

De la cellule biologique à la cellule urbaine : retour sur trois expériences de modélisation multi-échelles à base d'agents

J. Gil Quijano^a G. Hutzler^a T. Louail^{a,b}
quijano@ibisc.fr hutzler@ibisc.fr tlouail@ibisc.fr

^a Laboratoire IBISC (Informatique, Biologie Intégrative et Systèmes Complexes),
FRE 3190, Université d'Evry, CNRS, France

^b Laboratoire Géographie-Cités,
UMR 8504, Université Paris 1, Université Paris 7, CNRS

Résumé

Lorsque l'on simule des systèmes complexes, il est souvent nécessaire de considérer des entités à différentes échelles d'organisation, et prendre en compte la façon dont celles-ci s'inter-influencent. Le choix des échelles est conditionné par des contraintes méthodologiques qui incluent les objets manipulés par les thématiciens, la disponibilité des données, la question d'intérêt et également des contraintes algorithmiques. Différentes manières d'aborder les structures multi-échelles et les influences entre niveaux peuvent être identifiées. Dans un premier temps, nous présentons trois modèles multi-agents de dynamiques multi-échelles existants, centrés sur la croissance d'une tumeur cancéreuse, l'évolution de la distribution spatiale d'une population urbaine et l'évolution d'un réseau de villes. Nous proposons ensuite une grille de lecture de manière à comparer les différentes formes prises potentiellement par la notion de multi-échelles.

Mots-clés : simulation orientée agent, dynamiques multi-échelles, architectures multi-niveaux, comparaison d'approches agents

Abstract

Simulating complex natural or social systems often implies to consider entities at several levels of organisation, and take into account the way they influence each other. The choice of scales is driven by both methodological and thematic constraints which among others include the objects of significance for the specialists, availability of relevant data for comparison, question of interest and algorithmic considerations. These multi-scale structures and the inter-levels influences can be modeled in several ways. Firstly we present three existing examples

of multi-scale modeling, addressing the growth of a tumor cell, the evolution of a urban system over a long period, and the spatial repartition of urban population in Bogotá. Then we generalize the underlying agent-based architectures in order to illustrate the possibility of application on other modeling questions.

Keywords: agent-based simulation, multi-scale dynamics, multi-level architectures

1 Introduction

Les systèmes complexes naturels et sociaux sont souvent caractérisés par de nombreuses entités hétérogènes, de nature et de dimensions diverses, à différents niveaux d'organisation, interagissant de manière complexe les unes avec les autres, à des échelles de temps et d'espace elles aussi très diverses. Dans ce cadre, l'aspect « multi-échelles » ou « multi-niveaux » d'une modélisation correspond le plus souvent au choix d'un niveau particulier de description, et à l'observation de structures émergentes de plus haut niveau. Le choix d'un niveau conditionne la sélection des entités du système réel qui sont réifiées sous forme d'agents dans le modèle. Le système simulé peut alors être analysé à au moins deux niveaux d'organisation : celui des agents d'une part, via une analyse de leurs trajectoires, et celui du système d'autre part, via des mesures permettant de caractériser les structures produites au cours de la simulation. Le système peut aussi être mesuré à des niveaux intermédiaires qui peuvent faire sens pour le modélisateur, correspondant par exemple à des groupes d'agents partageant des caractéristiques communes, même si en pratique ce raffinement est plus rarement mis en place.

Il nous semble dans ce cas difficile de parler de modélisation véritablement multi-niveaux dans la mesure où seul le niveau « du bas » est présent dans le modèle et spécifié par le modélisateur. Le ou les niveaux supérieurs sont dans ce cas des niveaux qui peuvent être observés par l'expérimentateur qui « regarde » et mesure les simulations, mais ils ne sont pas réifiés en tant que tels dans le modèle. Le choix du niveau est particulièrement important car il doit satisfaire un certain nombre de contraintes qui s'imposent au modélisateur, sous peine de ne pas être pertinent. Ce dernier doit en particulier s'interroger sur la question à laquelle son modèle est censé répondre, sur les ontologies manipulées par les experts du domaine, sur les données disponibles aux différents niveaux de description, et enfin sur la puissance de calcul impliquée par le nombre d'entités et la période temporelle à simuler. Concernant la disponibilité des données, il faut noter que l'on possède rarement les données permettant de caractériser entièrement l'état et la dynamique d'un système à un niveau particulier. Les données collectées sont en effet fragmentaires et hétérogènes.

Cette approche émergentiste, mono-niveau dans la conception et bi-niveaux dans l'analyse des comportements produits, est de loin la plus courante actuellement en simulation orientée agents (voir par exemple les modèles présentés dans [16]). Elle n'est pourtant pas toujours applicable de manière satisfaisante pour des systèmes réels. En effet, un certain nombre de systèmes complexes ne peuvent se comprendre sans intégrer plusieurs niveaux d'organisation et des dynamiques qui ne peuvent se réduire à une vision purement émergentiste. Les systèmes sociaux en particulier intègrent de manière forte un rôle structurant des niveaux supérieurs sur les niveaux inférieurs. De ce fait, les théories du domaine correspondantes, permettant de décrire le fonctionnement du système, sont rarement mono-niveau, soit par manque de connaissance à un niveau donné, soit par le fait que les ontologies manipulées mixent des entités et des acteurs de niveaux d'organisation différents. Il est par exemple difficile d'imaginer reproduire l'évolution d'une ville en partant de la seule spécification du comportement de ses habitants. Par ailleurs, il serait trop coûteux d'un point de vue computationnel de simuler l'ensemble d'un système de grande taille au niveau le plus fin : simuler au niveau atomique, ou même au niveau mo-

léculaire, l'ensemble d'une cellule constituée de plus de 4.10^{13} molécules (cellule hépatique de rat) est ainsi parfaitement irréaliste.

Au final, la question posée nécessite la prise en compte d'entités à différents niveaux (moléculaire et cellulaire par exemple, ou ceux des individus, du quartier et de la ville) ce qui, comme le rappellent Servat et ses collègues dans [15], fait partie de la gymnastique intellectuelle scientifique : faire cohabiter des points de vue et descriptions complémentaires d'un même système. Et de fait, notre perception de la réalité et des objets qui nous entourent se fait à plusieurs niveaux d'échelle.

