

## **Une démarche expérimentale pour la simulation individus-centrée**

Guillaume Deffuant\*, Frédéric Amblard\*, Raphael Duboz\*\*, Eric  
Ramat\*\*

\*Laboratoire d'Ingénierie pour les Systèmes Complexes (LISC)  
Cemagref  
24, avenue des Landais BP 50085  
63172 Aubière Cedex  
France

\*\*Laboratoire d'Informatique du Littoral (LIL)  
Maison de la Recherche Blaise Pascal  
50, rue Ferdinand Buisson - BP 719  
62228 Calais Cedex  
France

email: [guillaume.deffuant@cemagref.fr](mailto:guillaume.deffuant@cemagref.fr)

### **Introduction**

Les interrogations sur le statut de la simulation sont souvent liées à la pertinence du lien entre le modèle et ce qu'il est censé représenter (Varenne 2001), et donc au crédit que l'on peut prêter aux résultats d'une simulation. Le problème est donc alors de rapprocher une simulation d'une expérimentation sur l'objet qu'elle représente, et de comparer les statuts de ces deux types d'expérimentations. Le remplacement d'essais de bombes nucléaires réelles par des simulations dans plusieurs pays par exemple pose légitimement ce type de questions.

Nous nous plaçons cependant dans une perspective sensiblement différente : nous nous interrogeons sur le statut de la simulation en tant qu'expérimentation directe sur le modèle computationnel. Dans cette perspective, les simulations ont pour objectif premier de faire progresser la connaissance du fonctionnement du modèle, avant de faire progresser la connaissance du phénomène représenté.

Une telle démarche peut probablement paraître choquante, au moins pour deux raisons. Tout d'abord, ne faisons-nous pas des modèles pour nous aider à comprendre des phénomènes que nous ne comprenons pas ? Alors quel intérêt d'y substituer des modèles que nous ne comprenons pas non plus (puisqu'il faut les étudier pour les comprendre) ? Mais, plus profondément, un modèle est totalement spécifié par son concepteur, donc les moindres mécanismes en sont connus. Comment se peut-il qu'il faille encore l'étudier pour le comprendre ?

Nous tentons de montrer que la démarche expérimentale dans l'objectif de mieux comprendre le fonctionnement de certains modèles, notamment certains modèles individus-centrés, est parfaitement fondée. En effet, la connaissance totale des mécanismes les plus intimes de ces modèles ne permet pas d'en comprendre le comportement global. Cette caractéristique, souvent associée à une certaine conception de la complexité, permet d'envisager pleinement le modèle comme objet d'expérimentation, au même titre qu'un phénomène naturel.

Nous exposons ensuite le projet SimExplorer, qui a pour objectif de faciliter la mise en œuvre concrète d'une démarche expérimentale sur des modèles individus-centrés.

## **Le modèle computationnel et les défis de la complexité**

### *Machine de Turing et complexité de Von Neumann*

L'un des séismes intellectuels du XXI<sup>ème</sup> siècle a été provoqué par le théorème de Gödel et son incarnation dans la machine de Turing. Cet événement nous intéresse particulièrement, car il fournit l'un des principes fondamentaux justifiant une démarche expérimentale sur des modèles computationnels. Ce principe est lié à une certaine forme de complexité, dont J.-P. Dupuy attribue la paternité à Von Neumann.

Résumons les principaux points de la lecture des événements faite par Dupuy : la machine de Turing a un fonctionnement dont les règles sont simples, et aisées à expliciter. Cependant, l'ensemble de ses comportements possibles<sup>1</sup> est infiniment plus difficile à expliciter, puisqu'ils échappent à toute formalisation récursive. La complexité au

---

<sup>1</sup> Notamment le fait qu'elle s'arrête ou pas en prenant certaines entrées

sens de Von Neumann est fortement inspirée par cette découverte. En effet, selon cette définition, une machine est complexe si l'ensemble de ses comportements possibles est infiniment plus difficile à caractériser que ses règles de fonctionnement. Ainsi, une machine très compliquée comme une centrale nucléaire, n'est pas complexe dans le sens de Von Neumann car l'ensemble de ses comportements est réduit : provoquer la réaction de fission contrôlée afin de produire de la chaleur.

La complexité, en ce sens précis, de certains modèles computationnels justifie leur statut d'objets d'expérimentations : leurs règles de fonctionnement, bien que totalement spécifiées par le programmeur, ne permettent pas de déduire leur comportement. Il est donc indispensable d'observer ce comportement par l'expérimentation, pour tenter éventuellement de produire une forme de théorie, plus compacte, de ce fonctionnement. Cet objet créé par l'homme est donc opaque à son propre créateur, qui doit appliquer une démarche scientifique pour comprendre sa propre création. Sa création devient pour lui équivalente à un phénomène naturel qu'il cherche à comprendre.

Cette situation prend l'allure d'un curieux retournement lorsqu'on l'observe à la lumière de la thèse de Hannah Arendt sur « l'Homo Faber ». Pour résumer grossièrement cette thèse, les sciences expérimentales (qui font la spécificité des sciences modernes) correspondent à une situation où l'homme ne comprend que ce qu'il fabrique. Il lui faut fabriquer la nature pour la comprendre. C'est ainsi que le laboratoire scientifique, en tant que structure, est le lieu de fabrication de phénomènes particuliers, contrôlés, qui peuvent par-là même être compris. C'est ainsi que les sciences fabriquent des modèles des phénomènes naturels, et les rendent alors accessibles à la compréhension humaine. L'idée centrale de cette thèse est que la découverte de règles de fonctionnement intime des phénomènes, nécessaire pour les fabriquer (voire les répliquer) permet d'accéder à leur compréhension et à la prédiction de leur comportement. Or, maintenant que l'homme fabrique des objets complexes au sens de Von Neumann, ce schéma est mis en défaut. En effet, la fabrication même de l'objet, et de ses règles de fonctionnement intimes, ne donne pas forcément accès à la compréhension de son comportement.

