

---

# Certaines propriétés d'un réseau social facilitent la propagation de l'extrémisme

## Etude du modèle d'influence proportionnelle à l'accord relatif sur des réseaux Small-World

*Frédéric Amblard,  
Guillaume Deffuant*

### 1. Introduction

Les travaux sur l'influence sociale, particulièrement ceux qui proviennent des sciences sociales, soulignent l'importance de la structuration sociale, et en particulier la forme des réseaux sociaux, de la population considérée. Pour étudier cette hypothèse dans le cadre particulier du modèle d'influence proportionnelle à l'accord relatif (Deffuant et al., 2002), il est important de générer des structures réalistes de réseaux sociaux. Des données empiriques concernant la topologie de ces réseaux étant particulièrement difficiles à obtenir, nous pouvons nous appuyer sur des travaux récents de modélisation qui proposent différents modèles de réseaux, en particulier les modèles de type small-world. Le but de cet article est d'étudier l'influence de la structure de ces modèles de réseaux sur la dynamique des opinions proposée dans (Deffuant et al., 2002).

Dans ce dernier article, nous avons proposé un modèle de simulation individus-centré de dynamique des opinions, le modèle d'influence proportionnelle à l'accord relatif (ou modèle RA pour Relative Agreement). Ce modèle peut être considéré comme une variante du modèle de confiance limitée (Deffuant et al., 2001; Weisbuch et al. 2002a, 2002b; Hegselmann et Krause, 2002) dans lequel un individu en influence un autre seulement si leurs opinions sont suffisamment proches (ou, pour reformuler, si la distance entre leurs opinions est inférieure à un certain seuil). Dans le modèle « classique » de l'influence limitée, l'influence varie linéairement en

fonction de la distance entre les opinions, jusqu'au seuil en question, où l'influence cesse brutalement.

Le modèle d'influence proportionnelle à l'accord relatif (modèle RA) s'appuie sur une représentation sous forme de segments d'opinions, définis par une opinion moyenne et une incertitude associée ; l'opinion et l'incertitude étant deux nombres réels, le segment d'opinion est donc simplement [*opinion moyenne* – *incertitude* ; *opinion moyenne* + *incertitude*]. La dynamique d'influence du modèle RA prend alors en compte le recouvrement entre les segments d'opinions des deux individus en interaction et évite la discontinuité présente dans le modèle de confiance limitée (voir plus bas). De plus, une des caractéristiques saillantes de ce modèle est que plus un agent est certain, plus son segment est réduit, plus il est influent ou convaincant au cours de l'interaction.

Nous modélisons les extrémistes dans cette population comme des agents qui ont une incertitude très faible et une opinion située aux extrémités de la distribution d'opinions initiale. Lorsque l'on introduit des extrémistes initialement dans la population et que chaque paire d'individus peut entrer en interaction, les simulations font apparaître trois types de convergence: la "convergence centrale" pour laquelle les extrémistes attirent seulement une partie très limitée de la population qui était initialement proche des extrêmes ; la « convergence vers les deux extrêmes » pour laquelle la population est divisée en deux groupes, chacun convergeant vers un des deux extrêmes ; « la convergence vers un seul extrême » pour laquelle l'ensemble de la population converge vers un seul des deux extrêmes. Ce dernier type de convergence, inattendu de prime abord, apparaît même lorsque le nombre initial d'extrémistes de chaque bord est le même. Dans cet article, nous étudierons l'influence de certaines structures de réseau social sur l'occurrence de la convergence vers un seul extrême.

Dans le premier paragraphe, nous présentons brièvement le modèle RA avec extrémistes et les principaux résultats dans le cas d'un réseau social complet. Nous exposerons alors le cas où le réseau social est une grille régulière avec un voisinage de Moore. La convergence vers un seul extrême n'apparaît pas dans ce cas<sup>1</sup>. Nous explorons alors le comportement du

---

<sup>1</sup> En fait, nos travaux récents montrent que lorsque le nombre d'extrémistes est très faible (plus faible que les pourcentages explorés ici), pour certaines valeurs de l'incertitude des modérés, une forme particulière de convergence vers un seul extrême peut apparaître. Nous la négligeons dans ce papier car elle est traitée en détails dans (Weisbuch et al. 2005).

modèle sur des réseaux Small-World dont le modèle nous permet de modifier à la fois la connectivité moyenne du réseau social et la régularité de ses connexions. Les simulations montrent qu'il existe un niveau critique de connectivité au-delà duquel la convergence vers un seul extrême apparaît, et que la régularité des connexions tend à défavoriser ce comportement du modèle.

## **2. Le modèle d'influence proportionnelle à l'accord relatif (RA)**

### **2.a Définition**

Soit une population de  $N$  individus. Chaque individu  $i$  est caractérisé par deux variables, son opinion  $x_i$  et son incertitude  $u_i$ , chacune étant un nombre réel. Nous appelons segment  $s_i = [x_i - u_i, x_i + u_i]$  sur l'axe des opinions, le segment d'opinion de l'individu  $i$ . Dans la suite, nous initialisons les opinions de la population à partir d'une distribution uniforme entre  $-1$  et  $+1$ .

Des interactions aléatoires par paire d'individus sont simulées, au cours desquelles les individus peuvent modifier leur segment d'opinion. Dans le modèle de « confiance limitée » un individu  $j$  peut influencer un agent  $i$  seulement si son opinion est comprise dans le segment  $s_i$ . Nous utilisons ici le modèle RA qui conduit à une variation continue de l'influence sur l'axe des opinions et où l'intensité de l'influence de  $j$  sur  $i$  dépend de l'incertitude  $u_j$  de l'agent  $j$ .

Plus précisément, si l'on considère les segments d'opinions  $s_i = [x_i - u_i, x_i + u_i]$  et  $s_j = [x_j - u_j, x_j + u_j]$ , nous définissons l'accord de l'agent  $i$  avec  $j$  (l'accord n'est pas symétrique) comme le recouvrement de  $s_i$  et  $s_j$ , moins la partie non-recouverte de  $s_i$ . Le recouvrement  $h_{ij}$  est donc :

$$h_{ij} = \min(x_i + u_i, x_j + u_j) - \max(x_i - u_i, x_j - u_j)$$

La largeur non-recouverte est donnée par :

$$2.u_i - h_{ij}$$

L'accord est alors calculé comme le recouvrement moins la partie non-recouverte :

$$h_{ij} - (2.u_i - h_{ij}) = 2 \cdot (h_{ij} - u_i)$$

L'accord relatif est l'accord calculé divisé par la longueur du segment  $s_i$ , soit :

$$\frac{2 \cdot (h_{ij} - u_i)}{2 \cdot u_i} = \frac{h_{ij}}{u_i} - 1$$

Si  $h_{ij} > u_i$ , alors les modifications de  $x_j$  et  $u_j$  suite à l'interaction avec  $i$  sont multipliées par l'accord relatif :

$$x_j := x_j + \mu \cdot \left( \frac{h_{ij}}{u_i} - 1 \right) \cdot (x_i - x_j)$$

$$u_j := u_j + \mu \cdot \left( \frac{h_{ij}}{u_i} - 1 \right) \cdot (u_i - u_j)$$

Où  $\mu$  est la vitesse de la dynamique.

