

GÉNÉRALISATION CARTOGRAPHIQUE AUTOMATIQUE : APPROCHE ET MÉTHODES

Par J-F HANGOUËT, Laboratoire Cogit, Institut Géographique National Jean-Francois.Hangouet@ign.fr

Version française, corrigée et légèrement remaniée, de la communication intitulée *Automated Cartographic Generalization : Approach and Methods*, présentée par Jean-François Hangouët et Sylvie Lamy à la 19^{ème} conférence internationale de l'ACI.

Résumé

Traditionnellement, la généralisation cartographique est l'activité du cartographe qui dessine une carte dérivée à partir d'une carte de référence. Aujourd'hui, les cartes trouvent leurs équivalents numériques dans les bases de données géographiques, des outils existent qui permettent à un utilisateur de dériver interactivement d'autres bases, mais la pleine automatisation de la généralisation est loin d'être achevée. Irréalizable ou inutile pour certains, elle inspire néanmoins de nombreux travaux confiants. En cherchant à formuler la difficulté de la généralisation non comme une vérité vaine (telle : « la généralisation est un art, point »), ni comme un ensemble de problèmes contingents, mais en tant qu'imitation de la géographie, la première partie de notre communication se conclut sur la distanciation accrue des données numériques à la réalité géographique. Pour préserver la dénotation, le second chapitre préconise de mener les recherches (sauf à les multiplier) par le questionnement assidu des phénomènes géographiques qui justifient le sens latent contenu dans les données numériques. L'approche proposée, nommée *phénoménologie (proche mais distincte de la phénoménologie philosophique)*, impose notamment de reconstruire les relations contextuelles entre les objets séparés dans la base par des méthodes respectueuses de l'intégrité des données. De telles méthodes existent : la troisième partie décrit le diagramme de Voronoï sur les segments, construction mathématique qui préserve la géométrie des objets qu'elle structure (au contraire des partitions de Voronoï habituelles sur les points). Ainsi fondée méthodologiquement et pratiquement, la généralisation automatique est possible et déterminée. Les manières de faire courantes, d'ailleurs, ne sont ni balayées ni rejetées par nos nouvelles approche et méthodes : elles trouvent au contraire un cadre où s'objectiver et s'évaluer. La quatrième et dernière partie décrit l'application de nos principes à la programmation d'un opérateur de typification de données qui représentent des habitations en paysage péri-urbain.

1. DE LA DIFFICULTÉ DE LA GÉNÉRALISATION CARTOGRAPHIQUE AUTOMATIQUE

La difficulté de l'automatisation de la généralisation est souvent exprimée en termes informatiques : « la généralisation automatique est (...) un problème « NP-complet » (c'est-à-dire qu'il n'y pas une solution algorithmique) » (Müller et al. 1995 : p.4) ; ou « la complexité des règles et des connaissances impliquées dans la rédaction cartographique (et la généralisation) est un vrai casse-tête... » (Morehouse 1995 : p.24), ou encore : « il est très difficile de concevoir que la théorie de l'information puisse jamais fournir un modèle pour tous les aspects de la généralisation » (Bjørke 1993 : p.51). Un autre argument consiste à mettre en contraste le systématisme de l'ordinateur et l'art du cartographe : « avec la généralisation, l'art entre dans la réalisation de la carte » [une citation précède de Max Eckert donnée dans (McMaster 1989 : p.1)], et l'art ne se laisse pas modéliser. Le but de cette première partie de l'article est de reformuler la difficulté de l'automatisation de la généralisation, en tentant de s'affranchir des méthodologies inspirées par le projet (l'automatisation) ou par les moyens (l'ordinateur), et en se recentrant sur l'objet en jeu, à savoir la généralisation par elle-même.

1.1 La généralisation cartographique

La généralisation cartographique peut être définie par l'ensemble des activités impliquées : *simplification, clas-*

sification, induction, symbolisation (Robinson et Sale 1969), ou *sélection, schématisation, harmonisation* (Cuenin 1972 : pp.168 sqq.). Elle peut être définie aussi par un ensemble de nécessités : le « *Fallen unter dem Maßstab* » (le « tomber sous l'échelle ») : les outils du cartographe ne lui permettent pas de tout dessiner), la *lisibilité* (les signes doivent être visibles sur la carte), l'*expressivité* (les objets doivent être rendus pour ce qu'ils sont) (Eckert 1921 : pp.330-1). Elle peut encore être définie par un ensemble de motivations : motivations *économiques*, motivations de *qualité*, motivations de *multi-fonctionnalité*, motivations de *communication* (Müller 1991 : pp.458-9). Une définition scrupuleuse, visant à prendre en compte tous les cas explicitement, devrait se transformer en manuel.

En fait il peut suffire de dire que la généralisation cartographique est le dessin d'une carte à partir d'une autre carte ou de données terrain.

Cette définition relativement immédiate englobe évidemment ce qu'évoquent les activités mentionnées ci-dessus, à l'exception de l'« *induction* » – dont on doit admettre cependant qu'il s'agit d'une dénomination assez peu précise. Renvoyant globalement à « *l'application en cartographie des processus d'inférence de la logique* » (Brassel 1990 : p.7), même aux commentateurs les plus avisés elle apparaît comme « *un élément plus abstrait* » (McMaster et Shea 1992 : p.19), ou elle tend à échapper à toute description : « *l'induction ne peut être entièrement appréhendée de manière raisonnée* » (Morrison 1974 : p.119). Ainsi, l'induction, en tant que processus créatif d'établissement de lois

à partir de faits, ne s'applique que de manière détournée à ce qu'évoque intuitivement l'expression «dessiner une carte à partir d'une carte». En fait, les efforts pour rendre l'«induction» applicable à la cartographie sont peut-être des constructions sur une confusion démarrée par le jeu des mots, qui a dû suggérer, promptement, d'importer le terme «induction» en cartographie : la duplicité du terme «généralisation» lui-même. Ses deux sens, ce qu'il recouvre en cartographie et ce qu'il recouvre en logique sont deux activités tout autant différentes l'une de l'autre que le sont la copie et la création. Si le pouvoir évocateur intrinsèque de l'expression «induction» ne nous éclaire guère sur la généralisation cartographique, son inadéquation même nous signale que la généralisation n'est pas une activité simple. La difficulté de la généralisation, nous allons le voir à présent, s'explique cependant de manière naturelle et confiante quand on la considère comme une *imitation*.

1.2 Généralisation cartographique et Imitation

Dessiner une carte à partir d'une carte est faire une imitation... Un tel constat, où *le dessin d'un dessin à partir d'un dessin* est décrit comme une *imitation*, n'apporte à première vue guère d'information - c'est presque une lapalissade. Et pourtant en rangeant la généralisation parmi les processus de copie il insiste sur sa qualité «déterministe» : le produit ne peut s'éloigner trop de la matière première. On pourrait rétorquer que le produit s'éloigne quand même parfois beaucoup de la matière première - mais l'imitation est affaire de copie, pas de copie conforme. Nous allons clarifier ce point en explicitant ce qui est imité effectivement, mais auparavant, un commentaire est nécessaire pour dissiper les appréhensions qui peuvent survenir à voir un terme chargé de connotation dépréciative employé à l'égard d'une activité aussi prestigieuse que la cartographie.

