent variable assuming value one, given the explanatory variables ($OCS_{mi} = 1|Z_{mi}$), is equal to the cumulative distribution function of ε_i (the Extreme Value Type I distribution in our case), evaluated at $\beta \cdot Z_{mi}$, that is:

$$\Pr[OCS_{mi} = 1 \mid Z_{mi}] = F(\beta \cdot Z_{mi}) = \exp\left(-e^{-\beta \cdot Z_{mi}}\right) \quad \textbf{(14)}$$

The effects that the explanatory variables have on the dependent binomial variable are not linear, because they get channelled through a cumulative distribution function. Thus, by observing the values and signs of the estimated coefficients, we can say something about the effect that explanatory variables have on the probability that the dependent binomial variable takes value one (wage worker), relatively to the probability that it takes value zero, but not in a linear way.
For instance, expected real wage and qualification seem to influence in a positive way the probability that the dependent variable takes value one (the more qualified the individual is, the higher is the probability for him her to be employed), as well as the presence of children under 6 does, which is the opposite of what was expected, but anyway this result is not significant. Moreover, for men the probability of being employed is higher than for women, as the variable *SEX*, which takes value zero for men and one for women, shows a negative coefficient. The same can be said about the region of residence: people living in the first region have a higher probability of being employed than people living in the second one. The variable referring to education, instead, seems to have a negative influence on the probability of being employed, which is the opposite of what we expected, and anyway it is not highly significant.
However, with the estimated coefficients we cannot perfectly predict the true labour market statuses that are actually observed in the survey. Thus, following the procedure described in Duncan and Weeks [30], we drew a set of error terms ε_i for each individual from the extreme value distribution, in order to obtain an estimate that is consistent with the observed activity or inactivity choices. From these drawn values, we select 100 error terms for each individual, in such a way that, when adding

[11] The Extreme Value distribution (Type I) is also known as Gumbel (from the name of the statistician who first studied it) or double exponential distribution, and it is a special case of the Fisher-Tippett distribution. It can take two forms: one is based on the smallest extreme and the other on the largest. We will focus on the latter, which is the one of interest for us. The standard Gumbel distribution function (maximum) has the following probability and cumulative density functions, respectively:
pdf: $f(x) = \exp\left(-x - e^{-x}\right)$
CDF: $F(x) = \exp\left(-e^{-x}\right)$.

it to the deterministic part of the model, it perfectly predicts the activity status that is observed in the survey. In other words, the residual term for an individual that is observed to be a wage earner in the survey should be such that:

$$\hat{\alpha} + \hat{\beta}_1 \cdot Log(RW_{mi}) + \hat{\beta}_2 \cdot SEX_{mi} + \hat{\beta}_3 \cdot Q_{mi} + \hat{\beta}_4 \cdot AREA_m + \hat{\beta}_5 \cdot CH6_{mi} + \hat{\beta}_6 \cdot SCH_{mi} + \varepsilon_{mi} > 0,$$

while, for an individual that is observed to be inactive in the survey, the same inequality should be of opposite sign (≤).

After a policy change, only the deterministic part of the model is recomputed. Then, by adding the random error terms previously drawn to the recomputed deterministic component, a probability distribution over the two alternatives (being a wage worker or being inactive) is generated for each individual. This implies that the model does not assign every individual from the sample to one particular choice, but it gives the individual probabilities of being in one condition rather than in the other. This way, the model does not identify a particular choice for each individual after the policy change, but generates a probability distribution over the different alternatives[12].

Table 3. Binary logit model for labour status' choice

Dependent Variable: Activity Status				
	Coefficient	Std. Error	z-Statistic	Prob.
ln(real wage)	0.197215	0.046458	4.245037	0.0000
sex	-1.894812	0.407759	-4.646894	0.0000
qualification	1.440805	0.425709	3.384482	0.0007
region	-0.718504	0.329501	-2.180586	0.0292
children under 6	0.269124	0.297251	0.905378	0.3653
education	-0.763275	0.671696	-1.136341	0.2558
Mean dependent var	0.664706		S.D. dependent var	0.473488
S.E. of regression	0.376673		Akaike info criterion	0.901535
Sum squared resid	23.26880		Schwarz criterion	1.012210
Log likelihood	-70.63049		Hannan- uinn criter.	0.946446
Avg. log likelihood	-0.415473			

3.2 The CGE Model

The CGE model for the fictitious economy is characterized by a representative household who maximizes a Cobb-Douglas utility function with three arguments: leisure and two consumption goods. These commodities are also used as inputs, together with capital and labour, in the production process, which is operated by two

[12] This procedure is also described in Creedy and Kalb [23]. See also Creedy *et al.* [25].

firms following a Leontief technology in the aggregation of value added and the intermediate composite good, a Constant Elasticity of Substitution (CES) function for assembling capital and labour into value added, and a Leontief function in the aggregation of intermediate goods. Both factors of production, capital and labour, are mobile among sectors. The capital endowment is exogenously fixed, while labour supply is endogenously determined through household's utility maximization (subject to fixed time endowment). The wage elasticity of labour supply is estimated from the household survey, in order to have consistency in labour supply behaviour between the two models. Investments are savings-driven, while government maximizes a Cobb-Douglas utility function to buy consumption goods and uses labour and capital. The public sector also raises taxes on household's income and tariffs on imported goods, while it pays transfers to the representative household. For the foreign sector we have adopted the Armington assumption of constant elasticity of substitution for the formation of the composite good (domestic production delivered to domestic market plus imports) which is sold on the domestic market (see [4]). Domestic production is partially delivered to the domestic market and partially exported, according to a Constant Elasticity of Transformation (CET) function. The balance of payments is required to be in equilibrium, and the small country hypothesis is assumed (the economy is price taker in the world market).

Table 4. SAM of the Economy

	C_1	C_2	S_1	S_2	K	L	H	G	SI	RoW	Tot
C_1			58	15			95	61	30	24	283
C_2			17	24			313	48	14	77	492
S_1	283										283
S_2		492									492
K			72	23				13			108
L			83	354				116			553
H					108	553		40			701
G			12	18			249				279
SI							44				44
RoW			41	59							100
Tot	283	492	283	492	108	553	701	279	44	100	

C_q: consumption of good *q*; ***S_q***: sector *q*; ***K***: capital account; ***L***: labour account; ***H***: representative household; ***G***: public sector; ***SI***: savings-investments account, ***RoW***: Rest of the World account.

In the model there are in total 49 variables and 42 equations, which, with the 8 exogenous variables (capital endowment, *KS*, time endowment, *TS*, public transfers, *TF*, the four world prices PWE_q and PWM_q, and the numeraire, *ER*), fully determine the model and allows for satisfaction of Walras' law (we have a redundant equation).

The calibration of the parameters for the CGE model is done on the basis of a Social Accounting Matrix (SAM) for the economy, in such a way that the benchmark situation is consistent with that of the microsimulation module (for instance, in the benchmark of the two models we have the same average income tax rate, the same average marginal propensity to save, the same budget shares for consumption of the two goods, and so on).
The SAM for the economy under study is reported in Table 4. The data in the SAM are in millions of the monetary unit we have used for the survey.

3.3 Linking the Models

The basic difficulty of this approach is to ensure consistency between the micro and macro levels of analysis. For this reason, one may introduce a system of equations to ensure the achievement of consistency between the two models[13]. In practice, this consists in imposing the macro results obtained with the CGE model onto the microeconomic level of analysis. In particular:

1) changes in the commodity prices, P_q, must be equal to those resulting from the CGE model;
2) changes in average earnings with respect to the benchmark in the micro-simulation must be equal to changes in the wage rate obtained with the CGE model;
3) changes in the return to capital of the micro-simulation module must be equal to the same changes observed after the simulation run in the CGE model;
4) changes in the number of wage workers in the micro-simulation model must match those observed in the CGE model.

