FORMALISMES POUR L'AUTOMATISATION DE LA GÉNÉRALISATION: MÉTRIQUE, TOPOLOGIE, PARTITIONNEMENT HÉRITARIQUE ET TRIANGULATION LOCALE

Par Anne RUAS
IGN-Laboratoire COGIT - 2 avenue Pasteur, 94160 Saint-Mandé

RÉSUMÉ

Afin d'automatiser les processus de généralisation, nous remettons tout d'abord l'idée de considérer l'information contenue dans les SIG et de sa manière d'être permettant de résoudre des tâches complexes. Nous avons donc implanté en Orientabilité différentes façons d'automatisation spatiale, choisis de façon à satisfaire les exigences de la généralisation et de faciliter le déploiement des opérations de généralisation. L'objectif est de partir d'une base de données vectorielle et d'utiliser un système à base de règles nous permettant de créer, en fonction de l'évolution de l'objet de la base et des contraintes attachées à cet objet, l'opération suivante à réaliser.

1. Introduction

1.1 La généralisation de base de données géographiques

La généralisation est un processus utilisé pour réduire et simplifier une information. Dans le cas de données géoréférencées, la généralisation doit permettre de dériver une base de données à partir d'une base de données source plus riche, en modifiant la sémantique (classes, attributs) et la géométrie des objets [3]. Pour généraliser une base de données, il est donc nécessaire de:
- définir un schéma de données (contenu théorique de la base de données à généraliser : classes, attributs, domaines de définition);
- définir les liens de passage entre les 2 schémas;
- réaliser le passage des données d'un schéma à l'autre en respectant la spécification de la nouvelle base de données : nouvelle résolution métrique, réduction du volume de données.

Le fait de définir une nouvelle résolution métrique induit non seulement la simplification de la géométrie des objets mais d'autres actions beaucoup plus complexes telles que le déplacement, l'agglomération d'objets (quelle est la signification de la nouvelle information créée ?) ou l'élaboration d'objets, si dans un espace il est géométriquement impossible de conserver toute l'information initiale en respectant les nouvelles contraintes de taille minimale des objets (ex : élimination d'une partie d'un réseau de rues en centre ville).

1.2 Automatisation des processus


Le passage à l'automatisation nécessiterait la connaissance:
- du séquencage des opérations,
- des données à traiter en priorité,
- des choix d'algorithmes adaptés aux données: pour une opération il existe toujours un ensemble d'algorithmes ayant des caractéristiques légèrement différentes, il faut donc trouver l'algorithme qui conviendra le mieux à cette partie particulière à traiter. Ces modalités font l'objet de plusieurs recherches au COGIT, dont le choix des algorithmes de simplification des lignes en fonction de leurs caractéristiques géométriques [11],
- des mécanismes de maintien de cohérence : chaque opération peut avoir des effets de bord qui il faut pouvoir corriger pour maintenir l'intégrité des données.
des mécanismes de validation de chaque opération, afin de contrôler si le nouveau état converge au non vers les objectifs de généralisation initialement fixés.

En théorie, pour automatiser un processus, on peut soit prédéfinir le séquençage des opérations, soit utiliser un système à base de règles qui fait le choix de l'opérations suivantes à réaliser. Une étude réalisée par R. Macdonald montre clairement que l'inversion de 2 opérations de généralisation peut donner des résultats très différents [8]. De plus ces choix permanents ne concernent pas uniquement l'opération à réaliser mais également le que les objets sur lesquels chaque opération doit être appliquée ; on ne peut pas généraliser de façon globale des données géographiques puisque chaque configuration a un caractère unique.

Il nous semble clair que la détermination a priori des opérations et des objets à traiter ne peut pas être satisfaisante : chaque opération modifie l'information et la configuration de l'opérateur. On est donc naturellement conduit à se tourner vers un système à base de règles qui évaluera, en cours de procédé, les nouvelles caractéristiques des données et permettra de choisir parmi les plus adaptées en fonction des spécialisations de la nouvelle base de données à réaliser. L'idée d'utiliser un système expert pour automatiser la généralisation n'est pas nouvelle, mais malheureusement aucun système n'est à l'état actuel satisfaisant. A noter aussi, l'objectif de la réalisation d'un système de généralisation à partir d'un système à base de règles n'est pas dû à l'outil mais au manque d'identification et formalisation des connaissances nécessaires aux choix et aux contrôles.

