

LES DÉFIS DU NUMÉRIQUE

Sylvie Lamy *

RÉSUMÉ. L'acquisition de données géographiques sous forme numérique devrait permettre la réalisation rapide de produits adaptés à chaque besoin. Toutefois, les systèmes d'information géographique (SIG) n'offrent actuellement que peu d'outils suffisamment souples. Le laboratoire de recherche Cogit de l'IGN travaille à l'élaboration de méthodes et d'outils adaptés à des tâches complexes telles que la généralisation, le positionnement de toponymes et la gestion des mises à jour des données.

ABSTRACT. The accessibility to geographic data in numerical form should allow for the rapid creation of products adapted to the specific needs of everyone. However, today's geographic information systems (GIS) offer but few tools which are sufficiently versatile. The research laboratory Cogit, located at the IGN (National Geographic Institute) is working towards the development of methods and tools adapted to complex tasks such as generalization, positioning of toponyms and the management of data updating.

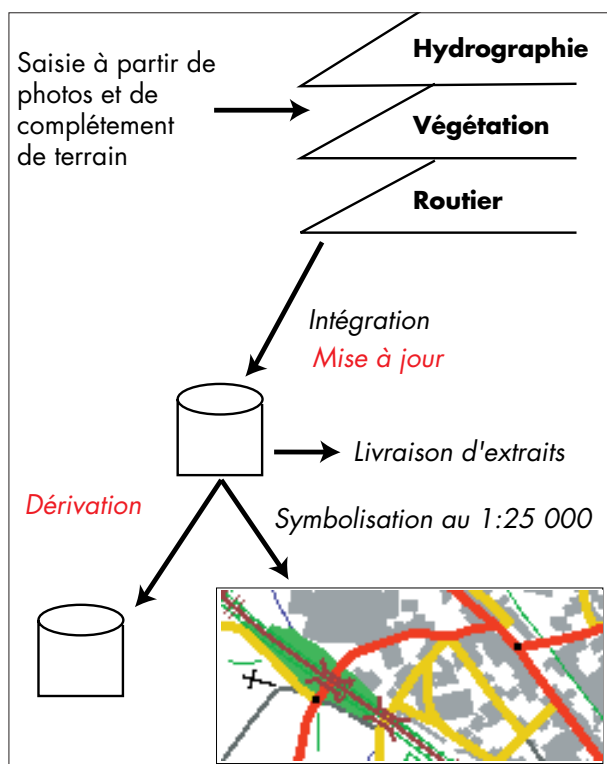
RESUMEN. La adquisición de datos geográficos numéricos deberían permitir la realización rápida de productos adaptados a cada necesidad. Sin embargo, los sistemas de información geográfica (SIG) ofrecen ahora solo pocos instrumentos bastante flexibles. El laboratorio de investigación Cogit del IGN (Instituto Geográfico Nacional) trabaja para elaborar métodos y instrumentos adaptados a tareas complejas tal como la generalización, el posicionamiento de topónimos y la gestión de actualización de los datos.

La carte a finalement peu évolué depuis plusieurs décennies. Si la gamme des produits cartographiques s'est considérablement élargie afin de répondre aux attentes des utilisateurs, une carte reste aujourd'hui un document figé, correspondant aux attentes générales du plus grand nombre et aux désirs particuliers d'aucun. Qui ne s'est pas agacé de voir une limite de carte tomber juste au milieu de « la zone intéressante » ? Chacun voudra accéder à une cartographie « à la carte », où l'utilisateur définira sa zone géographique, les éléments qu'il désirera voir figurer, l'échelle de sortie. Si cela reste aujourd'hui du domaine de l'utopie, les systèmes d'information géographiques (SIG) et les bases de données géographiques constituent certainement les premières bribes de réponse technique pour assurer cette évolution. Dès à présent, un SIG permet de réaliser une carte, avec des échelles et un contenu (géographique et thématique) variables. Cependant, si aucune précaution n'est

prise, les résultats sont décevants. Aussi l'un des premiers défis du numérique est-il de pouvoir dériver des cartes sur papier d'une base de données !

Les bases de données géographiques sont spécifiées pour répondre à certain niveau d'abstraction ou, par référence aux cartes, à une gamme d'échelles. Si, techniquement, il est possible de le faire, pratiquement, on ne peut utiliser directement la même base de données pour faire une sortie au 1:10 000 et au 1:100 000 ; cela correspondrait à photocopier un document en le réduisant d'un facteur 10. De plus, l'ensemble des informations n'est pas utile à tous les niveaux d'abstraction. Ainsi, lorsque vous allez de Paris à Marseille en voiture, vous allez choisir un itinéraire en consultant une carte au 1:1 000 000 et non les quelques 100 cartes au 1:25 000 correspondant à votre trajet ! L'excès d'informations est souvent pénalisant. Il y a par conséquent des choix à

* Laboratoire COGIT (Conception objet et généralisation de l'information topographique), 2 av. Pasteur, BP 68, 94160 Saint-Mandé, 01 43 98 80 00



1. Traitement des données géographiques

faire dans la sélection des éléments que l'on veut représenter, des simplifications de géométrie à réaliser, des regroupements d'éléments à effectuer. Aujourd'hui, c'est le cartographe qui maîtrise l'ensemble de ce processus, que l'on appelle *généralisation cartographique*.