For our model, these consistency conditions translate into the following set of constraints, which could be called linking equations:

Consumption levels:
$$C_q = \frac{CE_q}{\left(1+\Delta P_q^{CGE}\right)} \quad \textbf{(15)}$$

Logarithm of wage earnings:
$$Log\left(YL_{mi}\right) = Log\left(Y\hat{L}_{mi}\right) \cdot Log\left(1+\Delta PL^{CGE}\right) \quad \textbf{(16)}$$

Capital income:
$$YK_m = KS_m \cdot \left(1+\Delta PK^{CGE}\right) \quad \textbf{(17)}$$

Employment level:
$$\frac{\sum_{m=1}^{24}\sum_{i=1}^{NC_m}\hat{W}_{mi}}{\sum_{m=1}^{24}\sum_{i=1}^{NC_m}WA_{mi}} \cdot 100 = \Delta EMP^{CGE} \quad \textbf{(18)}$$

[13] This way, what happens in the MS module can be made consistent with the CGE modelling by adjusting parameters in the MS model, but, from a theoretical point of view, it would be more satisfying to obtain consistency by modelling behaviour identically in the two models.

The variables with no superscripts are those coming from the microsimulation module; those with the ^ notation correspond to the ones that have been estimated: in particular, $Log(Y\hat{L}_{mi})$ is the wage level resulting from the regression model for individual *i*, member of household *m*, while $\hat{W}_{mi}$ is the labour market status of individual *i* of household *m* deriving from the estimation of the binomial choice model.

ΔP_q^{CGE}, ΔPL^{CGE} and ΔPK^{CGE} indicate, respectively, the change in the prices of goods, the change in the wage rate and in the return to capital deriving from the simulation run of the CGE model, while parameter ΔEMP^{CGE} is the employment level percentage change from the CGE.

WA_{mi} is a dummy variable taking value one if individual *i* of household *m* is at working age (16-64), and zero otherwise. From equation **(18)**, the number of employed over the total number of individuals at working age resulting from the MS model must be equal to the change in the employment rate observed after the CGE run. This implies that the CGE model determines the employment level of the economy after the simulation, and that the MS model selects which individuals among the inactive persons have the highest probability of becoming employed (if the employment level is increased from the CGE simulation result), or either who, among the wage workers, has the lowest probability of being employed after the policy change (if the employment level is decreased)[14].

One possible way of imposing the equality between the two sets of parameters of system of equations **(15-18)** is through a change in the parameters of the selection and regression models. Following Bourguignon *et al.* [11], we restrict this change in the parameters to a change in the intercept of the two functions **(1)** and **(2)**. The justification for this choice is that it implies *neutrality* of the changes, that is, changing the intercepts *a* of equations **(1)** just shifts proportionally the estimated wages of all individuals, without causing any change in the ranking between one individual and the other. The same applies for the activity status choice equation: we choose to change the intercept *α* of equation **(2)**, and this will shift proportionally all the individual probabilities of being a wage worker, without changing their relative positions in the probability distribution, only to let some more individuals to become employed (or some less if the employment rate of the CGE model is decreased), irrespectively of their personal characteristics. This change in the intercept will be of the amount that is necessary to reach the number of wage workers resulting from the CGE model. Thus, this choice preserves the ranking of individuals according to their *ex-ante* probability of being employed, which was previously determined by the estimation of the binomial model. For this reason the change in the intercept parameter satisfies this neutrality property.

[14] And, in this case, his her new wage level will be determined by the regression model of wage earnings.

4 The Top-Down/Bottom-Up Approach

This approach was developed by Savard [46]. It allows overcoming the problem of the lack of consistency between the micro and macro levels of the Top-Down approach by introducing a bi-directional link between the two models: this is the reason why this approach is also called "Top-Down Bottom-Up". According to this method, indeed, aggregate results from the MS model (such as consumption levels or labour supply) are incorporated into the CGE model, and a loop is used to run both models iteratively until the two produce convergent results.

The value added of this approach is that it takes into account the feedback effects that come from the micro level of analysis, which are instead completely disregarded by the Top-Down model. The basic assumption behind this approach is that the microeconomic effects provided by the MS model run do not correspond to the aggregate behaviours of the representative households used in the CGE model, and that it is thus necessary to take these effects back into the CGE model to fully account for the effects of a simulated policy.

The bilateral communication between the two levels of analysis is achieved through a set of vectors of changes, as in the Top-Down approach: from the macro to the micro level of analysis the communication is guaranteed by the changes in the price, wage and return vector and in the employment levels, as before, while from the micro to the macro level the communication we apply two different strategies: in one version, we will use as input for the CGE model the labour supply level resulting from the MS labour supply model; in another version of the same model, the vector of the aggregate consumption levels from the MS model will be used as input for the CGE model[15]. The process is iterated as many times as it is necessary to come to a convergent point, that is, when convergence (at a certain number of decimals) is achieved in the aggregate levels of the two models.

5 Simulation

We will now run a policy simulation with each of the three models. The simulation will be an exogenous shock on the world price level of the good exported by sector 2, which is the labour intensive sector in our stylized economy. The world price of good 2 is reduced of 50 % from its initial value.

The simulation results for the most relevant macroeconomic variables are reported in percentage changes in Tables 6 and 7. In the table, also the two different strategies adopted for the TD BU approach are taken into account, so that we will compare the results coming from the introduction into the CGE model of, respectively, the consumption level and the labour supply coming from the microsimulation module.

In general, we can say that we have very similar results for most of the macro variables in all the four simulations. The shock has negative effects on the economy. Indeed, as we can observe in Table 6, the fall in the price of the exported good for

[15] The choice for consumption level as communicating variable is made following Savard [46].

sector 2 causes a reduction of the production level for this sector, which reduces its demand for both factors of production.
At the same time, exports for the other production sector become relatively more convenient, so that for this sector we observe an increase in the level of the exported good, an increase in the production level, and in the demand for capital and labour.
Anyway, in the economy as a whole, capital and labour demand are lower now, as also the public sector had to reduce its demand for these two factors after the reduction in the tax revenues. This reduction is due basically to the fall in imports levels and in the households' income level.
The reduction in the imported goods for both sectors is caused by the fact that we imposed to have equilibrium of the balance of payments.
As a consequence, sales on the domestic market will be lower, as well as their price, but this reduction is not enough to prevent the fall in the households' consumption level, caused by the decrease of their income.

However, a particular result needs further explanations: savings and investments in the TD BU-Consumption model decrease much more than in the other three models. The reason for this lays in the fact that, in order to be able to introduce exogenous consumption levels into the CGE model, we must endogenize one variable in the households' budget constraint to keep the equilibrium in this constraint. Savard's choice is for the marginal propensity to save, and we follow his approach. But the consequence of this will be a change in the household behaviour with respect to the initial assumptions made for the benchmark. Indeed, the marginal propensity to save of the household will decrease, and thus also households' savings. As in our model investments are savings-driven, this will generate a further reduction of investments.

