

COMPRENDRE LES DYNAMIQUES URBAINES : L'UTILISATION DE BASES DE DONNÉES TOPOGRAPHIQUES ET LA CRÉATION DE BASES DE DONNÉES SPATIO-TEMPORELLES

par Julien Perret, Anne Ruas

*Laboratoire COGIT, Institut géographique national
73 avenue de Paris, 94160 Saint-Mandé
Courriel : Julien.Perret@ign. ; fr Anne.Ruas@ign.fr*

et Annabelle Boffet Mas

*Laboratoire CEDETE, Université d'Orléans
10 rue de Tours, BP 46527, 45065 Orléans cedex 2, France
Courriel : Annabelle.Mas@univ-orleans.fr*

Comprendre les dynamiques urbaines est un défi d'une importance croissante qui peut être traité de façons différentes. Pendant ces dernières années, une littérature importante a couvert le sujet dans des domaines de recherche tels que la géographie urbaine et économique, la géomatique. Néanmoins, quel que soit le domaine de recherche, afin de mieux comprendre les dynamiques urbaines, les changements effectifs subis par les objets géographiques au cours du temps doivent être étudiés. La conduite d'une telle étude nécessite la constitution de bases de données spatio-temporelles contenant des données pour chaque étape des évolutions étudiées. La construction de ces bases de données est malheureusement un processus difficile qui requiert le suivi des objets géographiques dans les bases de données consécutives par la création de liens. Les évolutions des objets topographiques suivis peuvent être étudiées et ensuite simulées. Chaque objet est appelé agent. Les simulations effectuées à l'aide d'agents peuvent alors être exploitées afin de mieux comprendre les évolutions spécifiques prenant place sur un territoire donné et, à terme, afin d'aider l'aménagement du territoire et la prise de décision. La méthodologie présentée dans cet article comprend la construction de bases de données spatio-temporelles, la saisie de données topographiques historiques à partir de cartes topographiques, et un processus d'enrichissement des bases de données par la création de hiérarchies spatiales. Enfin sont présentées les bases de notre plate-forme de simulation à partir de données vectorielles.

1 Introduction : le projet GeOpenSim

Cet article présente le projet de recherche GeOpenSim, qui vise la conception d'un système open source permettant l'étude des évolutions de l'espace urbain. L'originalité de l'approche proposée concerne l'utilisation de bases de données topographiques vectorielles pour l'analyse de motifs d'évolution, et l'utilisation des fonctions d'évolution identifiées pour simuler les évolutions de l'espace à l'aide d'un système multi-agents. Dans cette

approche, la simulation est aussi utilisée afin de valider les connaissances identifiées sur les évolutions du territoire. Le projet GeOpenSim, financé par l'ANR, fait l'objet d'une collaboration entre quatre laboratoires de recherche français : deux laboratoires de géographie (le LIVE de Strasbourg et le CEDETE d'Orléans), un laboratoire d'informatique (le LSIIT de Strasbourg) et un laboratoire de géomatique (le COGIT de l'IGN qui coordonne le projet). Les développements du projet sont effectués sur la plate-forme GeOxygene, un SIG open source développé par le laboratoire COGIT.

1.1 Un aperçu des modules du projet GeOpenSim

Le système GeOpenSim doit tout d'abord permettre l'analyse de données topographiques afin de construire des fonctions d'évolution réalistes. En effet, les fonctions d'évolution sont construites par analyse de données topographiques à différentes temporalités et procurent des règles pour la simulation de l'espace. En comparant des jeux de données importants, nous cherchons à créer des règles de la forme suivante : *Si tel espace géographique possède telles propriétés, alors X% des objets de tel type sont susceptibles de changer de telle façon pendant telle période de temps.* Un instance de règle pourrait ainsi être : *dans une zone pavillonnaire, 15% de nouveaux pavillons sont construits entre 1980 et 2000.*

Une simulation est caractérisée par un espace géographique et une période de temps. Si le début de la simulation est dans le passé, le résultat de la simulation peut alors être comparé avec la réalité. Dans le cas contraire, la simulation constitue une nouvelle représentation de l'espace géographique qui doit être aussi réaliste que possible.