Le message délivré par une carte est en effet extrêmement complexe : l'esthétisme, la lisibilité, la compréhension restent en grande partie des valeurs humaines, que l'on a peine à coder, à qualifier, à mesurer avec des règles strictes. La totalité des virages d'une route sinueuse ne sera pas représentée sur une carte à petite échelle pour des raisons de lisibilité, mais sa symbolisation devra rendre compte de cette sinuosité. De même, le cartographe devra décider qu'une route doit être conservée pour une carte à une échelle inférieure, par exemple parce qu'elle est un accès obligé à un lieu touristique. Un traitement numérique supposera que l'on puisse qualifier une sinuosité et simplifier une représentation pour le premier exemple, mesurer (chiffrer) l'importance de la route pour le second. Il est illusoire de penser que, dès aujourd'hui, on peut s'affranchir des contraintes d'échelle, et qu'à partir d'une seule base de données très précise, on peut dériver des bases dont le niveau d'abstraction est plus élevé. L'automatisation des productions cartographiques passe

notamment par la compréhension et l'identification de l'information implicite contenue dans une carte, et dans la base de données originelle.

Un autre problème tient au placement des écritures. Dans les bases de données, les informations comme le nom d'un cours d'eau ou le nom d'une commune sont portées par les objets géographiques. Avec un système d'information géographique, on peut donc à tout moment, en désignant une partie quelconque d'une route, savoir de quelle départementale ou nationale il s'agit. Sur le papier, cette information doit figurer de manière explicite, non ambiguë, permanente et esthétique. Il faut éviter le recouvrement des écritures, répéter les numéros de route autant de fois que nécessaire. En outre, le placement des écritures n'est valable que pour une échelle, une zone géographique et des polices de caractères données. Changer l'un de ces paramètres oblige à les repositionner. Le processus actuel de placement des écritures est en grande partie manuel. Des recherches ont été engagées, afin de déterminer les positions optimales de placement, à partir des informations contenues dans la base de données. Une modélisation relativement simple donne dès à présent des résultats satisfaisants pour un nombre important d'écritures.

Si l'automatisation des processus de production de cartes est une application des bases de données, en aucun cas elle ne les justifie à elle seule, ne serait-ce qu'en égard à la difficulté d'automatiser la cartographie. L'apport des bases de données vient en grande partie de la multiplicité des applications liées aux possibilités de calculs : calcul d'un itinéraire, statistiques... Pour la plupart de ces applications, les informations doivent être justes. Si notre œil a l'habitude de corriger certains défauts, et si l'«on sait» que tout produit manufacturé a ses limites, un système informatique est généralement aveugle. Il ne calcule qu'avec les informations qu'il possède, et ne suppose pas qu'il lui en manque, ou que certaines informations sont erronées. Les bases de données doivent, par conséquent, être constituées et mises à jour avec des processus fiables ; en particulier, la cohérence des informations doit être en permanence vérifiée.

Plusieurs facteurs rendent cette tâche délicate. En premier lieu, les SIG ne sont pas aujourd'hui capables de gérer l'énorme volume de données que représenterait par exemple la couverture en numérique au 1:25 000 de la France (plusieurs milliers de Go). Pour cette raison, pour des bases de données à des moyennes ou grandes échelles, dès que la zone est étendue, il faut «découper la base» en dalles (par

exemple en rectangles, ou en sections cadastrales). Ces découpages imposent une vigilance particulière pour raccorder les dalles afin de pouvoir les utiliser avec des logiciels de calcul. Le raccord visuel assuré par la cartographie traditionnelle devient insuffisant. En second lieu, les utilisateurs de SIG ajoutent en général leurs propres informations à des bases de données exogènes (couverture au 1:100 000, cadastre numérique, données géologiques, etc.). Afin de maintenir leur système, ils se doivent d'intégrer les mises à jour de ces bases exogènes, sans perdre leurs propres informations. Ce processus d'actualisation fait l'objet de recherches, qui couvrent le domaine plus large de la gestion du temps dans les SIG. L'intégration et l'utilisation de données d'origines diverses sont par ailleurs d'autant plus délicates que chacun voit le monde et, par conséquent, conçoit sa base de données à sa manière. Ainsi une route sera modélisée différemment pour l'automobiliste (ce qui inté-

resse est la logique de circulation : comment aller d'un point à un autre), le responsable de l'entretien des routes (qui a besoin d'informations techniques), le piéton en ville (il lui faut les noms des rues). Un même phénomène sera vu de trois manières différentes. Certaines applications exigeant à la fois les informations « du promeneur » et de « l'automobiliste », il s'agira donc de savoir reconnaître les représentations différentes d'une même réalité géographique.

La cartographie du XXI^e siècle supposera non seulement une souplesse dans la conception (utilisation de données d'origines diverses, choix des utilisateurs plus importants), mais aussi certainement une intégration plus large des possibilités offertes par le numérique (animations à l'écran, visualisation de cartes en trois dimensions). Les bases sont jetées, mais nous n'en sommes aujourd'hui qu'aux balbutiements de l'utilisation de l'information géographique.