Table 6. Simulation Results: Percentage Changes

	Integrated Approach	Top-Down Approach	TD BU Approach (Cons.)	TD BU Approach (Labour Supply)
Government Budget	0.00	0.00	0.00	0.00
Wage Rate	-38.07	-37.94	-37.25	-37.92
Capital return	-20.91	-21.21	-25.01	-21.26
Consumer Price Index	-29.74	-29.72	-29.73	-29.71
Labour Supply	-0.91	-1.24	-2.12	-1.32
Employment Change	-0.91	-1.24	-2.12	-1.32
Government Use of Labour	-2.01	-2.58	-4.69	-2.67
Government Use of Capital	-23.27	-23.26	-20.24	-23.26
Income*	-32.94	-33.81	-34.30	-33.84
Disposable Income*	-32.94	-33.81	-34.30	-33.84
Consumption Expenditure*	-32.94	-33.81	-30.12	-33.84
Savings*	-33.40	-33.81	-72.60	-33.84
Utility Level*	-2.80	-5.02	-0.48	-5.07
Tax Revenues	-33.89	-33.90	-34.46	-33.93

Table 7. Simulation Results: Percentage Changes

	Integrated Approach		Top-Down Approach		TD BU Approach (Consumption)		TD BU Approach (Labour Supply)	
	S_1	S_2	S_1	S_2	S_1	S_2	S_1	S_2
Commodity Prices	-30.23	-29.59	-30.37	-29.52	-31.61	-29.15	-30.37	-29.51
Domestic Sales	-6.07	-8.14	-6.20	-8.37	-11.29	-6.91	-6.24	-8.43
Domestic Production	20.22	-8.56	20.25	-8.89	15.88	-7.68	20.22	-8.95
Labour Demand	30.57	-7.83	30.15	-8.18	22.89	-7.16	30.07	-8.25
Capital Demand	10.02	-18.44	10.12	-18.51	8.48	-15.07	10.13	-18.53
Investment Demand	-4.79	-5.67	-4.95	-6.08	-59.94	-61.33	-4.98	-6.14
Imports Demand	-26.99	-39.70	-27.19	-39.79	-32.00	-38.44	-27.23	-39.81
Exports Demand	152.18	-54.66	153.53	-55.01	152.18	-55.22	153.52	-55.07
Export Prices	0.00	-50.00	0.00	-50.00	0.00	-50.00	0.00	-50.00
Consumption Demand*	-3.11	-4.00	-4.95	-6.08	1.70	-1.23	-4.98	-6.14

* For the integrated model, these percentage changes are computed as average percentage changes across households.

With respect to the microeconomic results, and mainly the changes in poverty and inequality, we can observe in Table 8 and 9 that the differences are generally significant only for the case of the integrated model.

The underlying variable for the computation of the indices is per-capita real disposable income, obtained by dividing disposable income by the household specific consumer price index[16], and then dividing it again by the number of adult equivalents resulting by the "Oxford" or "Old OECD" scale (see OECD, 1982). This equivalence scale calculates the number of adult equivalents living in a household by assigning a value of 1 to the first household member, of 0.7 to each additional adult and of 0.5 to each child:

$AE = 1 + 0.7\cdot(\#Adults - 1) + 0.5\cdot(\#Children)$.

[16] The household specific price index is computed using households' consumption shares and the change in prices deriving from the CGE model, as follows:

$$CPI_m = \sum_{q=1}^{2} \eta_{mq} \cdot \left(1 + \Delta P_q^{CGE}\right)\cdot$$

Table 8. Inequality Indices on Disposable per capita Income

	Bmk. Values	Integrated Approach*	Top-Down Approach*	TD BU Approach (Cons.)*	TD BU Approach (Labour Supply)*
Gini Index	33.96	0.11%	-1.50%	-0.18%	-1.51%
Atkinson's Index, ε = 0.5	9.60	-2.03%	-4.49%	-2.34%	-4.51%
Coefficient of Variation	71.80	-0.55%	-2.66%	-1.97%	-2.67%
Generalized Entropy Measures:					
I(c), c = 2	25.78	-1.09%	-5.25%	-3.89%	-5.27%
Mean Log. Deviation, I(0)	19.93	-3.96%	-5.79%	-3.22%	-5.80%
Theil Coefficient, I(1)	20.55	-1.26%	-4.34%	-2.48%	-4.36%

* Percentage deviations from benchmark values.

Table 9. Poverty Indices on Disposable per capita Income

	Bmk. Values	Integrated Approach*	Top-Down Approach*	TD BU Approach (Cons.)*	TD BU Approach (Labour Supply)*
General Poverty Line					
Headcount Index, P_0	39.34	75.0%	-4.2%	-4.2%	-4.2%
Poverty Gap Index, P_1	9.88	153.3%	-2.7%	5.1%	-2.7%
Poverty Severity Index, P_2	0.00	184.5%	-7.0%	4.0%	-7.0%
Extreme Poverty Line					
Headcount Index, P_0	4.92	166.67%	0.00%	0.00%	0.00%
Poverty Gap Index, P_1	0.96	90.96%	-31.45%	-31.93%	-31.44%
Poverty Severity Index, P_2	0.00	35.34%	-43.04%	-42.67%	-43.02%

* Percentage deviations from benchmark values.

First of all, we observe that the Top-Down and the TD BU-Labour Supply approach show almost identical results for what concerns both poverty and inequality indices.
The TD BU-Consumption model we observe a smaller effect on inequality, but in the same direction as for the other two models, while for poverty we again observe almost the same effect as before.
The biggest difference in the microeconomic results is to be detected in the Integrated approach, where we observe similar results for the inequality indices, but not for the poverty measures. Apparently, from this model we observe a boost in the poverty measures after the fall of the price of the exported good for the second sector. This is probably due to the fact that the integrated approach is not able to capture the

behavioural responses of the agents to the policy reform, so that the labour supply behaviour of individuals remains basically unchanged, thus leading to a parallel shift of all incomes (and thus almost no changes in the distribution of incomes among the households, but a big increase in poverty rates, as the incomes are reduced with the shock) [17].
The TD and TD BU models, instead, while allowing for individually estimated labour supply responses, generate a lower change in the poverty indices.

6 Conclusion

In this paper we tried to give an assessment of the recent developments observed in methods that link together CGE and microsimulation models, with a special concern for the different linking approaches existing in the literature. Especially, we have focused our attention only on static models. By using data from a fictitious economy, we have built three models: one that follows the full integrated approach, as in Cockburn [16]; another one that follows the so called Top-Down approach, as it is developed in Bourguignon *et al.* [11], and the last one that follows the method developed by Savard [46], also known as Top-Down Bottom-Up model.
On one side we can say that a simple integrated approach like the one we have implemented in this paper is deficient on the side of the microeconomic specification and behavioural responses by individual agents. Anyway, the introduction of microeconometric behavioural equations into a CGE model looks of hard application and cumbersome for computational aspects.
On the other side, a Top-Down approach completely disregards the possible feedback effects coming from the microeconomic side of the economy, which could affect also the macroeconomic variables. Anyway, from the results of our analysis, these feedback effects apparently do not change results in a fundamental way. However, this specific aspect of linking CGE and microsimulation models needs further investigation, and we cannot say that there is a reason to believe *a priori* that the changes in the aggregate microeconomic behaviour do not affect the macroeconomic framework of the economy under study.
In our opinion, indeed, the TD BU approach looks the most complete approach, as on one side it can include all the possible microeconometric estimates to account for behavioural responses by individual agents, and on the other side it also takes into account the feedback effects from the micro to the macro level of analysis. Anyway, the way in which this approach is developed by Savard [46], that is, using consumption levels as communicating variable, does not look the most efficient one, as we have to introduce an exogenous consumption level into the CGE framework, and to do this it is necessary to endogenize another variable or parameter, thus changing the initial assumptions on agents' behaviour. If the change in consumption levels from the microsimulation model is small, then this change in agents' behaviour

[17] In the CGE model, we also have endogenous labour supply for each household, but in this case the labour supply response by households is the same for each agent, as it does not depend on individuals' personal and family characteristics, as in the microsimulation model.

could be very small and not significant for the results of the model, but if the shock is big, as it is in our model, then the change in agents' behaviour could lead to quite different results, at least in some important macro variables (in our example, savings and investments).
The choice of labour supply as communicating variable from the micro to the macro model looks a better choice, because it does not affect the initial assumptions that are made on individuals' behaviour, and at the same time it is able to account for individual behavioural responses that can affect the macroeconomic settings of the economy.