Ces auteurs pointent deux raisons supplémentaires à la prise en compte de niveaux d'organisation multiples [15]. La première est d'ordre informatique : beaucoup de systèmes distribués font preuve d'un comportement global qui devient remarquable lorsque le nombre de composants devient critiquelement grand. Substituer des « super-agents » à des groupes d'agents et leur associer un modèle comportemental plus global mais équivalent permet un gain de performance et éventuellement de pouvoir transposer les ressources gagnées pour modéliser plus finement d'autres parties du système. La seconde raison avancée est d'ordre plus thématique. La prise de conscience de sa similarité avec d'autres agents et de l'existence d'une structure « groupe » les rassemblant peut entraîner chez l'agent une modification intentionnelle de son comportement, soit pour renforcer son appartenance au groupe, soit pour s'en distinguer. Plus généralement, le fait de pouvoir prendre en compte plusieurs niveaux d'organisation dans un même modèle doit permettre d'aborder des questions originales par rapport à un modèle classique à deux niveaux, celui de la spécification et celui des structures collectives émergentes.

Il semble donc important de pouvoir intégrer, au sein d'un même modèle, les descriptions d'entités évoluant à différentes échelles spatiales et temporelles, correspondant à différents niveaux d'organisation du système. Cependant et comme le dit Lesne dans [8], « il ne s'agit pas tant de prendre en considération tous les éléments et leurs relations que de considérer *simultanément* le niveau des éléments et ceux de leurs propriétés émergentes et la façon dont ils sont couplés ». Pour le formuler autrement, c'est donc plus dans l'*articulation* de différents ni-

veaux d'organisation que dans leur seule juxtaposition que se situe l'apport d'un modèle multi-niveaux ou multi-échelles. Articuler les niveaux, cela implique de les prendre en compte explicitement dans le modèle et de définir la manière dont ils sont couplés, ce qui se traduira dans un modèle orienté agents par l'interaction entre agents/objets de niveaux différents. Cette question de l'articulation est d'actualité en simulation orientée agent. Treuil et ses collègues la présentent comme l'un des « chantiers » du domaine en conclusion de leur livre [16], et des projets actuels visent à construire des plateformes de simulation devant faciliter le multi-échelles (par exemple [4] en écologie, la plateforme Magéo en géographie¹).

Pour illustrer ces différents aspects et les potentialités d'une approche agent pour les aborder, nous présentons dans la partie 2 trois exemples de simulation de modèles multi-niveaux. Pour chacun nous essayons de mettre en évidence les différents niveaux du modèle correspondant aux différentes échelles représentées, ainsi que les mécanismes assurant leur couplage. Dans la partie 3, nous tentons une abstraction des organisations multi-agents sous-jacentes pour mettre en lumière leurs différences et discuter des conditions de leur déploiement. Enfin nous concluons ce papier en discutant des perspectives de recherche ouvertes par notre approche comparative.

2 De la cellule biologique à la « cellule urbaine » : trois exemples de simulation multi-échelles

Le premier exemple présenté dans cette section traite de la croissance d'une tumeur cancéreuse, abordée aux échelles moléculaire et cellulaire. Le second traite de la simulation des choix de logement des ménages dans la ville de Bogotá, traitée individuellement au niveau des ménages et des logements mais également au niveau de *groupes* de logements et de ménages. Enfin le troisième présente un modèle de croissance d'un système urbain, approché à l'échelle des quartiers, des villes et du système de villes.

1. <http://mageo.csregistry.org/tiki-index.php>

2.1 Modèle de migration de cellules cancéreuses

Lorsque l'on parle de cancer, l'un des facteurs de mauvais pronostic est lié à l'apparition de métastases, c'est-à-dire de tumeurs secondaires. Ces tumeurs secondaires peuvent se développer lorsqu'une cellule de la tumeur primaire passe d'un état prolifératoire à un état migratoire, s'échappant ainsi vers d'autres parties du corps. Quand les conditions micro-environnementales le permettent, la cellule arrête sa migration et revient à un état prolifératoire, conduisant au développement de la tumeur secondaire. Nous nous intéressons ainsi aux conditions micro-environnementales, présentes autour d'une tumeur cancéreuse, pouvant conduire à l'échappement métastatique d'une cellule. En particulier, nous étudions l'impact de la protéine PAI-1, suspectée d'être à l'origine de la transformation morphologique des cellules permettant l'échappement [10]. Cette molécule pourrait ainsi jouer un rôle clé dans l'initiation et le maintien de la migration des cellules cancéreuses. En effet, la protéine PAI-1 est produite par les cellules elles-mêmes et peut se fixer sur la matrice extra-cellulaire. Les cellules, disposant de récepteurs pour fixer puis internaliser PAI-1, produisent et internalisent simultanément cette molécule. Lorsque cette dernière est accrochée à la matrice extra-cellulaire, présente tout autour de la tumeur, elle fournit alors aux cellules des points d'adhésion qui leur permettent de migrer en se faulant à travers la matrice.

Nous avons développé un premier modèle à base d'agents dans lequel aussi bien les cellules que les molécules de PAI-1 sont modélisées de manière individuelle. Il s'agit d'un modèle à pas de temps discret implémenté sur la plateforme Netlogo. La modélisation de la croissance de la tumeur est obtenue par un comportement de division cellulaire : une cellule peut se diviser pour donner deux cellules filles au bout d'un temps variable qui dépend de son accès aux nutriments. Ce comportement prolifératoire est couplé à un comportement de répulsion entre cellules : deux cellules trop proches se repoussent. Le modèle intègre par ailleurs la dynamique de production et d'internalisation de PAI-1 par les cellules. Pour ce faire, une cellule produit des molécules de PAI-1 avec un certain taux τ_{prod} . Dans le même temps, elle peut internaliser le PAI-1 en suspension, ainsi que

le PAI-1 matriciel, avec une probabilité directement proportionnelle à son nombre de récepteurs nb_{upar} . L'accès aux nutriments définit trois états différenciés pour les cellules de la tumeur : une cellule ayant un accès satisfaisant aux nutriments (couche externe de la tumeur) est active et possède à la fois les comportements de prolifération et production/internalisation de PAI-1 ; une cellule ayant un accès insuffisant aux nutriments nécrose et meurt (couche interne de la tumeur) ; les cellules de la couche intermédiaire reçoivent quant-à elles suffisamment de nutriments pour produire et internaliser PAI-1 mais plus suffisamment pour proliférer ; ces dernières sont qualifiées de quiescentes.