Imitation

«Imitation» n'est vraiment pas un terme «positif». Dans le sillage de Platon qui disait de l'art de l'imitation qu'il est «*assurément loin du vrai*» (Platon : 598a), qu'«*il n'est le compagnon ni l'ami de rien de sain ni de vrai*» (Platon : 603b), on a considéré l'*imitateur* (non pas l'artiste qui peut faire «dans le style de») avec une défiance tacite ou ouverte dans la vie courante, dans les arts, en politique, en science... Approfondir les caractéristiques intrinsèques de l'«imitation», au lieu de nous contenter de l'expression, va nous permettre de révéler les conditions de l'imitation réussie et par là même de réévaluer la tâche de l'imitateur. La liste ci-dessous résume une discussion menée dans (Hangouët 1998 : pp.56 sqq.). Pour chaque caractéristique, nous donnons une transcription symbolique (Ê vaut pour la chose qui est imitée, Ê pour la chose qui imite) ; nous donnons aussi un court commentaire, et, en manière de fable, une citation de la littérature.

Ê-Ê et Ê-Ê : la ressemblance n'est pas symétrique.

Ce qui imite s'efforce de s'approcher de ce qui est imité, alors que ce qui est imité reste loin de ce qui imite.

Dans leur débat sur leurs mérites respectifs, le Renard explique soudain au Singe :

«*Là même où tu te dis, dans ton apparence physique, Être tout comme un homme, l'homme, lui, n'est pas comme un singe*»

(Spenser 1591 : vers 1041-2)

Ê et Ê : chacun a son sens.

Ce qui est imité a du sens par lui-même ; ce qui imite a du sens et par lui-même et par référence à ce qui est imité. «(La photographie - Ê - de l'œuvre d'art - Ê) *substitue l'œuvre significative au chef-d'œuvre et le plaisir de connaître à celui d'admirer.*» (Malraux 1951 : p.15)

Ê pour Ê : la généralité de ce qui est imité explique ce qui imite.

Ce qui est imité peut être pris pour quelque chose qui imite sans qu'une faute irréparable soit commise : très vite, l'intuition du référent plonge l'intercesseur supposé dans l'englobant qu'il est lui-même.

«*La science peut nous dire comment l'océan est né. Elle peut combiner le sel et l'eau (Ê est l'océan) mais nulle quantité d'eau salée ne fera jamais un océan sans cette étrangeté qui s'instaure lorsque l'homme y porte le regard (non, l'océan est Ê).*» (Gary 1962 : p.4)

Ê pour Ê : la généralité de ce qui est imité est limitée par ce qui imite.

Ce qui imite ne peut être pris pour ce qui est imité sans déception.

«*Le jeune chat joue avec des boules de papier parce qu'il les croit vivantes, semblables à lui, mais, un peu plus vieux, il comprend et les laisse de côté.*» (Schopenhauer 1844 : p.50)

Ê ou Ê? la comparaison effective résout la confusion éventuelle.

En cas de doute, seule la comparaison effective permet de distinguer ce qui imite et ce qui est imité.

«*Pour reconnaître si le tableau est vrai ou faux, nous devons le comparer à la réalité.*» (Wittgenstein 1921 : 2.223)

Il devient clair que la tâche de l'imitateur n'est pas simplement de fabriquer un produit qui ressemble à l'original, mais lors de cette fabrication même, d'être constamment attentif à la distance asymétrique qui s'instaure entre ce qu'il copie et ce qu'il fabrique (pour servir ou pour tromper : le contrefacteur s'évertuera à estomper tous les indices d'authentification, alors que l'imitateur à découvert cherchera à stimuler les utilisateurs à discerner le sens original à travers l'évidence qu'il propose). Notre imitateur n'imité pas par hasard et, contrairement à l'imitateur de Platon [«*le spécialiste de l'imitation ne connaît rien qui vaille aux choses qu'il imite*» (Platon : 602b)], il a une compréhension fine pour ce qu'il imite, et un profond respect : l'imitateur doit être admiré, en plus de ses talents évidents, pour sa rectitude éthique.

Cartographie

La carte généralisée vérifie les caractéristiques de l'imitation. En effet :

Ê-Ê et Ê-Ê. Entre la carte et sa généralisation, la ressemblance n'est pas symétrique. Qu'on la mesure de manière perceptuelle par le temps d'identification sur une carte d'un détail donné sur l'autre (Hangouët 1998 : annexe 3), ou de manière mathématique par les composantes de la distance de Hausdorff sur des détails homologues (Abbas-Hottier 1992 : pp.32 sqq.), la mesure de l'écart de la carte généralisée vers la référence est statistiquement plus faible que la mesure de l'écart de la référence vers la carte généralisée (cf. Figure 1). Ce que (Abbas-Hottier 1992 : pp.32sqq.) formulent ainsi : «*la carte est proche de la réalité (...) la réalité est loin de la carte.*»

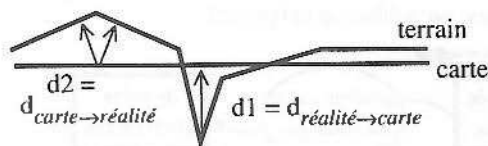


Figure 1 : Dissymétrie des distances mutuelle *carte* ↔ *réalité* sur un même détail

Sur des tests menés en zones urbaines, avec m moyenne de $d1$ et $d2$ (composantes de la distance de Hausdorff entre les deux lignes), l'indice $i = (d1 - d2) / m$ est positif et vaut 0.120 sur Mulhouse et 0.060 sur Montpellier [détails opératoires dans (Abbas-Hottier 1992 : pp.32sqq.)]

Ê et Ê. Chacune des cartes est adaptée aux besoins qui ont motivé ses spécifications.

Ê pour Ê. La carte de référence peut être utilisée pour ce pour quoi est utilisée la carte dérivée (moins rapidement peut-être, mais c'est un autre problème).

Ê pour Ê. La carte généralisée ne peut être utilisée pour toutes les applications de la carte de référence (on peut gâcher une petite promenade en forêt le dimanche après-midi en cherchant son chemin sur une carte à 1:100 000).

Ê ou Ê? Comparer les cartes entre elles aide à saisir le sens de chacune. " *La généralisation aiguise le voir*" (Eckert 1921 : p.339).

Quant au cartographe lui-même, il est bien conscient du statut représentatif de sa production, comme le prouve l'usage solennel du terme « fixation » (« Fixierung ») dans cette citation de Max Eckert : " *la fixation cartographique de tout tracé doit être le résultat d'une sérieuse réflexion préalable*" (Eckert 1921 : p.331). L'accomplissement de la « fixation » se décide aussi à partir de l'implication éthique du cartographe, comme le rappelle le titre provocant du livre de Mark Monmonier : *Comment faire mentir les cartes* (qui, loin d'être un manuel de malhonnêteté à l'usage des cartographes, rappelle constamment qu'il est facile, par manque d'application, de trop éloigner la carte de ce qu'elle représente). Engagement éthique qu'atteste aussi,

dans cette citation de Jean Barbier, l'adverbe inattendu « honnêtement » : " *Il appert donc que le travail du cartographe (...) est de comprendre complètement le sujet, d'en maîtriser tous les aspects, d'en montrer tous les mécanismes et de reconstruire pour le lecteur une image graphique qui illustre honnêtement tout ce qu'il a pu comprendre de la réalité jusque dans les détails que lui permettent l'échelle et le niveau de la carte*" (Barbier 1965 : p.168).