Vers la généralisation automatique

La comparaison attentive de cartes topographiques à différentes échelles montre que les thèmes représentés diffèrent et qu'un même objet topographique peut avoir des représentations distinctes. La carte au 1:100 000 est certes produite à partir de la carte détaillée au 1:25 000, mais le processus sous-jacent n'est pas plus une photographie réduite du document de référence qu'un nouveau dessin approximatif. Le passage de l'une à l'autre est le fruit d'une des activités les plus complexes de la cartographie : la généralisation.

Sur l'exemple de Ponthieux, on peut observer plusieurs différences expliquées par la généralisation :

- *Sélection due aux spécifications* : au 1:100 000, les thèmes « borne géodésique » ou « terrains de sport » ne sont pas représentés ; les courbes de niveau sont de 20 m en 20 m, contre tous les 5 m au 1:25 000.
- *Sélection due aux contraintes de rendu graphique* : réduction du nombre de rues et des alignements de maisons à l'ouest de la zone, et des toponymes.
- *Simplification des formes* : notamment sur chaque courbe de niveau retenue, dont les sinuosités mineures sont lissées ; pour les mêmes raisons de lisibilité, les bâtiments aux contours complexes ne sont plus repré-

sentés que par des rectangles.

- *Déplacements* : la disposition en fond de vallée de la rivière, des deux routes et de la voie ferrée est remarquable et doit être préservée sur la carte au 1:100 000. Mais, parce que ces éléments parallèles sont trop proches, le cartographe a été obligé de les espacer légèrement. On peut s'en rendre compte en mesurant la largeur entre les routes de part et d'autre de la rivière, au niveau de la gare : rapportée au terrain, cette largeur est de moins de 250 m si on la mesure au 1:25 000, et de 300 m si on la mesure sur la carte au 1:100 000.

- *Exagération des tailles* : pour être visibles, les petites dimensions doivent être exagérées ; ainsi le dessin de la maison à l'est du virage sous Bellevue est moins de deux fois plus petit au 1:100 000 qu'au 1:25 000. Le rapport d'échelles étant de 4, la maison est donc relativement deux fois plus grosse.

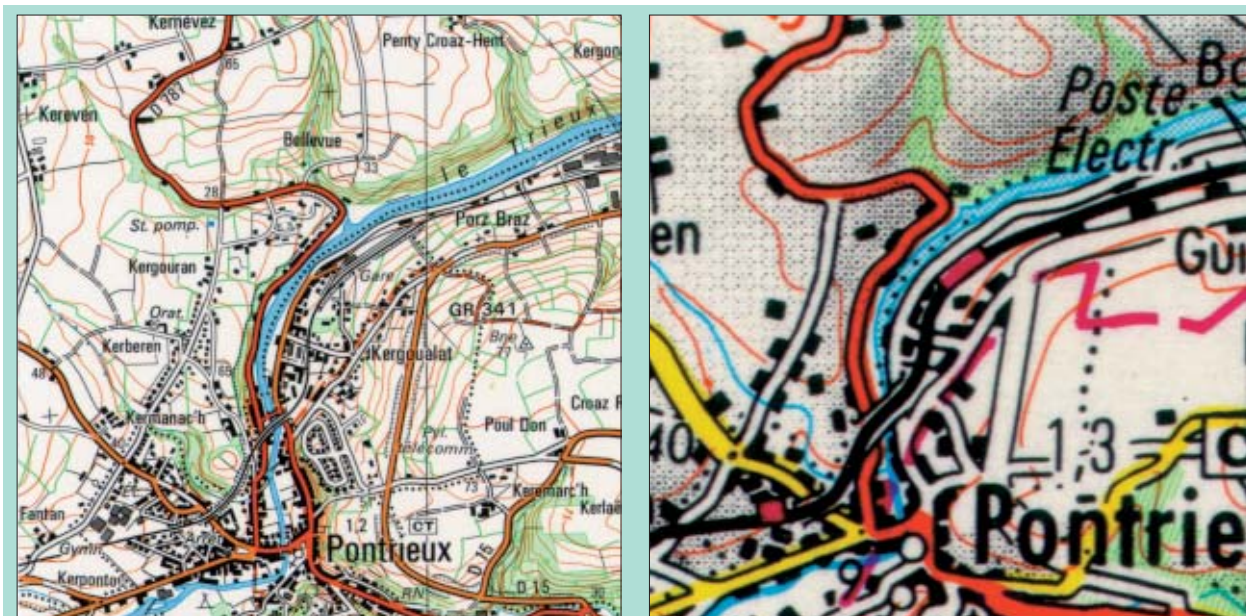
- *Agrégations* : dans certains cas, l'exagération des tailles des bâtiments conduirait à la formation de pâtés amorphes. Quand l'habitat est moyennement dense, il est préférable de représenter des groupes par un bâtiment unique. Il faut alors faire attention à préserver les rapports de densité d'une zone à une autre.

- *Changements de représentation* : la toute petite gare, au 1:100 000, est symbolisée par un rectangle fuchsia, alors que les voies sont encore représentées au 1:25 000. Le petit cimetière, au sud de Ponthieux, décrit au 1:25 000 par son emprise et une trame de croix sur sa surface, n'est plus représenté que par un symbole cruciforme au 1:100 000.