References

1. Adelman I., Robinson S. (1978), Income Distribution Policy in Developing Countries: A Case Study of Korea. Oxford, Oxford University Press.
2. Agénor P.R., Izquierdo A. and Fofack H. (2001), IMMPA: A uantitative Macroeconomic Framework for the Analysis of Poverty Reduction Strategies. Washington, World Bank.
3. Ahuja V., Bidani B., Ferreira F., Walton M. (1997), Everyone's miracle? Revisiting poverty and inequality in East Asia. Directions in Development, Washington D., World Bank.
4. Armington P.S. (1968), A Theory of Demand for Products Distinguished by Place of Production. IMF Staff Paper n° 16, pp. 159-176.
5. Arntz, M., Boeters S., Gürtzgen N., Schubert S. (2006), Analysing Welfare Reform in a Microsimulation-AGE Model: The Value of Disaggregation, EW Discussion Paper No. 06-076, Mannheim.
6. Atkinson A.B. (1970), On the measurement of inequality. «Journal of Economic Theory», 1970, 2: 244-263.
7. Bacharach M. (1971), Bi-proportional Matrices and Input-Output Change, Cambridge University Press, Cambridge.
8. Bargain O. (2005), On Modelling Household Labour Supply with Taxation, I A DP No. 1455.
9. Boccanfuso D., Cissé F., Diagne A. and Savard L. (2003), Un modèle CGE-Multi-Ménages Intégrés Appliqué à l'économie Sénégalaise, miméo, CREA, Dakar.
10. Bourguignon F., Pereira da Silva L., eds. (2003), The impact of economic policies on poverty and income distribution: evaluation techniques and tools, The World Bank.
11. Bourguignon F., Robilliard A.S. and Robinson S. (2003), Representative versus real households in the macro-economic modelling of inequality, DT 2003 10, DIAL Unité de recherche CIPRE.
12. Bourguignon F., Spadaro A. (2006), Microsimulation as a Tool for Evaluating Redistribution Policies, «Journal of Economic Inequality», vol. 4, No.1, pp.77-106.
13. Bussolo M. and Lay J. (2003), Globalisation and Poverty Changes: A Case Study on Colombia, Working Paper No. 226, OECD Development Centre, Paris.
14. Byron R.P. (1978), The Estimation of Large Social Accounting Matrices, «Journal of the Royal Statistical Society», Series A, 141 (3).
15. Cloutier M.-H., Cockburn J. (2002), How to build an integrated CGE microsimulation model: Step-by-step instructions with an illustrative exercise, mimeo, CIRPÉE, Université Laval.

16. Cockburn J. (2001), Trade Liberalization and Poverty in Nepal: A Computable General Equilibrium Micro-simulation Approach. Working Paper 01-18, CRÉFA, Université Laval.
17. Cogneau D. (2001), Formation du revenue, segmentation et discrimination sur le marché du travail d'une ville en développement: Antananarivo fine de siècle, DIAL DT 2001-18.
18. Cogneau D., Grimm M., Robilliard A.-S. (2003), Evaluating poverty reduction policies – the contribution of micro-simulation techniques, in Cling J.-P., Razafindrakoto M., Roubaud F. (eds), The New International Strategies for Poverty Reduction, London: Routledge.
19. Cogneau D., Robilliard A.-S. (2001), Growth, Distribution and poverty in Madagascar: Learning from a Microsimulation Model in a General Equilibrium Framework, IFPRI TMD Discussion Paper No. 61 and DIAL DT 2001-19.
20. Cogneau D., Robilliard A.-S. (2004), Poverty Alleviation Policies in Madagascar: A Micro-Macro Simulation Model, DT 2004 11, DIAL é Unit Recherche CIPRÉ.
21. Colatei D., Round J. (2000), Poverty and Policy: Some Experiments with a SAM-Based CGE Model for Ghana, mimeo, Warwick University.
22. Cororaton B., Cockburn J. (2005), Trade Reform and Poverty in the Philippines: a Computable General Equilibrium Microsimulation Analysis, CIRPÉE, Working Paper 05-13, Université Laval.
23. Creedy J. and Kalb G. (2005): Discrete hours labour supply modelling: specification, estimation and simulation, «Journal of Economic Surveys» 19, 697-734.
24. Creedy J., Duncan A. (2002), Behavioural Microsimulation with Labour Supply Responses, «Journal of Economic Surveys» n. 16, pp. 1-38.
25. Creedy J., Duncan A., Harris M. and Scutella R. (2002), Microsimulation Modelling of Taxation and The Labour Market: The Melbourne Institute Tax Transfer Simulator, Cheltenham: Edward Elgar.
26. Davies J.B. (2004), Microsimulation, CGE and Macro Modelling for Transition and Developing Economies, WIDER Discussion Paper No. 2004 08.
27. Decaluwé B., Patry A., Savard L. and Thorbecke E. (1999), Poverty Analysis Within a General Equilibrium Framework, Working Paper 99-09, African Economic Research Consortium, CRÉFA 99-06.
28. Decaluwé B., Dumont J. and Savard L. (1999), Measuring Poverty and Inequality in a Computable General Equilibrium Framework, Université Laval, Cahier de recherche du CRÉFA n° 99-20.
29. Dervis K., de Melo J., Robinson S. (1982), General Equilibrium Models for Development Policy. New York, Cambridge University Press.
30. Duncan A. and Weeks M. (1998), Simulating transitions using discrete choice models, «Papers and Proceedings of the American Statistical Association» 106, pp.151-156.
31. Ferreira F.H.G. and Litchfield J.A. (2001), Education or inflation?: The Micro and Macroeconomics of the Brazilian Income Distribution during 1981-1995, «Cuadernos de Economia», 38, pp. 209-238.
32. Foster J., Greer J., Thorbecke E. (1984), A Class of Decomposable poverty Measures, «Econometrica», 52(3), pp. 293-312.
33. Franses P.H. and Paap R. (2001), uantitative Models in Marketing Research, Cambridge University Press, Cambridge, UK.
34. Heckman J. J. (1976), The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. «Annals of Economic and Social Measurement», 5(4): 475-492.
35. Heckman J. J. (1979), Sample Selection Bias as a Specification error, «Econometrica», Vol. 47, No.1, pp.153-162.

36. Hérault N. (2005), A Micro-Macro Model for South Africa: Building and Linking a Microsimulation Model to a CGE Model, Melbourne Institute Working Paper Series, WP No. 16 05.
37. Lysy F.J., Taylor L. (1980), The General Equilibrium Income Distribution Model, in Taylor L., Bacha E., Cardoso E., Lysy F.J. (eds.), Models of Growth and Distribution for Brazil, Oxford: Oxford University Press, Ch.6.
38. Maddala G.S. (1983), Limited Dependent and uitative Variables in Econometrics, New York, Cambridge University Press.
39. Mitton L., Sutherland H., Weeks M., eds. (2000), Microsimulation Modelling for Policy Analysis. Challenges and Innovations, Cambridge University Press: Cambridge.
40. Mookherjee D., Shorrocks A. (1982), A Decomposition Analysis of the Trend in UK Income Inequality, «Economic Journal», Vol. 92, n° 368, 886-902.
41. Müller T. (2004), Evaluating the economic effects of income security reforms in Switzerland: an integrated microsimulation-computable general equilibrium approach, paper presented at the IIOA-Ecomod Conference, Brussels, September 2-4 2004.
42. OECD (1982), The OECD List of Social Indicators, Paris.
43. Orcutt G. (1957), A new type of socio-economic system, «Review of Economics and Statistics», Vol. 58, pp. 773-797.
44. Robilliard A.-S. and Robinson S. (2003), Reconciling Household Surveys and National Accounts Data Using a Cross Entropy Estimation Method, «Review of Income and Wealth», vol. 49, No. 3, pp. 395-406.
45. Robinson S., Cattaneo A., El Said M. (2001), Updating and Estimating a Social Accounting Matrix Using Cross Entropy Methods, «Economic Systems Research» 13 (1): 49-64.
46. Savard L. (2003), Poverty and Income Distribution in A CGE-Household Micro-Simulation Model: Top-Down Bottom-Up Approach, CIRPÉE Working Paper 03-43, Université Laval.
47. Savard L. (2004), Poverty and Inequality Analysis Within a CGE Framework: a Comparative Analysis of the Representative Agent and Micro-Simulation Approaches, CIRPÉE, Working Paper 04-12, Université Laval.
48. Shoven J.B., Whalley J. (1984), Applied General Equilibrium Models of Taxation and International Trade. «Journal of Economic Literature» 22: 1007-51.
49. Stone R. (1978), "Forward" to G. Pyatt, A. Roe et al., Social Accounting for Development Planning, CUP, (xvi-xxxi).
50. Van Soest A. (1995), A Structural Model of Family Labour Supply: A Discrete Choice Approach, «Journal of Human Resources», n.30, pp.63-88.

