

**Formalismes pour l'automatisation de la généralisation:
métrique, topologie, partitionnement hiérarchique et triangulation locale**

Anne Ruas
IGN- Laboratoire COGIT
2 avenue Pasteur, 94160 Saint-Mandé- France
ruas@cogit.ign.fr

Résumé

Afin d'automatiser les processus de généralisation, nous pensons qu'il est fondamental d'enrichir considérablement l'information existante dans les SIG et de se munir d'outils permettant de réaliser des tâches complexes. Nous avons donc implémenté en Orienté Objet différentes couches d'information: sémantique, topo-métrique, des triangulations de Delaunay - pour représenter des relations de proximité entre objets non connexes - et un partitionnement hiérarchique de l'espace pour faciliter le séquençement des opérations de généralisation. L'objectif est de partir d'une base de données enrichie et d'utiliser un système à base de règles nous permettant de choisir, en fonction de l'évolution de l'état de la base et des connaissances attachées aux objets, 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 les spécifications 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'agrégation d'objets (quelle est la signification de la nouvelle information créée?) ou l'élimination 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 du processus

L'automatisation du processus de généralisation concerne la réalisation automatique du passage des données du schéma source au schéma final. On appellera, dans ce texte, opérations les actions géométriques élémentaires permettant de modifier l'information (ex: simplification géométrique, caricature, agrégation). Si les opérations de base commencent à être bien connues - tant dans leur définition [9] que dans le développement d'algorithmes associés - l'automatisation de ce processus reste problématique. Il existe actuellement suffisamment d'algorithmes pour réaliser une généralisation interactive (ex: système MGE/MG d'Intergraph [6], plate-forme développée au laboratoire COGIT de l'IGN [4]), mais aucun système ne peut se passer d'un opérateur cartographe averti.

Le passage à l'automatisation nécessiterait la connaissance:

- du séquençement des opérations,
- des données à traiter en priorité,
- des choix des 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 le mieux adapté au cas particulier à traiter. Ces adéquations 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 qu'il faut

pouvoir corriger pour maintenir l'intégrité des données,

- des mécanismes de validation de chaque opération, afin de contrôler si le nouvel état converge ou non vers les objectifs de généralisation initialement fixés.

En théorie, pour automatiser un processus, on peut soit pré-définir le séquençement des opérations, soit utiliser un système à base de règles qui fait le choix de l'opération suivante à réaliser. Une étude réalisée par Mackaness 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 ou les objets sur lesquelles 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 le choix de l'opération suivante. On est donc naturellement tenté de se tourner vers un système à base de règles qui évaluerait, au cours du processus, les nouvelles caractéristiques des données et permettrait de choisir pas à pas la résolution la mieux adaptée en fonction des spécifications 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'heure actuelle satisfaisant. A notre avis, l'échec 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 de formalisation des connaissances nécessaires aux choix et aux contrôles.

Dans ce texte, nous essaierons :

- d'analyser ce qui déclenche le besoin de réaliser des opérations de généralisation: les conflits,
- de proposer processus type,
- d'identifier l'information sémantique et géométrique minimale,
- de proposer un moyen pour représenter et d'utiliser les relations de proximités entre objets non connexes,
- de proposer un partitionnement géographique de l'espace afin d'optimiser le choix du séquençement 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 plate-forme de généralisation où les données sont représentées en mode objets et où le séquençement des actions est géré à l'aide d'un système à base de règles.

2- Les conflits

Un conflit correspond à une situation où des faits ne correspondent pas aux règles que l'on s'est fixées. Dans le cas de la généralisation, on considérera qu'il y a conflits 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. A l'heure actuelle très peu d'études ont été réalisées sur les conflits en généralisation [8]. Nous décrivons donc simplement quelques principes permettant d'appréhender cette notion. Un conflit:

- est dû à la résolution géométrique de la base de données. Il devient plus "important" si l'on affecte aux objets une symbolisation (dans le cas de la généralisation cartographique),
- peut exister entre objets de nature différente (ex: entre une maison et une route),
- peut exister entre objets de type de géométrie différent (point, ligne, surface),
- est potentiellement transitif: si Oa et Ob sont en conflit et Ob et Oc sont en conflit, Oa et Oc sont potentiellement en conflit. Lors de la résolution du conflit entre Oa et Ob il faudra donc contrôler l'effet de la résolution sur l'objet Oc.