Les résultats obtenus à l'aide de ce modèle [7] montrent que la dynamique globale de croissance de la tumeur est correctement reproduite, et que l'on peut observer par ailleurs un dépôt de PAI-1 matriciel tout autour de la tumeur, conformément aux observations biologiques concernant les tumeurs les plus invasives. Le problème est que lorsque les tumeurs simulées atteignent quelques milliers de cellules, il faut alors gérer plusieurs centaines de milliers de molécules de PAI-1. On se trouve ainsi rapidement limités par la taille des tumeurs que l'on peut simuler. La solution proposée consiste à abstraire certains détails du modèle dans les zones où ils ne sont pas utiles, ce qui permet au choix de simuler des tumeurs plus grandes ou d'ajouter des détails dans les zones d'intérêt. Sachant que la zone d'intérêt principale se situe à l'interface entre la couche externe de la tumeur et la matrice extra-cellulaire, nous avons donc proposé de remplacer les couches internes de la tumeur par un modèle agrégé, constitué à la fois de cellules et de molécules de PAI-1.

Le coeur de la tumeur et son modèle orientés agents sont ainsi abstraits par un modèle global de calcul de flux entrant et sortant. Du point de vue spatial, ce modèle est délimité par l'ensemble des cellules nécrosées ou quiescentes, à l'exception d'une couche de cellules quiescentes qui constitue la bordure du modèle agrégé. Ces cellules étant immobiles ou très peu mobiles, cela permet de s'abstraire des mouvements de répulsion entre cellules, ceux-ci étant principalement dus au comportement de division cellulaire dans la couche de cellules actives. Le modèle agrégé intègre toutes les cellules ainsi que toutes les molécules de PAI-1

situées dans le périmètre ainsi défini. Il s'agit alors de calculer les interactions entre ce modèle agrégé et les cellules ou molécules situées à l'extérieur. Il est relativement aisé de déterminer les cellules et molécules qui devront être intégrées au modèle agrégé : dans le cas des cellules, il s'agit d'examiner lesquelles sont quiescentes et entièrement entourées d'autres cellules quiescentes ; concernant les molécules, il suffit d'examiner lesquelles, dans leur mouvement de diffusion aléatoire, vont percuter la frontière du modèle agrégé. Pour déterminer combien de molécules seront au contraire « relâchées » par ce modèle, il est nécessaire d'évaluer la proportion de cellules nécrosées par rapport aux cellules quiescentes, afin de déterminer le nombre de cellules produisant ou internalisant ces molécules. Il est ensuite aisé de mettre à jour la quantité de molécules de PAI-1 à l'intérieur du modèle agrégé, en considérant que l'activité d'internalisation et d'externalisation de ce modèle agrégé est directement proportionnelle à la quantité de cellules quiescentes qu'il contient. En assimilant les molécules diffusant à l'intérieur du modèle agrégé à un gaz parfait, on peut alors calculer la « pression » à l'intérieur du modèle, et donc déterminer la quantité statistique de molécules de PAI-1 devant quitter le modèle agrégé. Celles-ci sont alors distribuées aléatoirement à l'extérieur du modèle agrégé, au voisinage immédiat de la frontière.

2.2 SimulBogota : Logements et Ménages dans la ville de Bogotá

L'objectif du modèle SimulBogota que nous avons proposé dans [6] est de modéliser l'évolution de la distribution spatiale de la population de la ville de Bogotá sur plusieurs décennies. Cette évolution dépend principalement des migrations à l'intérieur et à l'extérieur de la ville ainsi que de l'évolution de la population de ménages et du parc de logements. L'évolution de la population de ménages est le résultat d'évènements individuels (mariage, divorce, émancipation, mort, etc.) qui produisent des changements de niveau socio-économique, ainsi que la création et la disparition de ménages. L'évolution du parc de logement dépend principalement du renouvellement et du vieillissement du bâti. Dans le cas de Bogotá, ces processus sont le résultat de dynamiques très complexes qui font non seulement interve-

nir la construction contrôlée par les politiques d'aménagement, mais aussi des mécanismes informels comme l'auto-construction et le lotissement illégal.

Pour la réalisation de notre modèle, nous disposons uniquement de descriptions socio-économiques des ménages et des logements ainsi que de leur distribution spatiale pour les années 1973 et 1993. En l'absence de données suffisantes, une modélisation centrée sur une représentation explicite de mécanismes individuels de décision et d'évolution s'avère inadéquate voire impossible. Dans ces conditions, le parti pris du modèle que nous avons proposé est alors de considérer des groupes de ménages et des groupes de logements comme entités principales de modélisation. Il s'agit de réduire la complexité des simulations par la prise en compte d'un niveau agrégé et d'y placer le processus de décision conduisant les ménages à déménager. Toutefois, il est nécessaire de considérer les populations de ménages et de logements afin de pouvoir modéliser l'évolution des groupes.

Dans ce modèle nous considérons de manière explicite deux niveaux de modélisation, le *niveau microscopique* qui contient les données qui décrivent la population de ménages et le parc de logements et le *niveau mésoscopique* qui contient les groupes de ménages et les groupes de logements. Le modèle est composé essentiellement de trois mécanismes exécutés en séquence à chaque pas de simulation :

1. Formation de groupes : cela permet de passer du niveau microscopique (ménages et logements) au niveau mésoscopique (groupes de ménages et groupes de logements). Il s'agit d'un mécanisme de classification automatique² effectué sur les données qui représentent les ménages et les lo-

2. L'algorithme de classification utilisé [5] possède les caractéristiques suivantes :

- Il se base sur une mesure de similarité entre données des ménages entre elles et entre données des logements entre elles. Les données des ménages décrivent le statut socio-économique des chefs de ménages (âge, sexe, niveau éducatif, type d'emploi, originaire de, , nombre d'enfants à charge etc.). Pour les logements, les données décrivent la qualité des matériaux de construction, la taille et l'interconnexion aux services publics.
- Le nombre de classes (groupes) est déterminé de manière automatique sur un intervalle donné en paramètre. Le réglage de ce paramètre permet de modifier la granularité des groupes.

Il faut noter que la localisation spatiale ne fait pas partie des variables considérées dans la classification.

gements.

