

GÉNÉRALISATION ASSISTÉE PAR ORDINATEUR

Jean-Philippe GRELOT, Ingénieur Géographe

La généralisation cartographique vise à réduire le nombre de données transmises au lecteur tout en conservant autant que possible les informations qu'elles véhiculaient. Cette opération s'applique soit en conservant le caractère général de la carte et en diminuant l'échelle (cas des cartes topographiques), soit en changeant le caractère de la carte dont on diminue éventuellement l'échelle (cas de la production d'une carte routière à partir d'une carte topographique).

La généralisation est un acte global : comme celle d'une caricature, la qualité d'une généralisation ne se juge que lorsqu'elle est achevée ; de ce fait on lui reconnaît une essence statistique difficilement compatible avec les contraintes de l'informatique et l'atomisation des tâches qu'elle requiert.

Sélection, schématisation, harmonisation

On distingue classiquement trois étapes dans un processus de généralisation.

La **sélection** consiste à choisir les objets qui seront conservés mais aussi et par voie de conséquence, ceux qui seront éliminés. Elle ne s'applique pas de manière arbitraire, ni uniforme : le nombre d'objets conservés dépend de leur densité initiale et de leur nature, et des critères de choix définissent des priorités. Une formulation mathématique est nécessaire pour opérer en cartographie automatique ce que le cartographe fait de manière presque naturelle lorsqu'il a sous les yeux l'ensemble d'une zone.

La **schématisation** est destinée à supprimer les détails graphiques nuisibles à la lecture et à renforcer les traits caractéristiques. Les lignes sont redessinées et non pas simplement réduites comme le ferait un appareil photographique, les signes conventionnels sont également modifiés. Les règles de lisibilité gouvernent cette étape.

La sélection et la schématisation ont essentiellement un caractère local. Pour compenser la tendance à l'égalisation et à l'uniformisation qu'elles induisent, on procède à une étape d'**harmonisation**, en rétablissant les rapports de densité entre les catégories d'objets, en assurant les raccords avec les coupures voisines, en ajustant les emplacements relatifs des objets les uns par rapport aux autres (rivières au fond des thalwegs dessinés par les courbes de niveau,...).

Généralisation conceptuelle et généralisation structurelle

Si l'on considère la carte comme une représentation de l'espace géographique, alors la généralisation doit être

vue comme conduisant à une nouvelle carte. Il s'agit d'abord d'un processus intellectuel de définition, que l'on nomme **généralisation conceptuelle**, suivi d'une phase de mise en œuvre par le dessin, ou **généralisation structurelle**. La généralisation conceptuelle intègre la sélection et une grande partie des opérations d'harmonisation, tandis que la généralisation structurelle est assez proche de la schématisation.

Cependant, cette approche est plus satisfaisante dans la mesure où elle se rattache davantage au concept de représentation de l'espace qu'au résultat de l'expérience ineffable de l'artiste praticien. En particulier, elle fait mieux saisir l'aspect taxonomique de la généralisation, tout à fait parallèle à son aspect géométrique universellement connu. Les objets cartographiques à une échelle réduite ne sont pas de même nature que ceux de la carte initiale, même si on leur donne en légende le même nom, ce qui est quelque peu abusif. Prenons l'exemple d'un point pris sur une carte à l'intérieur d'une zone boisée : si la carte est à grande échelle, le point correspondant du terrain sera réellement dans un bois, avec une probabilité de l'ordre de 99 % ; si maintenant la carte est à petite échelle, le point correspondant à celui relevé sur la carte sera réellement dans un bois avec une probabilité réduite à 85 % (par exemple), si le signe adopté pour la zone boisée a été appliqué à une étendue composée d'un bois dont les clairières couvrent le septième de la surface ; la notion de bois ne recouvre pas la même réalité géographique quelle que soit l'échelle.