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

- un conflit global représente une zone comprenant 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 différencier 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 ou zone de conflit entre 2 lignes),
- les conflits intra-objets sont qualifiables par des critères de taille (surface trop petite, trop fine, lignes trop petites, trop détaillées...).

La connaissance des conflits a plusieurs utilités lors du processus de généralisation:

- l'identification des conflits permet de déterminer les zones à traiter,
 - le suivi local du nombre de conflits permet d'évaluer la qualité d'une résolution choisie,
 - la qualification d'un conflit permet de choisir l'opération ou le séquençement 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 inutile 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 conflits intra-objets (simplification et amplification de la surface) se fera 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:

- A- détection et qualification des zones de conflits,
- B- sélection d'une zone de travail,
- C- identification des conflits locaux:
 - 1- choix d'un conflit à résoudre (donc choix d'un ou plusieurs objets à traiter),
 - 2- sélection d'une opération de généralisation,
 - 3- choix de l'algorithme et de ses valeurs paramétriques en fonction des objets,
 - 4- transformation des objets,
 - 5- validation interne en fonction de l'évolution du conflit entre le ou les objets retour à l'état précédant (C-2 ou C-3) si le résultat n'est pas satisfaisant,
 - 6- diffusion de l'effet de la transformation sur l'environnement pour garder les cohérences spatiales (géométriques, topologiques...),
 - 7- validation contextuelle en vérifiant les effets de la propagation retour à une autre diffusion (C-6) ou une autre résolution (C-2) si le résultat n'est pas satisfaisant,
 - 8- retour en C,
- D- retour en A.

Ce processus suit toujours la même logique répétitive : évaluation de l'état des données, choix d'une opération ou d'un algorithme, réalisation de la transformation choisie, validation interne et externe: retour à un état précédent ou progression.

On peut constater qu'il y a toujours plusieurs niveaux de qualification des données:

- globalement, il faut arriver à faire converger les données de sorte que les spécifications globales de la base de données soit respectées,
- localement, il faut progressivement résoudre les conflits sans en créer de nouveaux.

On voit qu'il est indispensable de disposer:

- d'outils permettant de qualifier les objets (identification des caractéristiques des objets),
- d'outils permettant d'identifier et de représenter les conflits,
- d'outils permettant de contrôler les déformations réalisées (ex: comparaison entre la nouvelle géométrie d'une route et sa géométrie initiale).
- de mécanismes permettant de revenir à un état précédent.

4- Représentation de l'information de base

4.1: L'information sémantique

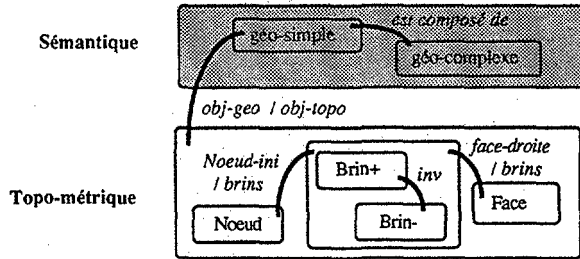
Le schéma de données permet de décrire les attributs des objets et les liens entre objets de classes différentes. Ces informations ne seront pas suffisantes pour la généralisation. Il va falloir associer à chaque classe des connaissances permettant de faire les choix décrits au chapitre précédent.

Parmi les attributs d'ores et déjà identifiables on peut citer:

- des opérations potentiellement réalisables,
- des algorithmes utilisables pour chaque type d'opérations (on n'utilisera pas les mêmes algorithmes pour agréger des maisons et des surfaces naturelles),
- des pré-ordre 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 décrire 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 connexité. Pour représenter ces relations topologiques, nous utilisons la modélisation sous forme de cartes topologiques [1] qui permet de se déplacer facilement sur le graphe et de connaître les relations d'inclusion. Cette couche peut dans notre cas être appelée topo-métrique, parce que ce sont les brins (arcs orientés) qui portent également la localisation des objets. Nous pourrions constater par la suite que cette couche est au coeur de la modélisation choisie puisque toutes les autres couches d'information lui sont liées, et que c'est elle qui va permettre de rétablir la cohérence entre les différentes connaissances.