Dans ce texte, nous essayerons:

- d'analyser ce qui dépend du besoin de réaliser des opérations de généralisation, les conflits;
- de proposer des processus types;
- d'identifier l'information sémantique et géométrique minimale,
- de proposer un moyen pour représenter et utiliser les relations de proximité entre objets non connexes,
- de proposer un partage géographique de l'espace afin d'optimiser le choix du séquençage des objets à traiter,
- d'intégrer ces nouvelles connaissances dans un processus global de généralisation.

Toutes ces parties correspondent à des développements réalisés sur une plateforme de généralisation utilisation données sont représentées en support et où la résolution des actions est gérée à l'aide d'un système à base de règles.

2 Les conflits

Un conflit correspond à une situation où deux ou plus ne correspondent pas aux règles que l'on s'est fixées. Dans le cas de la généralisation, on considère qu'il y a conflit lorsque les données ne répondent pas aux critères de résolution géométrique. De façon générale on peut dire qu'il y a conflit lorsqu'un objet ou un ensemble d'objets ne sont pas perceptibles à l'œil actuellement trop d'informations ont été réalisées sur les conflits en généralisation [8]. Nous énonçons donc simplement quelques exemples permettant d'apprendre cette notion. Un conflit:

- est donné à la résolution géométrique de la base de données. Il doit être le plus important à l'angle affecte à un objet une symétrie (dans le cas de la généralisation cartographique).
- peut émerger entre objets de nature différente (ex : entre une maison et une route).
- peut émerger entre objets de type géométrique différent (point, ligne, surface).
- est potentiellement transitoire, si Or et Ob sont en conflit et Ob est Ocupé quand conflit, Or et Ob sont potentiellement en conflit. Lors de la résolution du conflit entre Or et Ob l’objectif doit être de résoudre la résolution suivante.

On est amené à distinguer des conflits globaux (entre un ensemble d'objets), des conflits inter-objets (les objets en conflit sont identifiés), des conflits intra-objet (due à la géométrie propre de chaque objet).

- un conflit global représente un ensemble d'objets en conflit de proximité. Il pourra être qualifié par un critère de densité géométrique ou quantitatif (ex : nombre de représentants d'une classe dans une surface donnée).
- les conflits inter-objets sont principalement liés à des critères de proximité. On peut les distinguer selon la dimension géométrique des objets concernés. Dans le cas des lignes et surfaces, les conflits de proximité peuvent être plus ou moins locaux (ex : point au bord de conflit entre 2 lignes). Les conflits inter-objets sont qualifiables par des critères de taille (surface trop petite, trop fine, lignes trop petites, trop détaillées...).

La connaissance des conflits a plusieurs utilisations dans le processus de généralisation :

- l'identification des conflits permet de déterminer les zones à traiter.
- le suivi local du nombre de conflits permet d'éviter la qualité d'une résolution choisie.
- l'identification d'un conflit permet de choisir l'opération ou le séquençage d'opérations à utiliser. Par exemple, si un ensemble de maisons est en conflit global et que chaque maison est en conflit de surface-trop petite, il est mal défini de généraliser les maisons une à une. Il faut soit changer de type d'information (passer en zone urbaine) soit éliminer les objets les moins caractéristiques de la zone. La résolution de conflit intra-objet (simplification et amplification de la surface) se fait alors sur un ensemble réduit d'objets.

3 Un premier processus

Si l'on considère que l'on dispose d'un système à base de règles et d'outils de détection et de qualification de conflits, la démarche globale de généralisation peut être la suivante :
4 Représentation de l'information de base

4.1 L'information sémantique

Le schéma de données permet de décrire les attributs des objets et les liens entre objets de classes différentes. Ces informations ne sont pas suffisantes pour la généralisation. Il va falloir associer à chaque classe des connaissances permettant de faire les choix décrits au chapitre précédent.

Parmi les attributs d'ores et déjà identifiables on peut citer :
- des opérations potentiellement réalisables,
- des algorithmes utiles pour chaque type d'opération (on n'utilisera pas les mêmes algorithmes pour agréger des maisons et des surfaces naturelles),
- des pré-ordres pour définir des priorités entre objets de classes différentes pour certaines opérations de base (ex : priorité de déplacement si 2 objets linéaires sont en conflit de proximité).