Un autre cas de généralisation montre un aspect structurel qui dépasse la seule schématisation. Lorsque l'échelle se réduit, une ville passe successivement de sa représentation par des objets ponctuels (les constructions) à une autre par des objets zonaux (les zones bâties, agrégations de constructions), puis par un signe conventionnel ponctuel (la ville), avant d'être intégrée à un ensemble zonal (l'agglomération urbanisée), puis de disparaître. Ou encore, un barrage figuré initialement comme objet linéaire devient ensuite ponctuel conventionnel.

Prêter à la généralisation un aspect conceptuel et un aspect structurelle est une vue théorique intellectuellement satisfaisante. Elle requiert toutefois, et ceci est inscrit dans les termes eux-mêmes, que l'on puisse, de la lecture des données initiales, tirer le schéma conceptuel ayant présidé à l'élaboration de ces données et en déduire un nouveau schéma conceptuel. Cette tâche de conceptualisation, déjà difficile en cartographie, est encore compliquée par les restrictions des modèles de données informatiques. La pratique fait revenir au triptyque sélection-schématization-harmonisation, mais en replaçant en quelque sorte la carte ou les données à obtenir au point de départ du processus : on définit ce que l'on veut montrer (objets, relations entre objets) avant de voir comment traiter les données initiales.

Tant que le facteur de réduction d'échelle est inférieur à 5, la généralisation reste essentiellement structurelle ; si ce facteur est supérieur à 5, alors l'étape conceptuelle devient déterminante. De fait, ce ratio est fréquemment le seuil de passage d'une série cartographique à une autre, chacune des séries ayant ses techniques de rédaction à partir de ses sources de données, son rythme de révision, ses utilisations et ses utilisateurs. Ceci produira des bases de données constituées assez indépendamment les unes des autres, quitte à accrocher ensuite des passerelles pour, en particulier, faire bénéficier les unes des mises à jour effectuées sur les autres, avec les précautions géométriques et taxonomiques exigées par les différences de précision.

Formules de sélection

La définition conceptuelle de la carte étant faite, en particulier pour ce qui concerne sa taxonomie, on procède à la sélection des objets. Dans chaque unité de traitement, qui prend commodément l'aspect d'une dalle rectangulaire dans le système de coordonnées choisi, on va reporter certains des objets présents dans l'espace correspondant de la carte initiale, triés selon les critères retenus.

Intéressons-nous au nombre d'objets conservés. Dési-

gnons par N le nombre d'objets, par $\frac{1}{E}$ l'échelle ; par l

l'épaisseur moyenne des objets linéaires et par s la surface moyenne des objets zonaux, mesurées sur la carte et respectivement par les indices i et f les données relatives à la carte initiale et à la carte finale. En conservant les mêmes caractéristiques graphiques, on obtiendrait par strict respect de la densité des objets figurés, des valeurs du rapport $\frac{Nf}{Ni}$ égales à $(\frac{Ei}{Ef})^2$ pour les objets ponctuels,

$\frac{Ei}{Ef}$ pour les objets linéaires et 1 pour les objets zonaux

dont la superficie peut être réduite dans le rapport des surfaces. Cependant, par observation de cartes réelles, Töpfer et Pillewizer ont établi une formule générale $\frac{Nf}{Ni} = \sqrt{\frac{Ei}{Ef}}$ devenue ensuite $\frac{Nf}{Ni} = Cb Cz \sqrt{\frac{Ei}{Ef}}$ par l'introduction de deux paramètres (tableau 1).

- Cb , paramètre d'exagération symbolique, traduisant l'amplification nécessaire pour respecter les règles de lisibilité ;

- Cz , paramètre de forme symbolique, traduisant l'importance donnée à certains types d'objets qui donne un aspect thématique à la carte généralisée.