2. Interaction entre groupes : cela permet de relocaliser les ménages dans l'espace. Il s'agit d'un mécanisme basé sur des enchères qui permet l'échange de logements entre groupes de ménages. Ces interactions sont régies par des listes de préférences de logements construites dynamiquement³ et par une matrice statique de coûts de déménagement entre secteurs urbains. Dans ce mécanisme, on fait l'hypothèse que les ménages cherchent à se rapprocher des ménages de leur groupe social et à occuper les mêmes groupes de logements. Il s'agit donc d'un mécanisme de ségrégation spatiale auto-renforcé.
3. Evolution de la population : il s'agit d'un mécanisme basé sur l'exécution de règles globales d'évolution de la population de ménages et de logements. Ces règles permettent la création ou la suppression d'entités au niveau microscopique (ménages ou logements). Chaque règle associe un profil d'entité avec un nombre d'entités à traiter. Ces règles représentent des tendances globales d'évolution et non pas des événements individuels très difficiles à prendre en compte et très gourmands en données.

Pour pouvoir tracer l'évolution des groupes, un dernier mécanisme permet de mettre en relation les groupes trouvés au cours de deux pas de simulation successifs.

Les résultats des simulations ont été validés à partir du calcul de la corrélation entre les distributions spatiales réelle et simulée pour la dernière année de simulation sur des modes statistique et cartographique. Ces résultats démontrent notamment une bonne stabilité des mécanismes développés et une bonne capacité de reproduction de la distribution spatiale de la population à la fin de la période de simulation.

3. Les préférences sont construites à partir de la distribution spatiale des ménages qui composent chaque groupe. Chaque groupe de ménages possède une liste de préférences par secteur urbain qu'il occupe. La préférence d'un groupe de ménages (GM) pour un groupe de logements (GL) dans un secteur urbain (SU) est proportionnelle au nombre de logements de GL occupés par GM dans SU. Les préférences évoluent au gré des changements dans la distribution spatiale des groupes.

2.3 Simpop3 : simuler l'évolution urbaine à trois niveaux d'échelle

Les modèles Simpop, conçus et développés au sein de l'équipe P.A.R.I.S du laboratoire Géographie-Cités, sont des modèles multi-agents de croissance urbaine sur plusieurs siècles. Le premier d'entre eux, le modèle Simpop1 [1] fut l'une des toutes premières applications des SMA à la géographie. Dans cette section nous présentons successivement deux modèles indépendants, SIMPOP2 et SIMPOP-NANO, puis leur couplage en SIMPOP3, un modèle multi-niveaux qui permet d'aborder de nouveaux questionnements géographiques.

SIMPOP2 : simuler la genèse et l'évolution des systèmes de villes SIMPOP2⁴[14] est un projet démarré en 2002 qui s'est donné comme objectifs d'une part de construire un modèle générique encapsulant les processus communs à l'oeuvre dans la croissance des systèmes de villes dans le monde, et d'autre part de raffiner ce modèle en plusieurs modèles appliqués, permettant de reproduire des systèmes particuliers : Europe, Etats-Unis, Inde, Afrique du Sud. Ceux-ci présentent des régularités remarquables (loi rang-taille) tout en ayant connu des morphogénèses sur des temporalités historiques très différentes (émergence continue en Europe et Inde, brusque aux Etats-Unis et en Afrique du Sud). Ils présentent des ordres de grandeur différents (en termes de dimensions spatiales, population, richesse). SIMPOP2 s'appuie sur la théorie évolutionnaire urbaine [13], dans laquelle les systèmes de villes sont analysés comme des systèmes complexes auto-organisés dont les villes sont les composantes principales. La question sous-jacente au projet est la compréhension et l'explicitation des mécanismes qui à la fois ont présidé à la genèse de ces systèmes organisés et hiérarchiques, et qui sont aujourd'hui le moteur de leur évolution.

Les *villes* sont les agents principaux du modèle. Elles interagissent en se vendant et en s'achetant des biens associés aux *fonctions urbaines* qu'elles possèdent. Les interactions spatiales entre villes sont de plusieurs types, et dépendent de l'agent fonction urbaine assurant la communication : proximité spatiale, délimitation spatiale administrative ou réseau privi-

légié à grande échelle. Les villes sont également en compétition pour l'acquisition de l'innovation, qui est modélisée par l'apparition de nouvelles fonctions urbaines au cours de la simulation. Une fonction urbaine caractérise le rôle joué par une ville au sein du système dont elle fait partie, et elle est associée à une certaine période temporelle d'activité (fonction industrie lourde, industrie automobile, nouvelles technologies par exemple, mais aussi fonction capitale ou chef-lieu, etc.) A l'exécution, la dynamique interactionnelle produit une spécialisation et une différenciation progressives entre les villes. Au niveau global, cela se traduit par l'apparition dynamique d'un *système de villes*, dont on caractérise les propriétés par des mesures d'indicateurs-clés (population totale, courbe rang-taille, indices de primatie, etc.) qui peuvent être comparées aux valeurs dont on dispose pour les systèmes réels. Les résultats obtenus sur les modèles Europe et Etats-Unis ont confirmé la pertinence à la fois d'une lecture des dynamiques des systèmes de villes à l'échelle des villes, et ont aussi confirmé le potentiel explicatif des mécanismes intégrés dans le modèle. Utilisé comme un filtre, le modèle permet d'isoler les processus génériques des processus géographiques et historiques propres à chacun des systèmes simulés⁵.

SIMPOP-NANO : simuler l'évolution de l'organisation interne d'une ville sur le long-terme SIMPOP-NANO [9] fait suite à SIMPOP2 et a pour objectif de reproduire l'évolution de l'organisation spatiale d'une ville sur la même période que celle de SIMPOP2.

Le modèle intègre deux types d'agents : des agents *quartiers*, qui représentent des portions de l'espace intra-urbain, et des agents *fonctions* qui représentent les types d'activités socio-économiques possédées par la ville, et qui sont les mêmes que celles qui sont présentes dans SIMPOP2. Les *fonctions* possèdent des actifs à localiser dans la ville, ainsi qu'un solde économique permettant de payer pour leur localisation. Les quartiers sont reliés dans un réseau à structure dynamique et se différencient par leur accessibilité (qui traduit leur centralité au sein du réseau et leur niveau de desserte par les transports) ainsi que par leur composition

4. <http://www.simpop.parisgeo.cnrs.fr>

5. <http://www.simpop.parisgeo.cnrs.fr/theInstantiatedModels/introduction.php>

fonctionnelle. Cela les rend plus ou moins attractifs pour les *fonctions*, mais également plus ou moins chers. Les *fonctions* se différencient par leurs moyens économiques et leur stratégie de localisation. La dynamique interactionnelle entre ces agents produit à l'exécution une organisation spatiale, une ville *émergente*, qu'il est possible de caractériser au moyen de plusieurs indicateurs : gradients centre-périphérie de densité d'occupation du sol et de prix, cartes de répartition spatiale des activités, indices de spécialisation fonctionnelle, etc. La comparaison des valeurs de ces indicateurs aux données issues des bases donne une mesure de la capacité du modèle à produire des structures cohérentes, ressemblant aux formes typiques des villes européennes ou nord-américaines.