4.2 L'information spatiale

Dans la mesure où la généralisation induit de nombreuses transformations géométriques, il est indispensable de décrire les relations spatiales entre objets [3]. Autant pour le choix des opérations, le maintien des relations spatiales ou la diffusion de transformations locales, il est indispensable de définir au mieux les relations spatiales entre objets. Une solution souvent choisie dans les SIG est de projeter la géométrie des objets sur un plan et de décrire toutes les relations d'inclusion et de connectivité. Pour représenter des relations topologiques, nous utilisons la modélisation sous forme de cartes topologiques [1] qui permet de se déplacer facilement sur le graphique et de connaître les relations d'inclusion. Cette couche peut dans notre cas être appelée topo-métrique, parce que ce sont les bords (frontières ouifices) qui portent également la localisation des objets.

Nous pourrons constater par la suite que cette couche est au cœur de la modélisation choisie puisque toutes les autres couches d'information lient les objets, et que c'est elle qui va permettre d'établir la cohérence entre les différentes connaissances.

![Diagramme de relations géométriques et spatiales](image)
5 Calcul et représentation des relations de proximité

5.1 Besoins et type de représentations

Les relations de proximité permettent de décrire le positionnement d'objets non connectés. Elles peuvent être utilisées pour propager des déformations, pour identifier des conflits de proximité et des structures géométriques caractéristiques (ex : alignements de maisons) [3].

Il faut différencier la méthode de calcul des relations de proximité, de leur représentation et de leur utilisation. A priori on peut penser qu'un découpage régulier et fin-éventuellement hiérarchique de l'espace permet de facilement calculer les relations de proximité (maillage de l'espace et repérage des objets faisant partie des mêmes mailles). Si ces calculs sont assez simples, le fait de savoir que 2 objets sont partiellement en même maille permet d'en déduire que les objets sont proches mais pas de décrire le type de proximité entre les objets. Il est nécessaire de définir des relations plus fines pour répondre à l'ensemble des problèmes à la fois internes et externes.

On peut également représenter les relations de proximité à l'aide d'un réseau permettant de répartir des points d'ancrage de chacun des objets [2]. En utilisant la triangulation de Delaunay pour décrire les relations de connectivité et de proximité entre données géographiques pour la généralisation de données grandes échelles. Nous utiliserons également les triangulations de Delaunay, mais dédiées à nos problèmes.

5.2 Les triangulations locales de Delaunay : principes et modélisation

Les triangulations ne sont pas calculées sur toute la zone à généraliser mais sur des zones réduites définies. On parlera donc de triangulations locales. Il peut y en avoir plusieurs à un instant donné.

Les TLD ne portent pas l'information topo-métrique, mais sont reliées. Elles sont calculées pour certaines opérations et déformations (fig). Une TLD est composée d'un ensemble de noyaux et d'arcs reliés par des relations de connectivité. Les noyaux représentent des points d'ancrage entre les objets topo-métriques et la triangulation. Chaque noyau aura donc un lien avec un objet topologique, lui-même en relation avec un objet géographique. Son comportement peut donc dépendre de l'objet géographique auquel il est lié et sa déformation peut être propagée sur l'objet topologique auquel il est lié. L'effet de la triangulation dépendra donc du choix des noyaux triangulation.

Ces choix dépendent du type d'action à réaliser (ex : type d'information recherchée (ex : alignements de maisons, conflits de proximité).

Les triangulations sont calculées par la méthode de Weiss [12] par un calcul de l'enveloppe convexe des noyaux initiaux et insertion des noyaux un à un. Les noeuds de la triangulation sont classés en fonction de la géométrie des objets géographiques qu'ils relient afin de pouvoir plus facilement différencier leur comportement lors de la propagation des déformations.

5.3 Les TLD pour la gestion des déplacements

On se propose ici d'utiliser une TLD pour propager la déformation d'un objet (créé par une opération quelconque) sur les objets proches non connectés, afin d'éviter des superpositions d'objets. On peut imaginer facilement le cas où l'opérateur applique une déformation d'une telle et la propagation des déformations sur les maisons proches. La propagation d'une telle déformation dans un nouveau contexte, il convient de corriger la déformation de l'espace de propagation (frontières de la zone devant rester fixe) et de prendre en compte tous les objets à l'intérieur de la zone de déformation.

