



Vers l'automatisation de la production de cartes

Guillaume Touya

► To cite this version:

Guillaume Touya. Vers l'automatisation de la production de cartes. Informatique [cs]. Université Paris Est, 2017. tel-02475520

HAL Id: tel-02475520

<https://theses.hal.science/tel-02475520>

Submitted on 12 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Vers l'automatisation de la production de cartes

Synthèse de travaux de recherche
en vue d'obtenir **l'habilitation à diriger des recherches**
délivrée par l'université Paris-Est
spécialité Sciences et Technologies de l'Information Géographique

Guillaume Touya

IGN/SRIG/Equipe COGIT



Présenté le 4 mai 2017 devant le jury composé de :

Mme Paule-Annick Davoine, Rapporteur, Maître de conférences, HDR, Université de Grenoble
Mme Cécile Duchêne, Examineur, IDTGCE, HDR, IGN
M Christophe Hurter, Rapporteur, Professeur, Ecole Nationale d'Aviation Civile
M Emmanuel Pietriga, Examineur, Directeur de recherche, Université Paris Sud, INRIA
Mme Anne Ruas, Examineur, Ingénieur en chef des ponts, eaux et forêts, HDR, IFSTTAR
M Robert Weibel, Rapporteur, Professeur, Université de Zürich (Suisse)

Table des matières

Chapitre 1.	La production de cartes, aujourd'hui et demain	4
1.1	Quelques notions de cartographie	8
1.1.1	Bases de données géographiques vectorielles et bases de données cartographiques ..	8
1.1.2	Notion d'échelle et de niveau de détail	9
1.1.3	Généralisation cartographique et son automatisation	12
1.1.4	Autres processus de cartographie automatique	17
1.1.5	Une Chaîne de production automatique de cartes	18
1.2	Cartes en série et carte à la carte	19
1.2.1	Production de cartes en série et mise à jour	20
1.2.2	La Carte à la carte ou cartographie à la demande	20
1.3	Carte papier et application web de cartographie	22
1.3.1	La Carte sur un écran	22
1.3.2	Les Portails cartographiques	23
1.4	IGN, données hétérogènes et collaboratives	24
1.4.1	Quelles différences entre toutes ces données ?	24
1.4.2	Difficultés liées à l'hétérogénéité sémantique	25
1.4.3	Difficultés liées à l'hétérogénéité de qualité	25
1.4.4	Difficultés liées à l'hétérogénéité de niveau de détail	26
1.5	Synthèse et objectifs scientifiques	26
1.5.1	Où sont les verrous scientifiques pour l'automatisation des processus cartographiques ?	26
1.5.2	Plan du mémoire	28
Chapitre 2.	Comment automatiser ces processus cartographiques ?	29
2.1	Recalage géométrique de données géographiques	29
2.2	Harmonisation de niveau de détail pour les cartes à grande échelle	32
2.2.1	Inférence du niveau de détail pour des données hétérogènes	33
2.2.2	Opérations d'harmonisation de niveau de détail	37
2.2.3	De l'algorithme au processus, parallèle avec la généralisation	41
2.3	Formalisation des connaissances en cartographie	42
2.3.1	Connaissances nécessaires à un processus de généralisation automatique	42
2.3.2	Vers une ontologie des relations spatiales	43
2.3.3	Des ontologies partagées par les chercheurs en généralisation ?	45
2.4	Meta-modèles de généralisation	46
2.4.1	Opérateur, algorithme, modèle et meta-modèle de généralisation	46
2.4.2	Le ScaleMaster2.0	47

Chapitre 3. Amélioration des processus de généralisation automatique.....	51
3.1 Modélisation multi-niveaux dans la généralisation par agents	51
3.1.1 Généralisation cartographique et modélisation multi-niveaux	51
3.1.2 Le modèle DIOGEN	53
3.1.3 Application aux cartes de randonnées à la manière des plans de bus	56
3.1.4 Patterns d'analyse pour les SMA multi-niveaux.....	59
3.2 Génération de niveau intermédiaires dans une pyramide d'échelles	61
3.2.1 Etude des portails cartographiques multi-échelles existants.....	62
3.2.2 Dérivation de niveaux intermédiaires entre deux niveaux préexistants	66
3.2.3 Généralisation « consciente » des niveaux multiples	68
Chapitre 4. Conclusions et perspectives de recherche	71
Résumé des contributions pour l'automatisation des processus cartographiques.....	71
Besoins futurs en cartographie automatique	71
Perspectives en généralisation cartographique.....	72
Démontrer l'utilité de la généralisation cartographique	74
Vers une meilleure intégration de tous les processus cartographiques.....	74
Cartographier de gros volumes de données	75
Bibliographie.....	77
Eléments de CV.....	90
Parcours professionnel.....	90
Activités de recherche.....	90
Projets collaboratifs :.....	90
Organisation d'évènements, société savantes :.....	90
Autres :	90
Travaux encadrés	91
Thèses dirigées par dérogation	91
Thèses encadrées	91
Post-doctorants/ingénieurs d'étude encadrés.....	91
Stages encadrés (ou co-encadrés).....	91
Publications	91
Revue à comité de lecture	91
Actes de conférence (papier long)	92
Chapitres de livre.....	93
Ateliers ou conférence sur papier court.....	93
Autres (posters, revues sans comité de lecture, rapports techniques...).....	95

Chapitre 1. La production de cartes, aujourd'hui et demain

Le processus de création d'une carte, notamment dans le cas d'une carte topographique (une carte qui décrit les caractéristiques et la géographie du terrain), est souvent long et fastidieux, même lorsqu'on dispose d'un système d'information géographique (SIG). En effet, il faut choisir avec soin les symboles et la légende, placer de manière lisible les textes de la carte, et pour que les symboles de la carte soient lisibles lorsque l'échelle diminue, il faut les grossir ce qui augmente leur emprise et rend indispensable une simplification du contenu, que l'on appelle généralisation cartographique. Très tôt, les chercheurs en cartographie ont cherché à utiliser l'informatique pour automatiser les processus de création d'une carte, souvent en essayant d'imiter les cartographes humains (Mackaness et Fisher 1987), (Buttenfield et Mark 1991), (Buttenfield et McMaster 1991). Ces techniques se sont perfectionnées avec les avancées de l'informatique, empruntant notamment au domaine de l'intelligence artificielle, comme le montre l'évolution des recueils sur les avancées en généralisation cartographique (McMaster et Shea 1992), (A. Ruas 1999), (Mackaness, Ruas et Sarjakoski 2007) et (Burghardt, Duchêne et Mackaness 2014).