Cependant, ce comportement d'un système complexe artificiel peut faire l'objet d'une enquête scientifique classique, et ainsi donner lieu à l'élaboration d'un second modèle, qui donne à l'expérimentateur des clés plus accessibles, et permet notamment de faire des prédictions correctes du comportement du premier modèle. Les créations humaines dont les règles de fonctionnement les plus intimes sont totalement connues comme des modèles computationnels, peuvent donc devenir des objets d'investigation expérimentale. La simulation acquiert alors le statut d'expérimentation au même titre que l'expérimentation classique de laboratoire. Elle a simplement lieu dans un laboratoire « virtuel ».

Cette situation est particulièrement nette pour des modèles de phénomènes collectifs, c'est à dire fondés sur l'interaction de nombreux éléments.

### ***Complexité des phénomènes collectifs***

Une analogie, soulignée fortement par J.P. Dupuy, apparaît entre la complexité de Von Neumann, qui est « l'échappée » du modèle à son propre créateur, et la complexité depuis longtemps observée dans les phénomènes collectifs réels (en économie, en sociologie ou en écologie).

En économie, la fable des abeilles de Mandeville exprime parfaitement une sorte de déconnexion entre les mécanismes individuels et les mécanismes collectifs. On retrouve cette déconnexion dans la fameuse « main invisible » d'Adam Smith, ou sous sa forme moderne dans l'optimisation réalisée sur les lois du marché (par exemple chez Hayek). La figure ressemble à celle que nous avons identifiée dans la complexité de Von Neumann : même des hypothèses simplificatrices faites sur la connaissance intime des mécanismes de fonctionnement du système (ici le comportement d'individus rationnels) ne permettent pas de comprendre son comportement global. Bien sûr, le grand triomphe de l'économie classique, avec la théorie de l'équilibre général, est d'avoir construit un objet théorique qui permet de rendre compte des mécanismes constatés aux deux niveaux. Cette réussite est bien sûr remarquable, mais il n'est pas sûr qu'elle aplanisse totalement le fossé entre les deux niveaux. En effet, ce fossé est une hypothèse de la théorie elle-même, puisque les acteurs économiques sont supposés considérer

qu'ils n'ont pas d'influence individuelle sur les prix, alors que c'est bien la somme de ces influences individuelles qui fixe leur valeur. On garde donc, même dans les cas les plus favorables, une forme de déconnexion entre les mécanismes individuels, et les phénomènes collectifs qu'ils produisent.

La sociologie, dès sa fondation, se situe au cœur de cette complexité. En effet, personne ne va contester que les « faits sociaux » au sens de Durkheim sont des produits de l'activité humaine. Cependant, il est postulé qu'ils se situent à un niveau indépendant de celui des comportements individuels, ce qui justifie que leur étude fasse l'objet d'une démarche scientifique particulière, et justifie donc la création d'une discipline scientifique indépendante. Ici encore, une figure similaire à celle de la complexité de Von Neumann apparaît. La compréhension du phénomène n'est pas recherchée dans ses mécanismes les plus intimes (les comportements individuels), mais à un autre niveau (celui du phénomène collectif lui-même).

Nous nous garderons bien de pousser cette analogie jusqu'à en faire équivalence. Ici, nous nous distinguons de J.P. Dupuy qui semble parfois ne pas différencier les deux formes de complexité<sup>2</sup>. En effet, il nous semble que la définition de Von Neumann, s'appuyant sur le théorème de Gödel et la machine de Turing universelle se justifie par une difficulté intrinsèque à caractériser l'ensemble des comportements de cette machine (à cause du théorème sur l'incomplétude des systèmes axiomatiques de Gödel). Or, cette caractéristique ne se retrouve pas forcément dans des phénomènes collectifs complexes. Il se peut très bien que l'ensemble des phénomènes collectifs produits par certains types de comportements individuels soient faciles à caractériser, et à comprendre, à leur niveau, alors même que le lien direct entre ces phénomènes collectifs et les comportements individuels garde une certaine opacité. On peut le constater sur des exemples dans le domaine de la physique : on a pu très tôt établir des lois entre la pression, le volume et la température d'un gaz (donc au niveau collectif), tout en ayant des théories, satisfaisantes à l'époque, au niveau du comportement des particules, sans être à même d'établir un lien direct entre les deux niveaux. Il a fallu pour cela développer

---

<sup>2</sup> la complexité de Von Neumann et la complexité des phénomènes collectifs

des outils de mathématiques statistiques sophistiqués. Même dans les cas où la déconnexion entre les deux niveaux est plus profonde, il se peut que les comportements collectifs puissent être caractérisés de manière simple. Nous considérons donc que la complexité des phénomènes collectifs a sa spécificité, même si elle présente une analogie formelle avec celle de Von Neumann.