La construction d'un tel système repose sur un schéma de données spécifique adapté à la simulation urbaine et un jeu de six modules. Le schéma de données utilisé par GeOpenSim permet la représentation de bases de données spatio-temporelles et leur simulation. Nous distinguons, dans un premier temps, la représentation d'une entité géographique comme un agent possédant une durée de vie, et la représentation d'objets topographiques qui sont des instantanés des agents à un instant donné. Ainsi, chaque entité géographique est représentée par un agent géographique et possède un ensemble de représentations topographiques (cf. section 3). Ce schéma inclut, par ailleurs, des objets méso tels que les îlots construits à l'aide de fonctions spécifiques (cf. module 2 ci-dessous).

1.2 Les modules du projet GeOpenSim

Module 1, Chargement des données : Ce module charge les données d'une base de données topographique dans le schéma de données GeOpenSim (cf. section 3).

Module 2, Analyse spatiale et enrichissement des données : Ce module crée les objets méso tels que les aires urbaines, les îlots et les groupes de bâtiments à partir des données initiales. Il permet aussi de caractériser les objets micro et méso à l'aide d'outils d'analyse spatiale. Par exemple, les îlots sont caractérisés par leur densité, leur taille, leur

forme et la nature de leurs bâtiments. La caractérisation est utilisée pour associer la fonction d'évolution adaptée à une situation urbaine donnée. Elle repose sur des outils d'analyse classiques ainsi que sur des algorithmes d'apprentissage supervisés pour la classification, notamment des îlots urbains.

Module 3, Simulation : Ce module permet la simulation des évolutions urbaines à une période et pour une durée données. Il utilise un schéma de données temporelles, des règles d'évolution construites à l'aide du module 5 et des fonctions de peuplement permettant de modifier l'espace topographique. Une fonction de peuplement peut, par exemple, servir à ajouter une nouvelle maison dans un îlot pavillonnaire tout en respectant l'organisation existante de l'îlot.

Module 4, Construction de bases de données spatio-temporelles (BDST) : Afin de construire des fonctions d'évolution, les données topographiques à différentes dates sont analysées. Comme de telles bases de données existent rarement, nous proposons des méthodes pour les construire à partir de bases de données existantes et d'autres sources de données géographiques telles que des photographies aériennes et des cartes.

Module 5, Construction de règles d'évolution : Une fois que les bases de données spatio-temporelles sont construites, ce module analyse les changements tels que la densification des îlots et le prolongement des routes. Des algorithmes d'apprentissage sont alors utilisés pour analyser et généraliser les exemples afin de produire des règles d'évolution.

Module 6, Évaluation des états : L'objectif de la simulation est d'étudier (et ainsi d'améliorer) nos connaissances sur les évolutions temporelles de l'espace urbain. Il est ainsi nécessaire de développer un module qui permette la comparaison des états : soit la comparaison des résultats de la simulation avec des données réelles, soit les résultats de plusieurs simulations effectuées avec des règles d'évolution différentes.

La section 2 présente la méthode que nous proposons afin de construire des bases de données spatio-temporelles (module 4). La section 3 décrit notre approche multi-agents pour la simulation (modules 2 et 3). Les conclusions et perspectives de ce travail sont discutées en section 4. Les règles d'évolution, l'évaluation des états et les fonctions de peuplement ne sont pas présentées dans cet article et feront l'objet de prochains papiers.

2 Création de bases de données spatio-temporelles à partir de bases de données topographiques existantes

Cette partie présente la méthodologie de création de la base de données spatio-temporelle (STDM) à l'échelle de la ville. De nombreux travaux de recherche dans le domaine :

(<http://www.fas.harvard.edu/~chgis/work/design/>) s'appuient sur l'analyse des évolutions des unités administratives à partir de données statistiques et/ou sur l'évolution de la géométrie de ces unités administratives. Le projet GeOpenSim vise à décrire le tissu urbain par l'analyse de tous les objets géographiques qui composent la zone d'étude, et ce, à différentes échelles d'analyse. Nous ne travaillons pas avec des données socio-économiques, mais avec des données topographiques (base de données et cartes papier) produites par l'Institut géographique national. Des données d'autres origines auraient pu être intégrées, mais elles ne le sont pas par souci de cohérence.