Si  $h_{ij} \leq u_i$ , il n'y a pas d'influence de  $i$  sur  $j$ .

Les principales propriétés du modèle RA sont d'une part qu'au cours des interactions, les agents s'influencent à la fois en opinion et en incertitude, d'autre part, que l'influence n'est pas symétrique quand les agents ont des incertitudes différentes : les agents « certains » (faible incertitude) sont plus influents. L'influence (les modifications de  $x_j$  et  $u_j$ ) varie de manière continue lorsque  $x_j$ ,  $u_j$ ,  $x_i$  et  $u_i$  varient de manière continue.

### 2.b Ajout d'extrémistes

Nous introduisons maintenant des extrémistes dans la population en supposant que ces agents ont des opinions situées aux extrémités de la distribution d'opinion et qu'ils ont une incertitude moindre que les individus non-extrémistes (ou modérés). Nous incluons dès lors deux variables supplémentaires dans notre modèle, l'incertitude initiale des extrémistes  $u_e$  et l'incertitude des modérés de la population  $U$  que nous posons comme supérieure à  $u_e$ .

Nous définissons également  $p_e$  comme la proportion globale d'extrémistes dans la population.  $p_+$  et  $p_-$  sont respectivement les proportions d'extrémistes situés aux extrémités positives et négatives de l'axe des opinions.

La différence relative entre les proportions d'extrémistes positives et négatifs est notée  $\delta$ :

$$\delta = \frac{|p_+ - p_-|}{p_+ + p_-}$$

En pratique, nous tirons tout d'abord aléatoirement les opinions de  $(1-p_e)N$  agents de la population à partir d'une distribution uniforme entre  $-1$  et  $+1$ . Nous initialisons alors  $Np_+$  agents à  $+1$  et  $Np_-$  agents à  $-1$ , de plus, nous les initialisons avec l'incertitude  $u_e$ , et les premiers (modérés) avec l'incertitude  $U$ .

### **2.c Résultats dans le cas d'un graphe complet**

#### *Trois types de convergence*

Ce modèle très simplifié de l'extrémisme donne naissance à trois régimes dynamiques en fonction des paramètres : la "convergence centrale" pour laquelle les extrémistes ont une faible influence sur le reste de la population; la "convergence double extrême" pour laquelle l'ensemble de la population devient extrémiste avec un nombre relativement égal d'extrémistes à chaque extrémité ; et la « convergence vers un seul extrême » où l'ensemble de la population est attiré vers une seule des deux extrémités

Les figures suivantes, obtenues à partir de simulations, décrivent ces trois régimes. L'axe des abscisses représente le temps (en nombre d'itérations), l'axe des ordonnées représente les opinions, le niveau de gris représente l'incertitude (faible pour gris clair, forte pour gris foncé). Chaque trajectoire permet de suivre l'évolution de l'opinion et de l'incertitude d'un agent de la population. Les paramètres communs aux quatre simulations sont  $\mu = 0.5$ ,  $\delta = 0$ ,  $u_e = 0.1$ ,  $N = 200$ . Le paramètre d'incertitude de l'ensemble de la population  $U$  augmente de la Fig.1(a) à la Fig.1(d).

La Fig. 1(a), obtenue pour  $U = 0.4$ , illustre la convergence centrale. Dans ce cas, seule une partie marginale de la population initialement modérée devient extrémiste (4%). La Fig. 1(b), obtenue pour  $U = 1.2$  est un exemple de bipolarisation. Dans ce cas, les agents modérés sont attirés par l'une des extrémités en fonction de leur position initiale. La Fig. 1(c) obtenue pour  $U = 1.4$  illustre le cas de convergence vers un extrême. Dans ce cas, la majorité de la population est attirée par l'une seulement des deux extrémités. Ce comportement est observé même lorsque le nombre initial d'extrémistes est

exactement le même des deux cotés. Cependant, pour un autre exemple tiré à partir de la même distribution initiale, les autres paramètres étant les mêmes, on peut observer des cas de convergence centrale (cf. Fig. 1(d)). La sensibilité aux conditions initiales et à la sélection aléatoire est une caractéristique générale du modèle dans les zones de transitions.

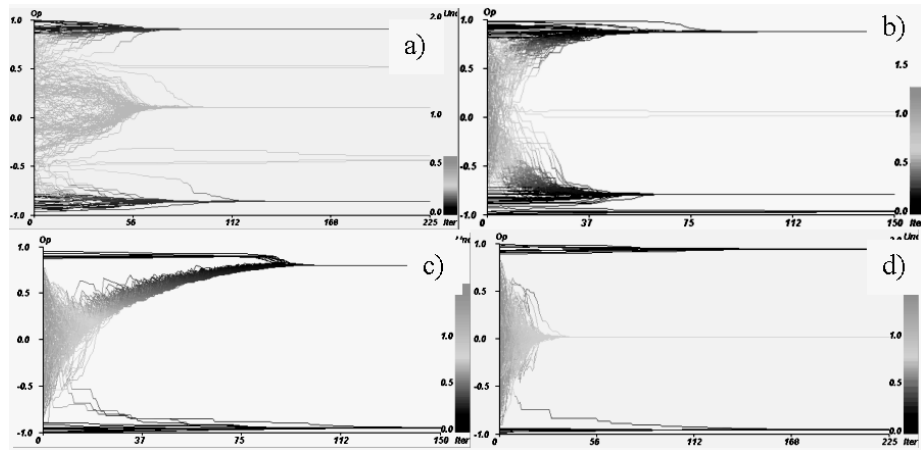


Fig. 1. (a) Exemple de convergence centrale. Abscisses: itérations. Ordonnées: opinions. Niveaux de gris: incertitudes.  $p_e = 0.2$ ,  $U = 0.4$ ,  $\mu = 0.5$ ,  $\delta = 0$ ,  $u_e = 0.1$ ,  $N = 200$ . La majorité (96%) des agents modérés ne sont pas attirés par les extrêmes. (b) Exemple de convergence double extrême.  $p_e = 0.25$ ,  $U = 1.2$ ,  $\mu = 0.5$ ,  $\delta = 0$ ,  $u_e = 0.1$ ,  $N = 200$ . Les agents initialement modérés se séparent et deviennent extrémistes (43% du côté positif, 56% du côté négatif). (c) Exemple de convergence simple extrême.  $p_e = 0.1$ ,  $U = 1.4$ ,  $\mu = 0.5$ ,  $\delta = 0$ ,  $u_e = 0.1$ ,  $N = 200$ . La majorité (98.33%) des agents initialement modérés est attirée par l'extrême positif. (d) Convergence centrale pour les mêmes paramètres qu'en (c), la majorité reste au centre (seul un agent rejoint l'extrême négatif).