1.3 La généralisation cartographique comme imitation de la réalité géographique

Cette dernière citation, incidemment, révèle ce qui est imité effectivement avec la généralisation : non pas vraiment la carte de référence, mais la *réalité géographique*, c'est-à-dire ce qui, dans le *monde réel* (« réalité »), vient ou peut *graver la Terre* [« geo-graphique »]. Précisons néanmoins d'un point de vue formellement linguistique que l'analogie entre « graver » et « graphique » est peut-être fortuite. Selon le Robert, Dictionnaire historique de la langue française, entrée « graphique » : le verbe grec « graphèin » est " *apparenté à des termes qui, en vieux slave et haut allemand, signifient couper, entailler, et est rattaché à une racine indo-européenne "gerbh- de même sens"*. Mais selon l'Oxford English Dictionary, entrée « grave » : entre le haut allemand « grafan » et le grec « graphein », " *la connexion n'est plus acceptée par les philologues*". Réalité, carte et carte généralisée s'organisent ainsi selon le schéma d'imitation donné en figure 2.

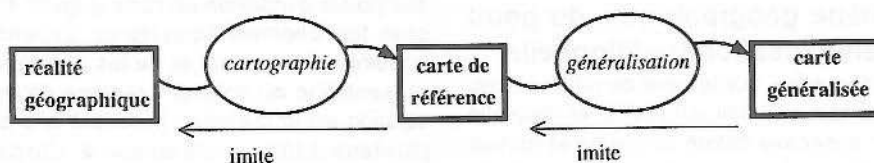


Figure 2 : Généralisation en tant qu'imitation

1.4 Difficulté de la généralisation cartographique

Il apparaît désormais de manière évidente que la difficulté de la généralisation cartographique réside dans la soumission éthique à la réalité géographique. La complaisance et la créativité du cartographe doivent être mises au service de la recherche de la réalité à travers la carte donnée ou les données terrain, et de sa restitution dans la carte à produire.

1.5 Difficulté de la généralisation cartographique automatique

Après plusieurs décennies de recherche, la faisabilité de la généralisation automatique est toujours sujette à caution. Existera-t-il des systèmes automatisés capables de faire ce que font les cartographes ? Ou les systèmes imposeront-ils leurs propres standards et règles du jeu,

comme le suggèrent (Müller 1991) ou (McMaster-Shea 1992) ? Toujours est-il que c'est l'observation du travail des cartographes qui a guidé jusqu'à présent l'élaboration des algorithmes et des méthodes. Les indices sont dans les termes et dans les intentions :

- les algorithmes se répartissent en classes d'opérateurs déduites de l'observation attentive des cartes traditionnelles : *sélection, simplification, caricature, agrégation, typification* etc. (Müller *et al.* 1995) ;
- les opérateurs élémentaires sont orchestrés par un processus de tâtonnement itératif inspiré par les essais/petites corrections du cartographe (Ruas-Plazanet 1996), (Ruas 1999) ;
- et tant la transformation que la décision reposent nécessairement sur l'estimation de la situation qui se présente (Müller *et al.* 1995) (analyse idéalement automatique, pour l'heure encore réalisée de manière interactive).

Si «généralisation» trouve jamais une nouvelle acception dans un monde où les SIG remplacent les cartes, il restera vrai cependant qu'historiquement, c'est bien l'idée que la machine peut être conçue pour faire ce que l'homme

sait faire qui aura motivé l'automatisation de la généralisation. Autrement dit, l'automatisation de la généralisation est menée comme l'imitation des pratiques traditionnelles, ce qu'illustre la figure 3.

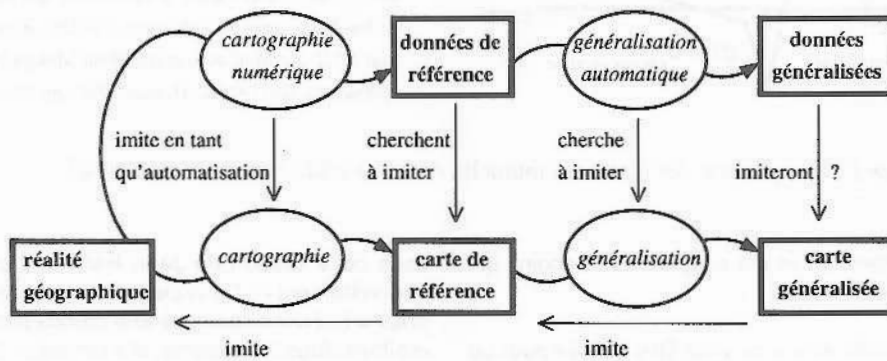


Figure 3 : Généralisation automatique en tant qu'imitation

Sur ce schéma se lit la difficulté de la généralisation automatique : les données généralisées sont l'imitation de l'imitation de la réalité. Distanciation qui rend incertaine la visibilité de la réalité. Comment s'assurer cependant que la visibilité soit possible et valable ?

2. LES PHÉNOMÈNES GÉOGRAPHIQUES

Dans les paragraphes qui suivent, nous décrivons ce qui emplit de sens géographique les signes que voit le cartographe comme les *phénomènes géographiques*. La notion est étendue aux données géographiques numériques, et nous suggérons que le manque évident d'implication éthique chez l'ordinateur, qui le rend aveugle aux données qu'il est amené à manipuler et traiter et qui éloignent celles-ci de leur référent géographique, peut être compensé par une attitude rigoureuse de la part du *chercheur*, qui consiste à penser en termes de *phénomènes* au lieu de *données*. Cette attitude ou méthode est décrite sous le nom de *phénoménologie*.

2.1 Le phénomène géographique, du point de vue de la généralisation traditionnelle

Ce que voit le cartographe, sur la carte de référence ou dans les données terrain qu'il doit retravailler en nouvelle carte, n'est pas un ensemble coloré de lignes et formes ou points nommés. Ce n'est pas non plus un ensemble de signes, mais ce vers quoi ces signes pointent. Nous appellerons ce qui avec le signe rattache sa perception et son usage au sens géographique souhaité, le *phénomène géographique* du signe. En fait le signe, en tant que dessin sur une feuille de papier qui hérite de la légende, est

contingent, alors que le phénomène, enraciné dans le monde réel et dont le signe n'est qu'un écho, est inconditionnel : si la notion de «phénomène géographique» peut être introduite à partir de la notion familière de signe, elle mérite néanmoins d'être définie pour elle-même, plus indépendamment.

Le «phénomène géographique» est le couple constitué par :

1. l'objet du monde réel, homologue du signe
2. l'immersion de cet homologue dans une explicabilité géo-graphique. «Géo-graphique» avec un tiret : pour insister sur le fait que l'adjectif doit être pris dans le sens plein de «qui grave la surface de la Terre».