Poster Session

Governance Based on Reputation

Rosaria Conte and Mario Paolucci

LABSS, Institute for Cognitive Science and Technology, Rome, ITALY

The *governance* of social, economical and cultural transformations aims to move beyond the contrast between institutional, centralised *government* and complete deregulation. In these pages we will endeavour to show that reputation constitutes an essential aspect of the governance of economy and of public institutions.

1 What is governance

Regrettably, the term governance is vague and ill defined. Governance is understood as a new style of management, different from hierarchical, top-down approaches, and characterised instead by a larger degree of cooperation and interaction between the State and other actors inside mixed, public-private decisional networks.

In a more concrete sense, governance is a set of regulation mechanisms that applies to complex social systems, and that are

decentralised
dynamical
bidirectional both from the top and from the botton, from citizens to institutions and the other way around
mixed
decentralised
both deliberate and spontaneous
realised by the intervention of several agencies both public and private

The main thesis of the paper is that reputation is the most important mechanism to satisfy such requirements: the most spontaneous, and probably the most efficient too.

2 What is reputation

Reputation is a socio cognitive artefact, evolved to deal with different social order problems that human societies have encountered. It is especially suited for social control and partner selection (see Conte & Paolucci, 2002) Reputation interacts with social evaluation, but is distinct from it . Evaluation is a value judgement formed from the evaluator about some definite target (individual, firm, institution etc..), based on direct or indirect experience with the target itself. Making such a judgement explicit in the form of direct feedback, the evaluator commits itself on the truth value of the judgement. As a consequence, the evaluator take responsibility both towards the target, especially for negative judgements, and toward the information recipients, more so in the case of positive judgements.

Not so for reputation, that is born from reporting others' evaluation as gossip; in this case, the informer is not an evaluator; it does not give a personal judgement on the target, and does not take responsibility toward the information recipients.

3 Reputation for the governance of economy

Reputational systems are already appreciated by the market; they constitute an asset with a precise economic value (compare the large literature on corporate reputation). In addition, reputation for electronic markets, and in particular in electronic auctions (eBay being the most striking case), is an undeniable reality. Finally - and more generally - we must account how reputation has become a technological artefact, from its origins as a spontaneous artefact. Reputation technology is a business, bound to grow in the next years as the business exploits the complex effects of reputation, modifying and them and making them extreme. Moreover, business plays the double role of reputation exploiter, for market regulation, and reputation guarantee, as the proliferation of "reputation defender" systems demonstrates.

However, this proliferation not always gives the desired effects. For example, the diffusion of online reputation - especially in electronic auctions - has not generated a decrease of frauds (see Marmo, in press), that instead tend to grow year after year. What is also growing is the market of auctions itself: eBay has shown an impressive growth rate, without pauses, from its beginnings. Do we have to deduce that reputation, instead of acting as a regulatory instrument, is simply a change to reap new profits? To the contrary, reputation is a regulatory instrument. But what is implemented in online markets has very little to do with reputation, apart from the name. Let's see why.

4 Courtesy and Prudence

The first reason of the courteous variant of reputation that we find on the internet, so different from the natural experience of reputation that we share from participating in gossip, hides behind the lack of distinction between evaluation and reputation. Online reputation, as the current state of the art, is so courteous and so useless because evaluators simply cannot gossip; they cannot hide behind the screen of impersonal judgements.

In addition, often the system itself is characterised by target-oriented benevolence. Evaluators are at the same time targets and beneficiaries of the evaluation, and thus tend to operate in collusion.

To the contrary, systems that favour an effective social control are characterised by benevolence towards the beneficiary, to the expenses of targets (imagine the gossip of students about teachers). Targets are met with a deterrent given by the costs of bad reputation, and an incentive to obey rules, reciprocate, collaborate.

Reputation, quite obviously, can operate to reduce frauds in markets: but depending on how it is built and applied. Not a rigid instrument, reputation is a flexible artefact. In natural societies, reputation has evolved working in two directions: decentralized norm enforcement, or social control, in the form of punishment by isolation on one side; network enlargement, thanks to the availability of informations beyond the familiarity network, on the other. These function, however, appear as results of mechanism that are still under study - chain like communication, narrowcast, low intersection between targets and reputation users, anonymity, and, principally, distinction between supported feedback and reported feedback. If we want to apply reputation for governance and fraud control, we have somehow to reorient the reputational system in a richer and different way, getting rid of the simplifications introduced by current online systems. In short, to operate reputation based governance, we need to

- acquire a detailed knowledge on how reputation works in nature
- build theory-driven technologies, in function of the specific objectives to pursue.

5 Reputation for the governance of public institutions

An important aspect of the current notion of governance is the interleaving of public and private aspects. Financial crises in the 80's and 90's let emerge the idea of 'steering', that redefined public intervention.

Independently for any attribution of value, this mix of public and private favoured two complementary directions:

- a set of new rules and new institutions (economic laws, documentation etc.) aimed to the firms and to the market
- the "New Public Management", a new management of public sector aimed at
  - the reconfiguration of public structures as enterprises, often simplified in the shape of a set of efficiency indicator to be respected
  - decentrated and bottom-up control of public structures

We see this last point, in particular, with interest; this approach is based on the idea that institutions, even public ones, are accountable and subject to evaluation from several categories of subjects: users, operators, other institutions. These evaluation call, in turn, for a theory of institutional reputation.

Bibliography

Conte, R., Paolucci, M. (2002): Reputation in artificial societies: Social beliefs for social order. Kluwer Academic Publishers (2002)

Sabater, J., Paolucci, M., Conte, R. (2006): Repage: Reputation and image among limited autonomous partners. Journal of Artificial Societies and Social Simulation 9(2)

Modelling upstream-downstream problems using the IAD framework

Eva Ebenhöh[1] and Gert Becker[2]

[1] Institute of Environmental Systems Research
University of Osnabrück
49069 Osnabrück, Germany
Eva.Ebenhoeh@usf.uos.de
[2] Institute for Environmental Studies (IVM)
University of Amsterdam
1081 HV Amsterdam, The Netherlands
Gert.Becker@ivm.falw.vu.nl

Abstract The model presented in this paper deals with different motivations of upstream and downstream users in a river system. It contrasts three stylized regimes of transboundary flood protection: decentralised, centralised, and collective action. Although the model is based on findings of a case study conducted in the Rhine, the model makes no claim at being predictive. The Rhine is modelled as a simplistic environment which serves as a learning environment for decision making agents, which are the focus of the research.

1 Introduction

A research interest of a case study currently in process is the investigation of how cooperation in transboundary flood protection can be established. Different motivations can lead to transboundary cooperation. These are investigated not only empirically, but also in form of an agent-based model. Using the Rhine as an example, this model contrasts three stylized regimes of transboundary flood protection: decentralised, centralised, and collective action. The model is developed as one of several models of water management regimes [1] based on the Institutional Analysis and Development (IAD) framework [2,3]. Using the same conceptual framework aids model comparison and enables researchers to use agent-based modelling as an enhanced process tracing [4] technique. Data from case studies are used to design models, which are not primarily used to replicate data or make predictions, but rather for exploration, explanation and testing assumptions on real world processes. Consequently, the model makes no claim at being predictive or descriptive with regard to the Rhine.