Cette spécificité a été remarquée depuis longtemps. Elle s'exprime par des formules vagues telles que « la somme des parties n'est pas équivalente au tout » ou encore les différentes définitions, souvent tout aussi vagues, de l'émergence. Peut-être touchons-nous ici simplement à des limites de la compréhension humaine, qui ne peut pas prendre en compte, en même temps le comportement d'un grand nombre d'éléments qui s'influencent mutuellement. On voit bien que ce peut être le cas même lorsque les individus du système ont un comportement très simple. Ce seraient donc ces limites qui exigeraient de rechercher des explications directement au niveau collectif, en manipulant un nombre de descripteurs globaux plus restreint. Ainsi, selon (Atlan, 1986), le problème de la connexion entre niveau d'organisation est peut-être un faux problème. Les divisions et classifications de la nature viennent de la nécessité d'une construction réductionniste liée à la méthode scientifique elle-même. A chaque niveau d'organisation nous inventons donc un vocabulaire, un paradigme, un formalisme qui créent le contexte dans lequel nous raisonnons. Dans cette vision des choses, chaque niveau d'organisations est « un point de vue » sur la réalité qui développe ses propres méthodes pour comprendre et éventuellement prédire la nature à son propre niveau, dans son propre langage. Pour synthétiser la thèse d'Atlan, le changement de niveaux se fait par changement de paradigme, la connexion entre niveaux étant réalisée par l'observateur mais n'ayant pas d'existence en soit. Il reste que la compréhension de ce niveau et de ses lois requiert une démarche particulière, d'identification de ces descripteurs et des lois qui les régissent.

Cependant, ce lien entre des niveaux différents peut parfois être plus profond que l'observation de niveaux de description différents. Dans certains cas en effet, le niveau global rétroagit sur les comportements individuels. Certains auteurs utilisent l'existence de cette rétroaction comme définition de la complexité. On constate évidemment l'existence de telles situations dans les phénomènes sociaux (existence

d'institutions qui modifient le comportement), mais aussi en physique (lorsque le comportement individuel de particules crée un champ qui agit sur le comportement individuel).

Quel que soit l'angle de vision adopté, il apparaît que les mécanismes intimes du système (donc ici de ses éléments) ne permettent pas de déduire une compréhension de son comportement global. Une approche expérimentale pour établir les lois de ce comportement global est donc justifiée.

## **Une démarche expérimentale pour les simulations individus centrées**

### ***La simulation individus-centrée comme accès à la complexité des phénomènes collectifs***

Les modèles de simulations individus-centrés sont au centre de ce débat sur la complexité et la vision de la simulation comme expérimentation. En effet, la caractéristique première de ces modèles est de se fonder sur une représentation explicite de l'ensemble des individus du système ainsi que de leurs interactions. Il s'agit donc d'une description des mécanismes au niveau des individus qui composent le système, par opposition à des modèles plus classiques, qui expriment des dynamiques (par équations différentielles par exemple) sur des variables agrégées (donc directement à un niveau global).

Le terme multi-agents qui peut parfois être compris comme synonyme à individus-centré, car certains modèles dits multi-agents (Doran, 1997 ; Drogoul et al., 1995 ; Drogoul et Ferber, 1994) implémentent en premier lieu des mécanismes individuels, leurs concepteurs tentant ensuite d'observer des régularités à un niveau collectif. Cependant, certains auteurs restreignant le terme multi-agents à des modèles dont les entités respectent certains critères plus ou moins bien définis (autonomie ou sociabilité par exemple), nous utiliserons de préférence le terme individus-centré, qui nous paraît être d'une acception plus large et plus claire.

Le prototype du modèle individus-centré est l'automate cellulaire, développé à l'origine par les physiciens, pour étudier les problèmes liés à l'agrégation (Toffoli et Margolus, 1987 ; Wolfram, 1986). Il

s'agit de l'illustration parfaite de la complexité des phénomènes collectifs : des automates extrêmement simples acquièrent des comportements globaux extrêmement riches, difficilement déductibles directement des dynamiques individuelles. Ici, les modèles sont étudiés pour leurs propriétés intrinsèques, au niveau collectif.

Beaucoup de modèles individus-centrés sont utilisés pour simuler des écosystèmes (Grimm 1999) ou des systèmes sociaux (Ferrand, 1999 ; Ballot et Weisbuch, 2000). Il est alors nécessaire de comparer le comportement global du modèle avec des données équivalentes dans les systèmes représentés. Cela ne fait que renforcer la nécessité de soigner le dispositif expérimental permettant de caractériser le comportement global observé dans les simulations.

On pourrait donc s'attendre à ce que la question du dispositif expérimental à mettre en œuvre dans le cadre de ces simulations fasse l'objet d'une réflexion poussée dans la communauté de modélisation individus-centrée. En fait, à notre connaissance, ce n'est pas vraiment le cas. Au contraire, certains constats (Grimm 1999) dénoncent le manque de temps et d'efforts passés à tester le modèle après son élaboration : on ne fait que quelques simulations qui sont sommairement comparées à quelques données.

Cependant, on peut identifier des particularités de l'expérimentation sur des modèles computationnels (comparée à l'expérimentation sur des systèmes physiques par exemple). Tout d'abord, le modèle computationnel offre l'avantage énorme de permettre facilement d'une part une automatisation des expérimentations et d'autre part la mesure de résultats qui ne subissent pas les biais de mesures classiques lors d'expérimentations sur des systèmes réels. Par ailleurs, les expérimentations sont faites dans l'objectif d'observer des régularités dans le comportement global du modèle, et de les lier à certaines autres variables, facilement interprétables, décrivant le modèle. Cet objectif particulier requiert de pouvoir exprimer les résultats globaux obtenus selon différents angles.