2.1 Cartes et données existantes

L'IGN produit, entre autres, la BD Topo®, une base de données géographique vectorielle de précision métrique adaptée aux échelles allant du 1 : 5 000 au 1 : 25 000. Cette base de données est structurée en couches thématiques : *axes de communication (routes, chemins), voies ferrées, transport d'énergie (réseau électrique), hydrographie (rivières, lacs), bâtiments, relief (courbes de niveau), limites administratives, végétation*. Depuis 1996, les cartes au 1 : 25 000 sont produites à partir de la BD Topo®. Avant, les cartes papier au 1 : 25 000 constituaient la meilleure source de données IGN (en terme de précision géométrique).

Une des deux zones test du projet GeOpenSim est centrée sur Orléans, dont la base de données topographique a été mise à jour en 2007. Les données historiques ont donc été collectées à partir de cartes papier, il s'agit des années 1957, 1978, 1989 et 1999 (pour Orléans).

2.2 Création de la base de données spatio-temporelle (BDST)

Ce paragraphe présente la méthode de création de la BDST à partir de cartes et de base de données topographiques existantes. La méthode doit être simple, rapide et facilement reproductible dans la mesure où la création de la BDST n'est pas l'objectif final du projet GeOpenSim. La méthode s'appuie sur l'information disponible la plus récente : la dernière BD topographique (celle de 2007 dans notre cas). La

BDST est créée par un processus manuel qui remonte dans le temps : une copie de la BD de 2007 est modifiée en fonction de la carte de 1999 par suppression, création, modification des objets topographiques (les bâtiments et les routes dans un premier temps). Chaque BDST est créée séquentiellement à partir de la BDST qui a été créée précédemment, identifiant ainsi, pas à pas, les modifications de chaque objet entre deux dates successives.

Une autre alternative à ce processus séquentiel est la possibilité de créer systématiquement chaque BDST à partir de la référence la plus récente (2007). Le processus de production devient donc plus rapide (chaque BDST peut être produite simultanément) et évite la reconduction d'erreur. Néanmoins, ce processus simultané nécessite des post-traitements afin de reconstruire les liens entre les objets similaires présents dans les différentes versions temporelles intermédiaires.

2.3 Suivi des objets dans le temps

La création de liens entre objets géographiques de chaque version temporelle est nécessaire afin de suivre l'évolution de chaque objet de la ville. Dans le cas d'une unique base de données décrivant le type d'évolution et la succession des objets dans le temps (Mas 2008), les liens sont gérés par un objet virtuel auquel différentes géométries et différents statuts sémantiques sont rattachés. À l'opposé, l'approche proposée dans cet article s'appuie sur la multiplicité des BDST dans le but de les enrichir par de nouvelles données rapidement et facilement ; aucun lien n'existe a priori dans les différentes versions, il faut les créer. Ces liens sont basés sur l'identifiant (*idGeo*) donné aux objets de la BD de 2007. Du point de vue de la modélisation agent, cet *idGeo* est unique pour chaque objet (agent) durant sa vie : un bâtiment construit en 1980, qui existe toujours en 2007 aura le même *idGeo* dans toutes les BD successives (1987, 1999 et 2007).

Bien que Hornsby et Egenhofer (1997) présentent une classification complète des changements dans le temps et l'espace dans les SIG, six cas sont gérés dans le projet. Une attention particulière doit être portée à la nomenclature associée, d'une part, aux événements de la vie des agents et, d'autre part, au processus de création de la BD. Entre 1999 et 2007, les évolutions possibles sont (fig. 1) :

Stabilité d'un bâtiment : l'objet est le même dans les 2 BD – *idGeo* est conservé.

Création d'un bâtiment : l'objet est détruit sur la copie de la BD 2007 pour créer la BD 1999.

Destruction d'un bâtiment : un objet est créé

dans la copie de la BD 2007 pour générer la BD 1999 - Un nouvel *idGeo* voit le jour.

Modification d'un bâtiment : l'objet est manuellement modifié dans la copie de 2007 pour créer la BD 1999 (le type de modification est calculé en post-traitement par comparaison des géométries). Comme l'objet existe toujours, l'*idGeo* est conservé.

Restructuration d'ordre 1-n : (substitution d'un habitat collectif à plusieurs habitations individuelles) plusieurs destructions et une création d'objet sont réalisées sur la copie de la BD 2007 pour créer la BD1999 . Un nouvel *idGeo* est attribué à l'objet construit dans la BD 1999.