2 Model Outline

2.1 Environment

The Rhine is modelled as a simple three-bucket model of peak water levels. As the area of interest is the German-Dutch cooperation, the first bucket is the upper Rhine in southern Germany, the second the lower Rhine in North Rhine-Westfalia, and the third the Dutch part of the Rhine. As long as the current water level is below the height of the dikes, water simply flows from one bucket to the next. The second bucket receives an additional input from a tributary. If, however, the peak water level threatens to rise above the dikes, water is redirected into retention areas. We distinguish between retention areas that have a substantial impact inside the bucket and those that only have an impact on the discharge into the next bucket. Retention amounts correspond to decreases in peak water levels. If this is not enough to decrease it below dike level, overflow of dikes also reduces the discharge.

This is a simplistic environment, which nevertheless serves as a learning arena for decision making agents. Daily water peak levels can be varied as an input to the model, corresponding to climate change scenarios.

2.2 Management Regimes

Decision making agents are authorities in the areas corresponding to the three buckets. Their objective is to increase the water level at which no damage occurs, that is, no dikes are overflown. This can be achieved by raising or relocating dikes and/or by creating retention areas. Decision making agents make plans including these three different measures based on data about their own area. This data includes possible retention areas, but might also include expectations of how high water levels can become. Measures have costs and decision making agents aim for cost-effectiveness. Furthermore, creating retention areas may fail, because of potential vetos by people living in or owning the areas. This still entails costs.

2.3 Institutional Variation

In the decentralised version of the model, decision makers calculate for their respective areas only. No retention areas are planned, which would lower only the flow into the next bucket. Raising dikes, which effectively creates higher discharges into the next bucket, is an attractive measure. The external effect of raising dikes on downstream buckets is not taken into account.

In the centralised version, decision makers are obligated to calculate effects of measures by taking the whole river basin into account. However, it may occur, that this rule is not effective. There are several ways around this rule. Raising the dike in an upstream bucket can, for example, still happen on the expectation that in the downstream buckets dike heights are also increased. A possible retention area close to the border to the next bucket may not be build on the expectation that residents will veto anyway.

In the collective action version, decision makers meet to potentially exchange information and money, and by that build trust and other regard in the sense, that it may be included in their own preferences to take the whole basin into account. Agreements on compensatory payments, in order to build a floodplain in another bucket, are possible. Communication is modelled in a stylised way, allowing only exchanges of (a) money, (b) information on water levels and potential retention areas, and (c) suggestions on flood protection plans (to include specific potential retention areas in the plan), which may be coupled with an exchange of money. The process of increasing other regard does not happen consciously.

3 Expected Results

It seems obvious, that an *effective* centralised regime fares better than the decentralised regime. Interesting aspects are reasons for following a rule or not, including individual preferences and potential sanctions. Other factors will have to be further investigated in a case study.

More interesting than contrasting centralised and decentralised regimes is comparing them with the collective action situation. It is to be expected, that a collective action approach may, under certain conditions, work as well as an effective centralised regime and outperform the decentralised and ineffective centralised regimes.

This model is a tool to investigate conditions and mechanisms for effective collective action regimes based on empirical evidence gathered in case studies and on theory (e.g. [7]).

References

1. Ebenhöh, E.: Testing hypotheses on adaptive management using an agent-based model based on the institutional analysis and development framework. Deliverable 1.2.3 in the NeWater Project (2007) <www.usf.uos.de/∼eebenhoe/newater/Mileston1.2.3.pdf>.
2. Ostrom, E., Gardner, R., Walker, J., eds.: Rules, games, and common-pool resources. Ann Arbor, Michigan: The University of Michigan Press (1994)
3. Ostrom, E.: Understanding Institutional Diversity. Princeton University Press (2005)
4. Homer-Dixon, T.: Strategies for studying causation in complex ecological political systems. <http://www.library.utoronto.ca/pcs/eps/method/methods1.htm> (February 28th, 2007) (1995)
5. Crawford, S.E.S., Ostrom, E.: A grammar of institutions. American Political Science Review **89** (1995) 582–600
6. Crawford, S., Ostrom, E.: A grammar of institutions. [3] chapter 5 137–174
7. Ostrom, E.: Toward a behavioral theory linking trust, reciprocity, and reputation. [8] chapter 2 19–79
8. Ostrom, E., Walker, J., eds.: Trust and Reciprocity. Volume VI of Russel Sage Foundation Series on Trust. Russel Sage Foundation, New York (2004)

Modelling Primate Social Order: Ultimate Causation of Social Evolution

Hagen Lehmann and Joanna J. Bryson

Department of Computer Science, University of Bath, Bath BA2 7AY, United Kingdom

H.Lehmann@bath.ac.uk, J.J.Bryson@bath.ac.uk

The objective of our work is to understand the selective pressures that result in different social organisations in primates. We approach this by means of Agent Based Modelling. We here introduce a simulation which models the changes of social styles in the genus *macaca* depending on environmental factors such as predation and food distribution.

For many species of primates, arguably including humans, social behaviour can be characterised along an axis usually described as running from *egalitarian* to *despotic*. Despotic societies are characterised by a strict hierarchy with very few aggressive interactions, but where aggression occurs it is typically violent and usually unilateral from dominant to subordinate. Egalitarian societies have less well-defined hierarchies; frequent, bilateral, but less violent aggression, and a large repertoire of reconciliation behaviours.

We model the closely related species in the genus *macaca* and the environmental pressures working on their social organisation. To simplify the problem we have split the modelling into the ultimate and proximate cause of behaviour. In ethology ultimate causation explains an animal's behaviour based on evolution, while proximate causation explains an animal's behaviour based on trigger stimuli and internal mechanisms.

Our present simulation focuses on the ultimate causation. We represent two groups of macaques differing in their social style. One group is egalitarian, the other despotic. The social styles are defined by the way the animals interact and the average distance between them. Different experimental conditions require changing the environmental setups these groups are confronted with. Conditions represent either environments with high predation pressure and restricted food access, high predation pressure and high food access, low predation pressure and high access to food or low predation pressure and low food access. We then measure which group was more successful in which environment. The success of a group was defined as the number of offspring.

Our results correspond to empirical findings from primate field research. The group with egalitarian social structure did significantly better in conditions with predation pressure and high access to food and the despotic groups did better in environments with low predation pressure and low food access.

Our work provides additional evidence for currently discussed theories in evolutionary socio-biology, and provides a tool for testing these theories.

Linking Artificial Models and Reality: The Unnecessary Quest

Sébastien Liarte[1]

[1] IAE de Toulouse, Université de Toulouse 1, Rue du Doyen Gabriel Marty,
31042 Toulouse Cedex, France.
sebastien.liarte@univ-tlse1.fr

For management researchers, computer simulation is a particularly helpful research method. However, its recent introduction, in social sciences in general and in management sciences in particular, means that computer simulation has yet to be accepted in the researcher's "toolbox." To guarantee computer simulation's early acceptance by the scientific community, its users have paid particular attention to the validity of their results. In particular, external validity, through comparison of real and simulated data, has been observed. However, this focus on reality has diverted attention from the most fertile issues regarding computer simulation, and created some misunderstandings.

The aim of this paper is to clarify the three main misunderstandings. Imitation of reality is not the aim of computer simulation (Misunderstanding 1). Simpler computational models are no more useless than more complex ones (Misunderstanding 2). Artificial data simply do not have the same status as real data (Misunderstanding 3).

Here we do not discuss the willingness to establish links between the computational models and reality, but the utility of this correspondence. Instead of reducing criticism, the use of real data increases it. The question of the best use of the computer simulation remains. The purpose of this paper is to propose an alternative use of computer simulation. Researchers can use computational models as metaphors from which to establish analogies. Indeed, researchers can establish similarities between the simulated model and reality, and work from artificial data.