L'instabilité de l'attracteur est confirmée par l'analyse en équation maîtresse (Faure et al., 2003). Pour une distribution initiale parfaitement uniforme des opinions, l'équation maîtresse (qui est déterministe) fait apparaître un attracteur symétrique (convergence centrale ou convergence double) dans la région de la convergence vers un seul extrême pour le modèle multi-agents. Mais une légère asymétrie dans la distribution initiale change la convergence centrale en convergence vers un seul extrême (dont la taille est déterminée par l'asymétrie).

La représentation des incertitudes montre que dans chacun des trois cas, le regroupement a lieu non seulement sur les opinions mais également sur l'incertitude.

*Résultats généraux de l'exploration de l'espace des paramètres*

Indicateur du type de convergence

Nous avons réalisé l'exploration de l'espace des paramètres en nous appuyant sur un indicateur de convergence, noté  $y$ . Pour calculer cet indicateur  $y$ , nous considérons les opinions de la population à convergence, puis :

- Nous déterminons les proportions  $p'_+$  et  $p'_-$  des agents initialement modérés qui deviennent extrémistes (respectivement positifs et négatifs).
- L'indicateur est alors :  $y = p'^2_+ + p'^2_-$ .

La valeur de cet indicateur nous permet de catégoriser les simulations en fonction du type de convergence :

- Si aucun des agents modérés n'est devenu extrémiste, alors  $p'_+$  et  $p'_-$  sont nuls et  $y = 0$ .
- Si la moitié des modérés initiaux convergent vers l'extrême positif et l'autre moitié vers l'extrême négatif, alors  $p'_+ = 0.5$  et  $p'_- = 0.5$ , d'où  $y = 0.5$ .
- Si tous les agents modérés vont uniquement vers une extrémité (par exemple positive), nous avons  $p'_+ = 1$  et  $p'_- = 0$ , d'où  $y = 1$ .

Les valeurs intermédiaires de l'indicateur correspondent à des situations intermédiaires.

Formes typiques de  $y$

L'exploration du modèle peut être représentée de manière satisfaisante sous la forme de variations de  $y$  en fonction de  $U$  et  $p_e$ . Cette représentation conduit à une forme typique des  $y$  moyens pour  $\delta = 0$ , et une autre pour  $\delta > 0$ . Il y a une différence significative entre les cas où la proportion d'extrémistes positifs et négatifs est exactement la même et quand elle est légèrement différente. Quand les autres paramètres  $u_e$  et  $\mu$  sont modifiés, la forme globale reste la même : seules les positions des frontières entre les zones de convergence diffèrent.

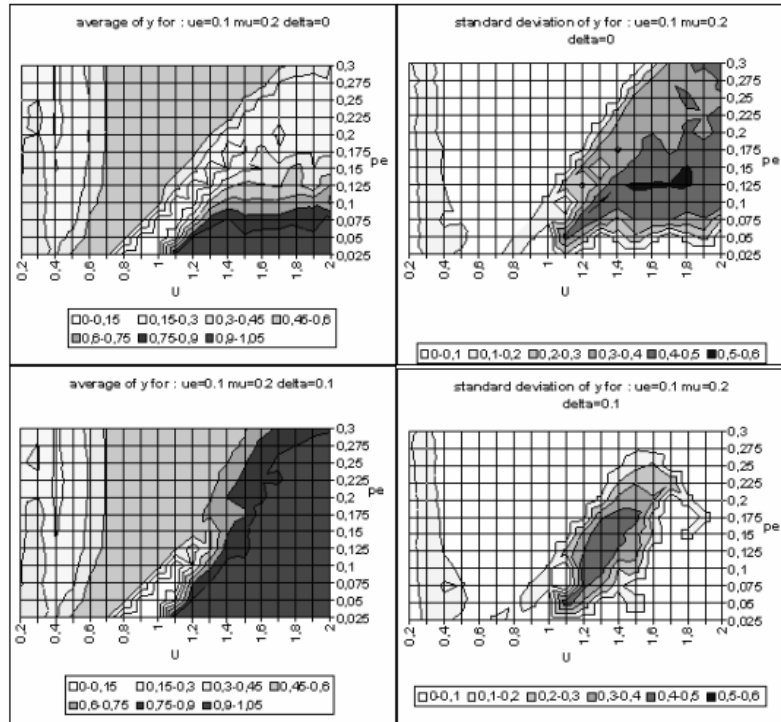


Fig. 2. Patterns typiques de la moyenne et écart-type de l'indicateur  $y$  (50 réplifications pour chaque point du graphe) en fonction de l'incertitude des modérés ( $U$ ) et de la proportion d'extrémistes ( $p_e$ ) pour  $\delta = 0$  (haut) et  $\delta = 0.1$  (bas). Les autres paramètres sont fixés : incertitude des extrémistes  $u_e = 0.1$ , intensité des interactions  $\mu = 0.2$ , différence relative initiale entre les extrémistes,  $\delta = 0.1$ . Sur le graphe des  $y$  moyens, les zones claires à gauche correspondent à des cas de convergence centrale, les grises, au milieu et en haut à de la convergence vers du double extrême et les foncées en bas à droite à de la convergence vers un seul extrême.

Les patterns typiques obtenus pour  $\delta = 0$  et  $\delta > 0$  sont présentés Fig. 2, chaque point de la grille correspond à 50 simulations avec 1000 agents. On peut identifier quatre régions avec différentes valeurs moyennes de  $y$  qui correspondent à chacun des trois cas de convergence: deux zones claires (une à gauche et une autre partant en diagonale à partir du milieu en bas) correspondant à des cas de convergence centrale, une zone grise (formant un triangle au milieu) correspondant à des cas de convergence vers les deux extrêmes et une zone foncée (en bas à droite) correspondant à des cas de convergence vers un seul extrême.

Les diagrammes de la Fig. 2 font apparaître de larges régions où  $y$  prend des valeurs intermédiaires qui sont dues à des distributions bimodales



d'attracteurs purs dépendant de la sélection aléatoire et de la distribution initiale. Ces distributions sont confirmées par une analyse plus détaillée (Deffuant et al., 2003).

### 3. Dynamiques d'opinions sur réseaux

#### 3.a Cas de la grille régulière comme réseau social

Nous remplaçons maintenant le réseau complètement connecté pour exécuter le modèle sur une grille régulière (un tore) avec un voisinage de Moore (connectivité  $k = 8$ ). Nous explorons alors de manière systématique l'espace des paramètres formé par l'incertitude des individus modérés  $U$  et la proportion d'extrémistes  $p_e$  (cf. Fig. 3).