Le fine tige végétale incurvée dans l'assemblage d'une carte de navigation des îles Marshall est un signe. L'objet du monde réel qu'il dénote est la houle dominante. La houle est la moins perturbée à quelque distance des îles, desquelles elle semble rayonner (non ?).

Prenons ces couples de lignes courbes qui se croisent aux points d'inflexion de l'une (Figure 4) : ce sont des signes formellement équivalents. Cependant, au jeu d'introduire un englobant et de les distinguer sur un axe *représentation du concret / représentation de l'abstrait*, la solution est immédiate : le couple *b/* a quelque chose de plus terre-à-terre que le couple *a/*. L'explication vient sans doute du référent voulu pour chaque signe : "*l'invisible de la subjectivité absolue*" (Henry 1988 : p.25), "*ce que Kandinsky appelle l'Intérieur*" (Henry 1988 : p.21) pour le couple *a/*, cas de lignes dessinés par Kandinsky ; route dans la région de Valence pour le couple *b/*, cas de généralisation du linéaire routier par Corinne Plazanet.

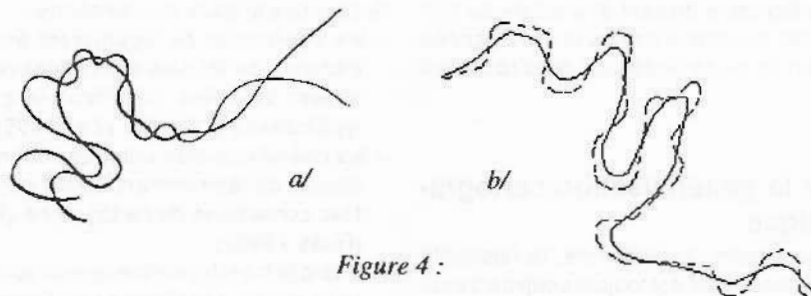


Figure 4 :

Des lignes courbes s'accompagnent
(Kandinsky 1926 : fig.63)

et Lignes cartographiques
et (Plazanet 1996 : p.41)

Le cartographe dans son activité ne mobilise pas seulement son art du dessin, son acceptation des spécifications de la carte à faire, et le donné dans la carte de référence - mais aussi toute sa *connaissance* et toute son *expérience* du monde réel.

2.2 Le phénomène géographique, du point de vue de la généralisation automatique

L'ordinateur n'a ni connaissance, " *ni à propos, ni expérience* " (Rimbert 1990 : p.76), ni non plus de fait cet engagement éthique qui permet de s'emparer pleinement du monde. L'ordinateur a pour fonction de ruminer des données, le sens des processus et le sens des données elles-mêmes restant de la compétence des concepteurs... Mais alors même qu'en chercheurs nous pensons naturellement en termes de phénomènes (la «route», la «maison», la «crête»...), nous pouvons parfois mésestimer l'étendue du compromis imposé par la simplicité d'esprit (s'il en est) du système auquel est confiée la géographie et être ainsi conduits à nous contenter d'approximations inutiles, difficiles voire même erronées. La situation en géographie par ordinateur peut être comparée à celle de l'arithmétique exacte, à ceci près cependant que ce domaine de l'informatisation des mathématiques est très étudié - alors qu'on considère toujours l'ordinateur comme le vecteur de la géographie renouvelée et qu'on oublie qu'il la limite aussi. Voilà pourquoi ce n'est pas un vain effort que de rappeler la teneur de la représentation des données numériques, c'est-à-dire d'insister encore sur les phénomènes géographiques qui leur procurent sens et fondement.

Le «phénomène géographique» est le couple constitué par :

1. l'objet du monde réel, homologue des données considérées de la base,
2. l'immersion de cet homologue dans une explicabilité géo- graphique.

Cette définition rassemble l'importance des données, du monde réel, et de la carto-graphie. Nul sacrifice ne peut être fait, à ce niveau, pour de quelconques raisons d'efficacité calculatoire. Au chercheur qui se contenterait rapidement de méthodes purement algorithmiques, cette définition rappelle la nécessaire loyauté au monde réel et la foi en ses propres connaissances et expériences. Au chercheur qui, voyant la profusion des cas du monde, désespérerait de la faisabilité de l'automatisation, elle rappelle l'imparabilité des données et la poignée qu'elles offrent sur le monde.

Notre définition de l'expression «phénomène géographique», que les chercheurs emploient fréquemment dans leurs communications pour justifier leurs travaux, mais sans jamais chercher à la disséquer ou à la justifier elle-même, objectifie la motivation de leurs recherches. Cependant, en l'état, elle laisse un goût d'inachevé. Et alors : Dois-je l'afficher dans mon bureau ? La faire clignoter à l'écran ? Ou pour les moins cyniques : Comment rester fidèle à mon allégeance à la géographie ?

2.3 Phénoménologie

Nous avons décidé d'appeler *phénoménologie* l'approche de l'automatisation de la généralisation qui se fonde sur l'acceptation de la définition qui vient d'être donnée du «phénomène géographique». La phénoménologie peut être considérée soit comme une méthode explicite (quand on la ressent comme une prescription) soit plutôt comme une attitude (quand on la reconnaît comme une description). Démarrant au moment même où le chercheur s'ap-

prête à aborder sa recherche, voire auparavant, elle inclut les cinq mouvements qui suivent :

1. L'éveil à la manifestation continue des phénomènes du monde réel. Tout signe du monde qui évoque le phénomène géographique général sous-jacent au cas de l'objet réel, qu'il soit incertain ou incontournable, est à retenir - il n'est pas encore nécessaire de réfléchir à partir de lui : l'essentiel est de ne pas le perdre.

2. Le questionnement volontaire et opiniâtre du phénomène pour mieux le voir et circonscrire son explicabilité. Tous les moyens sont bons ; plus on les multiplie, mieux c'est : la pratique «naïve» ou renouvelée du paysage où l'on vit, les photographies, les romans, les cartes, les manuels, les monographies scientifiques, les discussions avec les experts ou des non-experts, les peintures, les résultats d'autres recherches de généralisation sur des données équivalentes etc.

3. La cristallisation franche de l'essence du phénomène, c'est-à-dire, par les mots, le choix ou la construction d'une formulation expressive du phénomène considéré, qui soit à la fois descriptive (elle contient toutes les images accumulées ainsi que l'explicabilité du phénomène), indiscutable (nul cas réel du même phénomène ne lui est incompatible sans raison inhérente) et inconditionnelle (elle ne fait pas référence à la base de données).