Traffic Simulation with the TRASS Framework

Ulf Lotzmann

Institute of Information Systems Research, University of Koblenz,
Universitätsstraße 1, Koblenz 56070, Germany
ulf@uni-koblenz.de

In order to simulate route choice behavior of agents in complex and heterogeneous environments, we developed an agent-based TRAffic Simulation System (TRASS) suitable for traffic simulation with a wide range of freedom concerning agent design – particularly in terms of implementing agent behavior[1].

In TRASS, objects like traffic participants (e.g. cars, pedestrians, traffic lights) and topographical regions are regarded as agents.

The entities of the first kind of agent are classified into two subtypes. "Simple" agents are of the reactive type and feature properties such as shape (a freely definable set of circles), position, direction, velocity. They include custom code for reacting on requests by other agents or external events (triggered by timer or agents). A more complex type of agent extends the reactive type by a powerful sensor unit (a freely definable set of circular-sector-shaped perception areas) in combination with custom code for perception of the environment and implementation of cognitive behavior.

The second type of agent – the topographical region – embrace behavioral properties of the reactive agent type, but otherwise exhibit significant differences, as they are static (in terms of size and position) and polygon-shaped objects.

In principle, the customization of agent behavior has to be done by implementing Java methods, which are predefined by kernel interfaces. The proceeding is similar to the way of using other well-known simulation frameworks like Repast and Mason.

In the TRASS framework, the agent behavior is encapsulated within so-called strategy modules, which are Java objects that can be attached to respective agent objects.

In order for agents to navigate in complex environments, the agent-strategy design has to cover two distinct aspects. First, an agent must be able to interact within his physical environment directly. This includes recognition and classification of agents and topographic regions as well as conducting elementary actions (e.g. driving a curve, changing a lane) within this environment. Second, an agent needs some kind of reasoning to make strategic decisions with the aim to produce goal-driven behavior.

For this reason we use a two-layered approach. The first facet – recognition and basic action – is more of a technical nature and therefore packed into a so-called "robotics layer". The techniques used in this layer, such as state machines, are quite similar to those used to control autonomous robots. The other layer – called "AI layer" – comprises all aspects subsumed under the keyword reasoning. This layer manages the robotics layer and can be seen as the link to the social simulation domain.

[1] Lotzmann, U.: Design and Implementation of a Framework for the Integrated Simulation of Traffic Participants of All Types. In: Bruzzone, A. G. et al (eds.): Proceedings of the International Mediterranean Modelling Multiconference, Barcelona (2006)

Associative Memory Approach to Modeling Stock Market Trading Patterns

Makarenko A.*, Levkov S.**, Solia V.*

*) National Technical University of Ukraine "KPI",
Institute for Applied System Analysis, 37 Pobedy Ave., 252056 Kiev, Ukraine,
makalex@i.com.ua
**) New Jersey Institute of Technology,
Electrical and Computer Engineering, University Heights, Newark, NJ07102,
levkov@megahertz.njit.edu

We use the ideas of stochastic Theory of Social Imitation (W.Weidlich, E. Calen and D. Shapiro, T. Vaga), and of the associative memory approach to modeling the dynamical structure of polarization relationships (S. Levkov&A. Makarenko) for modeling the stock market trading patterns. The method potentially will allow us to forecast the offer and demand dynamics of a particular security, and lead to modeling of the assets' price behavior. In contrast to the existing ANN models, where the real process is considered as a “black box”, and ANN is trained on the sets of input and output data to simulate the nonlinear relationship between them without actually revealing the nature and structure of the prototype process, our approach is based on the attempt to utilize the principles of the certain classes of neural networks to reveal and model the underlying structure of the real dynamical process.

For simplicity we consider the idealized market of one security. The trade consists of discrete steps. Within each step we identify the sub steps, which describe the dynamic bidding and asking or decision-making processes for every individual. The market consists of N participants. With every trader we associate the state variable s_i which represents the number of shares that trader i is planning to buy (if s_i>0) or to sell (if s_i<0). With every pair of traders i and j we associate the variable $c_{ij} \in \mathbf{R}$ - the integral value of reputation that trader j has from the point of view of trader i. to be wrong in his judgment. As one of the basic characteristics of the system we introduce the concept of a vector field of influence Obviously, the best strategy for rational individual will be to adjust his own initial intentions to the filtered information about others. Thus, we formulate the evolution equation describing the trading dynamics. The initial conditions for this dynamic equation are the intentions of each individual to buy or sell at the beginning of the trading step. Given the initial conditions for s_i and known values of influence matrix, we may calculate the dynamics of the traders.

The next step in development of proposed models is to account the internal structure of agents (we named such agents as ‘intellectual’). Proposed approach allows developing the software and trying to understand some properties of market. All proposed internal representation may be considered as some correlate to ontology of market participant. Also the approach reminiscent usual multi-agent approach. But the prospective feature in the approach is the associative memory property.

The Good, the Bad and the Rational: Attraction and Cooperation

Elpida Tzafestas

Institute of Communication and Computer Systems
National Technical University of Athens
Zographou Campus, Athens 15773, GREECE
brensham@softlab.ece.ntua.gr

In this paper we are presenting a mechanism of attraction and its impact on cooperation in a society of agents. We adopt the benchmark setting of the Iterated Prisoner's Dilemma (IPD, [1]) and we implement attraction as a mechanism that changes the usual agent's behavior (strategy). More specifically, an agent follows its regular strategy unless it faces an attractive agent. In the latter case, she becomes unconditionally cooperative. This implementation of attraction makes sense in a diverse or noisy environment where many agents differing in behavior coexist and the outcomes of agent actions may be distorted by the environment.

We perform our experiments as tournaments in populations of agents using one or more of the extreme "irrational" models ALLC (allways cooperate), ALLD and the "rational" models TFT (Tit-For-Tat) and Adaptive TFT. Occasionally we use other irrational models as well. We find that in populations of 20 to 40 agents with an attraction factor from 5 to 20 (the number of other agents that an agent is attracted by) the average agent score indeed rises. The rise is more significant in the case of irrational agents, and reaches its maximum for the ALLD. These results should be expected because normally irrational agents achieve low scores and especially in front of the rational ones who retaliate consistently.

We have experimented with the type of attraction pairing by studying three variants: the average pairing (where on average every agent is attracted by N others), the exact pairing (where each agent is attracted by exactly N others) and the random pairing (where, unlike the two previous cases, the attraction relations are not necessarily reciprocal). We found that the latter type is the most performant, owing to its potential to drag rational but not-attracted opponents to cooperation with an attracted, even if irrational, agent. The improvement also increases with the attraction factor. In sum, all social variants that induce more interactions between agents with non-reciprocal attraction relations, are bound to lead to a better social average. Final experiments have been with extreme agents, namely, a "Don Juan" agent that is attracted by all others, and a "Sex Symbol" to whom all others are attracted. Very often, the introduction of a single extreme agent in a society with no other attraction relations may lead to the same result as a regular society with attraction (but without the extreme agent).

Our results suggest that often social systems with economic relations (exemplified by the IPD benchmark) can be performant and stable with the aid of external relations that act psychologically, such as attraction, despite apparent inconsistencies or irrationalities in individual agents behavior. Accordingly, we can sometimes manage some economic systems by manipulating an external social condition that interferes with normal economic behavior.

Keywords: Cooperation, Iterated Prisoner's Dilemma (IPD), Rationality, Attraction, Don Juan, Sex Symbol, Groups.