Le schéma est le suivant :

1. On choisit l'élément à déformer b (Fig. 3) sa géométrie avant et après déformation : (x, y) et (x', y').
2. On identifie tous les objets non connectés proches.

Pour ce faire, on repère les 2 faces topologiques (ou une autre zone de travail telle qu'une partition cf. 3).

5.2 Les triangulations locales de Delaunay : principes et modélisation

Les triangulations ne sont pas calculées sur toute la zone à généraliser mais sur des zones réduites définies. On parlera donc de triangulations locales. Il peut y en avoir plusieurs à un instant donné.

Les TLD ne portent pas l'information topo-métrique, mais y sont reliées. Elles sont calculées pour certaines opérations et déformations (ex : alignements de maisons, conflits de proximité).

Les triangulations sont calculées par la méthode de Weiss [12] par un calcul de l'enveloppe convexe des noyaux initiaux et insertion des noyaux un à un. Les noeuds de la triangulation sont classés en fonction de la géométrie des objets géographiques qu'ils relient afin de pouvoir plus facilement différencier leur comportement lors de la propagation des déformations.

5.3 Les TLD pour la gestion des déplacements

On se propose ici d'utiliser une TLD pour propager la déformation d'un objet (créé par une opération quelconque) sur les objets proches non connectés, afin d'éviter des superpositions d'objets. On peut imaginer facilement le cas où l'opérateur applique une déformation d'une telle et la propagation des déformations sur les maisons proches. La déformation d'une maison peut créer un nouveau conflit. Il convient de corriger la déformation de l'espace de propagation (frontières de la zone devant rester fixe) et de prendre en compte tous les objets à l'intérieur de la zone de déformation.

Le schéma est le suivant :

1. On choisit l'élément à déformer b (Fig. 3) sa géométrie avant et après déformation : (x, y) et (x', y').
2. On identifie tous les objets non connectés proches.

Pour ce faire, on repère les 2 faces topologiques (ou une autre zone de travail telle qu'une partition cf. 3), dont l'objet ima aise fait partie, et on repère tous les objets (faces a, b) délimitant les zones de travail ainsi que les objets contenus dans ces zones.

3. Pour chaque objet a ou b faisant partie des zones (ex : les maisons) on choisit un point d'ancrage (en pratique le centre constitue un bon point), et on projette l'ensemble de ces points sur les bords (bords de la). Évidemment on utilise la géométrie de l'objet initialement déformé avant sa déformation. On dispose ainsi d'un ensemble de noyaux de la triangulation face[s] provenant d'objet locaux (ex : les maisons) soit des bords délimitant les zones (ex : les routes),
4. À partir de cet ensemble de noyaux on calcule une triangulation locale de Delaunay pour chaque zone de déformation.
5. On calcule les vecteurs de déplacement sur l'objet initialement déformé b (Fig. 3). Pour cela on projette les points d'ancrage de b (Fig. 3) sur sa nouvelle géométrie (x', y')
6. On propage ces déplacements via les arcs de la triangulation sur les objets proches.

212
CFC (N° 145-147 - déc.95/mai 96)
Dans la mesure où l'on dispose d'une couche topométrique, l'identification des objets à déplacer est simple. Par contre, il convient de choisir une bonne méthode de propagation des déformations par la triangulation suivant que :
- il est évident de déplacer un objet éloigné,
- les déformations sont amorties en fonction de l'éloignement d'un objet à l'objet initialement déformé,
- les brins définissant les zones de travail doivent rester fixes,
- il faut éviter des phénomènes de cyclicité à travers la triangulation : les objets doivent converger vers une position stable,
- certains objets constituant la zone de frontière fixent les objets isolés dans une position quasi-stable. On peut considérer que l'on dispose donc au départ de forces de déplacements et d'un ensemble de forces de fixation : les libertés de déplacement des objets dépendant donc des directions : certains objets se déplacent plus facilement dans une direction que dans une autre.

5.4 Choix d'un modèle de force

On peut représenter les forces subies par chaque nœud comme étant la moyenne des vecteurs de déplacements réciproques pondérée par la distance entre le nœud considéré et les nœuds qui lui propagent un déplacement à l'aide des arcs de triangulation.

En un nœud n-centre on a donc :
\[
\text{Vect-Dep (n-centre)} = \frac{\sum (\text{Vect-Dep})_{\text{n-centre}}}{\sum (1/\text{dist}_{\text{n-centre}}^2)}
\]

en considérant que la somme ci-dessus est vectorielle.