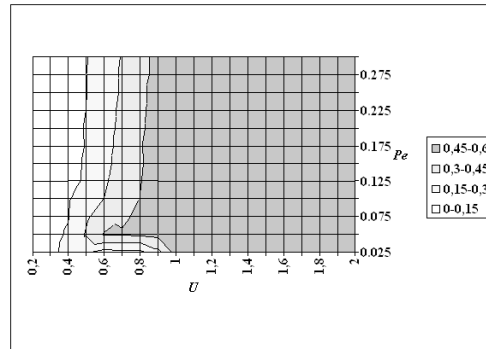


Fig. 3. a) Exploration de l'espace des paramètres ( $U, p_e$ ) avec un voisinage de Moore (connectivité 8) sur une grille régulière, les autres paramètres sont  $u_e=0.1$ ,  $\mu=0.2$  et  $\delta=0$ . Les écarts-type associés sont relativement faibles, exception faite de la zone de transition pour de faibles valeurs de  $p_e$ . Les principales observations sont une décroissance générale des valeurs de  $y$  et l'absence de convergence vers un seul extrême ( $y$  n'atteint jamais des valeurs proches de 1).

On peut remarquer que :

- $y$  est presque toujours inférieur à 0.6 ce qui implique que la convergence vers un seul extrême ne se produit que très rarement et pour des configurations et des valeurs de paramètres qui ne sont pas celles évoquées dans le cas complètement connecté. Nous observons uniquement la convergence vers un seul extrême pour des cas très particuliers de positionnement des extrémistes ou pour de fortes valeurs de  $\delta$  (différence entre les proportions initiales d'extrémistes de chaque bord). Quand le nombre initial d'extrémistes est le même pour chaque extrême, nous observons uniquement une proportion finale plus élevée

pour un extrême, mais jamais la convergence vers un seul extrême telle qu'observée aussi clairement dans le cas complètement connecté.

- La transition de la convergence centrale vers double extrême a une forme similaire à celle observée lorsque la population est complètement connectée.

Ces observations peuvent s'expliquer comme suit :

- Pour de faibles valeurs de  $U$  un nombre important de clusters apparaît du fait que les agents ont tendance à être isolés : il y a une forte probabilité que tous leurs voisins aient une opinion trop éloignée de la leur pour être influents ou être influencés. C'est également le cas des extrémistes qui ne sont pas particulièrement influents pour de faibles valeurs de  $U$ .
- Pour de fortes valeurs de  $U$ , les agents ont au contraire de grandes chances de trouver des interlocuteurs dans leur voisinage. L'influence des extrémistes se propage en suivant la topologie du graphe, tout d'abord en attirant leurs propres voisins puis les autres. La contamination s'arrête lorsque le cluster formé rencontre un cluster d'opinion opposée.
- Une explication possible de l'absence de convergence vers un seul extrême s'appuie sur cette propagation locale de l'extrémisme. Il ne se produit pas lorsque la connectivité est forte car la majorité "rattrape" toujours les agents qui sont occasionnellement attirés par un des extrêmes (conduisant au phénomène de virement global vers un extrême, lorsque la majorité perd le contact avec l'autre extrême). Avec un voisinage de Moore, pour de fortes valeurs de  $U$ , chaque extrémiste se propage rapidement dans son voisinage, ce qui induit une convergence vers deux extrêmes et empêche la convergence vers un seul extrême de se produire.

Si un voisinage de Moore avec une connectivité de 8 n'est pas suffisant pour observer des cas de convergence vers un seul extrême, nous devons augmenter cette connectivité moyenne. Nous avons donc besoin d'un modèle qui permette de modifier facilement la connectivité moyenne, ainsi que la régularité du réseau. Nous avons donc sélectionné le modèle  $\beta$  parmi les modèles de small-worlds (Watts et Strogatz, 1998) qui satisfait ces pré-requis.

### 3.c Exploration sur des small-worlds

Le modèle  $\beta$  de small-world est contrôlé par deux paramètres : la connectivité moyenne  $k$  et le bruit appliqué sur les connections  $p$ . En partant d'une structure régulière (dans notre cas un anneau régulier) de connectivité  $k$ , nous enlevons chaque lien avec la probabilité  $p$ , en les reconnectant aléatoirement. Dans notre cas  $k$  doit être paire puisque chaque individu sur l'anneau a  $k/2$  connections de chaque coté.

Nous étudions à partir de là le comportement du modèle pour une grosse incertitude de la population  $U=1.8$ , et un faible taux d'extrémistes  $p_e=0.05$  (en bas à droite de la figure 2), qui conduit à une convergence vers un seul extrême lorsque la population est complètement connectée. Nous exécutons 50 répliquions des simulations sur des réseaux obtenus pour  $p$  variant de 0 à 1, donc de réseaux réguliers jusqu'à des réseaux complètement aléatoires et  $k$  variant de 2 à 256 suivant les puissances de 2 donc de 0.2% à 25% de la population (1000 individus). Au-delà de connectivités de 25%, le comportement du modèle est le même que dans le cas complètement connecté, quelle que soit la valeur de  $p$ .

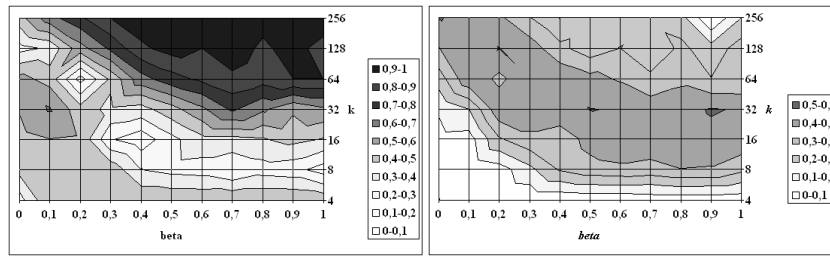


Fig. 4. Exploration de l'effet réseau avec une topologie small-world avec un anneau comme substrat dans l'espace de paramètres ( $\beta$  (noté  $p$  dans le texte),  $k$  paramètre de connectivité) sur la dynamique du modèle pour  $U=1.8$ ,  $u_e=0.1$ ,  $N=1000$ ,  $\mu=0.1$ ,  $\delta=0$ ,  $p_e=0.05$  (correspondant à une convergence vers un extrême dans le cas complètement connecté)

Nous observons (cf. Fig. 4) une transition d'une convergence double extrême vers une convergence simple extrême lorsque la connectivité ( $k$ ) augmente. Dans la zone de transitions, les écarts-types importants associés à  $y$  indiquent la co-existence (vérifiée par ailleurs) de cas de convergence centrale et de convergence vers un extrême. L'analyse des traces d'évolution des opinions pour plusieurs simulations confirme l'hypothèse exprimée en 3.a. Pour de faibles connectivités, chaque extrémiste influence son voisinage

qui devient rapidement extrémiste. Nous obtenons alors plusieurs clusters, chacun étant en général contrôlé par un extrémiste, l'ensemble conduisant à une convergence double extrême. Lorsque la connectivité atteint une valeur critique, la population tend à se regrouper au centre et les fluctuations du cluster d'opinion peuvent conduire à le déconnecter d'un extrême, conduisant ainsi à une dérive vers l'autre extrême. La convergence centrale a lieu lorsque la majorité centrale perd contact avec les deux extrêmes. Cette situation est favorisée par un faible niveau de connectivité. Lorsque l'on augmente la connectivité, cette situation de convergence vers un seul extrême apparaît régulièrement comme le confirme l'analyse des distributions des cas de convergence pour une valeur fixe de  $p$  (cf. Fig. 5).