Mais dans le même temps, ces évolutions technologiques ont entraîné une révolution pour le domaine de la cartographie. Avec la conjonction d'internet et de l'accessibilité des géo-navigateurs, comme le GPS, à petit prix, tout le monde, ou presque, est devenu un cartographe. De nouveaux acteurs, comme Google ou OpenStreetMap, sont apparus et ont bousculé les habitudes des producteurs traditionnels de cartes. Si ces nouveaux acteurs ont démocratisé la cartographie, à l'aide des techniques de cartographie sur le web, ils ont favorisé le développement et l'utilisation de cartes de mauvaise qualité, mais produites et diffusées rapidement. Il s'agit d'un changement de paradigme dû au changement de support cartographique, du papier à l'écran : il est désormais plus important d'avoir une carte à jour accessible rapidement, plutôt qu'une carte complète et parfaitement lisible, mais qui prend du temps à construire. Nous appelons ce phénomène la cartographie « *vite fait mal fait* » par opposition à la cartographie traditionnelle des agences de cartographie comme l'IGN, qui peut être vue comme « *lentement fait bien fait* ». On peut voir un certain nombre d'exemples représentatifs des différents types de cartes « vite fait mal fait » que l'on peut trouver sur internet dans (Gaffuri 2011), notamment des cartes qui superposent des données thématiques d'un utilisateur à un fond cartographique proche d'une carte topographique, ce qui est une autre des nouveautés amenées par ces changements technologiques. Mais nous illustrons d'abord ici le problème à travers des cartes topographiques sans superposition thématique : la Figure 1-1 montre quatre cartes au 1 : 25000 de la même zone, la ville d'Orthez, dont trois exemples différents du « vite fait mal fait » et une carte IGN, représentative d'une bonne qualité obtenue par un long travail par un cartographe de métier. La carte issue d'OpenStreetMap est complète mais n'a pas subi d'opération avancée de généralisation, ce qui fait que les bâtiments, par exemple sont souvent trop petits et trop proches les uns aux autres pour être vus comme séparés à cette échelle. La carte produite par la société ESRI, extraite du produit *World Topographic Map*, a, elle, bien été généralisée : les bâtiments ont été grossis et éloignés entre eux et des symboles de routes. Mais cette généralisation a été faite avec des opérations trop simples et donc peu efficaces, entraînant notamment beaucoup de superpositions entre bâtiments. Enfin, la carte extraite de GoogleMaps prend le problème de la généralisation autrement en enlevant les éléments susceptibles de créer des problèmes de lisibilité. Mais le vide de cette carte est aussi pour nous un signe de mauvaise qualité cartographique.

Notre objectif à long terme, qui est aussi celui de nombreux chercheurs en cartographie, est de permettre une cartographie « vite fait bien fait ». Une vue centrée sur l'IGN de cet objectif est de se rapprocher de la qualité de la carte IGN avec le plus possible d'automatisation des processus. Toutefois, devant la très forte utilisation de ces cartes « vite fait mal fait », on pourrait se dire que cet objectif n'est pas très important, les utilisateurs pouvant se débrouiller avec des cartes de mauvaise qualité. Mais les utilisateurs ne se rendent pas compte de la perte d'efficacité dans leur utilisation des cartes. Ils pourraient faire plus de choses, plus rapidement avec de meilleures cartes. A ce sujet, Ormeling plaide pour le développement de la cartographie de qualité (Ormeling 2011) :

"The digital generations have brought us ubiquitous cartography, that is spatial data that can be accessed anywhere, anytime. These generations will realise before long, that in order to be able to read and process and understand that spatial information, people want it to be designed in a form they can cope with. And they also will understand in time that people at large will need some coaching in map use, in order to understand what they see on their screens."