### ***Exemple de recherche de régularités globales dans un modèle individu centré***

Nous considérons l'exemple du modèle d'influence sociale développé dans le cadre du projet IMAGES (Deffuant, 2001), afin d'illustrer nos



propos. Nous en resterons à une description qualitative de ce modèle, qui est suffisante pour notre propos. En effet, nous nous focalisons sur la recherche d'un lien entre des valeurs de paramètres de la dynamique, et des caractéristiques globales de sa dynamique. Nous insistons sur la difficulté principale qui est de déterminer les variables agrégées pertinentes, et ensuite l'exploration de l'espace nécessaire pour établir les liens.

Décrivons donc rapidement ce modèle (voir Deffuant et al., 2002 pour les détails). Les individus sont caractérisés par une opinion continue, initialisée par une distribution uniforme entre  $-1$  et  $1$ , et un autre paramètre (continu lui aussi) de certitude (ou de conviction) en cette opinion. On suppose qu'aux extrêmes de l'axe d'opinion, un certain pourcentage  $p_e$  d'individus sont des extrémistes, c'est à dire qu'ils ont une certitude très forte (ils sont tous initialisés à  $u_e$ ), alors que le reste de la population est plutôt incertain (initialement, tous les individus ont la même incertitude  $U$ ). La dynamique met en œuvre des rencontres entre deux individus choisis aléatoirement dans toute la population. Les individus modifient leur opinion au cours de cette rencontre selon des règles mathématiques dont nous ne donnerons que les effets qualitatifs : tout d'abord une réelle modification des opinions n'a lieu que si elles ne sont pas trop éloignées l'une de l'autre (en tenant compte des paramètres d'incertitude) Par ailleurs, les individus convaincus sont plus influents que ceux qui ne le sont pas.

Nous nous intéressons à l'influence des extrémistes sur le reste de la population. Une première exploration aléatoire de l'espace des simulations fait apparaître trois types de dynamiques (voir figure 1). Ces dynamiques sont caractérisées par la convergence vers des groupes d'opinions homogènes et stables. Le premier type correspond à une influence faible des extrêmes, donc une population dont les groupes finaux sont en majorité hors des zones extrêmes définies à l'initialisation. Le second type présente une bipolarisation de la population vers les extrêmes. Enfin dans dernier type de dynamique, un seul des extrêmes attire la majorité de la population. En fait, le troisième type a été mis à jour très tard, car il correspond à des valeurs des paramètres un peu inhabituelles (très grande incertitude  $U$  de la majorité de la population).

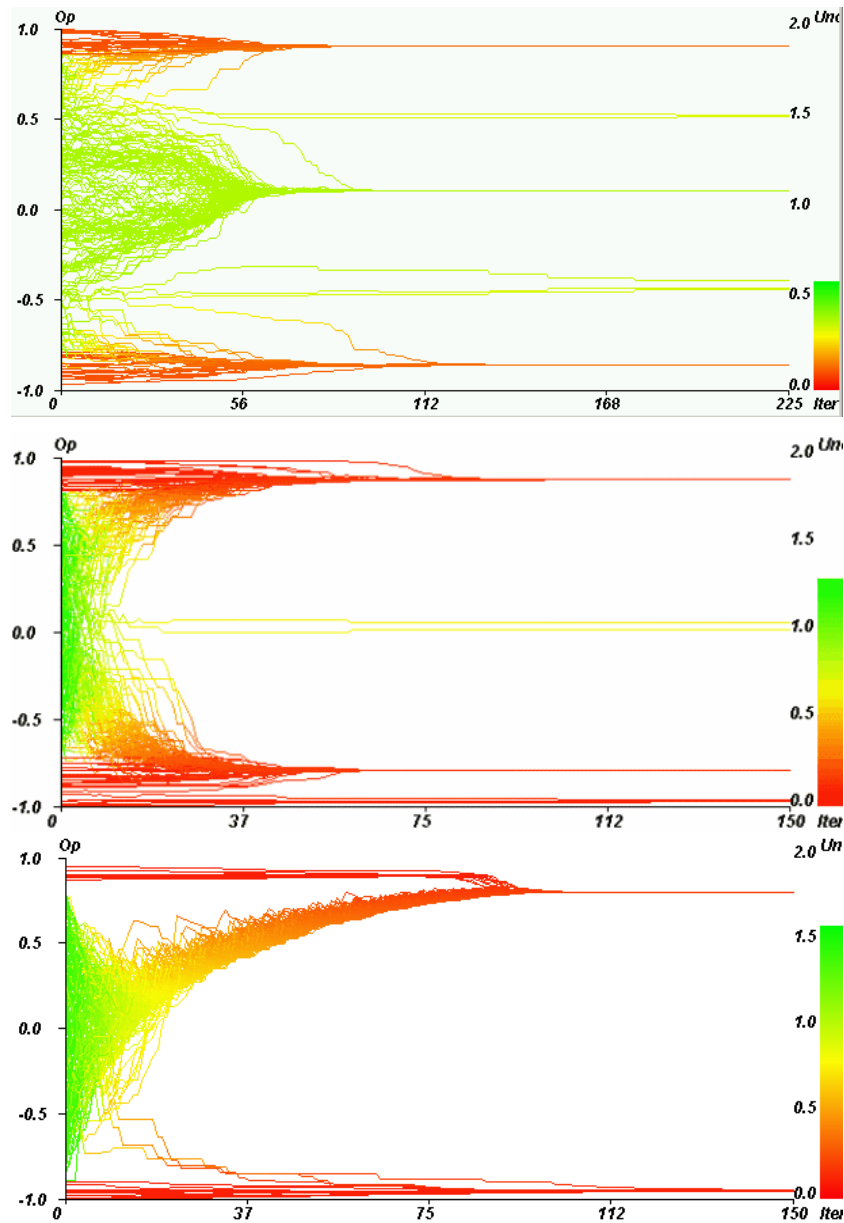


Figure 1: Les trois cas de convergence observés, de haut en bas, convergence centrale, convergence vers les deux extrêmes et convergence vers un seul extrême.