Simpop3 : coupler deux modèles pour le passage à l'échelle Avec SIMPOP3, l'objectif est de coupler les deux modèles précédents en un système multi-agent « multi-couches » qui permette d'étudier des dynamiques urbaines simultanément à trois niveaux de l'échelle spatiale : le quartier, la ville, et le système de villes. L'idée géographique sous-jacente est que ces niveaux exercent bien sûr les uns sur les autres des influences réciproques. Etant donné la multiplicité des facteurs causant ces influences inter-niveaux, il est difficile de formaliser et quantifier leur action et de la traduire en règles inter-niveaux. Le modèle sert alors à tester des hypothèses sur la nature de ces liens d'influence qu'entretiennent ces trois niveaux d'organisation spatiale.

SIMPOP3 inclue donc des agents à trois niveaux d'organisation spatiale :

- Au niveau le plus « bas », niveau *micro*, sont spécifiés des *agents quartiers* et les *agents fonctions* de SIMPOP2.
- Au niveau intermédiaire, niveau *meso*, on spécifie des *agents villes* et des *agents fonctions* qui sont ceux de SIMPOP2. Les *agents fonctions* sont les mêmes que ceux de SIMPOP2.
- Enfin au niveau du système de villes, niveau *macro*, on spécifie une entité qui contient des paramètres globaux. Ceux-ci sont nécessaires au vu de la période considérée. Ils s'appliquent à tous les agents de niveau inférieur et, sans influencer sur la différenciation produite par les interactions, injecte des taux de croissance démographique et économique

globaux qui assurent que les valeurs des variables soient d'un ordre de grandeur comparable aux données.

L'exécution du modèle consiste en l'exécution alternée d'une itération de SIMPOP2 et d'une itération de SIMPOP2.

- SIMPOP2 calcule à chaque pas de temps les interactions entre villes et détermine en conséquence l'état suivant de chacune : les fonctions qu'elles acquièrent, le nombre d'employés et la richesse de chacune.
- Ces sorties sont les données d'entrée de SIMPOP2, qui calcule la ventilation de ces effectifs des fonctions au sein des quartiers de la ville. Les fonctions payent leur implantation sur les quartiers grâce à l'argent généré au niveau SIMPOP2, lors des échanges avec d'autres villes. En sortie simpop2 produit un indicateur de la qualité de l'implantation des fonctions à l'intérieur de la ville.
- Cet indicateur est utilisé par les agents-fonctions au niveau SIMPOP2 pour décider de la ville où elles vont s'implanter

Le résultat de cette intégration de modèles est un nouvel artefact, permettant de tester des hypothèses sur les mécanismes assurant le « passage » entre les niveaux. Dans quelles proportions l'organisation spatiale interne d'une ville influe-t-elle sur sa position hiérarchique au sein du système de villes auquel elle appartient ?

3 Vers des organisations multi-agents multi-niveaux

Après avoir rendu compte de différentes expériences de modélisation à base d'agents à différents niveaux d'organisation ou échelles, il s'agit maintenant d'essayer de les systématiser. L'objectif est de proposer une grille d'analyse de ces modèles multi-échelles en vue de faciliter leur conception et leur mise au point. Nous examinerons ensuite comment ces différentes visions du multi-échelles peuvent s'articuler.

3.1 Des niveaux et leur couplage dans les modèles étudiés

Avant de proposer une grille d'analyse, nous allons revenir sur les trois modèles présentés pour répondre aux questions suivantes :

1. *Spécification des niveaux* : quels sont les différents niveaux identifiés, quels sont les

types d'objets et/ou d'agents définis à ces niveaux, et comment ces types sont-ils définis (sont-ils spécifiés a priori ou bien sont-ils découverts de manière dynamique ?) ;

2. *Instanciation des objets* : comment l'instanciation des objets et/ou agents est-elle effectuée (celle-ci s'effectue-t-elle de manière statique en définissant explicitement les agents du modèle ou bien ceux-ci sont-ils créés de manière dynamique dans la simulation en fonction du contexte ?) ;
3. *Couplage entre les niveaux* : de quelle manière les différents niveaux sont-ils couplés (de quelle nature sont les interactions entre agents de niveaux différents ?).

Tumeur cancéreuse et ceinture de PAI-1. Dans ce modèle, l'aspect multi-niveaux peut-être vu de deux manières complémentaires. La première est de constater que le modèle, composé de cellules d'une part, de protéines d'autre part, intègre des objets appartenant à des échelles très différentes. La deuxième est de proposer que des objets correspondant à des abstractions de plus haut niveau d'organisation puissent être introduits dynamiquement dans la simulation. L'introduction d'un modèle agrégé pour prendre en compte le coeur de la tumeur rentre ainsi dans cette catégorie.

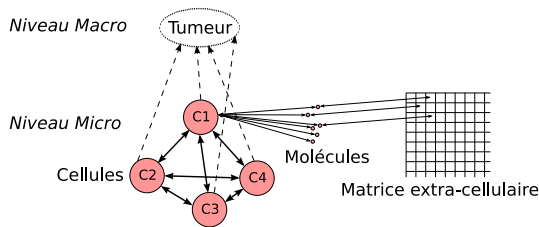


FIGURE 1 – Modélisation d'une tumeur à partir d'entités cellules et molécules

Dans le premier cas (voir figure 1), les niveaux correspondent aux cellules et aux molécules, ces types étant définis a priori. L'instanciation s'effectue en créant une cellule initiale, qui elle-même crée d'autres cellules par divisions successives, et produit et externalise des molécules de PAI-1. Les niveaux cellulaire et moléculaire sont couplés du fait de l'activité des cellules : les cellules produisent des molécules de PAI-1 avec une certaine vitesse ; dans le même temps, elles internalisent ces mêmes molécules avec une vitesse qui dépend du nombre de récepteurs disponibles sur la membrane de la cellule.