4. La mise en conformité des données de la base avec le phénomène du monde réel. D'un point de vue informatique, il s'agit d'une phase d'analyse (spatiale) pour rassembler ou pour éclater les données numériques selon les principes de l'explicabilité du phénomène. La plupart du temps, cela consiste à rendre leur pleine surface aux objets qui sont stockés par leur contour ou par leur axe, leur segmentation naturelle aux objets qui sont stockés d'un bloc, leur disposition relative aux objets qui sont stockés en listes au lieu d'images surfaciques ou volumiques. La restauration doit être cohérente avec ce que permet l'explicabilité du phénomène. Lorsque les outils de mesure employés ou conçus spécifiquement ratent quelque chose de l'explicabilité du phénomène (et ils manquent toujours quelque chose), cela doit être documenté afin a/ que les utilisateurs ne soient pas «trompés» par ce qui pourrait paraître comme un phénomène prétexte, b/ que les outils soient retravaillables ou prolongeables par les recherches subséquentes.

5. La conception d'outils de transformation qui se contentent de prendre les contraintes cartographiques en paramètres et qui sont véritablement fondés sur les lignes de force particulières à l'explicabilité géographique du phénomène.

Les deux premiers mouvements visent à sonder le monde par une contemplation intensive, en mobilisant une attention plus vaste que ce que la science cartographique peut généralement suggérer ; les deux derniers mouvements visent à rendre les données manipulables de manière plus intuitive. Le mouvement central, la *réduction* du phénomène à sa plus simple et à sa plus puissante expression, mais aussi la *raison* des processus automatisés, est l'articulation nécessaire et cruciale. Les cinq mouvements ne se succèdent pas les uns aux autres, au contraire, ils s'accumulent, chaque mouvement ajoutant, dans les mouvements qui le suivent, à son influence reconnaissable son déploiement continu. Les étapes ultimes vibrent ainsi de la tension de possibles remises en question - mais dans notre

contexte phénoménologique, celles-ci ne sont plus liées à l'implémentation fragile ou hâtive ni à la distraction algorithmique, mais à la seule fragilité naturelle de l'analyse spatiale, à savoir le risque de manquer quelque chose de ce que le monde sans cesse surgissant peut nous dire.

L'appellation «phénoménologie», qui dans notre contexte peut se lire comme «le discours sur les phénomènes géographiques en général et en particulier», est une élaboration naturelle à partir de l'expression «phénomène géographique». Dans notre décision d'adopter effectivement «phénoménologie», nous avons été confrontés soudain à la prégnance de l'acception du même terme en philosophie. Notre phénoménologie, de fait, participe de la phénoménologie de Husserl, dans la mesure où ce qui dans notre méthode pourrait être rejeté par les scientifiques cartésiens comme subjectivité déplacée (à savoir l'insistance sur l'implication du chercheur dans les phénomènes, au lieu de présumer l'abstraction de l'observation des objets - qualité pour laquelle, soit dit en passant, l'ordinateur fournit le parangon) est considéré par Husserl et les phénoménologues modernes comme la condition même de la connaissance et de la compréhension : " (Husserl reproche à Descartes) *d'avoir soustrait au doute méthodique les vérités de la géométrie* " (Desanti 1963 : p.25).

3. LES DIAGRAMMES DE VORONOÏ SUR LES SEGMENTS

Notre approche générale, conceptuelle, de l'automatisation de la généralisation ayant été exposée, il reste à montrer ou illustrer comment elle peut être conduite en pratique. Cette partie décrit une structure générale, mathématique, qui permet de reconstruire et de mesurer les relations de proximité entre les objets, restituant ainsi les listes numériques de données creuses en tout cartographique. La quatrième partie illustrera l'utilité de cette structure pour l'élaboration phénoménologique de l'opérateur de généralisation appelé *typification*.

3.1 Notions de base

Le diagramme de Voronoï d'objets séparés plongés dans un espace métrique (et appelés *sites*) est le lieu géométrique des points qui sont minimalement équidistants aux sites. Le diagramme réalise une régionalisation de l'espace en autant de cellules qu'il y a de sites ; il y a une cellule autour de chaque site ; tout point à l'intérieur d'une cellule est plus proche du site de cette cellule que de tout autre site. Dans nos applications, les sites sont les segments et les points qui composent la géométrie des objets tels qu'ils sont stockés dans nos bases de données vectorielles ; l'espace est le plan euclidien. Le diagramme de Voronoï est alors composé de segments de droites (équidistance entre deux points, équidistance entre deux segments) et de segments de paraboles (équidistance entre un point et un segment).

3.2 Les diagrammes de Voronoï en géomatique

Les diagrammes de Voronoï sur des sites *ponctuels* sont bien connus, et leur utilisation est très répandue pour l'identification des relations de proximité entre objets [soit tels quels, soit sous la forme duale de la triangulation de Delaunay ; voir par exemple (Peng *et al.* 1995), (Ware et Jones 1998), (Ruas 1999)]. En revanche, relativement peu de communications traitent des diagrammes de Voronoï sur des sites *segments*. La plupart d'entre elles sont du domaine de la géométrie algorithmique et s'intéressent au

calcul du diagramme [un problème de complexité $O(n \cdot \log n)$ que rendent délicat les erreurs d'arrondi du calcul par ordinateur, voir par exemple (Boissonnat et Yvinec 1995) ou le riche site Internet <http://www.ics.uci.edu/~eppstein/gina/voronoi.html>]. Ceci étant, certaines des propriétés des diagrammes de Voronoï sur les segments sont expliquées dans (Okabe *et al.* 1992) ; et des applications puissantes de reconstruction topologique dans les SIG sont décrites et démontrées dans une série de travaux de Christopher Gold (rassemblés sur le site Internet de l'auteur, entièrement consacré aux diagrammes de Voronoï : <http://www.voronoi.com/>). Plusieurs applications mathématiques utiles pour des opérations routinières d'exploitation de SIG ou de généralisation (incluant l'enveloppe convexe, l'arbre recouvrant minimal, le calcul de la distance de Hausdorff, de squelettes, de dilatation et d'érosion en mode vecteur) ont été exposées dans (Hangouët et Djadri 1997). Le lecteur est invité à consulter ces références pour faire plus ample connaissance avec les diagrammes de Voronoï sur les segments. Dans cette communication, nous aborderons ce qui rend cette structure si magique pour les données géographiques numériques d'un point de vue uniquement conceptuel, prosaïque.

3.3 Sur l'efficacité du diagramme de Voronoï sur les segments

Le diagramme de Voronoï sur les segments ne «rate rien» des objets tels qu'ils sont stockés dans les bases de données géographiques vectorielles, étant calculé sur les segments et les points qui sont les composants atomiques de leur géométrie. Comme il est expliqué dans (Hangouët 1998), les triangulations de Delaunay et les diagrammes de Voronoï habituellement utilisés sur des points sont au fond des approximations des diagrammes de Voronoï sur les segments. Dans ces applications, certains points sont choisis pour représenter les données, soit stratégiquement, soit empiriquement. Il peut ainsi résulter des triangulations ou des régionalisations égales pour des configurations géographiques différentes, si jamais les points représentatifs sont choisis aux mêmes endroits. À côté de cette pré-généralisation (pour laquelle il n'est certes pas impossible, mais néanmoins difficile, de justifier les principes géographiques qui l'autorisent : pourquoi choisir ici par exemple les centres des bâtiments et les sommets intermédiaires des rues, ou là les coins des bâtiments et leurs projections sur les rues ?), l'emploi des structures sur les points dépend étroitement des procédures particulières qu'on automatise, et nécessite des interprétations complexes, comme le montrent les descriptions de (Ware et Jones 1998).