Tableau 1 - Valeurs du rapport $\frac{Nf}{Nc}$

	$\begin{matrix} & & cb \\ & \diagdown & \\ cz & & \end{matrix}$	pas d'amplification (objets zonaux principalement) $\sqrt{\frac{Ef}{Ei}}$	amplification dans le rapport d'échelles (objets linéaires principalement) 1	amplification importante (objets ponctuels principalement) $\sqrt{\frac{Ei}{Ef}}$
pas de modification essentielle	1	1	$\sqrt{\frac{Ei}{Ef}}$	$\frac{Ei}{Ef}$
mise en évidence d'objets linéaires	$\frac{li}{lf} \sqrt{\frac{Ei}{Ef}}$	$\frac{li}{lf} \sqrt{\frac{Ei}{Ef}}$	$\frac{li}{lf} \frac{Ei}{Ef}$	$\frac{li}{lf} (\frac{Ei}{Ef})^{3/2}$
mise en évidence d'objets zonaux	$\frac{si}{sf} \frac{Ei}{Ef}$	$\frac{si}{sf} \frac{Ei}{Ef}$	$\frac{si}{sf} (\frac{Ei}{Ef})^{3/2}$	$\frac{si}{sf} (\frac{Ei}{Ef})^2$

On constate à la lecture de ce tableau une surdensification des objets linéaires et surtout ponctuels lorsque le caractère de la carte initiale est maintenu. Si l'on met en évidence une catégorie d'objets, ce qui donne un aspect thématique à la carte, on élague davantage jusqu'à éliminer l'effet de surdensification : il en est ainsi des objets linéaires lorsque la généralisation cherche à les montrer, ou des objets ponctuels lorsque le thème porte sur les objets zonaux.

Ces formules peuvent être nuancées en fonction de l'échelle, car la surdensification ne croît pas indéfiniment. La densité maximale est souvent atteinte pour une échelle de l'ordre de 1:1 000 000, après quoi le caractère de la carte change et induit une réduction de la densité de données.

Par ailleurs, des formules ainsi appliquées de manière

indépendante aux diverses catégories d'objets supposent que les proportions d'objets des différents types soient dans des valeurs proches de normes fixées par avance. Le respect des caractères locaux, qui répond à un souci d'harmonisation, demande l'analyse de corrélations fines entre les nombres d'objets de chaque type, ce qui vient ensuite amender les formules de sélection. Il reste à définir alors des modèles d'équations de corrélation adaptés aux paysages que l'on traite.

Généralisation structurelle

En théorie, la généralisation structurelle repose sur l'analyse de la forme de l'objet. Or la cartographie automatique achoppe précisément sur l'appréhension de la forme dans sa globalité, elle se contente d'une décomposition en micro-éléments désarticulés. Il faut donc user de moyens détournés.

Le premier de ces moyens est la description sémantique de l'objet, dont le rapprochement avec la taxonomie retenue indiquera, outre le mode d'implantation (ponctuel, linéaire ou zonal), une approche du traitement géométrique à mettre en œuvre.

Dans un certain nombre de cas, on décidera alors d'employer des signes conventionnels, se substituant à d'autres signes conventionnels ou à des descriptions spécifiques. Les variables visuelles : forme, couleur, taille, orientation donnent la capacité de concevoir une large palette de tels signes ; cependant, une surabondance nuit à la lecture et l'équilibre doit être trouvé entre la vision synthétique qu'offre une carte généralisée et l'analyse que requiert la multiplicité des signes conventionnels ; l'une est une "carte à voir", l'autre une "carte à lire".