La généralisation automatique en zone urbaine

Le caractère contextuel de la généralisation apparaît intensément en zone urbaine. Le moindre déplacement, la moindre exagération locale pour faire ressortir un détail induit la confusion ou la superposition : une rue qu'on simplifie traverse des bâtiments, une maison qu'on agrandit empiète sur d'autres maisons ou sur la rue. En plus de ces opérateurs de déformation d'objet (simplification, équarrissage, amplification), il faut donc des opérateurs capables de traiter des ensembles d'objets et de les simplifier en bloc, de façon à éviter les effets d'entraînement.

Les opérateurs contextuels (déplacement propagé, amalgamation, structuration), consacrés au traitement de l'espace libre ou libérable entre les objets, ont chacun des



2. La généralisation : Pontrieux (22), cartographie traditionnelle, 1:25 000 et 1:100 000 déduite (ici grossie quatre fois)

conditions d'application particulières. Ainsi, la propagation de déplacement ne s'effectue que sur des objets qu'un espace libre peut recevoir. En cas d'impossibilité, on peut tenter d'agrèger, en représentant dans un seul contour les objets rapprochés. Une suite d'objets régulièrement espacés ne subira pas nécessairement une agrégation : la barre créée ne serait pas représentative de l'espacement régulier. On lui préférera alors une structuration : la représentation de quelques éléments seulement, dont on réorganise la localisation de sorte que soit conservée l'idée de régularité.

Un cartographe sur poste interactif peut facilement utiliser ces opérateurs automatiques, à partir du moment où il dispose d'outils de sélection des objets. Son juge-

ment lui permettra de choisir les groupements représentatifs, ainsi que le type d'opérateur adapté ; éventuellement, de reprendre des résultats peu satisfaisants. Mais ces opérateurs restent de bas niveau (fig. 3) et l'automatisation de la généralisation est en fait plus ambitieuse : il s'agit de réduire autant que faire se peut la part d'interactivité avec la machine. Les recherches en cours portent donc sur deux points : l'identification des proximités et des groupements constitutifs du paysage de la ville, l'émulation optimisée du cycle tentative-validation du cartographe.

Pour reconnaître automatiquement les proximités et les groupements, nous utilisons à l'IGN trois approches complémentaires : l'approche *géométrique*, qui consiste

à quantifier les proximités par le calcul de triangulations entre les éléments ; l'approche *perceptuelle*, qui consiste à déceler des groupements visuellement frappants à partir des critères de formes, d'espacement, d'orientation, etc., définis dans cette branche de la psychologie qui s'appelle le gestaltisme ; l'approche *phénoménologique*, qui consiste à associer les objets entre eux par des critères géographiques simples comme le critère de *desserte* entre bâtiments et rues.

L'activité la plus complexe à automatiser est le cycle de contrôle du processus. Il faut pouvoir apprécier une situation, c'est-à-dire déterminer les groupements, mesurer leurs dispositions, et être capable de choisir l'opérateur adapté. Il faut pouvoir mesurer les écarts des transformations à la carte de référence pour ne pas trop détériorer la précision. Il faut pouvoir juger de la conservation de la forme de la ville. Il faut pouvoir juger de l'uniformité des traitements d'une zone à l'autre de la carte : par exemple, les rapports de densité entre le centre-ville et la périphérie doivent être conservés. Il faut pouvoir identifier et reprendre des résultats mauvais, revenir en arrière. Les problèmes semblent s'accumuler sans fin. La recherche a encore de beaux jours devant elle. – Jean-François Hangouët, COGIT



3. Structuration automatique

Données initiales, et généralisation de la structure du bâti dans un rapport de résolution 5/7.

Généralisation automatique des objets géographiques linéaires : le cas des routes

Une route symbolisée par un ruban de 1 mm de large sur la carte de France au 1:1 000 000 aurait 1 km de large sur le terrain. Une série de 10 virages en épingle à cheveux serait engloutie par son symbole. Le cartographe qui généralise supprime des virages, lisse des détails et amplifie des accidents ou des formes importantes. Les algorithmes existants ont des effets différents en fonction de la géométrie des objets. Pour automatiser la généralisation, il s'avère nécessaire d'enrichir le niveau géométrique des modèles classiques dans les SIG.

Pour décrire la géométrie d'un objet linéaire, on considère son homogénéité et les types de formes qui le composent (portions droites, virages mous, virages très serrés). Nous proposons de décrire les objets linéaires sous la forme d'un arbre comportant des attributs descriptifs : qualification des formes et mesures morphologiques de plus en plus fines. L'arbre est construit en deux étapes : 1) une segmentation des objets linéaires en portions auxquelles on pourra appliquer les mêmes opérations, 2) une analyse des portions pour choisir les opérations adéquates. Ces phases sont récursives pour permettre une analyse de plus en plus fine des formes.