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 connexes. 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 les calculs sont assez simples, le fait de savoir que 2 objets font partie d'une même maille permet d'en déduire que les objets sont proches mais pas de décrire la façon dont ils sont positionnés l'un par rapport à l'autre. De plus les maillages posent toujours des problèmes à leurs bords.

On peut également représenter les relations de proximité à l'aide d'un réseau permettant de relier des points d'ancrage de chaque objet. Jones [2] utilise une triangulation de Delaunay pour décrire les relations de connexité 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

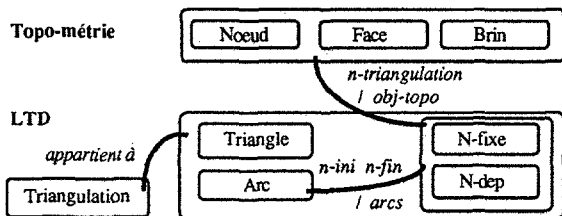
Nos triangulations ne sont pas calculées sur toute la zone à généraliser mais sur des zones réduites pré-définies. On parlera donc de triangulations locales. Il peut y en avoir plusieurs à un instant donné.

Les TLD ne portent pas l'information topo-métrique, mais y sont reliées. Elles sont calculées pour certaines opérations et détruites aussitôt.

Une TLD est composée d'un ensemble de noeuds et d'arcs reliés par des relations de connexité. Les noeuds représentent des points d'ancrage entre des objets topo-métriques et la triangulation. Chaque noeud 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'efficacité de la triangulation dépendra donc du choix des noeuds de la triangulation. Ces choix dépendront du type d'action à réaliser ou du type d'information recherchée (ex: alignements de maisons, conflits de proximité).

Les triangulations sont calculées par la méthode de Tsai [12] par un calcul de l'enveloppe convexe des noeuds initiaux et insertion des noeuds un à un.

Les noeuds de la triangulation sont classés en fonction de la nature des objets géographiques qu'ils relient afin de pouvoir plus facilement différencier leur comportement lors des propagations des déformations.

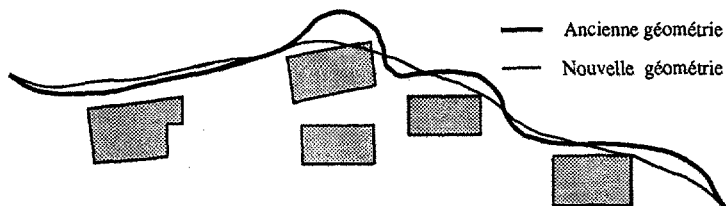


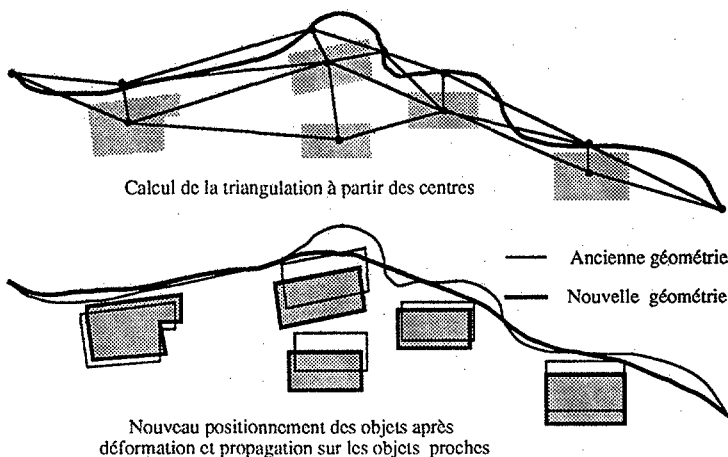
5.3: Les TLD pour la gestion des déplacements

On se propose ici d'utiliser une TLD pour propager la déformation d'un objet (créée par une opération quelconque) sur ses objets proches non connexes, afin d'éviter des superpositions d'objets. On peut imaginer facilement le cas de la simplification de la géométrie d'une route et la propagation des déplacements sur les maisons proches. Le déplacement d'une maison pouvant créer un nouveau conflit, il convient de correctement délimiter l'espace de propagation (les frontières de la zone devant rester fixes) et de prendre en compte tous les objets à l'intérieure de la zone de déformation.