Une fois que chaque vecteur de déplacement est calculé, on peut considérer les forces de fixation induites par les nœuds de la triangulation qui ne peuvent pas bouger (ce sont les nœuds appartenant au contour de la zone), et qui devront empêcher les nœuds libres de trop se rapprocher d'eux. Pour cela, il faut simplement tester la distance séparant un nœud libre d'un nœud de fixation et interdire ou de limiter le déplacement à l'aide d'un calcul d'angle :

Soit le nœud n-fixé lié à nœud n-centre. On a donc :
\[
\text{Vect-Dep final (n-centre)} = \text{Norme} \left( \text{Vect-Dep (n-centre)} \right) \times \cos \left( \text{angle} \left( \text{Vect-Dep (n-centre)}, \text{Vect (n-centre - n-fixé)} \right) \right) / 2
\]

L'inconvénient de cette méthode de calcul est que les nœuds libres sont liés entre eux et se propagent des déplacements (puis que certains nœuds de l'extémité sont reliés à des nœuds de l'extémité). Selon l'ordre du calcul des propagations on obtiendra des résultats assez différents et il est impossible de décrire a priori le meilleur séquençage des calculs. Du plus les forces de fixation ne pouvant avoir un effet que sur les nœuds avec lesquels ils sont directement reliés. Nous nous sommes donc tournés vers un modèle de forces plus complexes mais plus réalistes, inspiré des modèles utilisés en mécanique. A l'heure actuelle seule la formalisation du problème sous forme d'équations a été réalisée. Nous traversons sur la recherche de méthodes simples de résolution.

Nous considérons un système composé de n nœuds soumis à des liaisons dont le comportement est connu, et nous cherchons l'équation du système. Les nœuds de la triangulation représentent les nœuds du système.
Les arcs de la triangulation représentent les liens dont certains exercent des forces dont nous pouvons dériver le potentiel U. Nous sommes donc dans le cas de la résolution d'un système linéaire à n inconnues, que l'on peut réduire à l'aide du Lagrangien L :

\[ L = C - U \] (1) avec \( C \) = somme de l'énergie cinétique en chaque point et \( U \) = somme des énergies potentielle.

Si \( q_1, q_2, q_3, ..., q_n \) est un système de paramètres indépendants définissant la position du système, le lagrangien, note alors L(q, q', t) est une fonction de q1, q2, ..., qn, des dérivées q1', q2', ..., qn' et du temps. Le système d'équations différentielles :

\[ \dot{q} = (\partial L / \partial \dot{q}) - \partial L / \partial q = 0 \] (2)

est appelé le système d'équations de Lagrange du mouvement.

Dans notre cas, nous disposons de noeuds découpés en 2 ensembles :
1. E-centre = \{n-centres\}, composé des noeuds fixes, qui exercent des forces entre eux et pour lesquelles nous connaissons leur position initiale et nous recherchons leur position finale. Les équations de Lagrange sont réalisées sur cet ensemble.
2. E-zone = \{n-zones\}, composé des noeuds sur la zone de travail, qui n'exercent pas de forces entre eux, pour lesquels les positions initiales et finales sont connues et qui exercent des forces sur les noeuds de E-centre.

L'ensemble E-zone se décompose en 2 sous-ensembles :
- E-zone-fix = \{n-z-fix\}, composé des noeuds qui ont été initialement déplacés. On considère donc qu'ils ont une nouvelle position connue et une masse initiale nulle. À ce stade on ne met pas en cause l'action initiale qui déclenche la recherche d'une solution. Chaque déplacement est soumis à des forces de déplacement sur les noeuds de E-centre auxquels ils sont liés.
- E-zone-m = \{n-z-m\}, composé des noeuds qui restent fixes. On doit les considérer puisqu'ils exercent des forces de fixation sur les noeuds de E-centre.