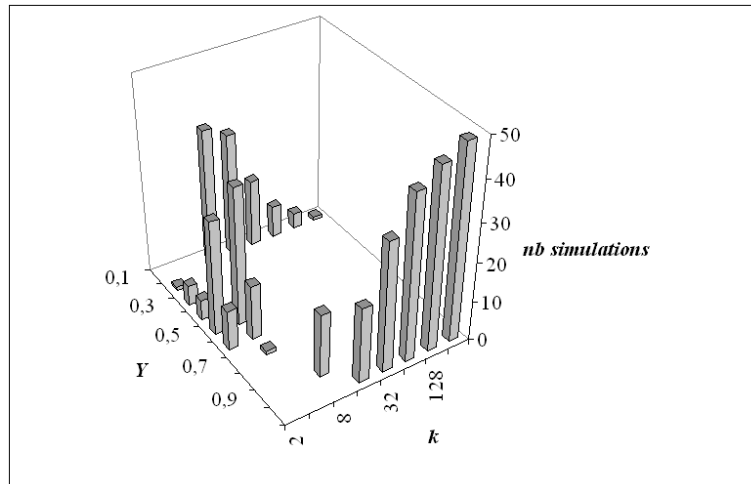


Fig. 5. Distribution de  $y$  pour une topologie small-world avec  $k$  paramètre de connectivité sur une échelle logarithmique, pour  $p = 0.8$ ,  $U=1.8$ ,  $u_e=0.1$ ,  $N=1000$ ,  $\mu=0.1$ ,  $\delta=0$ ,  $p_e= 0.05$ . Nous observons une transition de phase pour des valeurs de connectivité proches de 8 et à partir de ce point nous observons également que le mélange entre convergence vers un extrême et convergence centrale tend à devenir uniformément de la convergence vers un seul extrême.

De plus, la transition a lieu pour des connectivités plus élevées lorsque  $p$  décroît, i.e. lorsque les réseaux deviennent plus réguliers. Notre hypothèse est que la régularité du réseau renforce l'effet local qui favorise une propagation locale rapide de l'influence de chaque extrémiste initial, conduisant à une convergence vers les deux extrémités.

Ces résultats sont robustes relativement au substrat choisi pour construire le réseau, les mêmes phénomènes apparaissant pour une grille régulière plutôt qu'un anneau (Fig. 6).

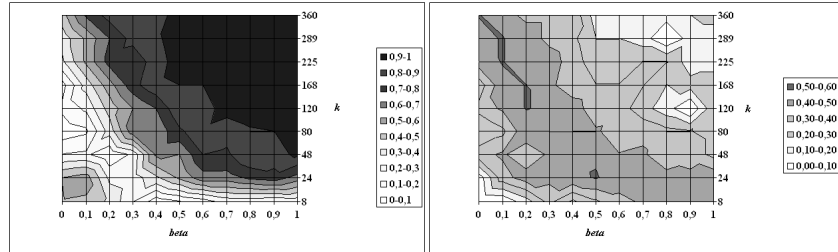


Fig. 6. Exploration de l'effet réseau avec une topologie small-world construite sur un substrat grille en appliquant un voisinage de Moore généralisé, sur la dynamique du modèle pour  $U=1.8$ ,  $u_e=0.1$ ,  $N=1000$ ,  $\mu=0.1$ ,  $\delta=0$ ,  $p_e=0.05$  (convergence vers un seul extrême dans le cas complètement connecté). Les résultats qualitatifs sont les mêmes comparés à un substrat en anneau même si des distorsions sont observées du fait de valeurs de connectivité qui devraient être testées dans le cas de l'anneau.

Nous sélectionnons alors trois couples typiques  $(U, p_e)$  correspondant à différents types de convergence lorsque la population est complètement connectée.

Tous les couples choisis conduisent à des dynamiques similaires lorsque  $p$  et  $k$  varient. Pour de faibles connectivités nous observons une majorité de convergences double extrême et pour de fortes connectivités nous observons le cas de convergence observé dans le cas complètement connecté. Nous observons l'occurrence systématique de convergence centrale dans la transition entre ces comportements. De plus, lorsque nous augmentons  $p$ , nous nous rapprochons de réseaux aléatoires et la transition a lieu pour des connectivités plus faibles.

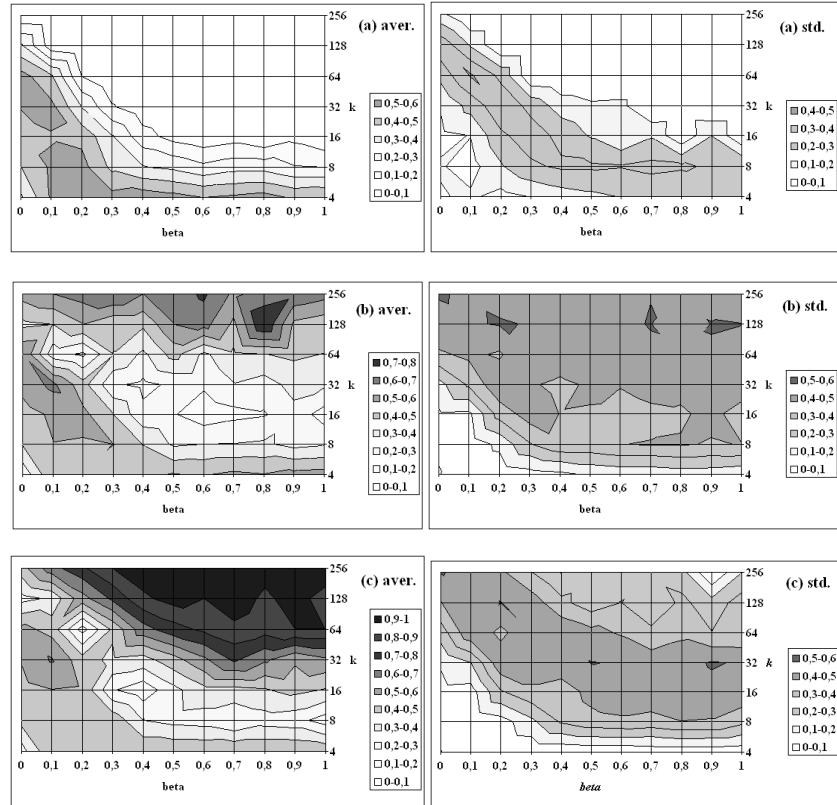


Fig. 7. Représentation des valeurs moyennes de  $y$  sur 50 répliques pour chaque point à gauche et de l'écart type associé à droite pour différents couples de valeurs de  $(U, p_e)$  les autres paramètres sont gardés constants par rapport à la Fig. 4. Les trois points  $(U, p_e)$  considérés correspondent pour le cas complètement connecté aux (a) cas de convergence centrale ( $U=1.0$  et  $p_e=0.05$ ) (b) cas de convergence double extrême ( $U=1.2$  et  $p_e=0.05$ ) et (c) cas de convergence vers un seul extrême ( $U=1.4$  et  $p_e=0.05$ ) comme nous pouvons le remarquer lorsque la connectivité augmente.