Restructuration d'ordre n-1 : (substitution de plusieurs bâtiments individuels à un seul bâtiment collectif) une destruction et plusieurs créations d'objets sont réalisées sur la copie de la BD 2007 pour créer la BD1999. Plusieurs nouveaux *idGeo* sont attribués aux objets construits dans la BD1999.

3 Représentation de données spatio-temporelles et simulation

Cette section présente notre schéma d'application pour les BDST et la simulation des dynamiques urbaines. L'intégration du temps dans les systèmes d'information géographique a été largement adressée (Langran 1993). Néanmoins, une attention particulière y est portée depuis que les agences nationales de cartographie ont commencé à créer leurs premières bases de données versionnées. La solution présentée consiste à étiqueter chaque représentation topographique avec la date d'acquisition des données utilisées pour saisir les données. Chaque base de données topographique peut ainsi être considérée comme un instantané du système étudié à une date donnée. Ces instantanés, ainsi que leurs identifiants, peuvent être utilisés pour identifier les éléments constituant le système : nous les appelons des agents géographiques. Chaque agent géographique est ainsi lié à une ou plusieurs représentations qui peuvent servir à étudier son évolution, sa dynamique. La figure 2 illustre un agent géographique (le Louvre) et 4 représentations topographiques le représentant à 4 dates différentes (1589, 1643, 1871 et 2007).

3.1 Représentation du temps et de l'espace

En plus de son identifiant *idGeo* (cf. section 2), un agent géographique possède plusieurs représentations topographiques, une date de construction et de destruction estimées (fig. 3). Les représentations topographiques peuvent être simulées (créées par

simulation) ou pas (extraites de bases de données topographiques). Chaque représentation porte un identifiant *idRep*, une géométrie, une date source, et un lien vers l'agent géographique qu'elle représente. Les représentations sont des Features au sens de la définition de l'OGC (OGC, 1999).

3.2 Une représentation hiérarchique et sa construction

La création des différents niveaux géographiques est un processus important qui permet l'étude des changements et des dynamiques urbaines à différents niveaux. Pour chaque base de données topographique, trois principales échelles sont considérées : micro, méso et macro (Ruas 2000). Les micro-représentations concernent les bâtiments, les routes, les cours d'eau, les voies ferrées, etc. Les méso-représentations sont des agrégations de micro-représentations, telles que les groupes de bâtiments, les îlots, les quartiers, les aires urbaines, le réseau de communication (construit à partir de micro-représentations linéaires telles que les routes, les cours d'eau et les voies ferrées), etc. Finalement, les macro-représentations sont des populations de micro et de méso-représentations : elles contiennent tous les objets d'un type donné (une population de bâtiments contient tous les bâtiments d'une base de données topographique).

La plupart des micro-représentations sont directement extraites des bases de données topographiques. D'autres sont créées par analyse spatiale (cf. module 2 dans la section 1). De tels objets peuvent être des espaces vides (construits à partir des îlots et des bâtiments), des carrefours (construits à partir du réseau de communication), etc. La structuration de l'espace urbain à partir des objets micro est un processus composé de 5 principales étapes (de micro à méso) illustrées par la figure 4.

Étape 1 : Les aires urbaines sont construites à l'aide de zones tampon sur les bâtiments.

Étape 2 : Les aires urbaines sont découpées en îlots à l'aide du réseau de communication.

Étape 3 : Les îlots similaires adjacents sont regroupés en quartiers à l'aide de leur caractérisation (densité, fonction, etc.).

Étape 4 : Les îlots sont décomposés en structures (les groupes de bâtiments).

Étape 5 : Les espaces vides sont créés dans les îlots par soustraction des bâtiments à la surface des îlots.

Un exemple de processus de création de hiérarchie est illustré par la figure 5 sur des données topographique d'Orléans en 2007.

3.3 Représentation du changement

Une fois les hiérarchies créées, les évolutions des agents géographiques sont analysées. Une liste des changements les plus courants est donnée en section 2. La liste de tous les changements ayant lieu entre deux bases de données successives constitue un différentiel permettant au système d'analyser les changements de chaque type d'objet, mais aussi d'identifier les spécificités des évolutions des différentes zones géographiques (cf. module 5 dans la section 1). À partir de ses représentations topographiques et de ses changements, les dates de construction et/ou de destruction d'un agent géographique peuvent être analysées. Par exemple, dans l'exemple illustré par la figure 2, la création du bâtiment peut être déterminée comme ayant eu lieu avant 1589 et sa destruction comme n'ayant pas lieu avant 2007.