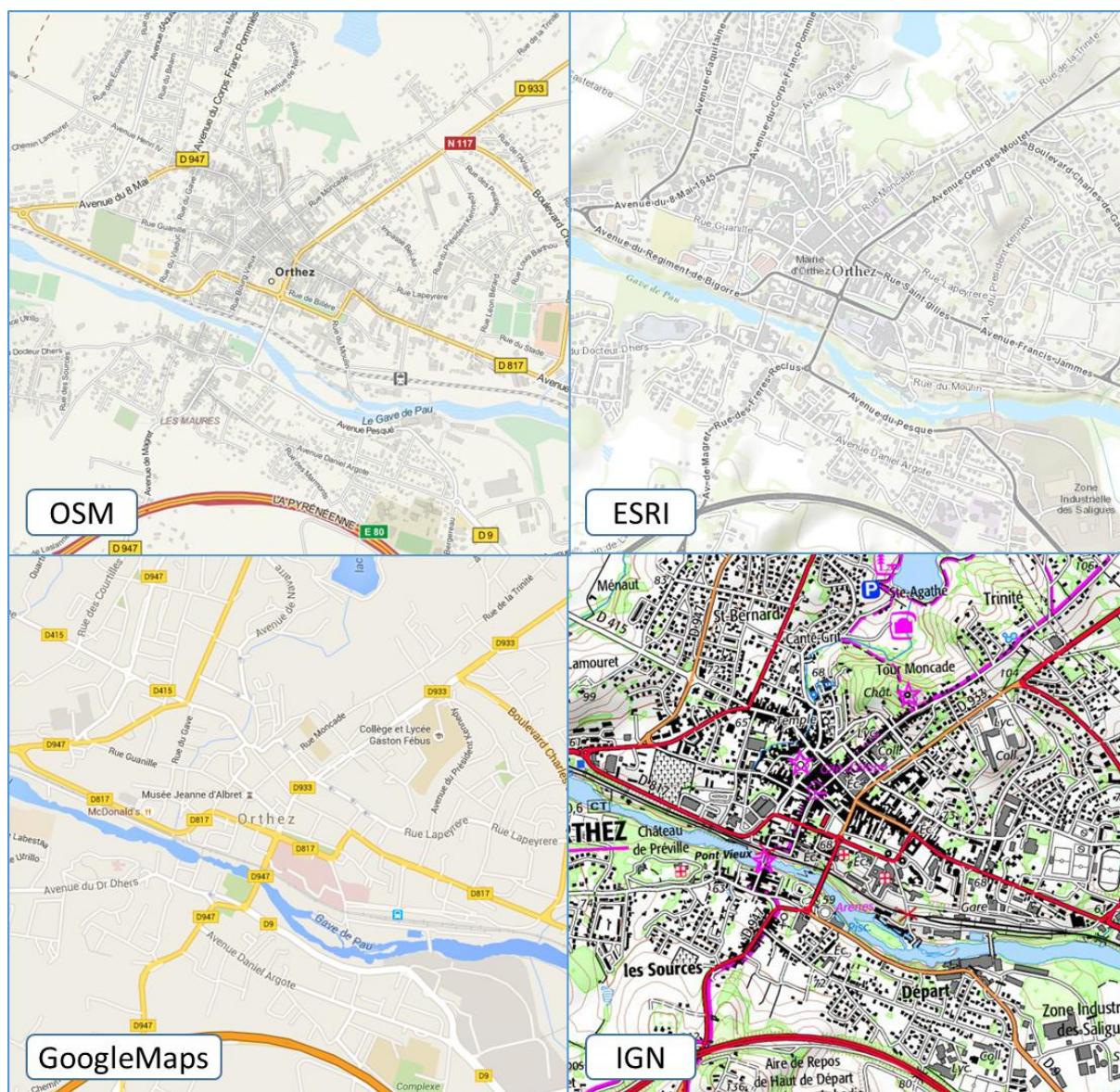


Figure 1-1 Quatre cartes topographiques de la même zone et à la même échelle (1 : 25000) qui présentent différentes qualité cartographiques.

Même si nous ne sommes pas vraiment d'accord avec les solutions proposées par Ormeling, c'est-à-dire former le public à la « vraie » cartographie, nous partageons le constat sur le manque d'efficacité des cartes de mauvaise qualité. Mais plutôt que de former le grand public à la cartographie traditionnelle, faisons en sorte de proposer des méthodes et outils pour que ces cartes accessibles à tout le monde soient de meilleures qualité. C'est cette envie, ce besoin qui sert de fil conducteur à l'ensemble des recherches présentées dans ce mémoire.

Parcours de recherche

Afin de bien comprendre, pourquoi mes recherches se concentrent sur cet objectif d'automatisation de la création de cartes de qualité, je fais un petit aparté pour décrire mon parcours de recherche, qui est assez atypique par rapport à un parcours universitaire classique. Après une formation à l'ENSG, école d'ingénieur de l'IGN et un DEA en géomatique, j'ai réalisé toute ma carrière de chercheur à l'IGN et dans l'équipe COGIT où je suis arrivé en septembre 2004. Les équipes de recherche de l'IGN mixent des permanents IGN affectés à la recherche, dont je fais partie, des chargés/directeurs de recherche du ministère de l'environnement et, comme n'importe quel laboratoire de recherche, des doctorants et post-doctorants. Ce statut particulier permet aux chercheurs de l'équipe d'être au contact des autres agents de l'IGN pour permettre le transfert d'innovations ou le recueil de besoins.

Autre particularité, j'ai travaillé pendant un peu plus de trois ans sur des projets de recherche de l'équipe, avant de commencer une thèse en janvier 2008, ce qui m'a permis d'apprendre le métier de chercheur tout en approfondissant un sujet qui m'a suivi jusqu'à aujourd'hui, l'automatisation de la généralisation cartographique. Mes travaux de thèse (soutenue en juin 2011) ont porté sur la généralisation collaborative, c'est-à-dire l'automatisation du recours à plusieurs méthodes de généralisation, non conçues pour fonctionner de concert, appliquées à différentes parties de la carte à généraliser. Par exemple, les zones urbaines sont généralisées avec la méthode la plus efficace sur ce type de zone, les zones rurales avec une autre méthode, et les zones de montagne avec une troisième méthode, tout en gérant les effets de bords dans ces zones qui peuvent se chevaucher.

Après ma thèse, j'ai eu l'opportunité de continuer à travailler sur les perspectives de ma thèse en généralisation, tout en démarrant de nouvelles recherches autour de la cartographie de données issues de projets de crowdsourcing comme OpenStreetMap. Ces différents travaux sont illustrés dans ce mémoire. Ce parcours m'a finalement amené au montage du projet MapMuxing¹ (ANR-14-CE24-0011-01), financé par l'ANR lors de l'appel générique 2014, dont je suis le porteur principal. Ce projet est réalisé en partenariat avec l'équipe ILDA de l'INRIA, spécialisée dans les interfaces homme-machine (IHM) pour naviguer dans des grands jeux de données, et avec des géographes de l'UMR GRED de Montpellier spécialisés dans la gestion des risques naturels. Il s'agit d'améliorer l'expérience de navigation dans les applications cartographiques multi-échelles telles que le géoportail ou GoogleMaps, dans lesquels des changements de niveau de zoom ou de type de représentation cartographique (de la carte à la photo par exemple) demandent un effort cognitif important à l'utilisateur, notamment dans des situations de stress comme une alerte avant un tsunami. Nous proposons de coupler des IHM innovantes à des techniques de cartographie automatique avancées pour améliorer cette navigation et de valider ces propositions par des tests utilisateurs rigoureux. La dernière partie de ce mémoire abordera les premiers résultats de ce projet devant se terminer en 2018. Ces différents projets de recherche m'ont permis d'aborder l'automatisation des processus cartographiques sous différents aspects que j'ai essayé de résumer dans ce mémoire.