Dès que l'on ne représentera pas un objet par un signe ponctuel ou un signe ponctuel orienté, on cherchera à simplifier la forme en travaillant la plupart du temps sur des éléments linéaires : implantation elle-même pour les objets linéaires, ou contours des objets zonaux. Les outils sont les algorithmes de lissage, séparés en trois catégories :

— les algorithmes d'élimination de points, comme l'élimination systématique de a points tous les b points avec $\frac{a}{b} = p \frac{E_i}{E_f}$ où p est soit constant (égal à 1), soit pris aléatoirement entre 0 et 2 ;

— les algorithmes de rééchantillonnage, comme celui de l'abscisse curviligne : sur une ligne décomposée en segments rectilignes de longueur moyenne λ_i , on choisit les points définissant des segments de longueur curviligne $\lambda_i \frac{E_f}{E_i}$, et ces points sont les sommets de la nouvelle ligne polygonale ;

— les algorithmes de sélection, comme celui de la corde glissante, qui procède par itérations : le premier point de la ligne est pris comme point d'ancrage et le dernier comme point flottant, et ils déterminent une corde ; on mesure la distance euclidienne de tous les points intermédiaires à cette corde ; si le point le plus éloigné est à une distance supérieure à la tolérance fixée (dans le cas de la généralisation, écart graphique retenu comme significatif pour le type de ligne étudié), il devient le nouveau point flottant ; on remonte ainsi par étapes vers le point d'ancrage, jusqu'à trouver un point flottant et une corde tels que les points intermédiaires soient à une distance de la corde inférieure à la tolérance ; ce point flottant devient point d'ancrage, et le dernier point de la ligne redevient point flottant ; la ligne généralisée a pour sommets les points d'ancrage successifs.

Parmi les algorithmes de rééchantillonnage doivent être mentionnées les fonctions splines. Ce sont des polynômes du troisième degré, définis avec des conditions aux limites préservant la continuité et la dérivabilité (et la courbure éventuellement), et qui minimisent la somme des carrés des valeurs prises par la fonction polynomiale aux points approximatifs : ainsi, les fonctions splines sont des estimateurs au sens des moindres carrés.

En mode maillé, il faut appliquer une convolution à chaque cellule de pixels donnant le pixel "généralisé". La fonction de convolution donne un poids plus ou moins prépondérant à l'un des pixels, et répartit les poids restants sur les pixels voisins.

La complexité relative d'un algorithme de lissage n'est pas, du point de vue de la généralisation, gage d'une plus grande efficacité. On sait que l'algorithme de l'abscisse curviligne amollit les courbes et en gomme les aspérités, moins toutefois que les fonctions splines, alors que l'algorithme de la corde glissante conserve en priorité ces aspérités ; l'élimination systématique a un comportement plus aléatoire. C'est donc en fonction de la morphologie de la ligne que l'on choisira un algorithme plutôt qu'un autre, et la description mécanique est utile en cela.

Le second moyen détourné pour appréhender la morphologie d'une ligne particulière est de calculer des grandeurs caractéristiques qui lui sont attachées : distribution des longueurs de segments, amplitude et distribution des changements de direction, etc ; il peut être nécessaire de fractionner les lignes en portions plus homogènes. Il est alors possible de recoder les coordonnées des lignes polygonales, qui deviennent comparables après une normalisation opérée par application d'un facteur d'échelle approprié. Après avoir défini une distance entre lignes polygonales normalisées, on regroupe celles-ci en classes d'équivalence : tout individu de la classe est apte à représenter ses congénères ; en particulier, la ligne qui comporte le nombre minimal de segments est en quelque sorte la forme généralisée des lignes de la classe. Pour atténuer l'effet par trop simplificateur de ce regroupement, il faut associer à chaque individu de la classe une grandeur compensant l'effet de la normalisation, comme le facteur d'échelle appliqué ou la dimension fractale de la ligne ; cette grandeur est utilisée lors de la restitution.

Plusieurs formes de description géométrique peuvent coexister au sein d'une même classe : lignes polygonales, courbes polynomiales, fonctions, fractales, distributions statistiques ou probabilistes, fractales. L'étude a priori des effets des algorithmes de lissage est utile pour guider le choix d'une restitution généralisée : il semble par exemple que l'algorithme de l'abscisse curviligne diminue la dimension fractale d'une ligne polygonale alors que celui de la corde glissante l'augmente.