Dans la figure 7.b, on peut observer une zone de transition entre deux zones de convergence double extrême. Ces deux zones résultent en fait de processus différents : pour des faibles connectivités, cette convergence provient de l'agrégation de processus de convergences locales vers un extrême alors que pour des connectivités plus élevées, elle résulte d'une

convergence globale du cluster central, qui se divise en deux, chaque partie étant attirée par un des deux extrêmes.

#### **4. Discussion et perspectives**

Nous avons exploré le modèle d'influence proportionnelle à l'accord relatif avec extrémistes sur des réseaux small-worlds, en faisant varier la connectivité et le bruit appliqué au réseau. Nous avons mis à jour un niveau critique de connectivité, qui permet l'émergence de la convergence vers un seul extrême. Ce niveau critique de connectivité augmente lorsque la régularité du réseau augmente. Ce résultat peut être expliqué par le besoin d'une première phase de clustering central global pour permettre la réalisation d'une convergence vers un seul extrême. Une faible connectivité et une forte régularité du réseau favorisent une propagation locale rapide de l'extrémisme, qui empêche le clustering central global de voir le jour.

Nous avons conduit d'autres études pour comprendre le rôle des autres paramètres du modèle. Des simulations préliminaires laissent à penser que la taille de la population est un paramètre critique quand le réseau est régulier (en jouant indirectement sur le diamètre du graphe). La critique de ces résultats dans une perspective sociologique est également un défi majeur qui est en dehors du cadre de cet article. Cependant, l'idée de l'existence d'un niveau critique de connectivité et de régularité ou de redondance des liens dans le réseau permettant une dérive globale de la population vers une opinion extrême (par opposition à une bipolarisation) ne semble pas contre-intuitive, et pourrait trouver de riches interprétations dans des phénomènes sociologiques de la vie réelle.

Notons que nous présentons des travaux complémentaires dans (Weisbuch et al., 2005), où nous étudions des convergences vers un seul extrême différentes, qui ont lieu lorsque le nombre d'extrémistes est très faible. Dans ce cas en effet, pour certaines valeurs de l'incertitude, il se peut que tous les extrémistes d'un bord voient leur propagation bloquée, alors que certains de l'autre bord parviennent à emporter l'adhésion d'une large part de la population. Les deux phénomènes appellent des interprétations psycho-sociales différentes.

Par ailleurs, de nombreuses extensions sont possibles à partir de ce modèle pour prendre en compte plus finement certaines situations. Parmi celles-ci, Jager et Amblard (2005) ont introduit en plus de la dynamique d'influence, une dynamique de répulsion entre attitudes qui conduit, sous

certaines conditions à une radicalisation de celles-ci. D'autres, à venir, sont amenées à enrichir progressivement le modèle présenté ici.

## Remerciements

Les auteurs remercient chaleureusement Gérard Weisbuch pour ses réactions constructives sur l'article et pour sa collaboration bienveillante.

## Références

- 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*, vol. 5, n°4.
- Deffuant, G., Neau, D., Amblard, F. et Weisbuch, G., 2001. « Mixing beliefs among interacting agents », *Advances in Complex Systems*, vol. 3, pp.87-98.
- Hegselmann, R. et Krause, U., 2002. „Opinion Dynamics and Bounded Confidence Models, Analysis and Simulation”, *Journal of Artificial Societies and Social Simulation*, vol. 5, n°3.
- Jager, W. et Amblard, F., 2005. « Uniformity, bipolarisation and pluriformity captured as generic stylized behaviour with an agent-based simulation model of attitude change », *Computational and Mathematical Organization Theory*, vol.10, pp.295-303.
- Weisbuch, G., Deffuant, G., Amblard, F. et Nadal, J.-P., 2002. „Meet, discuss and segregate!”, *Complexity*, vol. 7, n°3, pp.55-63.
- Weisbuch, G., Deffuant, G., Amblard, F. et Nadal, J.-P., 2002. "Interacting Agents and Continuous Opinions Dynamics", *Lecture Notes in Economics and Mathematical Systems*, vol. 521, pp.225-242.
- Weisbuch, G., Deffuant, G. et Amblard, F., 2005. „Persuasion dynamics”, *Physica A*, vol.353, pp.555-575.
- Watts, D. et Strogatz, S.H., 1998. “Collective Dynamics of small world networks”, *Nature*, vol. 393, pp.440.



*Iadine Chadès :* Ce travail propose d'étudier les interactions et la dynamique d'influence des agents extrémistes dans un réseau social. Les auteurs utilisent un modèle de système multi-agents dit réactif. En ce sens, chaque agent représente un individu, qui est défini par une valeur d'opinion et une incertitude associée. Les agents interagissent entre eux selon une règle d'influence susceptible de modifier l'opinion et l'incertitude – ou la certitude – d'un agent. Cette influence suit le principe de l'accord relatif, dont une caractéristique est l'évolution de l'opinion et de l'incertitude d'un agent. Bien que cette évolution puisse se produire dans les deux sens, un agent plus certain est plus influent qu'un agent incertain. Il n'existe aucun contrôle central ou influence globale d'une possible « autorité » ou autre media d'information. Cette société d'agents interagit selon les caractéristiques individuelles qui leurs sont propres mais aussi selon leur position et leur relation dans un réseau social. Les auteurs étudient au cours de simulations sur différentes structures type de réseaux les convergences de populations d'agents vers des « majorités » et/ou « minorités ». La structure des réseaux n'évoluant cependant pas au cours de la simulation.

En tant que chercheur en Intelligence Artificielle mes réflexions porteront principalement sur le choix de ce modèle et peu sur les résultats de l'étude de l'influence des réseaux sociaux qui me paraissent complets, et bien traités. J'aurais cependant une remarque sur le fait que le segment d'opinion lié à l'incertitude  $U$  dépasse parfois le domaine de définition de l'opinion, est ce que cela biaise les simulations ? En effet, si  $U = 1.2$  il doit y avoir beaucoup de segments en dehors de  $[-1,+1]$ . Comment l'interprètent les auteurs ?