On distingue trois types de forces :
1. les forces \( F_e \) exercées entre les noeuds de E-centre qui se comportent comme des ressorts : \( F_k \) dep. \( k \) dépend de la distance entre les noeuds : plus la distance est grande, moins la propagation est forte.
2. les forces \( F_d \) exercées entre les noeuds de E-centre et les noeuds de E-centre, qui sont également des ressorts \( F_k \) dép. \( k_2 \)
3. les forces \( F_l \) exercées entre les noeuds de E-zone-fix et les noeuds de E-centre. Ces forces agissent à l'inverse d'un fil élastique : elles amortissent le déplacement lorsqu'un noeud de E-centre s'approche d'un noeud de E-zone-fix, mais laissent le déplacement beaucoup plus libre lorsqu'un noeud de E-centre s'éloigne d'un noeud de E-zone-fix. On peut prendre également une force du type ressort en limitant ses effets à l'aide d'un angle (cf. première méthode de résolution).

La triangulation initiale se transforme suivant les schémas suivants :
6 Partitionnement hiérarchique de l'espace

6.1 Présentation, choix des objets structurants

Le partitionnement hiérarchique de l'espace est composé d'un ensemble de partitions de l'espace à généraliser. Pour un niveau donné, chaque partie est décomposée en sous-parties. L'objectif de ces partitions est de se définir des zones de travail de différentes tailles afin de mieux maîtriser le schéma des opérations et de contrôler les propagations dues aux déformations. L'aspect hiérarchique du partitionnement permet d'adapter au mieux les zones de travail aux types de données et de complexité de la généralisation à réaliser. En pratique, 3 niveaux semblent être suffisants.

Cette stratégie de partitionnement est fréquente dans l'importer du traitement imitant en jeu un grand volume de données. L'originalité du partitionnement proposé ici est de fonder le partitionnement non pas sur un découpage régulier de l'espace (ex: quadrature) mais sur la géométrie de certains objets géographiques. Le principe est donc d'identifier des objets qui structurent l'espace et sont maintenus pendant le processus de généralisation. Le choix des objets structurants dépend de la base de données source et des spécifications de la généralisation.

Pour construire un partitionnement hiérarchique de l'espace à partir d'objets géographiques il faut s'assurer que les objets vérifient au mieux les propriétés suivantes:
- être adaptés à la création de cycles (pour la création de zones);
- être facilement hiérarchisables (par l'intermédiaire d'un code séquentiel ou possible);
- être maintenus pendant le processus de généralisation;
- avoir une densité liée à la densité des autres objets, afin que les parties soient d'autant plus petites que la densité d'objets est forte.

Afin de valider l'intérêt d'utiliser un partitionnement hiérarchique de l'espace pour la généralisation, nous avons recherché le type d'objet se stockant le mieux dans la base de données que nous souhaitons généraliser à l'IGN la BD Topo, qui est une base de données topographique à gravité large (ou de mémoire). En se basant sur une étude visuelle des objets conservés lors des généralisations manuelles des cartes IGN, nous avons constaté que les ruptures de niveau semblaient être bien adaptées à notre objectif:
- sur la base associée en commune naturellement des cycles,
- et que généralement maintenue, puisque l'information transverse montre que l'échelle est importante lorsque l'échelle n'est pas la même,
- et est déjà classée (sous au moins à carte, nationale, départementale)
- et sa densité est directement liée à l'activité humaine.

Il faut signaler que notre choix va dans le sens d'autres travaux de recherche en généralisation [7][10][5].

Une fois le partitionnement calculé, son utilisation est assez simple. On résout les conflits internes des bases de la partition de plus haut niveau (simplotification, cartographie...), puis on résout les conflits entre deux de même niveau (deplacements...), puis on considère que la structure est stable et on recommence le même démarchage pour les niveaux inférieurs. Enfin, on généralise les objets comprises à l'intérieur des partitions en considérant que l'espace décrit par chaque partition est fixe. Chaque modification géométrique donne lieu à un calcul de TLD pour recomposition des déformations géométriques afin de conserver le base dans un état cohérent. Cette méthode n'implique pas qu'il n'y aura pas de pertes en cause locales de ce schéma ou d'immobilisation de certaines base composant la structure (pour le niveau ou le bas). Elle n'impose pour autant de l'utiliser pour la résolution d'un nombre importants de conflits et de régler les autres conflits dans une étape ultérieure. À l'heure actuelle, il est encore difficile de se prononcer sur les limites d'utilisation du partitionnement hiérarchique.