¹ <http://mapmuxing.ign.fr>

Plan du chapitre

Ce premier chapitre est composé de cinq parties. La première présente des notions sur la cartographie et la généralisation cartographique, fondamentales pour la compréhension du mémoire. La deuxième partie présente les différences, pour l'automatisation de la cartographie, entre la production de cartes en série et la dérivation de cartes à la demande. La troisième partie est centrée sur les différences entre cartes imprimées sur du papier et cartes visualisées sur un écran, quelle que soit sa taille. La quatrième partie traite des spécificités de l'utilisation de données d'un utilisateur dans les cartes, ou de données issues de saisies collaborative par des citoyens. Enfin, la cinquième partie résume les verrous scientifiques abordés dans ce mémoire, et présente le plan des autres chapitres.

1.1 Quelques notions de cartographie

Afin de bien comprendre les enjeux des travaux de recherche décrits dans ce mémoire, mais aussi de mieux comprendre les travaux eux-mêmes, il est nécessaire de définir quelques notions de cartographie, et mais aussi de réaliser un rapide état de l'art du domaine. Nous allons tout d'abord définir les notions élémentaires autour des bases de données géographiques et des bases de données cartographiques, puis revenir sur les questions d'échelle et de niveau de détail. Nous présentons ensuite le problème de la généralisation cartographique, central dans ce mémoire, ainsi qu'un rapide état de l'art de la recherche dans le domaine. Enfin, nous décrivons ce que peut être une chaîne de production automatique de cartes puisqu'il s'agit de l'objectif à long terme qui guide nos recherches.

1.1.1 Bases de données géographiques vectorielles et bases de données cartographiques

Une base de données géographique (BDG) vectorielle est une base de données représentant des objets géographiques, spatialisés, dans laquelle une géométrie (souvent un point, une polygone ou un polygone) est attribuée à chaque objet. Les systèmes d'information géographique (SIG), au sens large, permettent de manipuler ces bases de données et d'en faire des cartes en appliquant des symboles cartographiques aux objets de la BDG (Figure 1-2). Mais, même si les symboles sont bien choisis, il n'est pas toujours possible de faire une carte lisible à une échelle donnée, surtout quand cette échelle est petite (Figure 1-2). Nous utilisons donc la notion de base de données cartographique (BDC) pour désigner une base de données géographique qui a été transformée pour être cartographiée de manière efficace à une échelle donnée, comme la BDC au 1 : 50000 de la Figure 1-2. Nous reprenons ici la définition d'une base de données cartographique issue de (A. Ruas 2004), qui s'appuie elle-même sur la distinction, en anglais et en allemand de Digital Landscape Model, équivalent de la BDG, et Digital Cartographic Model, équivalent de BDC (Grünreich 1985), (Brassel et Weibel 1988) :

"Une base de données cartographique regroupe des données géographiques dont la position de chaque objet est proche de la précision géométrique de saisie mais dont la taille et la position de chaque objet est conforme à l'échelle mathématique de représentation. Une base de données cartographique possède un espace de signification qui est constitué de concepts géographiques décrits dans la base. Ces concepts sont décrits par le schéma de données et les spécifications."

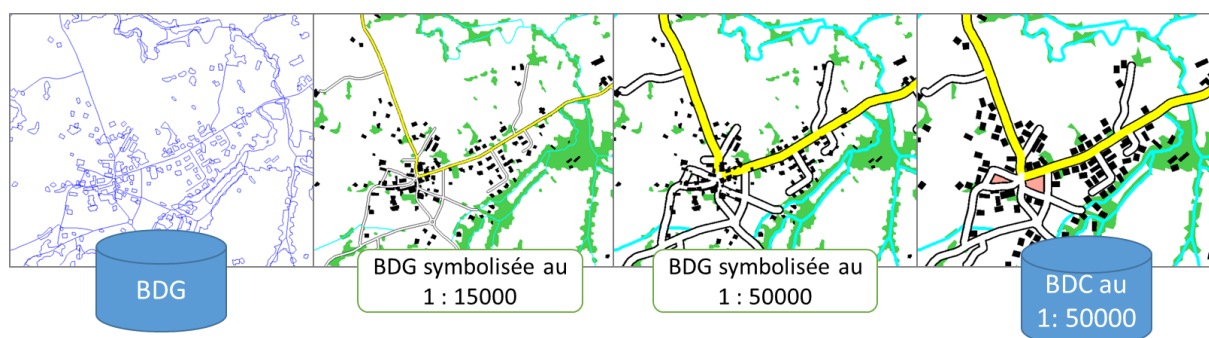


Figure 1-2 Exemple d'une base de données géographique qui peut être symbolisée à n'importe quelle échelle, mais toujours de manière lisible. La base de données cartographique est, elle, lisible à une échelle donnée si on lui applique des symboles adaptés à l'échelle.

On peut donc reformuler notre objectif global avec ces nouvelles notions : à partir d'une (ou plusieurs) base(s) de données géographiques, nous souhaitons obtenir le plus automatiquement possible une (ou plusieurs) base(s) de données cartographiques.

1.1.2 Notion d'échelle et de niveau de détail

L'Echelle d'une carte

Sur une carte imprimée sur papier, l'échelle est le simple rapport mathématique entre une distance sur la carte et la même distance sur le terrain. Cette échelle est donc une indication directe de la place dont le cartographe dispose pour représenter le terrain : au 1 : 25000, échelle historique des cartes topographiques de l'IGN, 1 mm correspond à 25 m sur le terrain. Au-delà de ces contraintes de dimension, l'échelle d'une carte transmet également des informations sur le contenu de la carte, issues des traditions cartographiques d'un pays, ou de l'échelle des phénomènes géographiques représentés (A. Ruas 2004), (W. A. Mackaness 2007) : par exemple, la carte au 1 : 25000 est une carte topographique utilisée pour la randonnée pédestre, et la carte au 1 : 250000 est une carte routière du fait de l'espace qu'elle couvre. L'échelle de la carte est donc une de ses caractéristiques fondamentales, et elle sera au centre de ce mémoire.