Combining Cognitive Plausibility with Social Realism

Kees Zoethout & Wander Jager

Faculty of Management & Organization
University of Groningen
k.zoethout@rug.nl

Abstract. By looking at studies that simulate social phenomena, we see a number of theories that are formalised into simulation programs (e.g. Troitzsch, 2005; see also the journals JASSS, SIMPAT, CMOT or AAMAS). Within the field of social simulation studies we can distinguish two approaches. The first approach involves the study of sociological systems, social networks, and social interaction (Axelrod, 1984; Zeggelink, 1993; Kitts et al, 1999; Back & Flache, 2006). The second approach involves the study of the behaviour of cognitive plausible agents (Carley & Prietula, 1994a; Helmhout, 2006; Sun, 2006). Whereas the first approach does not focus at a plausible description of the cognitive properties of the agents, the second approach does not apply sociological or social-psychological theories to its models (Zoethout, 2006). Only a very few scholars focus on both the individual and the social level (e.g. Conte & Paolucci, 2001). Exploring the links between the two approaches may provide a perspective on modelling and simulating processes of social interaction in a more plausible way (Zoethout, 2006; Sun, 2006). For instance, the emergence of social structures (Nowak & Latané, 1995) can very well be described without the underlying cognitive substratum. However, such a description does not answer question related to the influence of these higher order phenomena to the lower order cognitive structure, nor does it identify the causal mechanism actually resulting in the emergence of these phenomena. In order to increase the causal explanatory power of simulation models, we argue that social simulation sometimes needs cognitive plausible agents, especially when we are confronted with questions regarding lower order causes or consequences of higher order social phenomena. But the opposite is also true: Cognitive plausible agents may be able to mimic cognitive information processing principles, but what is their use as long as they are not capable to interact with the most significant part of their environment, i.e., their social world. Because of this, we see a tendency for cognitive models to become more social (e.g. Sun, 2006) and social models to become more cognitive (e.g. Conte & Paolucci, 2001). Further, as Sun's (2006) work illustrates, a growing interest emerges to relate both approaches. But although we applaud to this promising initiative, the examples suggest that are is yet much work to be done. Therefore, this paper focuses on possibilities for strengthening the relationship between both fields. First, we will elaborate on defining the problem itself by describing the need for integration. Second, we will discuss two ways of dealing with the problem. For each of these ways, without having the intention to give a

complete or even extensive literature review, we will describe some representing fields within MAS. On the basis of this description we will conclude that there are different venues to take in integrating both fields. We will end with some suggestions for integration based on each of these ways.

References

Axelrod, R., (1984), The evolution of cooperation. New York: Basic Books

Back, I. and Flache, A.(2006), The Viability of Cooperation Based on Interpersonal Commitment, Journal of Artificial Societies and Social Simulation 9(1) <http://jasss.soc.surrey.ac.uk/9/1/12.html>.

Carley, K.M. & Prietula, M.J. (1994), ACTS Theory: Extending the Model of Bounded Rationality, in Computational Organization Theory, K.M. Carley & M.J. Prietula (eds.) Lawrence Erlbaum Associates, Publishers, Hillsdale,etc.

Conte, R., and Paolucci, M. (2001), Intelligent Social Learning, Journal of Artificial Societies and Social Simulation vol. 4, no. 1, <http://www.soc.surrey.ac.uk/JASSS/4/1/3.html>

Helmhout, M. (2006), The Social Cognitive Actor: A Multi-agent simulation of organizations, Doctoral Thesis Groningen: University of Groningen, Faculty of Management and Organization, Ridderkerk, Labyrinth Publication

Kitts, J.A., Macy, M.W. and Flache, A., (1999). Structural learning: Attraction and Conformity in Task-Oriented Groups, Computational & Mathematical Organization Theory 5:2: 129-145.

Nowak, A & Latané, B. (1994) Simulating the emergence of social order from individual behaviour, in N. Gilbert and J. Doran, (eds.) (1994) - Simulating societies : the computer simulation of social phenomena, London [etc.] : UCL Press

Sun, R. (Ed.)(2006), Cognition and Multi-Agent Interaction: From Cognitive Modeling To Social Simulation, Cambridge, University Press

Troitzsch, K. (Ed.) (2005), Representing Social Reality, Pre-Proceedings of the Third Conference of the European Social Simulation Association, Koblenz, Verlag Dietmar Föllbach

Zeggelink, E., (1993), Strangers into Friends, The evolution of friendship networks using an individual oriented modeling approach, doctoral thesis, University of Groningen, Amsterdam, Thesis Publishers

Zoethout, K. (2006), Self-Organising Processes of Task Allocation: A Multi-Agent Simulation Study, Doctoral Thesis Groningen: University of Groningen, Faculty of Management and Organization, Ridderkerk, Labyrinth Publication

List of authors

Scientific Committee

Scientific Chair: Frédéric Amblard, *CNRS-IRIT, Toulouse University of Social Sciences, France*

Iqbal Adjali, Unilever Corporate Research
Luis Antunes, Porto University, Portugal
Arnaud Banos,University of Strasbourg, France
Olivier Barreteau, Cemagref, France
Alain Berro, CNRS-IRIT, Toulouse University of Social Sciences, France
Francois Bousquet, CIRAD, France
Rosaria Conte, Institute of Cognitive Sciences and Technologies, Roma, Italy
Gilles Daniel, ETH Zurich, Switzerland
Eric Daude, University of Rouen, France
Guillaume Deffuant, Cemagref, France
Julie Dugdale, Institut d'Informatique et Mathematiques Appliquees de Grenoble, France
Bruce Edmonds, Manchester Metropolitan University Business School, UK
Andreas Flache, ICS University of Groningen, The Netherlands
Nigel Gilbert, University of Surrey, Guildford, UK
Laszlo Gulyas, AITIA, Hungary
David Hales, University of Bologna, Italy
Kathrin Happe,LIADCEE, Halle, Germany
Richard Harrison, University of Texas, Dallas, USA
Dirk Helbing, University of Dresden, Germany
Cesareo Hernandez Iglesias, University of Valladolid, Spain
Wander Jager, University of Groningen, The Netherlands
Marco Janssen, Arizona State University, USA
Juergen Kluever, University of Essen, Germany
Christophe Le Page, CIRAD, France
Michael Macy, Cornell University, USA
Philippe Mathieu, University of Lille, France
Pierre Mazzega, CNRS OMP, Toulouse, France
Ugo Merlone, University of Torino, Italy
Scott Moss, Manchester Metropolitan University Business School, UK
Mario Paolucci, Institute of Cognitive Sciences and Technologies, Roma, Italy
Denis Phan, University of Rennes, France
Gary Polhill, The Macaulay Institute, UK
Matteo Richiardi, The LABORatorio R. Revelli, Turino, Italy
Camille Roth, University of Surrey, UK
Juliette Rouchier, CNRS-GREQAM, Marseille, France
Nicole Saam, Ludwig-Maximilians-Universitat Munchen, Germany
Stephane Sanchez, CNRS-IRIT, Toulouse University of Social Sciences, France
Alex Schmid, Savannah Simulation, Switzerland
Frank Schweitzer, ETH Zurich, Switzerland

Christophe Sibertin-Blanc, CNRS-IRIT, Toulouse University of Social Sciences, France
Jaime Simao Sichman, University of Sao Paulo, Brasil
Flaminio Squazzoni, University of Brescia, Italy
Pietro Terna, University of Torino, Italy
Guy Theraulaz, CNRS CRCA, Toulouse, France
Klaus Troitzsch, University of Koblenz-Landau, Germany
Gerard Weisbuch, ENS, Paris, France
Kees Zoethout, University of Groningen, The Netherlands

Organization Committee

Frédéric Amblard, CNRS-IRIT, Toulouse University of Social Sciences, France
Alain Berro, CNRS-IRIT, Toulouse University of Social Sciences, France
Michèle Cuesta, Toulouse University of Social Sciences
Véronique Desbats, IRIT
Pierre Mazzega, CNRS OMP, Toulouse, France
Stephane Sanchez, CNRS-IRIT, Toulouse University of Social Sciences, France
Christophe Sibertin-Blanc, CNRS-IRIT, Toulouse University of Social Sciences, France