3.3 Agents et simulation

La simulation urbaine est un domaine de recherche actif (Barros 2003, Batty 2005, Benenson 1997, Hammam 2007, O'Sullivan 2001). Alors que la plupart des approches existantes s'appuient sur les automates cellulaires, c'est-à-dire des approches raster, notre approche est une approche vectorielle. Elle est basée sur un système multi-agents où chaque agent géographique est autonome et possède un contexte (voisinage) qu'il perçoit. Cette perception est essentielle pour le processus de décision car elle influence la satisfaction des contraintes de chaque agent. Les agents ont deux types de contraintes : les contraintes macro concernent tous les agents du même type (les bâtiments ne doivent pas se superposer), alors que les contraintes méso sont locales et propres aux objets méso auxquels chaque agent appartient (la densité objectif d'un îlot est 0.5). Finalement, tous les agents doivent respecter les règles d'évolution présentées en section 1.

Chaque agent possède une liste d'actions qu'il peut appliquer. À chaque étape de la simulation, un agent évalue sa satisfaction (déterminée à l'aide de ses contraintes) et la liste de ses actions potentielles. Il sélectionne alors la meilleure action (s'il en existe une) et l'applique. Le cycle de vie d'un agent est illustré par la figure 6. Une simulation complète commence par l'activation des agents de plus haut niveau : les aires urbaines, qui, à leur tour, activent les quartiers qui les composent. Ce processus est appliqué de façon récursive jusqu'à ce que tous les agents soient activés. Après l'activation de tous ses agents, une fois qu'ils ont tous terminé leur cycle de vie, chaque agent méso vérifie les conflits potentiels entre les actions entreprises par ses agents composants. Les agents méso jouent alors un rôle

d'arbitrage et décident, en cas de conflit, quels agents sont autorisés à appliquer leurs actions. Finalement, les agents méso peuvent décider de la destruction de certains de leurs agents composants afin de permettre à d'autres agents d'appliquer leurs actions ou simplement pour satisfaire leurs propres contraintes (contrainte de densité pour les îlots).

4 Conclusion, perspectives et tâches futures

Dans cet article est présentée une approche originale pour la compréhension des dynamiques urbaines. Tout d'abord, nous proposons une méthodologie de création de bases de données spatio-temporelles à partir d'une base de données topographique et de plusieurs cartes. Cette méthodologie, appliquée sur des zones test autour d'Orléans, nous permet de suivre l'évolution des objets topographiques dans les bases de données consécutivement créées. De plus, nous introduisons un schéma d'application pour l'analyse et la simulation des dynamiques urbaines, et les bases de notre plate-forme de simulation. À terme, nous espérons que de telles simulations apporteront de nouvelles perspectives sur les dynamiques urbaines.

La création de la base de données spatio-temporelle présentée dans cet article montre que la construction de telles bases de données est un processus difficile et coûteux. Heureusement, grâce, entre autres, à la directive européenne INSPIRE (<http://inspire.jrc.ec.europa.eu/>), la qualité des bases de données topographiques augmente, notamment par l'intégration d'informations temporelles et d'identifiants uniques pour les objets topographiques. À terme, de tels identifiants faciliteront la création de bases de données spatio-temporelles. Néanmoins, ceci ne résoudra pas les problèmes liés à l'acquisition de données anciennes.

Dans le contexte du projet GeOpenSim, l'utilisation d'algorithmes d'appariement automatique afin de créer les liens entre les objets topographiques devra permettre d'accélérer le processus de création des bases de données spatio-temporelles. Finalement, un point important nous paraît concerner la calibration et la validation de règles d'évolution et de simulation à l'aide de données topographiques. L'utilisation d'outils d'apprentissage afin d'identifier et de généraliser les règles d'évolution à partir de bases de données spatio-temporelles nous fournira une riche variété de règles ainsi que les outils pour identifier les spécificités de différents espaces géographiques en termes d'évolution. Pour conclure, l'utilisation de bases de données topographiques pour l'analyse des dynamiques urbaines et leur

simulation est un réel défi qui ouvre des perspectives études géographiques et à des applications plus pour l'utilisation de telles bases de données à des vastes.