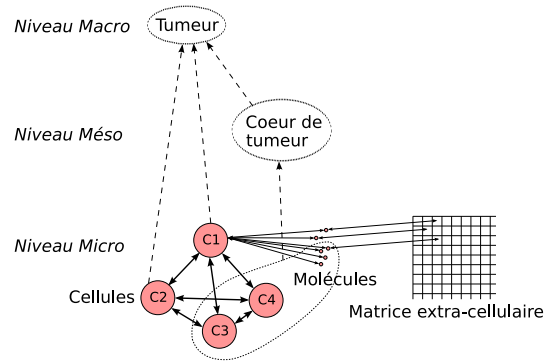


FIGURE 2 – Introduction d'un modèle agrégé pour le coeur de la tumeur

Il n'y a pas jusque là de passage d'échelle mais seulement l'interaction, dans un environnement commun, entre objets d'échelles différentes. Dans le second cas en revanche (voir figure 2), on identifie en cours de simulation une entité de niveau supérieur qui est réifiée en tant que telle. Il y a donc l'introduction d'un niveau supplémentaire de modélisation, celui d'un tissu multicellulaire correspondant au coeur de la tumeur. Ce type d'objet est spécifié a priori : on sait de manière certaine que la simulation du modèle va conduire au développement d'une tumeur et l'on peut donc anticiper l'apparition de cet objet. Son instanciation par contre dépendra du moment où un tel type d'objet pourra être observé dans la simulation. La dynamique de cet objet agrégé peut alors être vue, d'une certaine manière, comme la résolution numérique d'équations différentielles très simples, décrivant l'évolution des quantités de cellules respectivement nécrosées et quiescentes, l'évolution de la quantité de PAI-1 à l'intérieur du modèle, ou encore la quantité de PAI-1 externalisée.

SimulBogota. Dans ce modèle, afin de réduire la complexité des simulations et la quantité des données nécessaires pour représenter les dynamiques migratoires intra-urbaines, nous avons introduit un niveau intermédiaire composé par des structures « artificielles » : des groupes formés d'entités microscopiques semblables. Il s'agit de structures qui n'ont pas d'équivalent physique et qui résultent de l'analyse des données de ménages et de logements (niveau microscopique). Ce caractère « artificiel » des groupes conditionne la mise en place de mécanismes d'interaction entre groupes et des mécanismes d'interaction inter-niveaux (voir figure 3).

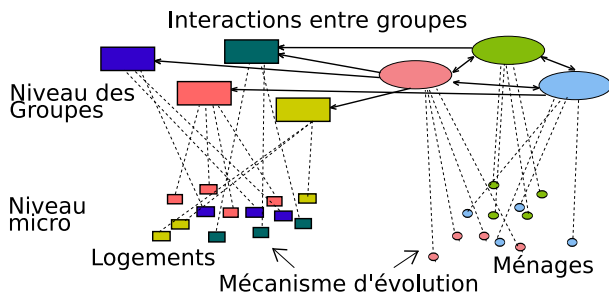


FIGURE 3 – Diagramme des interactions entre niveaux dans la simulation de l'évolution de la distribution spatiale de la population dans SimulBogota

Contrairement à la modélisation d'entités « physiques », dans ce cas-là, et en l'absence de règles de comportements de groupe bien établies ou facilement déductibles, le modélisateur est en effet *libre* dans la définition de mécanismes d'interaction. Egalement, par rapport aux modèles de phénomènes physiques, la question supplémentaire du niveau d'abstraction se pose : il s'agit de définir le nombre de groupes à considérer, sachant qu'un nombre réduit peut produire une sur-simplification du modèle, tandis qu'un nombre élevé peut rapprocher excessivement les groupes du niveau microscopique et entraîner la perte de leur capacité de synthèse.

Dans le cas présenté ici, le niveau des groupes ainsi que les mécanismes d'interaction entre groupes ont été *spécifiés* au préalable. Toutefois les groupes sont *instanciés dynamiquement* à partir de la classification des données de ménages et logements à chaque pas de simulation. Le choix de mécanismes de passage entre les niveaux (par exemple l'algorithme de classification utilisé pour passer du niveau micro au méso) et du niveau d'abstraction a par contre été guidé par le mécanisme de validation du modèle. Ces paramètres ont été définis en fonction de leur capacité à reproduire, le plus fidèlement possible, la distribution spatiale de la population dans l'année finale de simulation.

Etant donné que l'on connaît très peu sur le fonctionnement des groupes, le mécanisme choisi pour représenter l'interaction entre les groupes est aussi artificiel que les groupes eux-mêmes. Le mécanisme proposé basé sur les enchères de logement n'a pas de fondement réel, il s'agit simplement d'un artefact qui permet l'appariement dynamique et multi-critère entre

groupes de ménages et groupes de logements. Le choix de ce mécanisme ne relève que de l'intuition des modélisateurs. Grâce à ce mécanisme, les ménages sont rélocalisés, le passage entre le niveau des groupes et le niveau microscopique est ainsi représenté.

Du point de vue de l'évolution, des règles générales ont été construites à partir des données descriptives des ménages et des logements. Le moteur d'évolution est donc placé au niveau microscopique et ses effets sont propagés aux groupes par le mécanisme de passage entre le niveau microscopique et celui des groupes.

Dans ce modèle, il n'existe pas d'interaction directe entre entités microscopiques, les entités *actives* du système sont donc les groupes. Le niveau microscopique fait office de *point d'ancrage avec la réalité* tandis que le niveau de groupes offre *une vision synthétique de la réalité*, ce qui facilite la formulation et l'évaluation d'hypothèses sur le comportement du système permettant donc de mieux comprendre le phénomène modélisé.

SimPop3. Dans SIMPOP2 le niveau de spécification du modèle, celui où la connaissance géographique est injectée dans les agents et règles, est celui des villes : les villes sont des entités indivisibles, ce sont les agents du modèle. Le niveau système de ville constitue le niveau émergent. SIMPOP NANO travaille quant à lui à l'échelle inférieure : le niveau de spécification est celui des quartiers, dans lesquels doivent se localiser *ces mêmes fonctions urbaines* acquises par la ville au niveau système. Le niveau ville constitue le niveau émergent. Il ne s'agit donc pas de la « même » ville que celle de SIMPOP2 : celle de SIMPOP NANO est une entité qui n'existe que dynamiquement, à travers une superposition de mesures qui permettent au géographe de décider si la structure émergente de son modèle ressemble ou non à une ville. Pour SIMPOP3 nous avons donc été confrontés au besoin de construire un pont entre la ville spécifiée de Simpop2 et celle, émergente, de Simpop Nano. Ce besoin est illustré par la figure 4.