*Frédéric Amblard :* *En effet, en se limitant à une distribution des opinions bornées entre  $-1$  et  $+1$  et en utilisant cette même dimension pour l'incertitude, cela pourrait impliquer implicitement que l'interprétation que l'on peut construire se limite à cet intervalle. C'est un problème que nous ne nous étions pas posé jusqu'à présent, mais si l'on considère le modèle présent comme une abstraction très élevée (ce qui est le cas) de ce qui se passe réellement, comme un modèle abstrait donc, dont un des buts est d'être dérivé ou adapté pour être appliqué à la compréhension de cas concrets, une interprétation possible de ce « débordement » pourrait être prise en compte dans le cadre d'un élargissement du spectre des opinions possibles. Un débat pouvant permettre la construction de nouveaux arguments, ce débordement pourrait être interprété comme une sensibilité possible*

*des agents à des arguments plus extrêmes que ceux qui étaient présents à l'origine du débat. En dehors du problème posé par le cadre interprétatif, cela ne biaise pas les résultats.*

*Guillaume Deffuant : Pour moi, ce n'est pas un problème. La limitation des opinions entre  $-1$  et  $+1$  est simplement un choix pratique qui vient de la distribution initiale, mais n'est en rien une contrainte. Les opinions ne sont pas nécessairement bornées. Nous avons fait (beaucoup) de tests avec des distributions gaussiennes, qui ne sont pas bornées.*

IC : Mon autre remarque ne se rapporte pas tant au sujet même de l'influence sociale, mais plutôt à la première hypothèse qui n'a pas été énoncé dans ce travail : Qu'est-ce que l'opinion ? Mes réflexions sur ce travail sont principalement liées à cette première notion. Les auteurs ne donnant pas de définition à laquelle ils se rapportent dans cet article, j'en énonce une :

*« Avis personnel au sujet de quelque chose, avis que l'on tient pour vrai, sans quoi il ne serait pas le nôtre. L'opinion, c'est un énoncé qui prétend être vrai. »*

La nature de ce « quelque chose » me semble important pour le choix de la représentation d'une opinion. Il me paraît peu vraisemblable d'utiliser une représentation et un modèle général de l'influence sociale afin de traiter des sujets aussi variés sur lesquels un individu peut avoir une opinion (politique, religion, société,...). Selon l'importance du sujet choisi, il paraît intéressant d'une part de changer la distribution d'opinion initiale, la structure du réseau peut également en dépendre et enfin la dynamique d'influence entre individu (règle différente). Les auteurs ne semblent pas en discuter, qu'en pensent-ils ?

*FA : Là encore je ne peux qu'acquiescer à la remarque très juste qui nous est faite. Cependant, et pour resituer le cadre de ce travail, je tiens à souligner, que le modèle que nous proposons est un modèle générique, très abstrait, que nous nous garderions bien d'appliquer tel quel pour comprendre le comportement d'un système concret. Nous avons en quelque sorte cherché à produire un modèle qui pourrait servir de plus petit commun dénominateur à des travaux portant sur les dynamiques d'opinions que nous prenons par conséquent dans une acception plus large que celle que vous proposez, en cherchant à regrouper à la fois ce qui*

*relèverait de choix de consommation ou de problématiques liées au choix électoral par exemple.*

*GD : Tout à fait d'accord. D'où l'intérêt d'explorer les comportements du modèles pour différentes distributions, différents réseaux, etc...*

*IC : Dans la continuité du paragraphe précédent, une autre remarque concerne le choix de la représentation de l'opinion par ensemble de valeurs réelles comprise entre  $[-1,+1]$ . Ce domaine de définition est continu. Le choix d'un domaine de définition continu, l'utilisation de règle d'influence, induisent-ils une relation d'ordre dans l'ensemble des opinions? Si, oui comment interpréter cette relation d'ordre? Une relation d'ordre dans un ensemble  $E$  est une relation binaire dans cet ensemble qui permet de comparer ses éléments entre eux de manière cohérente. Un ensemble muni d'une relation d'ordre est un ensemble ordonné ou tout simplement un ordre.*

*GD : Oui il y a un ordre. Typiquement il s'agit d'un niveau d'approbation ou de désapprobation. Certains de nos collègues ont critiqué le choix d'une échelle continue comme peu réaliste, et ont proposé des versions discrètes du modèle, dont les propriétés sont peu différentes dès que le niveau de discrétisation est un peu fin.*

*FA : Pour moi la relation d'ordre est intrinsèque au choix de représentation, en l'occurrence la relation d'ordre existe sur un segment de nombres réels, pour autant, et en dehors des règles d'interactions qui utilisent de manière pragmatique cette relation d'ordre (le fait que la différence entre deux opinions soit inférieure à un seuil) nous ne faisons pas porter de signification particulière à cette relation d'ordre. Typiquement, en dehors du fait qu'il s'agisse des extrémités de la distribution, les opinions  $-1$  et  $+1$  sont dans ce modèle strictement équivalentes. Des extensions de ce modèle ont en revanche pris en compte cette polarisation de la distribution en interprétant de manière distincte les opinions positives et négatives et la prise en compte de la polarisation du segment dans la règle d'interaction est également envisagée (le fait que dans le cadre d'un jugement par un jury, la peine de mort n'est pas équivalente au non-lieu en terme de décision).*

*IC : Je précise que je trouve très intéressante l'utilisation de valeurs pour la modélisation d'une opinion. Elle permet entre autres d'exhiber et d'analyser des phénomènes de convergence des simulations.*

Les auteurs différencient 2 types d'agents. Les agents qui se situent à une extrémité du domaine de définition : « ils ont une incertitude moindre que les individus modérés».

Et les agents aux valeurs intermédiaires qui n'ont pas de « rôle » a priori, mais ils sont définis comme des « modérés » (incertitude de leurs opinions). Note: un modéré peut être un agent avec une opinion extrême s'il n'est pas certain de ses opinions.

Poursuivant la compréhension du modèle, un type d'agent me paraît jouer à un rôle particulier, ou pourrait être intéressant. L'agent "philosophe" ou l'agent "sage" serait par définition un agent sans opinion (agent avec une valeur d'opinion égale à 0?). Cet agent pourrait avoir un effet d'accroissement de l'incertitude au sein de la population d'agent. Suivant les principes des philosophes, l'accroissement de l'incertitude aurait des vertus de sagesse, si l'on considère que les agents recherchent à augmenter leur capacité de réflexion. Peut-être que la convergence centrale pourrait correspondre à ce type de phénomène.

***FA : L'hypothèse d'agents philosophes ou de modérateurs dans une population ou au cours d'un débat est très intéressante et sans avoir conduit les expériences correspondantes, je pense que la présence de ces agents qui auraient une opinion moyenne et une large incertitude (ou une large ouverture d'esprit puisqu'ils prendraient en compte par conséquent davantage d'opinions de la population) conduirait effectivement à un consensus, une convergence centrale. Ça fait réfléchir !***

***GD : Oui, cela fait réfléchir. Nous montrons aussi que l'accroissement de l'incertitude de la majorité modérée la rend plus vulnérable à l'influence des extrêmes (si malheureusement ces derniers ne sont pas sensibles à l'influence du philosophe)...***

*IC :* De par mes recherches, je fais un rapprochement entre opinion et croyance. Dans mon domaine d'étude –les processus décisionnels stochastiques–, les croyances (« beliefs ») apportent une représentation plus riche et plus complexe de la configuration d'un problème. Un état de croyance est défini par une distribution de probabilité sur les états possibles du système. L'objectif final est de trouver les décisions séquentielles optimales pour chaque état de croyance.