Enfin, et ceci différencie la généralisation assistée par ordinateur de celle opérée manuellement, la schématisation choisie dépend de l'utilisation de la carte. Ainsi, une visualisation sur console en temps réel optera pour une sélection systématique, rapide ; une carte routière nécessitera de conserver les traits marquants du réseau routier tels que changements de direction ou décrochements dans les carrefours, en recourant à l'algorithme de la corde glissante ; l'utilisation des mêmes données comme fond d'une carte thématique, par exemple une carte des flux de circulation, fera préférer un lissage extrême par l'algorithme de l'abscisse curviligne ou par fonctions splines.

En définitive, trois facteurs interviennent dans la généralisation structurelle : la nature sémantique des données, leur forme géométrique, et l'utilisation prévue des données généralisées.

Opérateurs élémentaires de généralisation

La généralisation ne porte pas sur un objet isolé de son contexte, mais intéresse l'ensemble des objets situés dans la zone traitée : en ce sens, elle présente un caractère plus global tout en se décomposant en un ensemble complexe d'opérations régies par des règles. Ainsi, il est possible de voir le résultat comme une fonction des objets initiaux dépendant de paramètres.

Tableau 2 - Opérateurs élémentaires de généralisation

Objets ponctuels	Objets linéaires	Objets zonaux
sélectionner/éliminer agrèger/classifier déplacer/écarter agrandir	sélectionner/éliminer classifier déplacer/écarter agrandir/rétrécir	sélectionner/éliminer classifier déplacer/écarter agrandir/rétrécir
	lisser	lisser le contour
	superposer	superposer
	réduire	réduire remplacer par objet linéaire/ponctuel
	convertir maillé- vectoriel/vectoriel-maillé	convertir maillé-vectoriel/ vectoriel maillé
	changer la topologie	changer la topologie
dessiner	dessiner	dessiner

Par exemple, pour une maison située en bordure d'une route : généraliser (maison, route) = réduire (déplacer, (intersecter (agrandir (maison) * agrandir (route))), ce qui signifie : agrandir les signes conventionnels maison et route, chercher leur intersection et si nécessaire déplacer les deux objets jusqu'à ne plus avoir d'intersection, puis réduire à l'échelle de sortie.

Les opérations varient selon le type d'objet (tableau 2).

On définit également des opérateurs pour les objets volumiques :

sélectionner/éliminer
classifier
lisser la surface
superposer
trouver les gradients/les pentes/les extrema locaux
reconnaître/éliminer des formes spécifiques
reconnaître des objets spécifiques
trouver/modifier les lignes caractéristiques
remplacer par des objets zonaux/linéaires/ponctuels
convertir maillé \Leftrightarrow isolignes \Leftrightarrow triangulation
dessiner

ainsi que pour les toponymes :

sélectionner/éliminer
déplacer/écarter
réduire/agrandir
changer la forme de la ligne de base
dessiner.

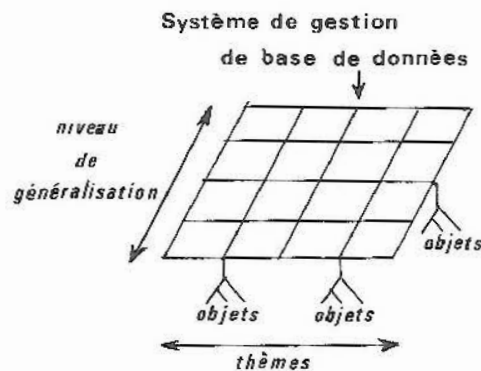
Système cartographique indépendant de l'échelle

Ces opérateurs font inmanquablement songer à constituer une base de données unique dont seraient dérivées des cartes généralisées, les réductions d'échelle et les accentuations thématiques s'effectuant quasiment en temps réel.