Pour réaliser cette articulation entre les deux modèles, nous avons choisi d'identifier une description commune aux deux niveaux, qui représente à la fois un effet de la dynamique du système des villes sur la ville, et qui puisse exprimer une dynamique interne à la ville sus-

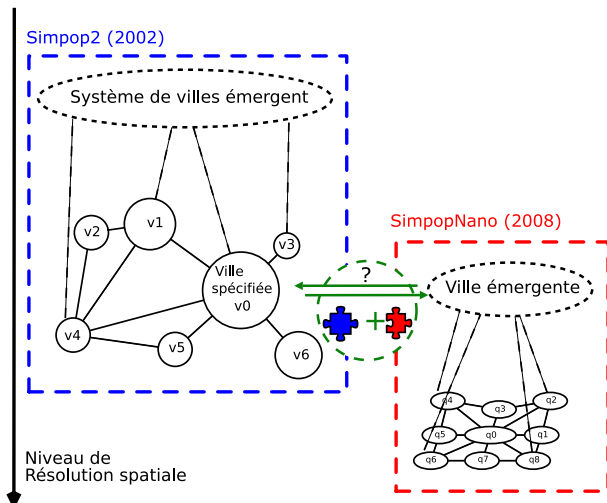


FIGURE 4 – SIMPOP3 : un SMA multi-niveaux pour tester des théories urbaines au passage d'échelles. Comment articuler l'agent ville *spécifié* de SIMPOP2 avec la ville *émergente* de SIMPOP3 ?

ceptible d'avoir des effets en retour sur le système. La notion de *fonction urbaine* a donc été instanciée dans le programme en agent « inter-échelles ». La fonction urbaine est à l'origine un concept abstrait défini par les géographes pour qualifier le rôle spécifique joué par une ville au sein du système, et qui trouve donc sa pertinence à l'échelle du système de villes. Pour réaliser le couplage, la notion a été infléchiée et l'hypothèse a été faite qu'elle pouvait être utilisée pour qualifier les types d'activités en interaction à l'intérieur de la ville. C'est en fait ce transfert d'un concept de l'échelle inter-urbaine à l'échelle intra-urbaine, qui permet d'utiliser l'objet fonction urbaine comme un agent « tampon » permettant de faire le lien entre les règles des deux modèles, et qui a permis de réaliser ce couplage.

3.2 Différentes approches de modélisation multi-niveaux

A partir des trois exemples présentés, ainsi qu'un ensemble d'autres modèles issus de la littérature, on peut tenter une catégorisation des modèles multi-échelles autour de trois aspects :

1. Le couplage de plusieurs modèles à l'origine indépendants : on considère des modèles à des échelles différentes en interac-

tion. Les sorties des modèles d'échelle supérieure servent à définir des paramètres globaux des modèles d'échelle inférieure et les sorties des modèles d'échelle inférieure sont considérées comme des flux d'entrée dans les modèles d'échelle supérieure. On peut distinguer encore deux sous-catégories :

- (a) les différents modèles s'exécutent de manière alternée (c'est le cas par exemple de Simpop3). Dans ce cas, le couplage entre les modèles est généralement faible mais des difficultés peuvent néanmoins apparaître si les différents modèles partagent des objets qu'ils peuvent chacun modifier (l'entité ville dans le cas de SIMPOP3).
 - (b) un modèle multi-agent est encapsulé à l'intérieur d'un agent de niveau supérieur. Si l'on pense au modèle de tumeurs, cela correspondrait au fait que l'on modélise la dynamique interne des cellules, plutôt que de considérer une production à taux fixe de molécules de PAI-1. Cela permettrait alors de prendre en compte les réseaux de régulation génétique contrôlant la production de PAI-1. Il faut faire attention, dans ce cas, à l'ordonnement de l'activité des agents aux différents niveaux, qui n'intervient pas nécessairement à la même échelle temporelle.
2. La cohabitation et l'interaction d'agents situés à des échelles spatiale, sociale ou temporelle différentes au sein d'un même modèle. On peut y parvenir d'au moins deux façons différentes :
 - (a) par spécification statique des entités simulées (ex. : des cellules qui interagissent avec des molécules dans le modèle de tumeur ; des personnes qui interagissent avec des quartiers dans le modèle Accessim [2], SIMPOP3). A nouveau apparaissent des problèmes liés à l'ordonnement des agents. Dans le cas de cellules et de molécules, les échelles de temps caractéristiques sont en effet très différentes. Cela impose de privilégier un mode de simulation à événements discrets, ou encore de gérer en parallèle plusieurs ordonneurs pour les différents types d'objets simulés.

- (b) par spécification dynamique des entités simulées : implémentation de mécanismes de type « loupe dynamique » qui permettent d'augmenter le niveau de résolution spatiale sur des zones critiques. (ex. : le modèle fractal de l'environnement proposé par [11]). La difficulté principale dans ce cas réside dans l'adaptation continue nécessaire pour toujours conserver le plus grand niveau de détail sur les zones d'intérêt, par exemple celles où se trouvent des agents.
3. La construction automatique d'agents d'échelle supérieure : il s'agit de modèles qui produisent des agents de niveau supérieur par « observation » automatique de la simulation, caractérisation de structures de plus haut niveau et réification de ces structures sous forme d'objets. On peut là encore distinguer deux cas distincts :
- (a) les agents des deux niveaux n'interagissent pas directement les uns avec les autres (ex. : simulBogota) car les entités des deux niveaux ne sont pas dans le même espace de modélisation. D'une certaine manière, ce cas ressemble au cas 1a, puisque l'on aura également une alternance entre l'exécution de modèles aux niveaux micro et macro : le simulateur exécute en effet en boucle l'évolution du modèle au niveau micro puis son analyse au niveau macro, puis l'évolution du modèle au niveau macro.
 - (b) les agents des deux niveaux peuvent interagir entre eux et avec les agents de l'autre niveau : les entités des deux niveaux sont dans le même espace de modélisation (ex. : modèle de tumeur avec modèle agrégé, RIVAGE [15], ou encore [3] qui se base sur des approches hiérarchiques de type holoniques). Cette approche, sans doute la plus riche, entraîne cependant une plus grande complexité puisqu'elle nécessite de pouvoir à la fois détecter l'apparition d'une structure émergente, caractériser sa dynamique et ses interactions avec les autres agents de manière à pouvoir la réifier, vérifier que la simulation reste valide, et enfin contrôler que les conditions de maintien de

la structure émergente sont vérifiées, faute de quoi, elle devra disparaître en tant que telle pour réapparaître sous forme d'agents individuels.