Nous voulons examiner la manière dont les différents types d'attracteurs se répartissent dans l'espace des paramètres. Il s'agit de prévoir le type de convergence en fonction des valeurs de ces paramètres (proportion d'extrémistes, incertitude des modérés notamment).

Nous exprimons donc tout d'abord les résultats de l'exploration à l'aide d'un indicateur de type de convergence (central, deux extrêmes

ou un extrême) noté  $y$ . Pour calculer l'indicateur  $y$ , nous considérons la distribution des opinions après convergence:

- Nous calculons les proportions  $p'_+$  et  $p'_-$  des agents initialement modérés qui sont devenus extrémistes positifs ou négatifs.
- L'indicateur est défini par:  $y = p'_+{}^2 + p'_-{}^2$ .

La valeur de cet indicateur indique le type de convergence (cf. table 1):

Valeur de $y$	0	0.5	1
Type de convergence	Centrale	deux extrêmes	Un seul extrême

*Table 1: Interprétation des valeurs de  $y$ . Les valeurs intermédiaires correspondent à des convergences intermédiaires.*

En utilisant cet indicateur, nous pouvons établir un lien entre des valeurs de paramètres  $U$  et  $p_e$  à la distribution des valeurs de l'indicateur  $y$  pour un ensemble de simulations faites à partir de ces valeurs. La figure 2 fait apparaître les moyennes de ces distributions de valeurs de  $y$  pour 50 répliques de simulations, ainsi que les écarts type. La figure 3 fait apparaître l'histogramme des valeurs de  $y$  pour les 50 répliques sur une ligne de la figure 2 (pour  $p_e$  constant). En fait, de telles cartes doivent être multipliées pour l'ensemble des valeurs des autres paramètres (notamment les paramètres de la dynamique).

Ce travail nous a permis d'identifier un ensemble de régularités dans le comportement global du système. Nous avons constaté que ces cartes présentent deux formes qualitatives différentes (selon que le nombre d'extrémistes est le même en chaque extrême ou pas). L'une de ces formes est présentée par la figure 2. Même un très petit déséquilibre entre les deux extrêmes provoque le passage d'une forme à l'autre. Ces formes typiques (voir figure 2) présentent des zones de l'espace dans lesquelles le comportement qualitatif du modèle est homogène. Dans certaines zones au contraire, on constate des changements brusques du fonctionnement qualitatif du modèle, ou encore des changements plus continus.

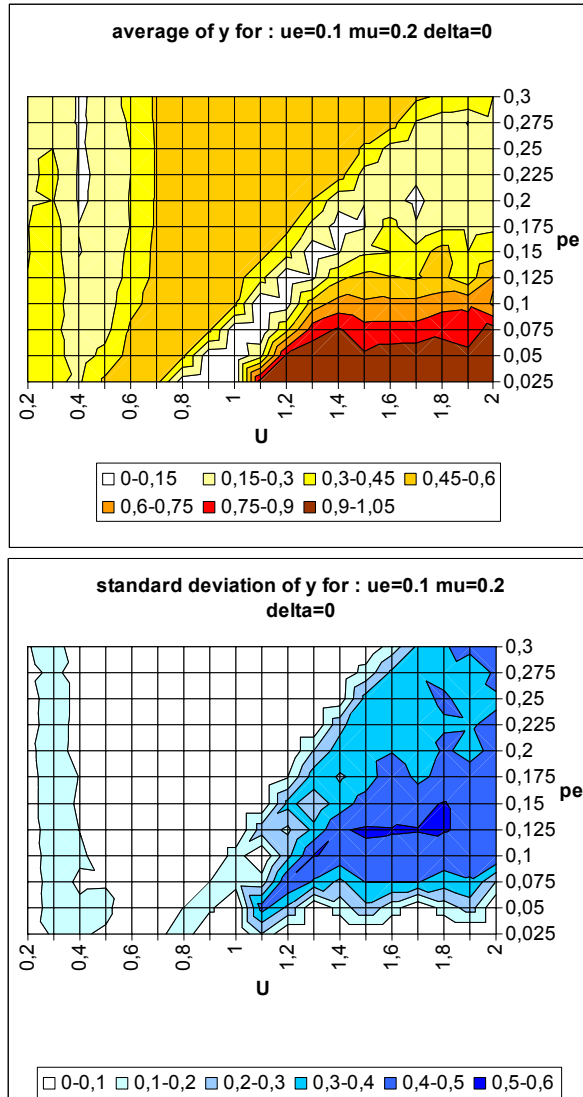


Figure 2: Formes typiques de la moyenne et de l'écart type de l'indicateur  $y$  (50 réplifications à chaque point du graphe) en fonction de l'incertitude des agents modérés ( $U$ ) et de la proportion globale d'extrémistes ( $p_e$ ). Sur le graphe représentant les moyennes de  $y$ , les zones jaunes et blanches à gauche correspondent à une convergence centrale, la zone orange au milieu à une convergence vers deux extrêmes et la partie marron en bas à droite à une convergence vers un seul extrême.