Bibliographie

J. X. Barros, 2003, 'Simulating Urban Dynamics in Latin American Cities', dans *Proceedings of the 7th International Conference on GeoComputation*, University of Southampton, United Kingdom.

M. Batty, 2005, *Cities and complexity: understanding cities with cellular automata, agent-based models, and fractals*, Cambridge MA, MIT Press.

I. Benenson, J. Portugali, 1997, 'Agent-Based Simulations of a City Dynamics in a GIS Environment', dans *COSIT '97: Proceedings of the International Conference on Spatial Information Theory*, pp. 501-502, London, UK. Springer-Verlag.

Y. Hammam et al., 2007), 'The dynamic geometry of geographical vector agents', *Computers, Environment and Urban Systems*, 31(5), p. 502-519.

K. Hornsby, M. J. Egenhofer, 1997, 'Qualitative Representation of Change', dans *COSIT '97: Proceedings of the International Conference on Spatial Information Theory*, p. 15-33, London, UK. Springer-Verlag.

G. Langran, 1993, *Time in geographic information systems*. Technical issues in geographic information systems. Taylor & Francis, London.

A. Mas, 2008, 'Temporal follow-up of morphological and social changes in urban zone', dans *AutoCarto'08: Proceedings of the 17th International Research Symposium on Computer-based Cartography*.

D. O'Sullivan, 2001, 'Graph-cellular automata: a generalised discrete urban and regional model', *Environment and Planning B : Planning and Design*, 28, p. 687-705.

OGC, 1999, *The OpenGIS Abstract Specification - Topic 5: Features, version 4*. OpenGIS Project Document Number 99-105r2.

A. Ruas, 2000, 'The roles of meso objects for generalisation', dans *Proceedings of the 9th International Symposium on Spatial Data Handling*, vol. 3b, p. 50-63.