CFC (N° 146-147 - déc. 95/mars 96)

215
6.2 Modélisation

Il existe une classe de partitions pour chaque niveau hiérarchique.
Une partie :
- est délimitée par un ensemble de brins,
- possède des sous parties (sauf pour le niveau le plus bas), et une sur partie (sauf pour le niveau le plus haut),
- contient des faces topologiques (systématique pour le niveau le plus bas, et occasionnel pour les autres niveaux).

6.3 Méthode de calcul

1- Acquisition interactive de la nature des objets géographiques structurants et affectation de la valeur de l'attribut « hiérarchie » pour tous les objets correspondants.
2- Propagation de la hiérarchie des objets géographiques aux objets topo-métriques,
3- Calcul des partitions : En commençant par le niveau le plus haut, calcul des cycles en utilisant les relations de connectivité au niveau topo-métrique et en limitant les candidats aux brins de hiérarchie correspondante. Les brins formant des impasses changent de niveau hiérarchique.
4- Calcul des relations d'inclusion entre les faces topo-métriques et les partitions : Chaque face est reliée à la plus petite partie dans laquelle elle est incluse. Le calcul de l'inclusion est réalisé par les liens qui relient un brin à une face et à une partition. Une face qui ne partage pas directement un brin (au minimum) avec une partition est affectée dans la même partition qu'une de ses faces connexes.
5- Calcul des relations d'inclusion entre les partitions : le principe est le même que celui décrit précédemment.

On peut décomposer le processus de généralisation en 2 étapes différentes :
- la pré-généralisation, qui représente la phase d'acquisition et de représentation des connaissances indispensables à la réalisation du processus,
- la généralisation à proprement parler qui représente la modification effective des données.

Un premier processus proposé est donc le suivant :

A- Enrichissement de la base de données :
1- Calcul et représentation des relations topo-métriques,
2- Détection et représentation des formes caractéristiques d'objets linéaires et surfaciques.
3- Détection et représentation des relations spatiales caractéristiques,
4- Saisie des comportements spécifiques associés aux différents types d'objets géographiques ; opérateurs et algorithmes utilisables, pré-order pour les opérations telles que le déplacement, ...
5- Calcul du partitionnement hiérarchique.

B- Généralisation des données :
1- Généralisation des objets structurants :
   Pour chaque niveau hiérarchique, on ne considère que les objets de ce niveau :
   1- Résolution des conflits intra-objets, avec propagation des effets de bords à l'aide des LTD,
   2- Résolution des conflits inter-objets, avec propagation des effets de bord à l'aide des LTD,
Les objets traités sont par la suite considérés comme fixes et intangibles.
2- Généralisation des autres objets.

7 Utilisation des différents formalismes dans le processus de généralisation

Dans les chapitres précédents, nous avons décrit un ensemble de connaissance qui nous paraît nécessaire pour automatiser le processus de généralisation. Il est évident que plus l'information est riche, plus elle va être difficile à gérer. Il ne faut pas que certaines traitements réalisés sur une couche d'information dégradent une information représentée sur une autre couche. C'est pourquoi, nous avons fait en sorte que toutes les cou-
Conclusion

Compte tenu de la complexité du processus de généralisation, il nous paraît indispensable d’enrichir considérablement l’information présente dans la base de données initiale et de se munir d’outils permettant d’identifier différents types de conflits et de guider le séquencement des opérations. Nous avons donc développé un système orienté objet différentes couches d’information : sémantique, topométrique, des triangulations locales (pour décrire les relations de proximité entre objets non connexes) et le partitionnement hiérarchique (pour faciliter le séquencement des opérations).

Il reste encore de nombreuses tâches à réaliser :
- l’étude du comportement des objets selon leur nature,
- le développement d’outils de détection et d’analyse de conflits,
- l’intégration dans le modèle des connaissances sur la géométrie des objets pour le choix des algorithmes.
Actuellement un système à base de règles nous permet de gérer la répercussion des déplacements entre objets. Nous pensons l’utiliser davantage pour des tâches plus complexes.

Références:

1. David Modélisation, représentation et gestion d’information géographique PhD Paris VI 1991
4. Locardi  Méthodes algébriques pour la généralisation cartographique Bulletin d’information de l’IGN N° 63 1995
5. Lee & Robinson Development of an automated generalisation system for large scale topographic maps GISRUK’93
6. Lee From master database to multiple cartographic representations ICA’93 pp 1075-1085
7. Leitner Prototype rules for automated map generalization Master of Geography Buffalo 1993