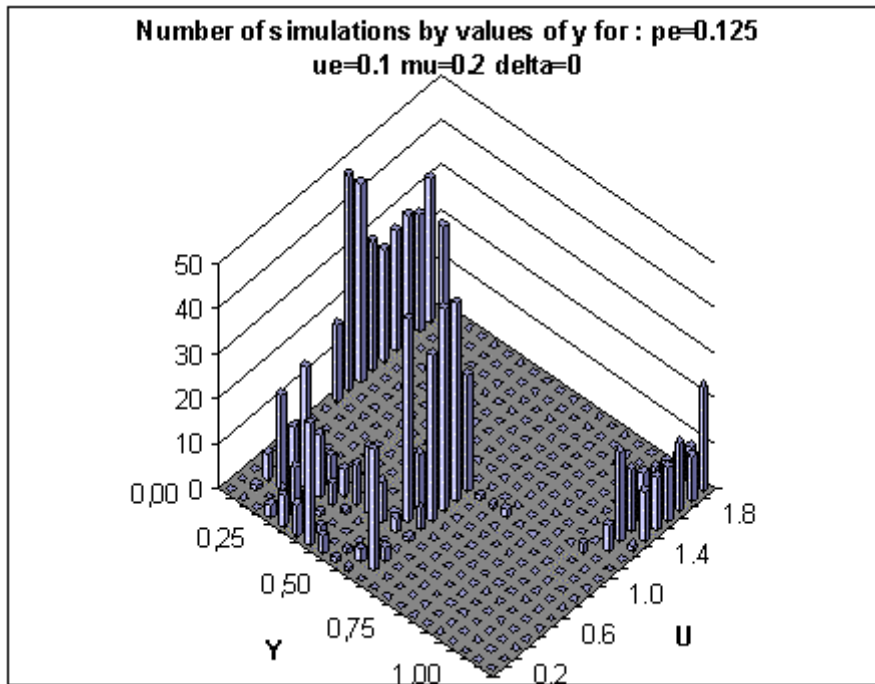


Figure 3: Histogrammes des valeurs de l'indicateur  $y$  pour  $p_e = 0.125$ . Pour  $U$  suffisamment grand ( $U > 1$ ) les mêmes paramètres conduisent suivant les répliques soit à un clustering central soit à un clustering vers un seul extrême ( $y$  proche de 0 ou proche de 1). Au milieu,  $U$  ( $0.5 < U < 1$ ) nous avons uniquement un clustering vers les deux extrêmes ( $y$  proche de 0.5). Pour  $U$  faible ( $U < 0.5$ ), l'histogramme montre différentes convergences qui sont intermédiaires entre une convergence centrale et une convergence vers deux extrêmes (avec des proportions différentes d'agents qui ont rejoint les extrêmes).

A partir de l'analyse de ces figures et de leurs régularités, il devient possible de proposer une théorie explicative du fonctionnement global du modèle (ce qui est esquissé dans Deffuant et al., 2002). On examine alors en détails l'évolution de certains exemples typiques de simulations dans les grandes zones identifiées, et on introduit des approximations de type champ local moyen.

Rien que pour ce petit modèle, on constate la nécessité de pratiquer un grand nombre d'expérimentations pour établir correctement une certaine compréhension de son comportement global. Insistons sur le fait que cette compréhension est nécessaire comme préalable à la question plus classique de l'adéquation du modèle à des phénomènes particuliers (sociaux en l'occurrence). On comprend sur cet exemple

l'intérêt d'un outil permettant de rendre ce travail plus systématique, et de capitaliser sur les expériences passées.

### **Le projet SimExplorer : vers un outil générique d'expérimentation sur des modèles individus-centrés**

Le projet SimExplorer est porté actuellement par le LISC (Cemagref), conjointement à d'autres équipes notamment l'Université du Littoral et le groupe de réflexion autour de MIMOSA (animé par J.P. Muller du CIRAD). Ce projet s'appuie sur une démarche expérimentale de l'exploration des modèles individus-centrés. L'objectif est de faciliter la mise en œuvre d'une telle démarche, afin de la rendre plus rigoureuse et plus systématique. Il s'agit aussi de tenir compte des particularités des modèles individus-centrés pour développer des types d'expérimentations qui leur sont propres. Nous présentons tout d'abord le noyau du projet SimExplorer, puis des projets de modules d'expérimentations plus particulières.

#### ***Le noyau du projet SimExplorer***

Le noyau de SimExplorer est actuellement en cours de développement au LISC. Ce logiciel doit à terme s'adapter facilement à différents modèles et permettre de lui appliquer une démarche expérimentale. L'outil doit intégrer la possibilité de créer des plans d'expérience classiques (Kleijnen, 1987 ; Kleijnen et Groenendaal, 1992), afin de parcourir au mieux l'espace des paramètres du modèle. Il doit permettre de lancer les simulations, de récupérer les résultats et d'en assurer certains traitements.

La définition des structures informatiques décrivant le modèle et les simulations prend en compte la spécificité des modèles individus-centrés. En effet, dans la structure descriptive du modèle, on distingue :

- la population des individus du modèle individus-centré, définis par des objets et leurs attributs,
- les variables régissant la dynamique du modèle, c'est à dire celles des individus, ainsi que des leurs interactions,
- des variables de scénarios, qui définissent des évènements venant de l'extérieur, à certains pas de temps de la simulation (par exemple

les élagages pour un modèle de peuplement forestier, ou les pêches pour un modèles de poissons).