La démarche est la suivante:

- 1- on choisit l'élément à déformer *b-def*, sa géométrie avant et après déformation: (vx,vy) (vx',vy') ,
- 2- on identifie tous les objets non connexes proches. Pour ce faire, on repère les 2 faces topologiques (ou une autre zone de travail telle qu'une partition cf. 6) dont l'objet linéaire fait partie, et on repère tous les objets *{brins-z}* délimitant les zones de travail ainsi que les objets contenus dans ces zones,
- 3- pour chaque objet isolé faisant partie des zones (ex: les maisons) on choisit un point d'ancrage (en pratique leur centre constitue un bon point), et on projette l'ensemble de ces points sur les brins *{brins-z}*. Évidemment on utilise la géométrie de l'objet initialement déformé avant sa déformation. On dispose ainsi d'un ensemble de noeuds de la triangulation *{noeuds-t}* provenant soit d'objets isolés (ex: les maisons) soit des brins délimitant les zones (ex: les routes),
- 4- à partir de cet ensemble de noeuds 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-def*. Pour cela on projette les points d'ancrage de *b-dep* sur sa nouvelle géométrie (vx',vy') .
- 6- on propage ces déplacements via les arcs de la triangulation sur les objets proches.





Dans la mesure où l'on dispose d'une couche topo-métrique, l'identification des objets à déplacer est simple. Par contre, il convient de choisir une bonne méthode de propagation des déformations par la triangulation sachant que:

- il est inutile de déplacer un objet éloigné,
 - les déformations sont amorties en fonction de l'éloignement d'un objet à l'objet initialement déformé, les brins délimitant les zones de travail doivent rester fixes,
 - il faut éviter des phénomènes de cycle 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 dépendent donc des directions: certains objets se déplaceront 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 noeud comme étant la moyenne des vecteurs de déplacements reçus, pondérée par la distance entre le noeud considéré et les noeuds qui lui propagent un déplacement à l'aide des arcs de la triangulation.

En un noeud n-centre on a donc:

$$\text{Vect-Dep (n-centre)} = \left\{ \sum_i (\text{vect-dep}_i / \text{dist}_{i \rightarrow \text{n-centre}}^2) \right\} / \left[\sum_i (1 / \text{dist}_{i \rightarrow \text{n-centre}}^2) \right]$$

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 noeuds de la triangulation qui ne peuvent pas bouger (ce sont les noeuds appartenant au contour de la zone), et qui doivent empêcher les noeuds libres de trop se rapprocher d'eux. Pour cela il faut simplement tester la distance séparant un noeud libre d'un noeud de fixation et d'interdire ou de limiter le déplacement à l'aide d'un calcul d'angle:

Soit le noeud n-fix, lié à noeud n-centre. On a donc:

$$\text{Vect-dep final (n-centre)} = \text{Norme (Vect-Dep (n-centre))} * \cos [\text{angle (Vect-dep (n-centre), Vect (n-centre - n-fixe))} / 2]$$

L'inconvénient de cette méthode de calcul est que les noeuds libres sont liés entre eux et se propagent des déplacements (puisque certains noeuds de E-centre ne sont reliés qu'à des noeuds de E-centre). Selon l'ordre du calcul des propagations on obtiendra des résultats assez différents et il est impossible de décider a priori le meilleur séquençement des calculs. De plus les forces de fixation ne pourront avoir un effet que sur les noeuds avec lesquels ils sont directement reliés. Nous nous sommes donc tournés vers un modèle de forces plus complexe mais plus rigoureux, inspiré des modèles utilisés en mécanique. A l'heure actuelle seule la formalisation du problème sous forme d'équations a été réalisée. Nous travaillons sur la recherche de méthodes simples de résolution.

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

Rappel:

$$L = C - U \quad (1) \quad \text{avec} \quad C = \text{somme de l'énergie cinétique en chaque point } 1/2 m_i v_i^2$$

$$U = \text{somme des énergies potentielles.}$$

Si $(q_1, q_2, q_3, \dots, q_n)$ est un système de paramètres indépendants définissant la position du système, le lagrangien, noté alors $L(q, q', t)$ est une fonction de q_1, q_2, \dots, q_n , des dérivées q'_1, q'_2, \dots, q'_n et du temps. Le système d'équations différentielles :

$$d/dt (\delta L / \delta q'_i) - \delta L / \delta q_i = 0 \quad (2)$$

est appelé système d'équations de Lagrange du mouvement.