Les outils de segmentation et de qualification s'appuient sur les points caractéristiques (points d'inflexion et sommets). Ces points sont hiérarchisés par une détection sur une ligne de plus en plus simplifiée pour permettre différents niveaux de segmentation et de mesure. Un exemple de segmentation automatique de premier niveau est proposé dans la figure 4. La méthode est fondée sur les variations des distances entre les points d'inflexion consécutifs. Chaque portion est ensuite qualifiée en fonction de sa sinuosité (plus ou moins forte) et en fonction d'une éventuelle répétition de formes types (par exemple une série de virages en épingle à cheveux).

Une fois l'arbre construit, il faut choisir, pour chaque nœud, les solutions de généralisation adéquates (les séquences d'opérations, d'algorithmes et les valeurs paramétriques) en fonction de l'arbre et de



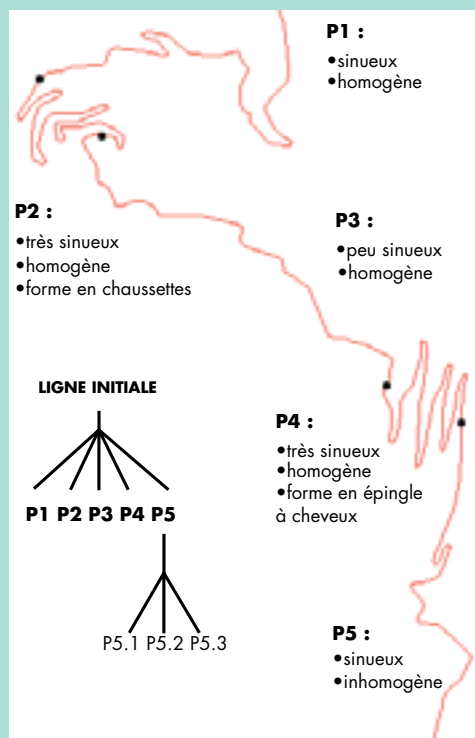
4. Extrait d'une route de montagne BDCarto : 1/au 1:100 000, 2/au 1:250 000, 3/mal généralisée

Au 1:250 000 les virages ne sont plus dissociables, il faut les généraliser. L'algorithme de Douglas (en bas), le plus célèbre pour la simplification, donne un résultat très insatisfaisant : il ne parvient pas à rendre les virages dissociables et les rend trop anguleux.

l'objectif visé (thème de la carte, échelle, symbolisation, etc.). Pour cela, des recherches sont en cours pour corréler types de géométrie et algorithmes de généralisation appropriés.

Proposer une description hiérarchique des objets linéaires fondée sur les points caractéristiques nous a également permis de réaliser un algorithme, baptisé l'accordéon, approprié à l'écartement longitudinal des séries de virages en épingle à cheveux dans les routes de montagne. Cet algorithme pallie un manque d'outils de caricature, y compris dans les systèmes interactifs. Le principe de l'algorithme repose sur un écartement calculé entre chaque point d'inflexion afin d'éviter les effets d'empâtement et une propagation des coordonnées de la ligne entre les points d'inflexion consécutifs.

L'automatisation du processus de généralisation repose donc sur une meilleure analyse de la géométrie des objets à traiter afin de pouvoir plus facilement corréler formes et algorithmes. De façon plus générale, l'effort de compréhension et de modélisation du spatial aura d'autres applications plus vastes dans les SIG. – **Corinne Plazanet, COGIT**

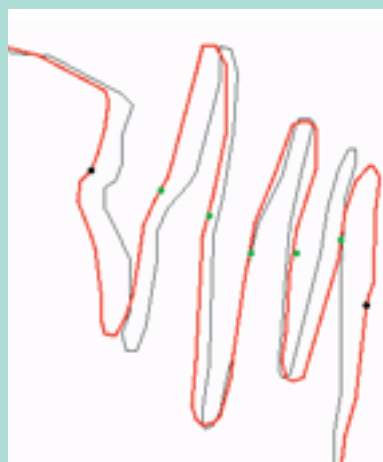


5. Arbres et modes de généralisation

Ci-dessus : segmentation et qualification d'une ligne par

- détection et hiérarchisation des points caractéristiques,
- analyse de la sinuosité dans chaque portion,
- appariement avec des formes types (épingles, chaussettes).

Ci-dessous : l'accordéon appliqué localement à une série de virages très sinueux, en « épingle ».



Le placement automatique des écritures

Les toponymes sont tous les noms présents sur une carte : noms de villes, noms de rivières, numéros de routes... Leur placement manuel sur une carte est coûteux. L'automatisation permet un gain appréciable. Le laboratoire COGIT mène donc depuis plusieurs années des recherches sur le placement automatique des toponymes. L'objectif est que la qualité cartographique de la plupart des placements automatiques soit aussi bonne que celle des placements manuels ; la petite partie des placements automatiques de qualité insuffisante serait corrigée interactivement.

- *Différents types de toponyme.* Les règles cartographiques de positionnement diffèrent selon le type de toponyme. Le COGIT a donc développé différents programmes de placement automatique : écritures horizontales, numéros de routes, kilométrages, noms de rues sur les plans de ville ; une recherche est en cours sur le positionnement des noms de rivière.