La structure descriptive d'une simulation (une expérience) comprend :

- L'ensemble des conditions initiales sur les valeurs des attributs définissant les individus. Souvent, ce sont les valeurs de paramètres de distributions statistiques qui définissent ces conditions initiales. Parfois ce sont des données qui sont issues d'un fichier de mesures.
- Les valeurs des paramètres de la dynamique, y compris le temps de simulation ou ses conditions d'arrêt.
- Un ensemble de valeurs des paramètres de scénario.

Les plans d'expériences peuvent porter sur l'ensemble des paramètres tous types confondus ou seulement sur l'une des catégories. Par ailleurs, des contraintes permettant de ne pas tester certaines combinaisons de paramètres peuvent être directement précisées dans l'interface.

L'utilisateur peut spécifier les différents types d'agrégation qu'il désire effectuer (et à quels pas de temps) sur les attributs de la population. Ces résultats peuvent être stockés sous différentes formes.

Une fois le plan d'expérience et les types de sorties souhaitées établis, SimExplorer lance les différentes expérimentations avec l'ensemble des spécifications de conditions initiales, de paramètres de la dynamique et de scénario, ainsi que les spécifications sur les sorties souhaitées. Il est prévu que l'outil gère également le lancement de ces simulations sur plusieurs machines en parallèle.

Reste ensuite à établir un lien entre les valeurs des paramètres et certains indicateurs agrégés pertinents. Une grande part du travail consiste à déterminer ces indicateurs pertinents et les liens statistiques entre leurs valeurs. Les analyses de sensibilité, peuvent être obtenues en faisant des différences sur les variables agrégées pour différentes valeurs de paramètres de simulation. Des outils conviviaux offrant le moyen de définir facilement ces variables agrégées sont donc d'un grand secours.

L'un des enjeux importants pour SimExplorer est de s'adapter facilement à de nombreux modèles. Ceci permettrait une diffusion

plus large de la pratique des simulations dans une démarche pleinement expérimentale.

En outre, un outil comme celui-ci ouvre la porte à des recherches sur des protocoles expérimentaux particuliers, adaptés à certains types de modèles. Les paragraphes suivants donnent quelques exemples que nous prévoyons d'investiguer.

### ***Expérimentations sur des distributions initiales et des structures spatiales***

La particularité principale des modèles individus-centrés est de représenter explicitement l'ensemble des individus du modèle. Une expérimentation requiert donc de spécifier les valeurs initiales des attributs de ces individus. Souvent cette initialisation est faite par des tirages selon des distributions de probabilité. Ces hypothèses d'initialisation peuvent avoir des conséquences importantes sur le comportement ultérieur du modèle. Il est donc important d'évaluer cette influence avec soin lorsqu'on cherche à comprendre un modèle. Nous prévoyons d'inclure dans SimExplorer la possibilité d'utiliser des distributions classiques pour initialiser les modèles.

Cependant, il est également important d'évaluer l'influence du nombre d'agents sur le comportement du modèle, notamment lorsque l'initialisation des attributs est faite par des tirages selon une distribution. En effet, plus le nombre d'individus est faible, plus la variabilité est importante, ce qui rend les tendances et les régularités plus difficiles à établir. Il peut donc être intéressant de tester les modèles avec de grands nombres d'individus afin de limiter cette variabilité. Il est donc nécessaire d'inclure la taille de la population comme paramètre à tester dans les expériences.

Par ailleurs, lorsqu'on utilise des distributions issues de données de relevés (pour des peuplements forestiers par exemple), l'utilisation d'indicateurs agrégés caractérisant la distribution spatiale peuvent être pertinents pour établir des régularités entre la distribution et la dynamique elle-même (Goreaud, 2000). De tels indicateurs peuvent être utiles également pour des initialisations par des distributions de probabilité artificielles. Il est donc intéressant d'intégrer de tels outils dans SimExplorer.



### *Expérimentations sur les structures d'interactions*

Nous considérons maintenant des modèles individus-centrés particuliers, fréquents en modélisation sociale, dans lesquels les interactions entre les individus sont définies par un graphe (le réseau social). Nous supposons ici que ce graphe est statique. Nous supposons aussi, bien sûr que les individus sont définis par un ensemble d'attributs dont les valeurs sont initialisées en début de simulation, et évoluent ensuite, notamment à cause des interactions entre les individus reliés par le graphe.

Un des problèmes intéressants posé pour ce type de modèle est d'évaluer l'influence de propriétés générales du graphe sur la dynamique globale du modèle. Par exemple, on peut s'intéresser à la formation de groupes homogènes d'agents, et à l'influence de la structure du graphe sur la formation et l'évolution de tels groupes.

Nous proposons différents tests à réaliser pour caractériser l'influence de la forme de la structure d'interactions sur le comportement du modèle. Nous adoptons la métaphore d'un graphe sur lequel on pose une distribution, chaque nœud prenant une valeur particulière de cette distribution.

Nous envisageons d'intégrer à SimExplorer un module permettant de tester des variations sur les graphes d'interactions afin de tester l'influence de leur structure sur la dynamique. Pour cela, différentes stratégies d'évolution du graphe sont possibles. L'utilisation d'indices caractéristiques issus de la sociologie des réseaux (Wasserman et Faust, 1994), de même que les indicateurs de distribution spatiale, fournit des indicateurs agrégés qu'il peut être intéressant de d'examiner conjointement aux caractéristiques de la dynamique. Il est donc important de pouvoir disposer de tels outils, qui peuvent grandement améliorer la compréhension du modèle au niveau global.

L'incorporation de telles méthodes dans SimExplorer permettraient de les utiliser systématiquement pour différents modèles. Cet outil devrait donc nous permettre d'améliorer notre démarche d'expérimentation sur les modèles individus-centrés, en la rendant plus rigoureuse et plus systématique.