Je me pose donc la question suivante : Est-il judicieux de considérer que l'étude de l'influence sociale sur les opinions peut se donner comme objectif la prédiction de futures prises de décision d'une population cible?

Suivant cette nouvelle hypothèse, l'échelle de valeur ainsi que l'hypothèse selon laquelle les extrémistes se placent aux extrémités de ce domaine peut favoriser l'idée que ce modèle concerne une opinion donnant lieu à une future décision de type vote référendaire « oui-non ». En ce cas, comment interpréter les valeurs d'opinion intermédiaire et médiane? Ces valeurs peuvent par exemple refléter une croyance: distribution de probabilité entre deux configurations/états possibles le "oui" et le "non".

Dans le cas où le domaine de valeur d'une opinion conduit à un ensemble fini de décisions possibles (un état pour chaque décision possible), le cas particulier des extrémistes définis comme des agents ayant très peu d'incertitude et par conséquent un fort potentiel d'attraction, ne devraient pas se limiter aux valeurs extrêmes d'une opinion. Les leaders charismatiques pourraient également figurer dans chaque classe possible de décision. Il serait donc intéressant de simuler la dynamique des opinions en incorporant des agents de type « leaders charismatiques » dans le modèle. Je pense que ce problème d'opinions/décisions multiples peut également être modélisé par l'utilisation de réseaux hiérarchiques interconnectés : chaque sous-réseau peut représenter un courant de pensée, et comporter ses agents caractéristiques – extrémistes et philosophes.

*FA : Pour réagir tout d'abord sur le point concernant l'utilisation de ce modèle à des buts prédictifs, il me semble que la volonté même de se focaliser sur un modèle générique dès le départ est incompatible avec la possibilité de réaliser des prédictions sur un domaine particulier, mais nous n'excluons pas qu'un modèle plus fin qui prendrait en compte les particularités liées à un contexte particulier puisse être utilisé à des fins de prédictions. L'interprétation proposée des opinions continues comme des probabilités de décision à choix discret (binaire ne l'occurrence) est très intéressante, en particulier, si l'on considère son application au modèle courant, cette extension pourrait correspondre à l'introduction d'une date de vote ou d'une itération de simulation correspondant au moment du choix, ce moment n'étant pas nécessairement choisi à convergence et pouvant intervenir à un moment où la population ne s'est pas clairement prononcée ou n'a pas suffisamment débattu. Le dernier point qui concernait l'introduction dans le*

*modèle de leaders charismatiques possédant des opinions intermédiaires nous l'avons étudié rapidement avant de nous pencher sur l'influence des extrémistes. Nous l'avons écarté pour des raisons davantage liées à la forme des résultats obtenus que sur des critères scientifiques. De fait, des leaders charismatiques avec des opinions intermédiaires conduisent à la formation de clusters d'opinions autour de ces leaders.*

*GD : Nous avons proposé un modèle de décision lié au modèle d'évolution des opinions (voir Deffuant et al. 2005). Dans ce modèle, la décision porte sur l'adoption ou non d'une innovation, et il y avait en outre des mécanismes de propagation d'information sur l'innovation. Nous avons mis en évidence des interactions particulières entre la dynamique des opinions et la propagation de l'information. Par ailleurs, des résultats expérimentaux montrent que les opinions sur les sujets qui tiennent à cœur tendent à se concentrer vers les extrêmes, alors que sur les sujets jugés peu importants, ces opinions forment plutôt une gaussienne centrée. Ces résultats suggèrent que les opinions fermes tendent à être situées aux extrêmes.*

*IC :* Le modèle d'influence proposé repose sur le principe de l'accord relatif. Il est intéressant de remarquer que ce modèle est déterministe dans sa phase «d'influence ». En effet, un agent migre petit à petit vers le plus influent des agents avec qui il est en relation ou connecté dans la structure de réseau choisie. En Intelligence Artificielle, nous avons l'habitude de considérer les effets stochastiques des décisions. En ce sens, il serait également intéressant d'enrichir le modèle d'influence en y ajoutant une composante stochastique qui trouverait une signification dans un échec possible d'une influence. Cela permettrait également d'insérer une hypothèse d'influence non symétrique selon l'appartenance originale d'un agent à une opinion (mémoire), ainsi qu'un phénomène de répulsion comme le précisent les auteurs dans leur conclusion.

*FA : Nous avons étudié récemment ces extensions. Le bruitage de l'influence suite à des arguments évoqués par Bruce Edmonds concernant la robustesse de la fonction d'influence, conduit notamment à une apparition plus fréquente de consensus (voir Deffuant 2006, à paraître dans JASSS). Le phénomène de répulsion (ou de radicalisation) est en cours d'étude avec un psychologue hollandais, Wander Jager et promet, dans un cadre d'application un peu différent, des résultats très intéressants.*

IC : On peut également imaginer d'autres formes de règle d'influence:

- rencontrer une opinion opposée en contradiction avec son opinion peut forcer un agent à "réfléchir", augmenter son incertitude,
- ou bien renforcer l'attachement à cette opinion et donc réduire son incertitude,
- certains agents peuvent également ne jamais être influencés/ ni vouloir influencer. Même s'ils ne sont pas certains de leur opinion. Leur rôle peut sembler mineur au premier abord mais peut s'avérer intéressant dans la construction de réseau dynamique, et règles d'influence via ce réseau.

Enfin, dans tout phénomène de propagation la structure du réseau est une composante importante. Une amélioration possible consisterait à intégrer une dynamique stochastique dans la structure de ce réseau. Cette dynamique peut refléter le changement d'opinion, mais aussi le changement de relations sociales qui peuvent se produire. Ce qui soulève l'intérêt d'études de simulations de modèles d'influence dans des réseaux hiérarchiques interconnectés.

*FA et GD : Cette dernière remarque nous permet de conclure avec enthousiasme cette série de questions tout en remerciant Iadine Chadès de ces interactions positives qui nous amènent à accroître notre incertitude au sujet de ce travail (ce qui est une bonne chose). Le modèle proposé ici est avant tout un modèle générique, envisagé comme une brique de base pour des modèles de dynamique d'opinions. Sa simplicité, tout en éveillant la curiosité, conduit fréquemment à de nombreuses suggestions pour l'améliorer ou l'appliquer à des cas particuliers. Nous avons nous même, en collaboration avec Gérard Weisbuch, réalisés certaines de ces extensions, ici en ajoutant une composante réseau social, avec Wander Jager en intégrant un mécanisme de répulsion ; d'autres aussi se sont appropriés ce modèle pour en proposer des extensions particulières, Diemo Urbig en intégrant des problématiques liées à la théorie de l'information, Jan Lorenz pour étudier le cas d'opinions multiples, D. Stauffer pour sa version utilisant des opinions discrètes et d'autres encore... Nous sommes à chaque fois très intéressés et très enthousiasmés par le fait que d'autres chercheurs s'approprient ce modèle et l'étendent dans des directions que nous n'avions pas forcément prévues..*