- *Programme de placement des écritures horizontales.* Le plus abouti de ces travaux est le programme de placement automatique des écritures horizontales : il place 90 % des écritures horizontales avec une bonne qualité cartographique. Le reste requiert une intervention manuelle. Ce programme a été industrialisé par la société Barco Graphics, au sein du logiciel de rédaction cartographique Mercator. Il sera utilisé pour la production des cartes au 1:1 000 000 et au 1:25 000 de l'IGN.

- *Principes des programmes de placement des toponymes.* L'approche choisie est d'implanter des algorithmes avec des langages informatiques procéduraux, sans recourir à l'intelligence artificielle. Le programme procède en deux étapes : 1)

pour chaque nom à placer, considéré isolément, le programme essaie plusieurs positions, et quantifie la qualité cartographique de chaque position ; 2) le programme sélectionne pour chaque nom une position, dont la qualité cartographique est la meilleure possible, et qui ne recouvre pas d'autre nom.

- *Quantification de la qualité cartographique d'une position essayée.* La qualité cartographique d'une position est le degré de conformité de cette position à l'ensemble des règles cartographiques. Pour chaque règle cartographique, le programme quantifie le degré de conformité d'une position par rapport à cette règle, au moyen d'un poids. Puis il additionne les poids des différentes règles cartographiques. Cette somme quantifie donc la qualité cartographique globale de la position essayée.

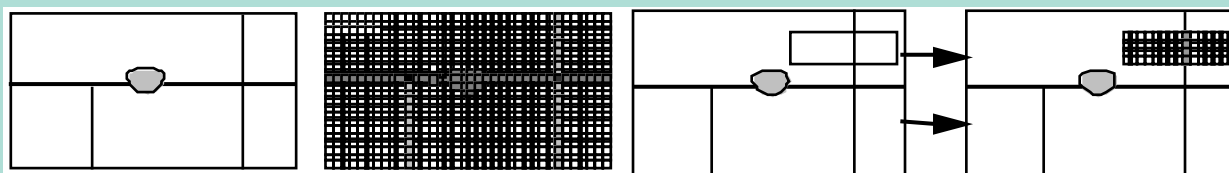
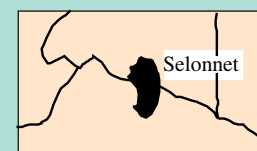
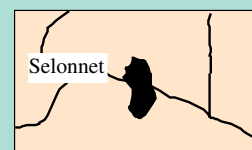
- *Mutilation des détails de la carte.* Une règle cartographique fondamentale est qu'un nom doit mutiler le moins possible les informations de la carte. Il est plus grave de mutiler certains détails que d'autres. Par exemple, mutiler un carrefour est plus grave que mutiler une route. Et mutiler une route importante est plus grave que mutiler une route secondaire (fig. 6).

Pour quantifier la conformité d'une position par rapport à cette règle, le programme établit une image maillée des détails de la carte. Chaque pixel de cette image a un poids, proportionnel à la gravité de la mutilation du détail cartographique se trouvant en ce pixel (fig. 6). La valeur des poids est définie *a priori* en fonction de nos connaissances cartographiques et du type de cartes réalisées à l'IGN. Ces valeurs sont évidemment modifiables pour la réalisation

d'autres produits. Ensuite, le programme somme les poids de tous les pixels recouverts par la position, ce qui donne la qualité de la position.

- *Mutilation entre noms* – Il faut à présent placer chaque nom en tenant compte des placements des autres. Avant tout, deux noms ne doivent pas se mutiler.

Le placement automatique des noms horizontaux ayant produit un résultat cartographique encourageant, le COGIT essaye d'améliorer le positionnement des noms de villes en recherchant une meilleure homogénéisation de la répartition des noms. De plus, le COGIT effectue des recherches sur le placements automatiques d'autres types de noms (noms de rues dans les plans de villes, noms de rivières). En effet, chaque type de toponyme requiert une recherche propre dans la mesure où les règles de positionnement dépendent des objets : les noms de rivières doivent suivre la courbure des rivières, coller à la ligne et être régulièrement répétés. – **François Chirié**, COGIT



6. Mutilation des détails de la carte

Gestion du temps dans une base de données géographiques

Le paysage évolue sous l'effet des actions humaines (construction d'autoroutes, de nouvelles résidences) et naturelles (érosion, tremblement de terre, glissement de terrain). Il est donc nécessaire de mettre à jour nos bases de données géographiques. À cette occasion, certaines informations doivent être retenues pour faciliter la gestion des bases de données : c'est le cas des informations de nature temporelle. Cet article s'attachera donc à présenter les caractéristiques des informations temporelles et leurs modes d'intégration dans une base de données.

Quelques notions temporelles

- *Unité du temps.* Le temps est par essence continu : entre deux dates, aussi proches soient-elles, on peut toujours trouver une date intermédiaire. Or, les outils informatiques ne peuvent pas supporter un tel concept : nous introduisons la notion de «chronon», c'est-à-dire du plus petit intervalle de temps en dessous duquel il n'y a pas de détermination possible. Par exemple, lorsque le chronon est égal à une seconde, aucun événement inférieur à la seconde ne sera répertorié. On peut assez aisément rapprocher la notion de chronon de celle de «résolution» pour ce qui concerne la saisie des données géographiques (la résolution est la taille du plus petit objet représentable dans la base de données géographiques).