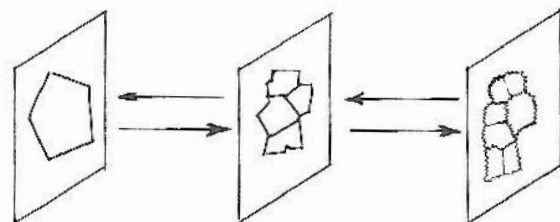
Ceci n'est guère envisageable, du moins en l'état actuel des connaissances. Au demeurant et comme il a déjà été souligné, une interprétation conceptuelle est nécessaire dès lors que le facteur de réduction dépasse 5, et rend

insuffisante la schématisation géométrique. Dans tous les cas, il semble difficile d'éviter des reprises sur un poste interactif.

Dès lors, on envisage plutôt la base de données "indépendante de l'échelle" sous forme d'un système général composé de plusieurs bases, chaque base correspondant à un niveau de généralisation et à un sujet particulier.



Chaque objet possède alors plusieurs descriptions, sémantiques et géométriques qu'il faut lier entre elles par des pointeurs.



Généralisation, prédicat, métrique

Au-delà de la technique cartographique, la réflexion sur la généralisation assistée par ordinateur est une réflexion sur l'information géographique et sur sa perception. Elle est essentielle du fait que la limitation des capacités des consoles de visualisation substituera à la lecture libre d'une carte conventionnelle détaillée, une lecture par approches successives allant du général au particulier, donc commençant par le niveau le plus généralisé pour accéder éventuellement aux données de base, détaillées.

La généralisation établit clairement la séparation formelle entre description sémantique et description géométrique. Lire dynamiquement une carte du généralisé au détaillé, c'est enchaîner des perceptions successives dans une activité de lecture : à chaque étape est mentalement élaboré un prédicat que la lecture suivante infirme ou confirme. Or ces prédicats de plus en plus précis reposent sur une description sémantique correctement assimilée et mémorisée, alors que la géométrie ne l'est qu'imparfaitement. Plutôt qu'aux positions individuelles des objets, le lecteur réagit à leur disposition relative, dénommée topologie par abus de langage.

Il s'ensuit une série de questions qui, outre le concepteur du système de visualisation accédant aux données, intéresse le concepteur de la base de données elle-même : quelles activités cartographiques nécessitent-elles une perception géométrique précise, autre que des listes de valeurs numériques (coordonnées, distances, longueurs...) ? Quels niveaux de généralisation adopter, pour quels niveaux de perception et quels prédicats (ou enchaînements de prédicats) ?

Les cartes conventionnelles donnent d'emblée la précision maximale des données, pour l'échelle retenue. Est-ce légitime, ou bien seulement une habitude imposée par la technique ? L'affinement, opération inverse de la généralisation, opéré par une visualisation dynamique au fur et à mesure que le lecteur requiert une plus grande précision, modifiera-t-il sensiblement le processus de lecture cartographique ? Je crois que ce sujet mérite d'être étudié au moment où de nouvelles bases de données vont amener à créer de nouvelles formes de cartes. Plus que jamais, la conception cartographique doit intégrer les activités et les moyens de communication.

BIBLIOGRAPHIE

BRASSEL K. 1985, Strategies and Data Models for Computer - Aided Generalization, *International Yearbook of Cartography* n° 25, pp.11-29

BRIEDESCA E. 1982, Shapes Classification on Digital Images, *ISPRS Commission IV Symposium 1982*, pp. 85-103.

CUENIN R. 1972, *Cartographie Générale*, Editions Eyrolles.

DEPUYDT F. 1984, La généralisation cartographique, *La Formation des Cartographes*, Séminaire ACI de Rabat, pp. 91-110.

SAALFELD A. 1985, Shape Representation for Linear Features in Automated Cartography, *1986 ACSM-ASPRS Annual Convention Vol. 1*, pp.143-152.

ZHU GUORI 1986, The Mathematical Model of Cartographic Generalization, *Seminar on Advanced Cartographic Education and Training*, Wuhan.