Dans un système d'information géographique (SIG) ou dans une application web de cartographie, il est possible de zoomer/dézoomer dans les cartes et l'échelle de la carte est plus difficile à définir. On distingue alors l'*échelle d'affichage* et l'*échelle de symbolisation* (Figure 1-3). L'*échelle d'affichage* est le rapport entre une distance à l'écran et la même distance sur le terrain. Cette échelle change donc à chaque fois que l'on zoome ou dézoome. L'*échelle de symbolisation* est l'échelle de définition des symboles, souvent définis en unité carte : par exemple, la largeur de certaines routes est définie de manière optimale à 0,7 mm sur la carte ; à une échelle donnée, par exemple le 1 : 100000, on peut définir le symbole en unité terrain (70 m dans le cas de notre symbole de route). Si on zoome ou dézoome dans une carte sans changer les symboles, l'échelle de symbolisation reste donc inchangée. Si l'on reprend le vocabulaire de la partie précédente, l'échelle de symbolisation peut être vue comme l'échelle de la base de données cartographique dont la carte est issue. Comme dans la Figure 1-3 extraite du géoportail français, les cartes peuvent avoir des échelles d'affichage et de symbolisation différentes.



Figure 1-3 Extrait du Géoportail de l'IGN : une carte dont l'échelle de symbolisation est le 1 : 100000, mais l'échelle d'affichage est le 1 : 34122.

Le Niveau de détail

S'il est possible de définir l'échelle d'une carte ou d'une BDC, les bases de données géographiques n'ont pas d'échelle. Toutefois, on attribue parfois une échelle à ces bases de données géographiques. Par exemple, on considère que la base de données géographique de l'IGN BD TOPO® a comme échelle le 1 : 15 000. On parle alors d'*échelle de référence* d'une BDG (Ruas 2004, p.29) pour « l'échelle la plus appropriée pour rendre sensible un phénomène en fonction de son étude ». De manière générale, on parle plutôt de *niveau de détail* pour une BDG, notion relativement floue qui englobe l'échelle de référence : il traduit le niveau de finesse dans le contenu de la BDG. Contrairement à l'échelle, le niveau de détail est complexe à définir et nous proposons de le faire en le décomposant en plusieurs propriétés, reprises de (A. Ruas 2004) :

- Le **schéma conceptuel**, c'est-à-dire la manière de représenter le terrain dans une base de données géographique. Par exemple, représenter une forêt par ses arbres individuellement est un schéma plus détaillé qu'une représentation de l'emprise de la forêt par un polygone.
- La **résolution géométrique**, c'est-à-dire la distance minimum entre deux points de la géométrie des objets, par analogie avec la résolution des images.
- La résolution **sémantique**, c'est-à-dire la quantité de détails dans les données attributaires associées à un objet géographique.
- La **précision géométrique**, c'est-à-dire la différence entre le positionnement dans la base de données et le positionnement sur le terrain.
- La **granularité**, c'est-à-dire la taille des plus petites formes d'un objet, ou même la plus petite taille des objets de la base de données.

Le niveau de détail est une notion également très utilisée dans le domaine de la modélisation en trois dimensions (voir le standard CityGML qui définit cinq niveaux de détails de LOD0 à LOD4 pour les modèles de ville en trois dimensions). Dans le cas des modèles de ville en trois dimensions, le niveau de détail peut être défini avec des composantes similaires (Biljecki, et al. 2014) : présence de certains objets (équivalent du schéma conceptuel), la complexité des objets (équivalent à la résolution géométrique), la dimensionnalité (qui se rapproche de la granularité), l'apparence ou résolution des textures, la cohérence spatio-sémantique et la résolution attributaire (équivalentes à la résolution sémantique). Pour parler de la granularité du contenu d'une base de données ou même d'une carte, certains parlent aussi de niveau d'abstraction ou de niveau de généralisation (Zhou 2014).

Lisibilité des cartes lors d'un changement d'échelle

Afin de définir des symboles lisibles et compréhensibles à une échelle de carte (ou de symbolisation) donnée, les cartographes se sont intéressés aux limites de la perception humaine (Cuenin 1972) (Salichtchev 1983). Cela a permis la définition de seuils (Figure 1-4) qui sont encore aujourd'hui très utilisés pour la définition des symboles d'une carte imprimée sur papier.

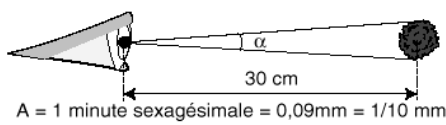
CONTRAINTES VISUELLES	REPRÉSENTATION GRAPHIQUE
Acuité visuelle de discrémiation	 $A = 1 \text{ minute sexagésimale} = 0,09 \text{ mm} = 1/10 \text{ mm}$
Acuité visuelle d'alignement	 0,02mm
Seuil de perception	Ponctuel ● 0,2 mm ■ 0,4 mm ○ 0,3 mm □ 0,5 mm △ 1 mm ~ 0,5 mm ~ 0,6 mm Linéaire ——— 0,1 mm
Seuil de séparation	Linéaire > 2/10 > 3/10 Ponctuel ■ ■ > 2/10
Seuil de différenciation	Ponctuel ● ● entre 2 paliers le rapport des surfaces doit être au moins de 2 Linéaire ——— Traits rapprochés écart d'épaisseur 0,1mm ——— Traits éloignés 0,3 mm minimum

Figure 1-4 Exemples de seuils de perception dans une carte, adaptés de (Cuenin 1972).

Maintenant que les cartes sont majoritairement visualisées sur des écrans de taille et résolution variable, ces seuils ne sont plus vraiment adaptés. Les travaux en visualisation d'information ont abordé ce problème à travers des seuils d'acuité visuelle sur un écran (C. Ware 2012) : suivant la résolution et la taille de l'écran, ainsi que la position du lecteur de la carte par rapport à l'écran, on peut déterminer si son acuité visuelle est aussi ou moins bonne que pour une carte imprimée sur papier : elle est généralement moins bonne avec les écrans utilisés aujourd'hui.