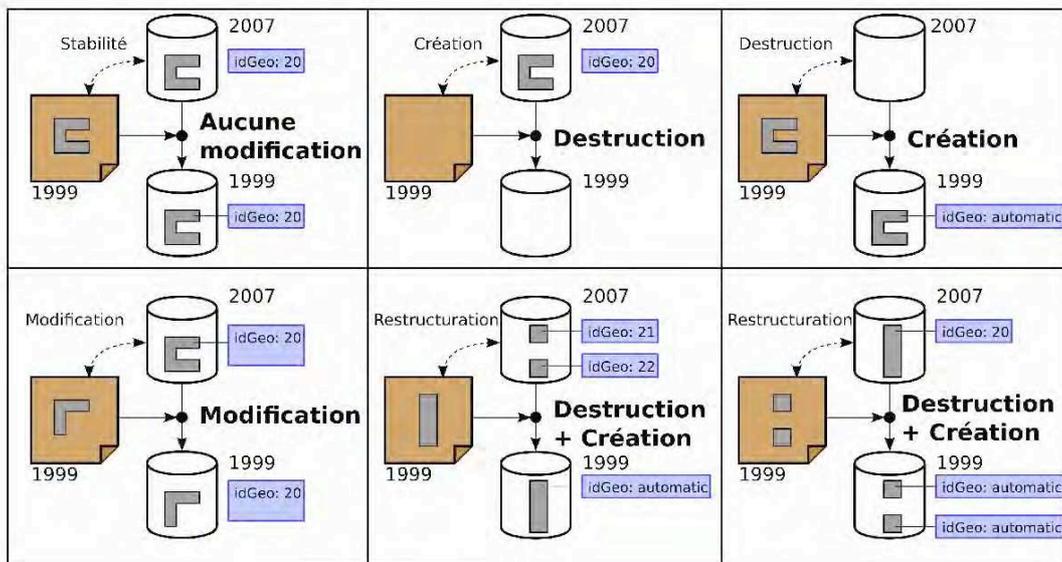


Figure 1 : Suivi d'un objet entre 1999 et 2007

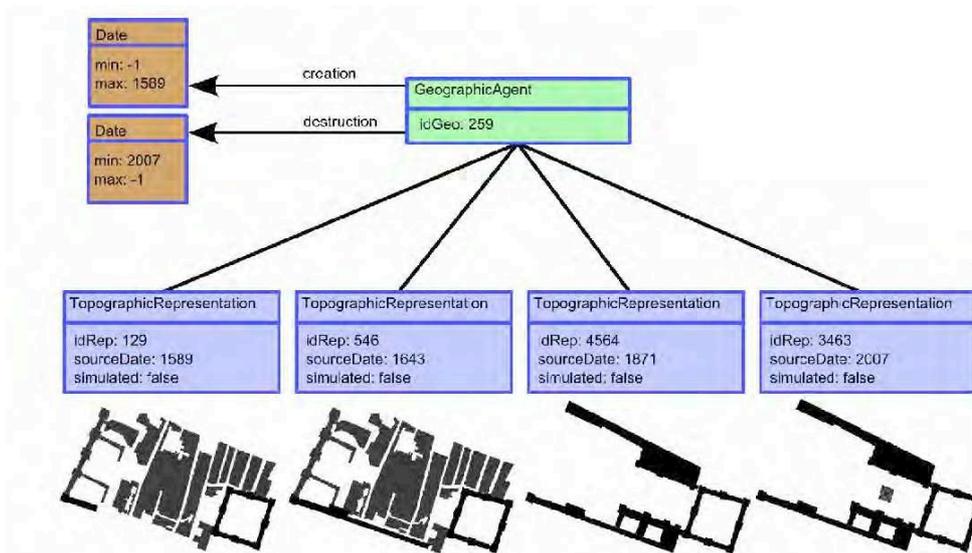


Figure 2 : Un agent géographique et ses représentations topographiques

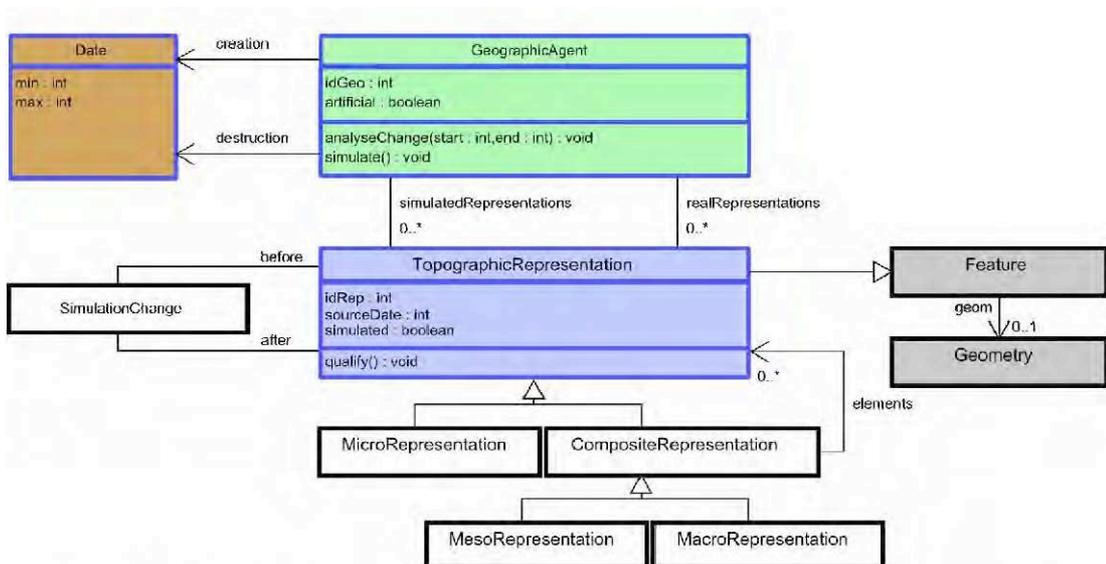


Figure 3 : Représentation des agents géographiques et de leurs représentations topographiques.