## **Conclusion**

Nous défendons une approche expérimentale de la simulation individus-centrée. Nous justifions cette position par la complexité des modèles computationnels, en particulier à cause de la déconnexion entre dynamiques individuelles et phénomènes collectifs. En effet, souvent la connaissance fine des dynamiques individuelles ne permet pas de déduire directement le comportement global du modèle. Une démarche expérimentale est donc nécessaire afin d'établir les liens entre des variables agrégées décrivant le modèle globalement. A partir de la constatation expérimentale de tels liens, des modèles théoriques au niveau global peuvent être proposés. Cependant, la littérature fait plutôt apparaître un déficit de rigueur et de systématisation dans les expérimentations numériques sur les modèles individus-centrés.

Le projet SimExplorer vise à faciliter et améliorer cette démarche expérimentale, tout en tenant compte des spécificités des modèles individus-centrés. Il s'agit de développer une interface de mise en œuvre d'expérimentations sur des modèles individus-centrés. Cette interface proposera un ensemble de fonctions permettant de définir un plan d'expériences sur le modèle, et les variables agrégées à observer. Nous prévoyons aussi le développement de modules spécifiques permettant d'explorer l'influence de la répartition spatiale des individus, ou encore l'influence de la structurer du graphe d'interactions entre les individus. Nous souhaitons la rendre adaptable à différents modèles le plus facilement possible, et ainsi en faciliter l'utilisation et la diffusion.

Nous pensons que cet outil permettra de faire progresser la réflexion sur les méthodes de traitement systématiques des résultats. On pense bien sûr aux méthodes de fouilles de données qui offrent des perspectives séduisantes pour déterminer des régularités dans ces résultats. En outre, la démarche d'expérimentation prend son sens dans une itération entre la définition du plan d'expérience et l'exploitation des résultats. L'automatisation partielle de ce processus itératif est un défi passionnant que cette démarche permettra d'aborder dans le futur.

## Remerciements

L'ensemble de l'équipe du LISC contribue à la définition et au développement de SimExplorer. Nous tenons donc à mentionner la participation active à ce projet de I. Alvarez, S. Bernard, M. Edwards, T. Faure, F. Goreaud, D.Hill, S. Huet et J.Truffot.

## Références

- Arendt, H. (1958). *Condition de l'homme moderne*. Calman-Lévy.
- Ballot, G. et Weisbuch, G. (2000). *Applications of Simulation to Social Sciences*. Hermès.
- Bousquet, F., Cambier, C., Mullon, C., Morand, P., Quensière, J. et Pavé, A. (1993). Simulating the interaction between a society and a renewable resource. *Journal of Biological Systems*, 1, 199-214.
- Deffuant, G. (2001). *Final report of project FAIR 3 CT 2092. Improving Agri-environmental Policies: A Simulation Approach to the Cognitive Properties of Farmers and Institutions*. <http://wwwlisc.clermont.cemagref.fr/ImagesProject/default.asp>
- Deffuant, G., Amblard, F., Weisbuch, G. et Faure, T. (2002). How can extremism prevail? A study based on the relative agreement interaction model. *Journal of Artificial Societies and Social Simulation*, 5, <<http://jasss.soc.surrey.ac.uk/5/4/1.html>>.
- Doran, J.E. (1997). From computer simulation to artificial societies. *Transactions of the Society for Computer Simulation International*, 14, 69-78.
- Drogoul, A., Corbara, B. et Fresneau, D. - Manta: new experimental results on the emergence of (artificial) ant societies. In N.Gilbert et R.Conte (eds.), *Artificial Societies: The Computer Simulation of Social Life*, UCL Press, 1995.
- Drogoul, A. et Ferber, J. - Multi-agent simulation as a tool for studying emergent processes in societies. In J.E. Doran et N. Gilbert (eds.), *Simulating Societies: The Computer Simulation of Social Phenomena*, UCL Press, 127-142, 1994.
- Dupuy, J.-P. (1992). *Introduction aux sciences sociales, Logiques des phénomènes collectifs*. Ellipses.
- Dupuy, J.-P. (1994). *Aux origines des sciences cognitives*. La Découverte.

- Ferrand, N. (1999). *Modèles et systèmes multi-agents pour la gestion de l'environnement et des territoires*. Cemagref Editions.
- Goreaud F. (2000). *Apports de l'analyse de la structure spatiale en forêt tempérée à l'étude et la modélisation des peuplements complexes*, Thèse de doctorat ENGREF.
- Grimm, V. (1999). Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future? *Ecological Modelling*, 115, 129-148.
- Kleijnen, J.P.C. (1987). *Statistical tools for simulation practitioners*. Marcel Dekker Inc. Pub.
- Kleijnen, J.P.C. et Groenendaal W.V. (1992). *Simulation a Statistical Perspective*. Wiley.
- Legay, J.-M. (1997). *L'expérience et le modèle, un discours sur la méthode*. INRA éditions.
- Lorrain and White (1971). Structural equivalence of individuals in social networks. *Journal of Mathematical Sociology*, 1, 49-80.
- Toffoli, T. et Margolus, N. (1987). *Cellular Automata Machines*. MIT Press.
- Varenne, F. (2001). What does a computer simulation prove? The case of plant modelling at CIRAD (France). In Proceedings of the ESS'01 conference, Marseille (France), 549-554.
- Wolfram, S. (1986). *Theory and Applications of Cellular Automata*. World Scientific.