Lorsque l'on dispose d'une BDG, on peut utiliser ces seuils de lisibilité pour construire une BDC dont l'échelle de symbolisation est l'échelle de référence de la BDG (le 1 : 15 000 dans la Figure 1-5). Mais lorsque l'échelle d'affichage diminue, comme au 1 : 50 000 dans la Figure 1-5, la carte n'est plus lisible, les différents symboles se retrouvant inférieurs aux seuils de lisibilité. La première chose à faire est d'adopter l'échelle d'affichage comme nouvelle échelle de symbolisation ce qui améliore clairement la lisibilité de la carte comme on peut le voir sur la Figure 1-5. Mais si on y regarde de plus près, de nombreux problèmes de lisibilité persistent. Certains objets surfaciques ne sont représentés que par une couleur de remplissage, donc le changement d'échelle de symbolisation ne les fait pas grossir (cas des bâtiments et des zones arborées dans la Figure 1-5). Ils sont donc toujours trop petits pour être vraiment lisibles. De plus, en grossissant les symboles des routes et des cours d'eau de manière proportionnelle à l'échelle, on génère des superpositions avec les objets voisins et parfois même des occlusions suivant l'ordre d'affichage des symboles. Il faut donc transformer les données de la carte, au-delà de la taille de leur symbole, pour la rendre la plus lisible possible : c'est ce que

l'on appelle la généralisation cartographique, dont les principes sont rapidement décrits dans la partie suivante.

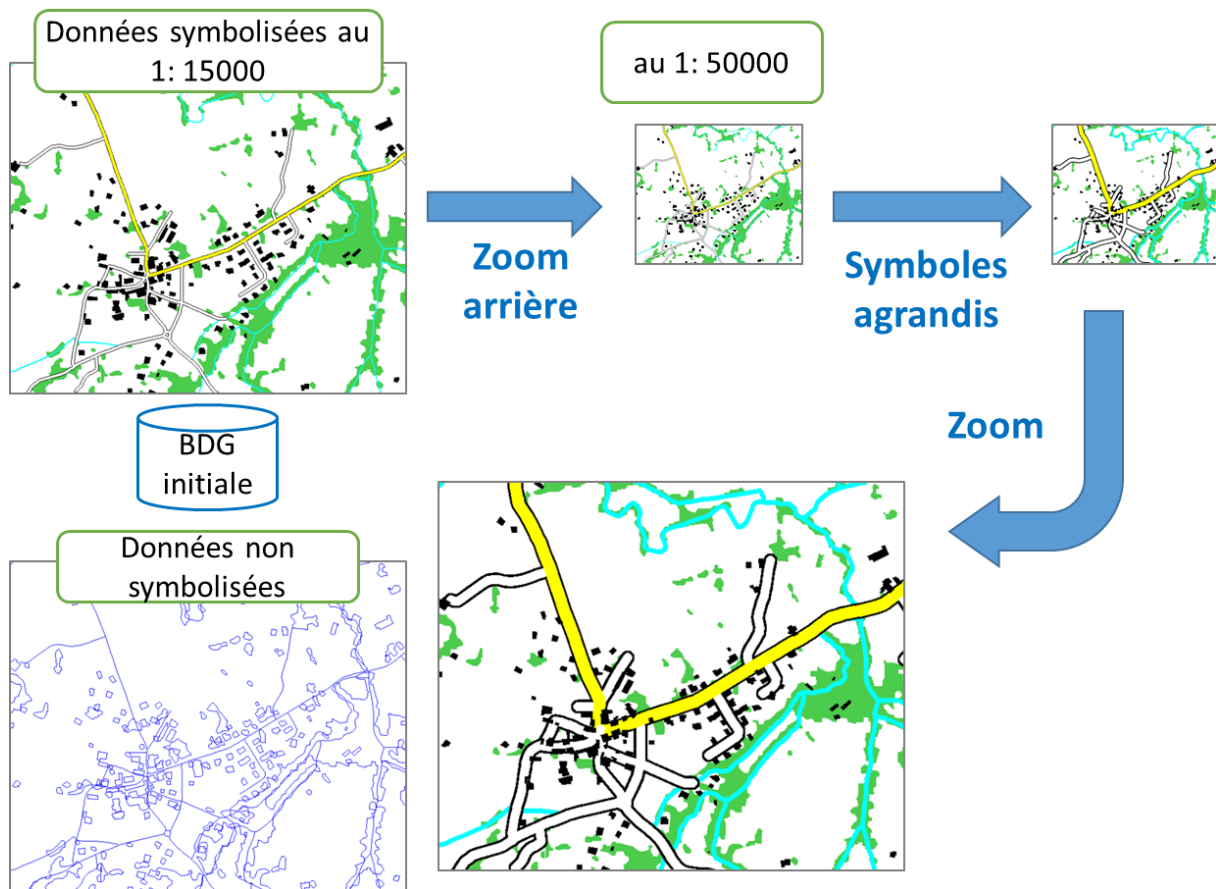


Figure 1-5 Problèmes de lisibilité causés par la visualisation des données géographiques à une échelle plus petite que l'échelle de symbolisation et l'échelle de référence.

1.1.3 Généralisation cartographique et son automatisation

Définitions

Nous avons donc vu que la constitution d'une base de données cartographique à une échelle donnée à partir d'une base de données géographique dont l'échelle de référence est plus grande pose un certain nombre de problèmes de lisibilité. Il faut donc transformer les données pour rendre la BDC lisible tout en préservant au mieux les caractéristiques des données géographiques utilisées. C'est cette opération d'abstraction et de transformation que nous appelons la généralisation cartographique (Figure 1-6).

Pour compléter cette définition minimale, nous reprenons celle, plus complète, donnée par Anne Ruas (2004) :

"La généralisation est un processus de synthèse d'information. Ce processus peut être comparé avec celui d'un résumé de texte dont les objectifs sont de réduire le nombre de mots, de retenir les idées principales, de ne pas faire de faux sens et si possible conserver le style de l'auteur (...). Pour bien généraliser des données géographiques, il faut également :

- *réduire la quantité d'information,*
- *mettre en valeur l'information la plus importante,*

- rester fidèle à l'information initiale,
et dans le cas d'une carte, respecter les règles de sémiologie qui permettent une bonne lecture de l'information."

La généralisation n'est donc clairement pas seulement une opération de simplification graphique pour plus de lisibilité, dans laquelle on manipule des géométries. Il s'agit surtout d'abstraire la géographie implicite de la carte, et les relations spatiales porteuses de sens dans la carte.