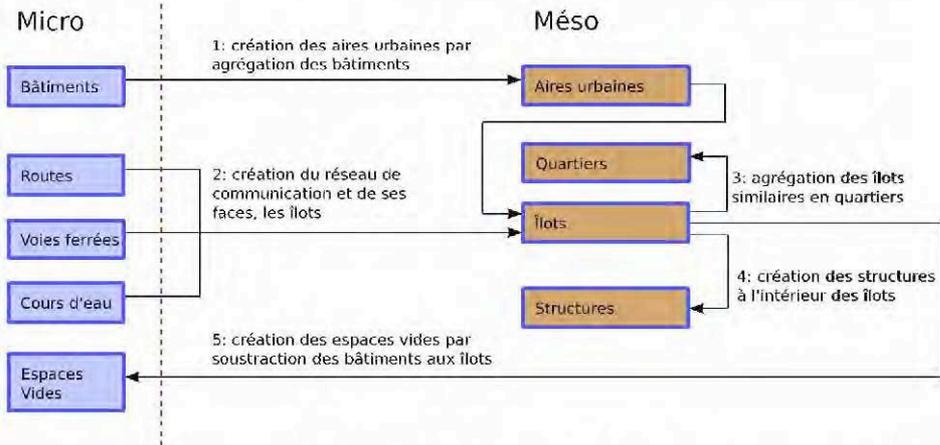


Figure 4. Le processus de création de la hiérarchie et ses 5 principales étapes.

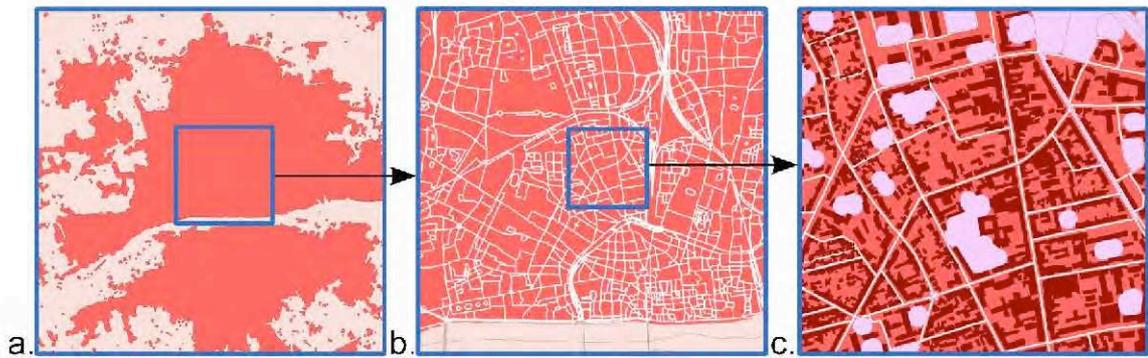


Figure 5 : Différentes étapes du processus de création de la hiérarchie : a. aires urbaines (en rouge) créées à l'étape 1, b. îlots (en rouge) créés à partir du réseau de communication (en blanc) à l'étape 2, c. îlots, bâtiments (en rouge foncé) et espaces vides créés à l'étape 5

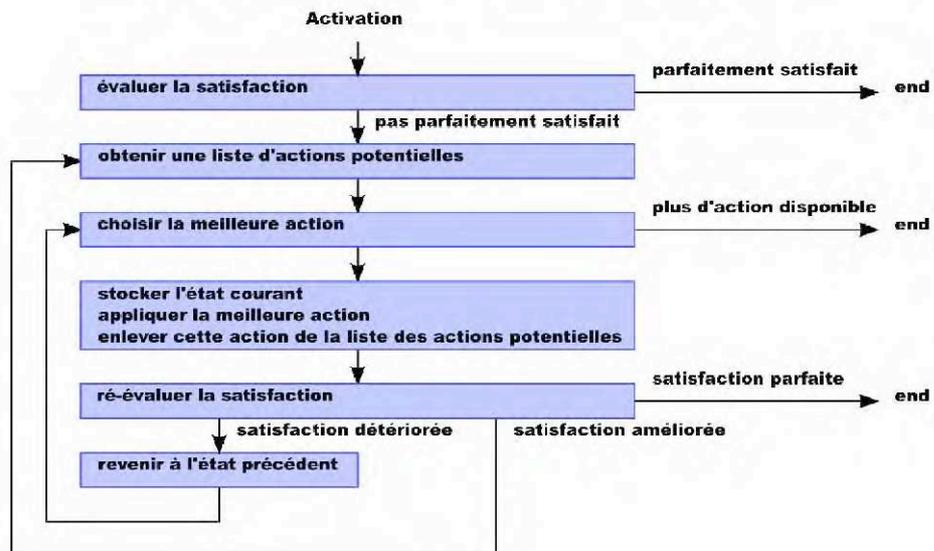


Figure 6 : Cycle de vie d'un agent