- *Temps valide et temps «base de données».* Après avoir choisi une unité de temps, il s'agit de déterminer quelle date retenir : est-ce la date «valide» (date à laquelle un changement s'est produit sur le

terrain) ou la date «base de données» (date à laquelle ce changement a été répertorié et est inséré dans la base)? Dans notre situation, la date «base de données» n'est que la somme de la date «valide» et du temps d'intégration des changements. Par exemple, si l'on décide que la dernière version de la base de données est du 1^{er} janvier 1996, et que six mois ont été nécessaires pour saisir les changements, on peut avancer que cette base reflète l'état du monde réel au 1^{er} juillet 1995. Une suppression de ce décalage temporel serait souhaitable mais s'avère impossible en pratique.

- *Rythme de mise à jour.* Une des caractéristiques des informations géographiques est sa quantité (le stockage des informations présentes sur une carte au 1:25 000 s'élève à environ 50 Mo, équivalant à 50 millions de caractères) et la maintenance sur tout le territoire nécessite de nombreux efforts. On introduit donc la notion de rythme de mise à jour : suivant des considérations techniques, économiques et organisationnelles, une nouvelle version de la base de données sera établie toutes les x années. Le rythme de mise à jour fait partie des spécifications de la base de données et peut influencer sur le contenu de la base (si l'on prévoit un rythme de mise à jour supérieur à un an, le type des cultures non pérennes ne sera pas saisi dans la base).

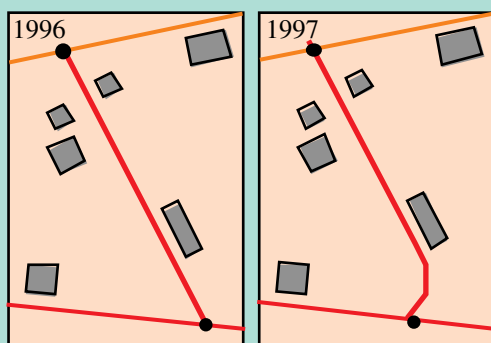
Gestion des mises à jour

La gestion des mises à jour est nécessaire, d'une part, pour assurer l'actualité de nos cartes et, d'autre part, pour assurer une mise à jour des bases de données chez nos clients.

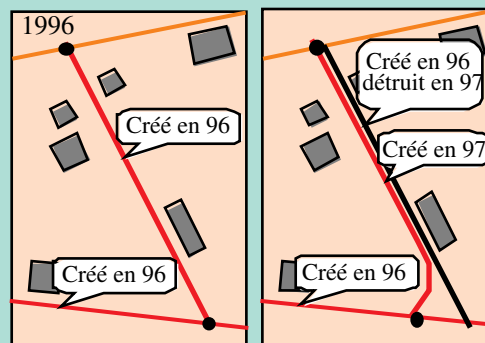
- *Archiver les données.* L'idée la plus simple, concernant la gestion des mises à jour, consiste à faire un archivage périodique de la base de données en procurant ainsi plusieurs clichés du monde réel à des dates différentes : on aura la base à la date $T, T_1, T_2...$ On dispose donc d'un ensemble de versions non corrélées (fig. 7). Un archivage de ce type ne permet pas de fournir des indications sur l'âge des données, sur la quantité de changement qui a été intégrée. En outre, il semble difficile pour le client d'actualiser sa base sans perdre les informations qu'il a lui-même saisies. Pour ce faire, il faut pouvoir reconnaître les objets des différentes versions et lui fournir les modifications qui sont intervenues, ce qui n'est pas aisé avec ce mode de gestion.

- *Dater les objets.* Le second enrichissement proposé, outre l'archivage aux dates butoirs, consiste à dater les objets, c'est-à-dire à renseigner la date de création et la date de suppression de chaque objet (fig. 8). Cela permet très simplement de connaître les objets créés entre deux mises à jour, les objets détruits et les objets inchangés. Au fur et à mesure que les mises à jour se répètent, des requêtes spatio-temporelles plus complexes peuvent être élaborées. Le volume de la base sera en constante augmentation mais il semble limité : du fait de la relative stabilité du paysage, on évalue à 10 % au maximum le nombre d'objets changés entre deux mises à jour. Enfin, l'actualisation chez le client semble beaucoup plus aisée à envisager.

- *Doter les objets d'un comportement.* Enfin, le dernier apport pour la gestion des