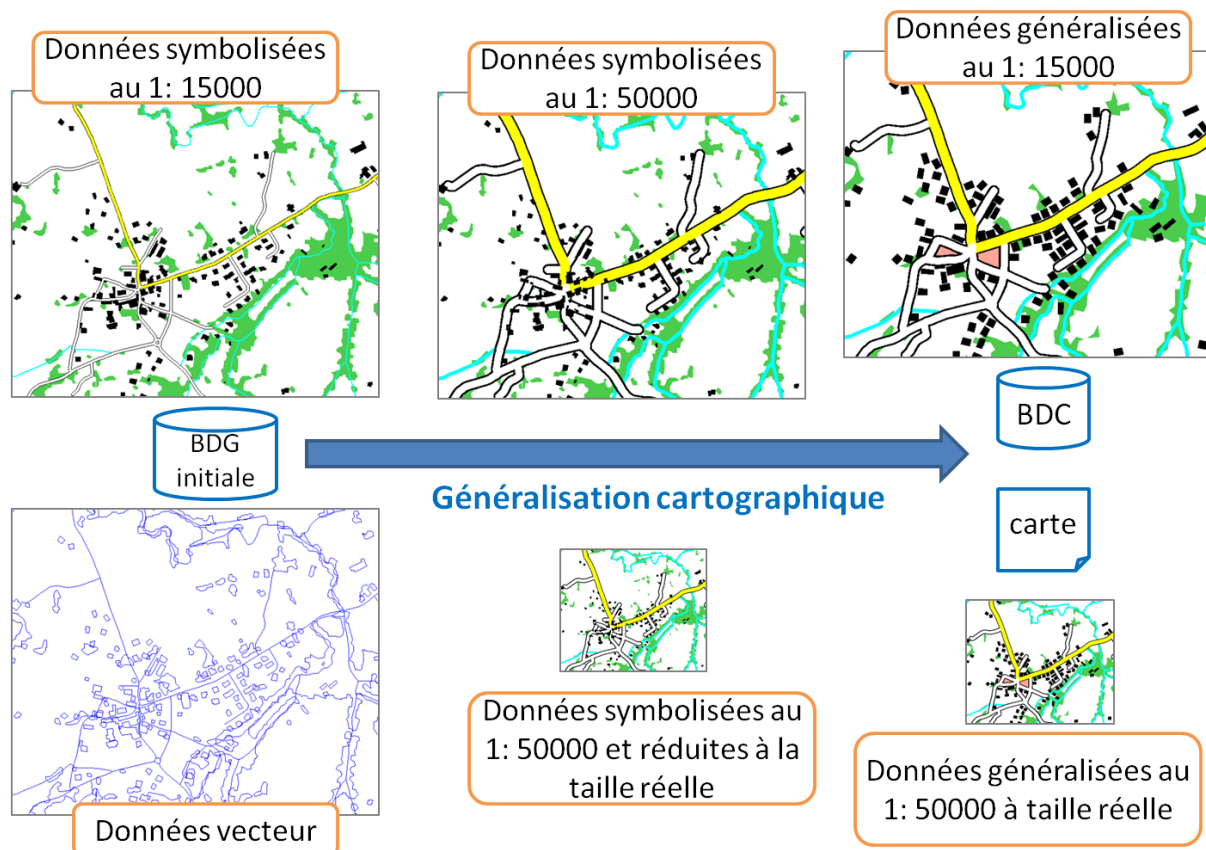


Figure 1-6 Définition illustrée de la généralisation cartographique : représenter les données d'une BDG à une échelle plus petite que l'échelle de référence engendre des problèmes de lisibilité et d'interprétation. La généralisation permet de rendre les données lisibles, même à taille réelle (en bas, à droite).

La généralisation peut être de différentes natures et les cartographes ont défini différents types de généralisation. Ainsi, Ruas (2004) distingue : la *généralisation cartographique*, dont le but est de construire une BDC ; et la *généralisation de base de données*, dont l'objectif est de dériver une autre BDG moins détaillée. Certains distinguent *généralisation cartographique* et *généralisation de modèle* (Weibel et Dutton 1999). Cette dernière recouvre les opérations de généralisation agissant sur le schéma de données, comme par exemple la transformation de cours d'eau surfaciques en cours d'eau linéaire. Enfin, Harrie et Sarjakoski (2002) proposent de distinguer la *généralisation graphique* qui consiste principalement à déplacer et déformer les objets de la carte pour des raisons de lisibilité, après différentes étapes de généralisation dans lesquelles les opérations peuvent être plus drastiques (élimination ou changement de primitive géométrique par exemple). Ces différents concepts ne sont pas synonymes mais proviennent de différents points de vue sur la généralisation, en fonction des objectifs (une carte papier ? une base de données ?) et de la manière de formaliser le processus. Dans les différentes variantes de la conception cartographique automatisée présentées dans la suite du mémoire, nous ferons appel à certains de ces concepts, car, par exemple, la généralisation de modèle seule peut être nécessaire dans un contexte donné.

Opérateurs et algorithmes de généralisation

Les opérateurs sont les briques de base d'un processus de généralisation, c'est-à-dire les transformations unitaires appliquées aux objets de la carte pour la rendre lisible suite au changement d'échelle : grossissement, simplification, déplacement, élimination, typification... Il existe un certain nombre de typologies de ces opérateurs, qui ne sont pas forcément concordantes (Mustière 2001) (Regnauld et McMaster 2007) (Foerster, Stoter et Köbben 2007) (Roth, Brewer et Stryker 2011). Ces opérateurs sont implémentés par divers algorithmes qui sont plus ou moins adaptés à l'ensemble des entités d'une carte ou des sauts d'échelle possibles. Par exemple, la simplification des lignes est un sujet largement traité car une grande partie des entités d'une carte possède une primitive géométrique linéaire, et de nombreux algorithmes existent : (Douglas et Peucker 1973) est plutôt adapté aux généralisations à but non cartographique, (Visvalingam et Wyatt 1993) est plutôt adapté aux lignes naturelles comme les cours d'eau ou le trait de côte car il préserve bien les grandes courbes, alors que (Lecordix, Plazanet et Lagrange 1997) est dédié à la généralisation des routes de montagne.