Le diagramme de Voronoï sur les segments, en revanche, est équivalent aux données géographiques impliquées : il n'existe qu'un diagramme pour une figure donnée, et qu'une figure pour un diagramme valide donné. En fait, le calcul du diagramme n'étant pas d'une complexité linéaire, mais supérieure, le diagramme de Voronoï ajoute quelque chose à la géométrie disponible : la cimentation même des objets dans un tout spatial et la possibilité de mesurer exactement les espaces vides entre eux. Quand les arêtes du diagramme de Voronoï sont «sémanticisées» en tant qu'arêtes à l'intérieur d'un même contour, arêtes sortant d'un même contour, arêtes séparant les mêmes deux objets (cf. Figure 5), des calculs de forme peuvent être effectués, ainsi que des calculs de distance et des calculs de proximité, exactement et intégralement : les données numériques, au départ stockées minimalement en listes d'axes ou de contours séparés, recouvrent enfin leur statut cartographique.

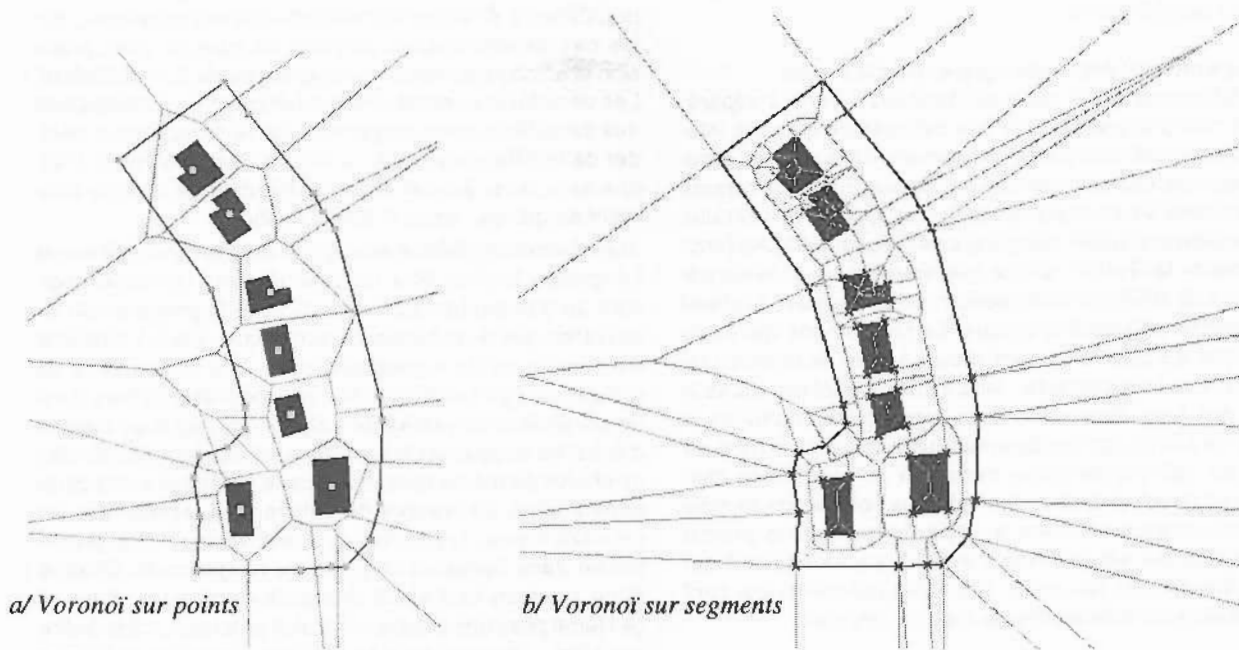


Figure 5 : Le diagramme de Voronoï sur segments est constitué de morceaux de droites et de paraboles qui limitent les aires d'influence (lieux des points du plan les plus proches) des points et des segments qui constituent la figure. Les arêtes à l'intérieur d'un contour en constituent le "squelette" ; les arêtes qui séparent deux objets différents les révèlent voisins et simplifient la mesure de leur proximité. Tout objet géographique participe intégralement à la construction du diagramme (ici, des bâtiments et des rues).

À gauche, le diagramme de Voronoï sur des points sur la même zone (centres de gravité des bâtiments et sommets intermédiaires des arcs routiers).

Le diagramme de Voronoï sur les segments, en prenant en compte les pleines géométrie et sémantique des données stockées dans la base de données géographique pour les porter dans leur disposition cartographique, est l'outil adéquat de l'approche phénoménologique, dont un des principes est de ne sacrifier en rien les données... Comment utiliser cet outil pour des buts de généralisation, sans non plus sacrifier la géographie, est illustré dans la partie suivante avec la description d'un opérateur de *typification*.

4. UNE APPLICATION : LA TYPIFICATION

La typification est l'un des opérateurs élémentaires de généralisation identifiés traditionnellement dans la littérature sur la généralisation automatique. Il s'agit de transformer un ensemble d'objets de même nature sans défaire l'apparence du groupe. Cette partie décrit comment notre opérateur a été élaboré pour des groupes de bâtiments «alignés», en insistant sur ses fondement; et rectitude phénoménologiques.

4.1 Le compromis cartographique

Avec la réduction d'échelle, les petits objets, s'ils ne sont pas relativement grossis, ne sont plus lisibles. Le grossissement relatif n'a pas d'effet induit lorsque le petit objet est isolé. Quand la densité en petits objets est très élevée, le grossissement de chacun conduit à un noircissement de la zone, ce qui est admissible voire sain visuellement

(nous parlons ici d'objets de même nature, entre lesquels ne surviennent pas d'autres objets) et qui, d'un point de vue algorithmique, se traduit facilement en objet agrégé pour de futurs traitements informatiques. Dans les densités intermédiaires cependant, le grossissement relatif individuel des objets conduit à un effet de sur-occupation : en fait, les espaces libres devraient eux aussi être élargis pour être lisibles - mais bien sûr, le grossissement simultané des objets et des espaces est physiquement impossible. La solution traditionnelle consiste à faire de la place en se débarrassant de quelques-uns des objets et en réorganisant les objets restant dans les limites de l'emprise originale. Cette manière de faire porte le nom de *typification*. Lacs harmonieusement distribués, virages qui rampent opiniâtrement l'un derrière l'autre sur le flanc pentu de la montagne, rues qui quadrillent la ville, maisons d'un quartier résidentiel : autant de candidats à la typification.

4.2 Typification, une définition

Rassemblant les diverses descriptions que donne la littérature [cf. (Hangouët 1998 : pp.226-7) pour une liste de références], la *typification* peut être définie comme l'opération qui *consiste à porter dans une représentation, qui en préserve la distribution reconnue, des objets de même nature qui se trouvent groupés localement par quelque processus géographique identifiable*. La typification est parfois, dans certains cas simples, sélection judicieuse (par exemple : «un objet sur deux»), le plus souvent cependant s'y ajoute ou plutôt s'y mêle le déplacement des objets retenus (leur «réorganisation»).