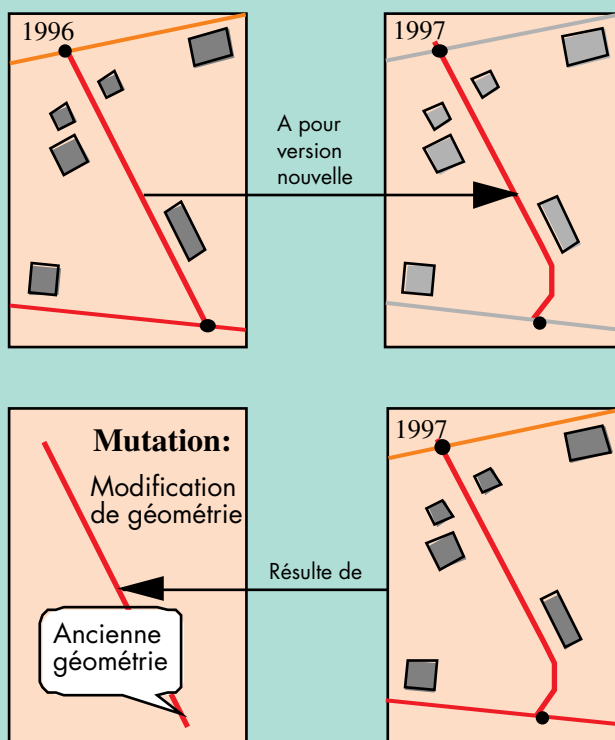
7. Archivage périodique des données



8. Datation des objets

mises à jour consiste à doter chaque objet d'un comportement; par exemple, on dira qu'une portion de route a été «en construction», puis «en service avec une voie de chaque côté», puis «élargie à deux fois deux voies», enfin devenue «une nationale». Deux approches peuvent être adoptées pour retenir un tel niveau d'information : une approche par état ou une approche par événement. L'approche par état consiste à garder chacun des états de l'objet et à les relier entre eux; on parle alors d'un modèle de version (fig. 9) : chaque objet a un numéro de version et ses versions sont reliées entre elles; un mécanisme de désignation logique permet d'éviter la duplication des attributs inchangés; on définit alors un identifiant unique pour chaque objet.

L'approche par événement privilégie l'aspect transformation et tente de décrire les changements qui ont eu lieu. Ce modèle de mutation (fig. 9) propose de stocker la dernière version de l'objet (la version à jour) ainsi que toutes les modifications apportées à l'objet. Ces modèles fournissent un apport sémantique important puisqu'on retient le cycle de vie de chaque



9. Version et mutation

objet : on est ainsi capable de présenter comment un lieu a évolué. Cependant, ces modèles imposent que l'on puisse définir de manière absolue ce qu'est un objet. En effet, quels critères doit-on adopter pour affirmer que la portion de route dont on a rectifié un virage est toujours la même, ou que la portion de route sur laquelle il y a eu un élargissement de voie n'est plus la

même portion de route ? Comment définir ces critères de modification ? On sent bien que cette décision est subjective et dépend de l'application et du SIG utilisés. Pour un producteur de données géographiques, il n'est pas possible de favoriser un point de vue applicatif plutôt qu'un autre. Aussi, ces modèles évolués semblent aujourd'hui réservés aux études spatio-temporelles sur l'évolution d'un territoire ou la propagation d'un phénomène, dans lesquelles des règles précises d'évolution peuvent être édictées.

Conclusion

Des procédures de gestion de mise à jour ont été implantées dans différents organismes producteurs de données géographiques (la BDCarto de l'IGN, le

Système d'information à la parcelle de l'Atelier parisien d'urbanisme, la base Voies de la Communauté urbaine de Lille). De plus, des systèmes intègrent un mécanisme interne de gestion de version. Ainsi, des éléments permettent d'assister l'utilisateur dans la gestion des mises à jour. Ils ne doivent pas masquer les efforts nécessaires à sa mise en place. – **Laurent Raynal**, COGIT

EN LIBRAIRIE

Le Pérou en cartes

Deux publications révèlent le bon cru 1997 de la coopération cartographique franco-péruvienne. Sur un plan général d'abord, le *Pérou en cartes* (1), sous forme papier et disquette PC Windows, repose essentiellement sur une exploitation fouillée du recensement national de 1993. Chaque carte est accompagnée de commentaires dont la conclusion générale aboutit à une reconstitution des diverses étapes de l'organisation de l'espace national. Auparavant, les cartes sont distribuées selon quatre chapitres principaux, population, éducation, activité économique et habitat. En dehors de l'actualisation des données qui corrobore bien des évolutions connues, le volume a le mérite de souligner la singularité du Sud du pays : la cartographie fait apparaître clairement une surprenante dynamique positive. L'*Atlas de la région du Cusco* (2) est évidemment beaucoup plus détaillé, plus illustré et va au-delà d'une cartographie, maintenant classique, purement statistique. Il utilise des

données de plusieurs origines, si possible diachroniques, ainsi que plusieurs échelles de représentation, situant la région Inka dans le Sud péruvien, dans le contexte national, voire dans l'ensemble limitrophe boliviano-brésilien. Malgré son intérêt touristique et ses richesses minières, la région du Cusco, séparée du pôle dynamique d'Arequipa, apparaît alors toujours située dans un angle mort et marginalisée. Concluons en soulignant que, dans les deux atlas, la conception et la présentation des cartes se rattachent très clairement aux productions estampillées RECLUS dont elles constituent de nouvelles retombées. –

Pierre Usselmann

(1) *Perú en mapas, Atlas en base al censo de población y vivienda*, 1997, 136 p., et disquette PC Windows, Lima : Instituto Nacional de Estadística e Informática et ORSTOM.

(2) Deler J.-P., dir., Hurtado I., Mesclier E., Puerta M., 1997, *Atlas de la región del Cusco*, Institut Français d'Études Andines, Centro de Estudios Regionales Andinos Bartolomé de Las Casas et ORSTOM, Cusco, 206 p.