Dans notre cas, nous disposons de noeuds décomposés en 2 ensembles:

E-centre = $\{n\text{-centre}_i\}_i \in [1, n\text{-c}]$, composé des noeuds mobiles, qui exercent des forces entre eux et pour lesquels nous connaissons leur position initiale et nous recherchons leur position finale. Les équations de Lagrange seront réalisées sur cet ensemble.

E-zone = $\{n\text{-zone}_j\}_j \in [1, n\text{-z}]$, 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:

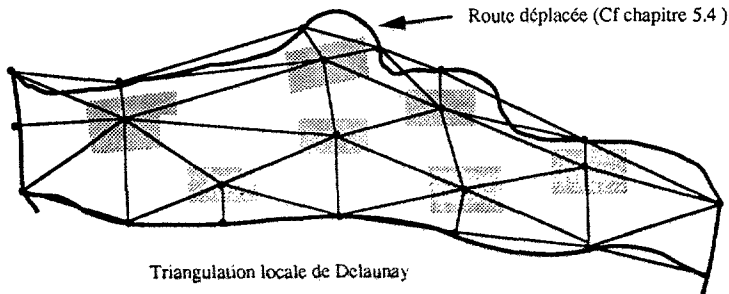
E-zone-dep = $\{n\text{-z-dep}_j\}_j \in [1, n\text{-z1}]$ composé des noeuds qui ont été initialement déplacés. On considère donc qu'ils ont une nouvelle position connue et une vitesse initiale nulle. A ce stade on ne remet pas en cause l'action initiale qui déclenche la recherche d'une solution. Chaque déplacement exerce des forces de déplacement sur les noeuds de E-centre auxquels ils sont liés.

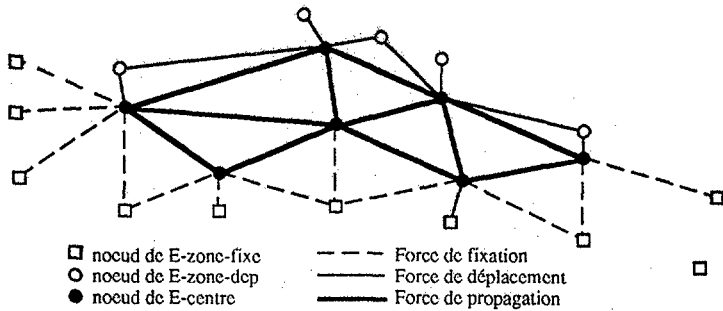
E-zone-fix = $\{n\text{-z-fix}_j\}_j \in [1, n\text{-z2}]$ composé des noeuds qui restent fixes. On doit les considérer puisqu'ils exercent des forces de fixation sur les noeuds de E-centre.

On distingue trois type de forces:

- 1- les forces F_c exercées entre les noeuds de E-centre qui se comportent comme des ressorts:
 $F = -k_c \text{ dep}$. 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-zone-dep et les noeuds de E-centre, qui sont également des ressorts $F_d = -k_d \text{ dep}$.
- 3- les forces F_f 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:





En chaque point de E-centre on doit donc résoudre l'équation différentielle (2) à partir du Lagrangien:

$$\begin{aligned}
 L = & \sum_{j; i \in E\text{-centre}} 1/2 m q_i'^2 - 1/2 \sum_{i \in E\text{-centre}} \sum_{j \in E\text{-zone-dep}} k c_{ij} * q_i^2 \\
 & - 1/2 \sum_{i; i \in E\text{-centre}} \sum_{j; j \in E\text{-zone-dep}} k d_{ij} * q_i^2 \\
 & - 2 \sum_{i; i \in E\text{-centre}} \sum_{j; j \in E\text{-zone-fix}} k f_{ij} * \\
 & [(q_i * \sin((q_i - q_{ij0}) / 2)) + 2 * \cos((q_i - q_{ij0}) / 2) - 2 * \cos(q_{ij0} / 2)]
 \end{aligned}$$

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 partitionnements est de se définir des zones de travail de différentes tailles afin de mieux maîtriser le séquencement 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 n'importe quel type de traitement mettant 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 : quadtree) 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:

les objets doivent:

- être adaptés à la création de cycles (pour la création de zones),
- être facilement hiérarchisables (par l'intermédiaire d'un code sémantique si possible),
- être maintenus durant 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'objets structurants pour la base de données que nous voulons généraliser à l'IGN: la BD Topo, qui est une base de données topographique à "moyenne échelle" (i.e. métrique). En se basant sur une étude visuelle des objets conservés lors des généralisations manuelles des cartes IGN il nous est apparu que le réseau routier semblait être très bien adapté à notre objectif:

- son graphe associé comporte naturellement des cycles,
- il est généralement maintenu, puisque l'information "communication routière" prend de plus en plus d'importance lorsque l'échelle diminue,
- il est déjà classé (au moins en Autoroute, Nationale, Départementale)
- sa densité est étroitement 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 brins de la partition de plus haut niveau (simplification, caricature...), puis on résout les conflits

entre brins de même niveau (déplacements...), puis on considère que la structure est stable et on recommence la même démarche pour les niveaux inférieurs. Enfin, on généralise les objets compris à l'intérieur des partitions en considérant que l'espace délimité par chaque partition est fixe. Chaque modification géométrique donne lieu à un calcul de TLD pour répercuter les déformations géométriques afin de conserver la base dans un état cohérent.

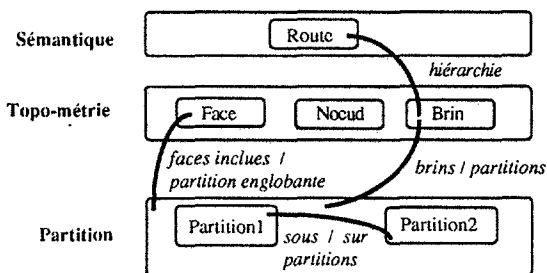
Cette méthode n'implique pas qu'il n'y aura pas de remises en cause locales de ce séquençement ou d'élimination des certains brins composant la structure (pour les niveaux les plus bas). Rien n'empêche pour autant de l'utiliser pour la résolution d'un nombre important de conflits et de régler les autres conflits dans une étape ultérieure. A l'heure actuelle, il est encore difficile de se prononcer sur les limites d'utilisation du partitionnement hiérarchique.

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.

7. Utilisation des différents formalismes dans le processus de généralisation:

Dans le chapitre 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 délicate à gérer. Il ne faut pas que certains 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 couches d'information soient reliées et nous avons développé des méthodes permettant ~~assurer~~ le passage cohérent d'une couche à une autre. En observant les différentes modélisations, on ~~peut constater~~ que la couche topo-métrique est le lien entre ces connaissances.

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-ordre 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 considérant que les objets de ce niveau:

- 1- Résolution des conflits intra-objets, avec propagation des effets de bords à l'aide des TLD,
- 2- Résolution des conflits inter-objets, avec propagation des effets de bords à l'aide des TLD,

Les objets traités sont par la suite considérés comme fixes et intangibles.

- 2- Généralisation des autres objets.

8- 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 séquençement des opérations. Nous avons donc développé sur un système orienté objet différentes couches d'information : sémantique, topo-métrique, des triangulations locales (pour décrire les relations de proximité entre objets non connexes) et le partitionnement hiérarchique (pour faciliter le séquençement 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
- [2] Jones, Ware & Bundy *Multiscale spatial modelling with triangulated surfaces* SDH'92 Vol. 2, pp 612-621.
- [3] Lagrange & Ruas *Geographic information modelling: GIS's and Generalisation* SDH'94 Vol. 2 pp 1099-1117.
- [4] Locordix *Méthodes algorithmiques 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] Leimer *Prototype rules for automated map generalization* Master of Geography Buffalo 1993
- [8] Mackaness *Issues in resolving visual spatial conflicts in automated map design* SDH'94 Vol. 1 pp 325-340.
- [9] McMaster & Shea *Cartographic generalization in digital environment: When and How to generalize* AUTOCARTO 9 1989
- [10] Peng *Automated generalization of urban road-networks for medium scale topographic data-bases* PhD ITC. 1992
- [11] Plazanet *Measurements, Characterization and Classification for Automated Line Feature Generalization*. AUTOCARTO 12. Vol. 4 pp 59-68. 1995
- [12] Tsai *Fast topological construction of Delaunay triangulations and Voronoi diagrams* Computers & Geosciences Vol. 19 N 10 pp 1463-1474. 1993