4.3 La typification de bâtiments

Notre opérateur de typification est conçu pour des alignements de bâtiments.

Groupement phénoménologique des bâtiments

Les bâtiments de la base de données viennent séparément : les groupes doivent être indiqués de manière interactive, ou calculés automatiquement. Alors que les groupes sont parfois calculés selon d'autres principes, comme notamment les principes gestaltiques qui visent à simuler la perception visuelle des groupements géométriques [«notamment» la Gestalt : parce que les résultats obtenus par (Regnauld 1998) sont très performants], nous avons choisi l'approche phénoménologique. Le phénomène qui explique que les bâtiments sont groupés dans le monde réel est celui de l'accessibilité : les bâtiments sont construits le long des rues. Pour cela, notre programme segmente la zone couverte par les données de la base en autant de cellules qu'il y a de cycles minimaux de rues. Dans chaque cellule (c'est-à-dire, dans chaque îlot), le programme associe chaque bâtiment au côté de rue le plus proche (opération qui est immédiate avec le Voronoï calculé sur les bâtiments et les rues). Enfin, les bâtiments qui sont associés à un même côté de rue sont groupés.

Le groupe qui vient d'être obtenu n'est pas d'une grande utilité cependant pour le traitement automatique que nous visons. Il faut aussi retrouver l'ordre de succession des bâtiments le long de leur rue. Ceci est à nouveau une opération simple avec le diagramme de Voronoï entre les bâtiments : il suffit de chaîner les bâtiments de proche en proche sur le critère «séparés par une arête du diagramme». Notre opérateur de typification ne fonctionne que sur des groupes «alignés» (c'est-à-dire en ligne droite ou courbe) mais pourtant tous les groupes de bâtiments ne sont pas linéaires... Ceci se lit à nouveau clairement et rapidement sur le diagramme de Voronoï : dès qu'un bâtiment a plus de 2 voisins dans le groupe, le groupe est déclaré «non typifiable par cette méthode».

L'«alignement» qui vient d'être obtenu n'est pas d'une grande utilité cependant pour le traitement automatique que nous visons. Comment décider des bâtiments à sélectionner pour la représentation à l'échelle finale ? Un grand bâtiment est un élément marquant du paysage, ou un bâtiment administratif, ou un bâtiment de forme bizarre... Sept qualités peuvent rendre un bâtiment différent : sa destination, sa taille, sa forme, son orientation absolue,

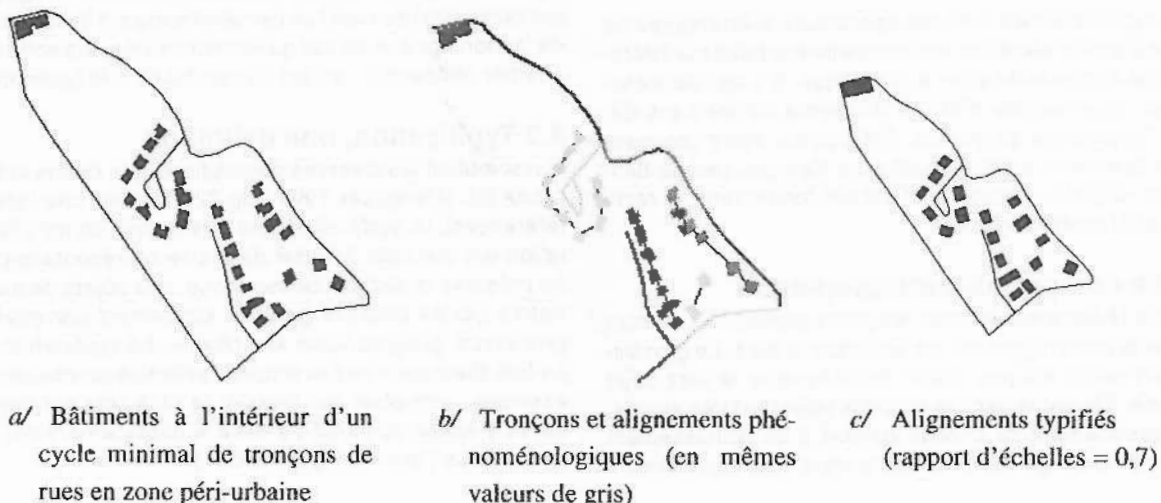
son orientation par rapport à la route, sa distance à la route, ses distances avec les voisins (Hangouët 1998 : pp.202sqq.). À l'exception de la qualité «destination», toutes ces caractéristiques peuvent se calculer avec précision et affinage au moyen du diagramme de Voronoï (*idem*). Les caractéristiques de chaque bâtiment sont comparées aux caractéristiques moyennes sur le groupe pour décider de la différence ou de la régularité du bâtiment. L'alignement, ainsi équipé de la contribution de chaque bâtiment au groupe, est prêt à être typifié.

Transformation phénoménologique du groupe de bâtiments

Le groupe final va être représenté dans l'emprise originale du groupe, tant pour des raisons de précision de localisation que de cohérence topologique - c'est-à-dire pour d'évidentes raisons géographiques. Les n bâtiments du groupe sont grossis par un facteur f pour être lisibles (f est un paramètre en entrée de l'algorithme, égal au rapport d'échelles lorsque les bâtiments sont représentés par des symboles de même taille sur la carte source et sur la carte généralisée). Le nombre de bâtiments à représenter, m , est calculé pour approcher au mieux le taux d'occupation terrain dans l'emprise originale de l'alignement. Chaque j ème position finale est interpolée entre les j ème et $(j+1)$ ème positions initiales «les plus proches», c'est-à-dire avec $i/m = j/n$ (cette manière de faire permet de préserver une large part de la dynamique de la configuration originale). L'algorithme est paramétré en entrée par les préférences cartographiques : les bâtiments qui présentent les caractéristiques remarquables indiquées seront préférés aux bâtiments de fond : les j ème et $(j+1)$ ème bâtiments initiaux sont mis en concurrence pour la représentation à la position finale i ; le bâtiment qui compte le plus de caractéristiques remarquables est préféré à l'autre. Si les deux sont également préférables, le bâtiment le plus proche de la position finale est choisi. Si à nouveau les deux sont à égale distance de la position finale, il n'y a ni moyen ni à vrai dire raison de distinguer l'un de l'autre et l'un comme l'autre fera l'affaire. Chaque bâtiment sélectionné est représenté soit dans sa forme originale relativement grossie, soit comme un assemblage ressemblant de rectangles (les bâtiments étant construits la plupart du temps - phénoménologie !- avec des murs à angles droits), le choix étant fonction de l'échelle de sortie, de la taille du bâtiment et de la complexité de sa forme, et d'un paramètre de «restitution cartographique préférentielle» indiqué à l'algorithme.

Ceci conduit à des rendus naturels, sans surprise (Figure 6).