4 Conclusion

Nous avons présenté dans cet article trois modèles à base d'agents qui peuvent tous, d'une manière ou d'une autre, être qualifiés de « multi-niveaux ». Ces modèles sont pourtant très différents les uns des autres, tant du point de vue du domaine d'étude et des échelles modélisées que des approches mises en oeuvre pour gérer la complexité entraînée par la gestion de différents niveaux d'organisation. Nous avons proposé une grille d'analyse permettant la classification de différents types de modèles multi-niveaux, à l'intérieur de laquelle nous avons positionné nos modèles. Celle-ci demande à être complétée et raffinée en y intégrant d'autres travaux de la communauté.

Ce faisant, nous ne prétendons pas proposer une méthodologie de conception de modèles multi-niveaux, qui puisse être réutilisée, étape par étape, pour la conception d'un nouveau modèle. L'objectif est davantage de montrer d'une part que différentes problématiques de modélisation peuvent conduire à des solutions très différentes et d'autre part d'aider le concepteur à identifier les points durs de chacune des approches. D'autre part, l'identification de ces difficultés nous conduit à proposer quelques pistes de réalisation de couplage inter-niveaux et orientés agent. Leur pertinence a été établie pour les trois exemples étudiés, ainsi que pour d'autres issus de la littérature.

Lorsque des modèles existent dans un domaine particulier, et ont prouvé leur efficacité, il est à la fois utile et important de pouvoir les réutiliser, ce qui implique de permettre le couplage de modèles à différents niveaux ou fondés sur des formalismes hétérogènes (discret/continu par exemple). Pour ce faire, d'une part des modèles microscopiques peuvent être encapsulés sous formes d'agents, d'autre part des modèles continus (par exemple des modèles à base d'équations différentielles) peuvent servir dans la définition de variables et dynamiques globales utilisables comme entrée de modèles d'échelle supérieure.

Et parce que les systèmes complexes évoluent de manière dynamique, il est également important que le découpage en agents puisse lui aussi évoluer dynamiquement. Pour ce faire, il est nécessaire de permettre la détection [12] et la réification automatique de structures émergentes, et réciproquement de permettre le découpage d'agents en organisations sous-jacentes, de manière interactive. En adaptant les niveaux de description du système, il ne s'agit pas tant de gagner en performances computationnelles, que de gagner en pouvoir explicatif en sélectionnant les niveaux d'organisation et de description pertinents, pour obtenir la plus grande compréhension possible des mécanismes à l'oeuvre dans le système complexe étudié.

Références

- [1] S. Bura, F. Guerin-Pace, H. Mathian, D. Pumain, and L. Sanders. Multi-agents system and the dynamics of a settlement system. *Geographical Analysis*, 28(2) :161–178, 1996.
- [2] M. Delâge, F. Le Néchet, and T. Louail. Accessim : modélisation et simulation d'accessibilité et d'inégalités géographiques dans la ville. *Mappemonde*, 92(4), décembre 2008.
- [3] N. Gaud, S. Galland, F. Gechter, V. Hilaire, and A. Koukam. Holonic multilevel simulation of complex systems : Application to real-time pedestrians simulation in virtual urban environment. *Simulation Modelling Practice and Theory*, 16(10) :1659 – 1676, 2008. The Analysis of Complex Systems.
- [4] J. Gignoux, I. Davies, and D. Hill. 3worlds : a new platform for simulating ecological systems. In *1st Open Inter. Conf. on Modelling and Simulation*, pages 49–64, Clermont-Ferrand, 2005.
- [5] J. Gil-Quijano and M. Piron. Formation automatique de groupes d'agents sociaux par techniques d'apprentissage non supervisé. In *Actes de EGC'07 Atelier Fouille de Données et Algorithmes Biomimétiques*, Namur Belgique, 23 Janvier 2007.
- [6] J. Gil-Quijano, M. Piron, and A. Drogoul. *Mechanisms of automated formation and evolution of social-groups : A multi-agent system to model the intra-urban mobilities of Bogotá city*, chapter 12, pages 151–168. Idea Group Inc., 2007.
- [7] J. Lepagnot and G. Hutzler. A multi-scale agent-based model for the simulation of avascular tumor growth. *Jour. of Biol. Phys. and Chem.*, 9 :17–25, March 2009.
- [8] A. Lesne. Biologie des systèmes - l'organisation multi-échelle des systèmes vivants. *Medecine/Science*, 2009.
- [9] T. Louail. Can geometry explain socio-economical differences between US and european cities ? In *Agent Based Spatial Simulation workshop*, Paris, Nov. 2008. Institut des Systèmes Complexes.
- [10] F. Maquerlot, S. Galiacy, M. Malo, C. Guignabert, D. A. Lawrence, M.-P. d'Ortho, and G. Barlovatz-Meimon. Dual role for plasminogen activator inhibitor type 1 as soluble and as matricellular regulator of epithelial alveolar cell wound healing. *Am. J. Pathol.*, 169 :1624–1632, 2006.
- [11] N. Marilleau, C. Cambier, A. Drogoul, E. Perrier, J.L. Chotte, and E. Blanchart. Environnement multi-échelles à base de fractales pour la modélisation agent d'écosystèmes. In *JFSMA'08*, Brest, 2008.
- [12] T. Moncion. *Modélisation de la complexité et de la dynamique des simulations multi-agents. Application pour l'analyse des phénomènes émergents*. Thèse de doctorat, Université d'Evry, 2008.
- [13] D. Pumain. Systèmes de villes et niveaux d'organisation. In P. Bourguine and A. Lesne, editors, *Morphogenèse. L'origine des formes*, chapter 13, pages 238–263. Belin, 2006.
- [14] D. Pumain, L. Sanders, A. Bretagnolle, B. Glisse, and H. Mathian. The future of urban systems. In D. Lane, D. Pumain, S. Van der Leeuw, and G. West, editors, *Complexity perspectives on innovation and social change*, Methodos Series, chapter 12. Springer, 2009.
- [15] D. Servat, E. Perrier, J.-P. Treuil, and A. Drogoul. When agents emerge from agents : Introducing multi-scale viewpoints in multi-agent simulations. In J.S. Sichman, R. Conte, and N. Gilbert, editors, *MABS*, volume 1534 of *LNCS*, pages 183–198. Springer, 1998.
- [16] J.-P. Treuil, A. Drogoul, and J.-D. Zucker. *Modélisation et simulation à base d'agents*. Dunod, 2008.