Figure 6 : Groupement et transformation phénoménologiques des bâtiments



CONCLUSION

Considérer la fabrication d'une carte comme l'imitation de ce qui est gravé ou gravable à la surface de la Terre nous a fourni une approche naturelle pour l'automatisation de la généralisation cartographique : le chercheur s'efforce de penser en termes de phénomènes géographiques, non en termes de données. Le principe implicite mais impérieux de la phénoménologie, «à tout phénomène géographique ses méthodes automatiques particulières», annonçant apparemment des tâches titanesques et sans fin,

inspire d'abord l'effroi. Il s'avère cependant que ces méthodes spécifiques et ces algorithmes s'appuient sur l'expression des relations spatiales, lecture que les diagrammes de Voronoï sur les segments rendent systématique et simple. Notre approche et nos méthodes ne sont pas limitées aux seules fins de la généralisation. La phénoménologie s'applique à la recherche dans tout domaine où l'on attend des données géographiques, dans leur ensemble ou en partie, qu'elles soient utilisables comme des cartes du monde.

Bibliographie

- Abbas, I. et Hottier, P. (1992). *Contrôle du tracé planimétrique d'une carte - Contrôle ponctuel et contrôle linéaire*. Bilan Recherche '92, Bulletin d'Information de l'I.G.N. (1993) pp.30-36
- Barbier, J. (1965). *Thematic Cartography. Problems Particular to Illustration*. Annuaire International de Cartographie, vol.V. 167-171 (passage cité traduit par nos soins)
- Bjørke, J.T. (1993). *Information theory as a tool to formalize cartographic knowledge*. In NCGIA Initiative 8 Specialist Meeting on Formalizing Cartographic Knowledge, October 1993. 45-52 (passage cité traduit par nos soins)
- Boissonnat, J-D. et Yvinec, M. (1995). *Géométrie algorithmique*. Édiscience international. Paris.
- Brassel, K. (1990). *Der Generalisierungsbegriff in der Kartographie und anderen Disziplinen*, Schweizerische Gesellschaft für Kartographie : Kartographische Publikationsreihe Nummer 10. pp.5-16 (passage cité traduit par nos soins)
- Cuenin, R. (1972). *Cartographie générale (tome 1 : Notions générales et principes d'élaboration)*. Eyrolles, collection scientifique de l'Institut géographique national
- Desanti, J.T. (1963). *Phénoménologie et praxis*. Editions Sociales, Paris
- Eckert, M. (1921). *Die Kartenwissenschaft, Erster Band (vol. 1)*. de Gruyter. Berlin und Leipzig (passage cité traduit par nos soins)
- Gary, R. (1962). *Here is might and reassuring promise*, Life, December 21, p.4 (passage cité traduit par nos soins)
- Hangouët, J-F., et Djadri, R. (1997). *Voronoi Diagrams on Line Segments : Measurements for Contextual Generalization Purposes*. In S.C. Hirtle and A.U. Frank (Eds.). *Spatial Information Theory. A Theoretical Basis for GIS (Cosit'97)*. Springer Verlag, 207-222.
- Hangouët, J-F. (1998). *Approche et méthodes pour l'automatisation de la généralisation cartographique ; application en bord de ville*. Thèse de doctorat, Université de Marne-la-Vallée
- Henry, M. (1988). *Voir l'invisible*, Éditions François Bourin
- Kandinsky, W. (1926) *Point et ligne sur plan*, Folio essais n.168 (traduction de Suzanne et Jean Leppien)
- McMaster, R. (1989). *Introduction to 'numerical generalization in cartography'*. *Cartographica* 26(1), 1-6. (passage cité traduit par nos soins)
- McMaster, R. et Shea, K.S. (1992). *Generalization in Digital Cartography*. Association of American Cartographers (passage cité traduit par nos soins)
- Malraux, A. (1953). *Les voix du silence*, Gallimard, Paris
- Monmonier, M. (1993). *Comment faire mentir les cartes*, Flammarion 1993 (traduction de D.A. Canal)
- Morehouse, S. (1995). *GIS-map compilation and generalization*. In J-C. Müller, J-P. Lagrange, et R. Weibel (Eds). *GIS and Generalization*, GISDATA n.1 (series Eds. I. Masser et F. Salgé), Taylor & Francis, pp.21-30 (passage cité traduit par nos soins)
- Morrison, J. (1974). *A Theoretical Framework for Cartography*, Annuaire International de Cartographie, vol. XIV, pp.115-127 (passage cité traduit par nos soins)
- Müller, J-C. (1991). *Generalization of spatial databases*. In Maguire, Goodchild, Rhind (Eds). *Geographical Information Systems vol.1 Principles*. Longman Scientific Technical. 457-475

Müller, J.-C., Weibel, R., Lagrange, J.-P. et Salgé, F. (1995). *Generalization : State of the Art and Issues*. In J.-C. Müller, J. P. Lagrange, et R. Weibel (Eds). *GIS and Generalization*, GISDATA n.1 (series Eds. I. Masser et F. Salgé), Taylor & Francis, pp.3-17

Okabe, A., Boots, B. et Sugihara, K. (1992). *Spatial Tessellations, Concepts and Applications of Voronoi Diagrams*. Wiley & Sons

Oxford English Dictionary (1971). Oxford University Press (passage cité traduit par nos soins)

Peng, W., Sijmons, K. et Brown, A. (1995). *Voronoi diagram and Delaunay triangulation supporting automated generalization*. Proceedings of the 17th International Cartographic Conference - 10th General Assembly of ICA. Barcelona, Sept. 1995. pp.301-310

Platon : *La République*. Folio Essais n.228 (traduction de Pierre Pachet)

Plazanet, C. (1996). *Enrichissement des bases de données géographiques : analyse de la géométrie des objets linéaires pour la généralisation cartographique (application aux routes)*. Thèse de doctorat, université de Marne-la-Vallée

Regnauld, N. (1998). *Généralisation du bâti : Structure spatiale de type graphe et représentation cartographique*. Thèse de doctorat, université de Marseille

Rimbert, S. (1990). *Carto-graphies*. Hermès.

Robert, *Dictionnaire historique de la langue française* sous la direction d'Alain Rey (1998), éditions Le Robert

Robinson, A. et Sale, R. (1969). *Elements of Cartography*. 3d edition. John Wiley and Sons

Ruas, A. (1999). *Modèle de généralisation de données géographiques à base de contraintes et d'autonomie*. Thèse de doctorat, université de Marne-la-Vallée

Schopenhauer, A. (1844). *Eis Eauton*, L'Anabase 1992 (traduction de Guy Fillion)

Spenser, E. (1591). *Mother Hubbard's Tale*, in *Daphnaïda* (passage cité traduit par nos soins)

Ware, J.M. et Jones, C.B. (1998). *Conflict Reduction on Map Generalization Using Iterative Improvement*. *Geoinformatica*, 2(4), 383-409

Wittgenstein, L. (1921). *Tractatus Logico-philosophicus*, Gallimard / Tel (traduction de Pierre Klossowski)

Note. Les thèses (Plazanet 1996), (Regnauld 1998), (Hangouet 1998), (Ruas 1999), et d'autres réalisées au laboratoire Cogit, sont disponibles sur le serveur FTP anonyme Sturm de l'Institut géographique national :

`ftp.sturm.ign.fr`

`login : anonymous`

`password : (adresse mel)`

`cd ign/generalisation/COMMUNICATIONS/THESES`