Spécifications de la carte et contraintes

Nous avons vu que les limites de perception des utilisateurs d'une carte entraînent des spécifications minimales pour la taille des symboles représentés sur la carte. Mais il existe d'autres types de spécifications qui complètent ou contrebalancent celles issues de la perception visuelle : les spécifications dérivées de l'objectif de la carte et les spécifications pour une bonne abstraction des objets géographiques sur la carte. Dans le premier cas, une carte touristique va nécessiter une emphase particulière sur les symboles de sites touristiques par exemple ; dans le deuxième cas, les spécifications peuvent réclamer la préservation de la forme générale des objets géographiques lors des opérations de simplification, pour qu'ils soient toujours reconnaissables.

Sous l'impulsion des travaux de Kate Beard (1991), ces spécifications sont maintenant communément décrites sous la forme de contraintes, contrebalançant la lisibilité de la carte (*les symboles des bâtiments doivent être séparés par au minimum 0.1 mm sur la carte*) et la préservation des caractéristiques géographique de la carte (*les bâtiments alignés le long d'une route doivent le rester*). Des travaux ont porté ces dernières années sur la formalisation et la standardisation de ces contraintes (Werder 2009) (Zhang, Stoter et Ai 2008) (Touya, Duchêne et Ruas 2010), et ce sujet sera abordé plus loin dans ce mémoire (partie 2.3).

Un processus de généralisation automatique

Un processus de généralisation automatique est un composant logiciel capable d'enchaîner l'application d'opérations de généralisation sur les différents objets de la carte en répondant aux questions *pourquoi* (quelle partie de la carte est mal généralisée ?), *où* (sur quel objet ?), *quand* (dans quel ordre ?), et *comment* (avec quel opération ?) généraliser (Mcmaster et Shea 1988) (Shea et McMaster 1989).

Pour concevoir un système capable de répondre à ces questions, il faut s'attaquer à différents problèmes bien connus des cartographes :

- *Importance, saillance et redondance* : la quantité d'information affichée sur la carte étant réduite, il faut y afficher les informations les plus importantes, les plus saillantes, et éviter les redondances. Mais il n'est pas toujours évident d'identifier les informations importantes, comme dans le cas de routes avec une classification similaire (Touya 2010).
- *Contexte géographique* : le choix d'une opération et de ses paramètres dépend très souvent du contexte géographique de l'objet généralisé. Par exemple, on va sélectionner les impasses

d'un réseau routier en fonction de leur longueur, sauf si elles mènent à un point d'intérêt important de la carte (Figure 1-7).

- *Espace d'analyse* : les opérations peuvent s'appliquer sur un objet, une partie de la géométrie d'un objet, plusieurs objets ou sur la relation spatiale entre deux objets, et ce choix de l'espace d'analyse s'applique aussi au contexte géographique. Par exemple, il faut découper les traits de côtes en parties d'objet homogènes en sinuosité pour bien les généraliser (Nakos, Gaffuri et Mustière 2008).

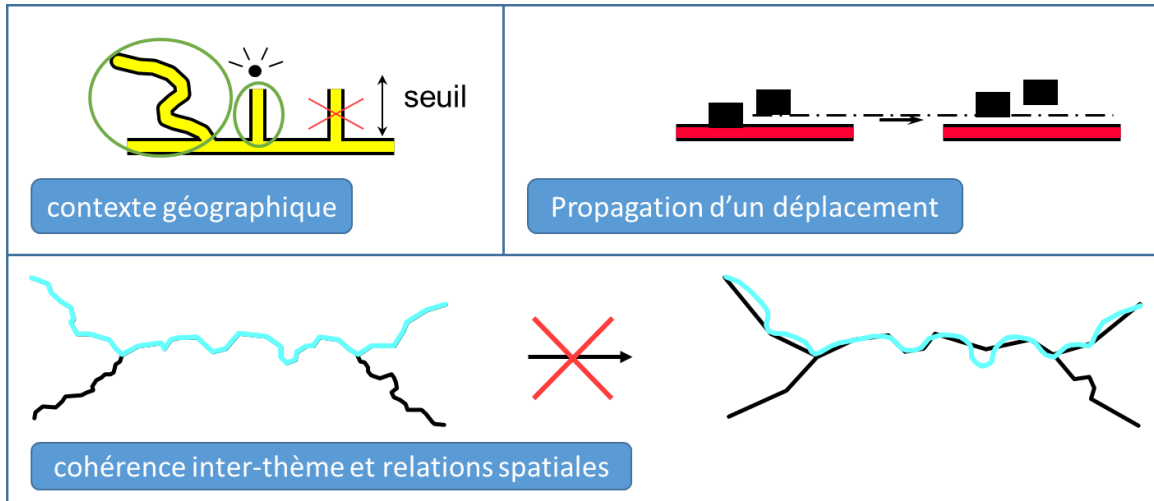


Figure 1-7 Illustration de certains enjeux de la construction d'un processus automatique de généralisation : la prise en compte du contexte géographique, la propagation des transformations et la cohérence inter-thème (source : Cécile Duchêne).

- *Choix des paramètres* : l'expérience montre que les opérations de généralisation sont souvent réglables à l'aide de plusieurs paramètres, qu'il est souvent difficile de régler pour une échelle et une carte donnée. Par exemple, les algorithmes demandent souvent de combien généraliser (grossir deux fois un polygone) alors que les contraintes donnent une information du type jusqu'où généraliser (grossir le polygone jusqu'à 600 m²).
- *Répercussion des opérations* : la plupart des opérations de généralisation déforment ou déplacent les objets de la carte ce qui peut entraîner des collisions avec les objets voisins non pris en compte par l'opération (Figure 1-7). Des opérations de propagation, avec amortissement pour éviter l'effet domino, sont donc souvent indispensables.
- *Cohérence et relations spatiales* : une carte est un ensemble de données géographique en relation spatiale implicite, ce qui peut se traduire par exemple par le partage d'une frontière entre une limite administrative et un cours d'eau (Figure 1-7). La généralisation doit donc éviter d'endommager ces relations en transformant les données sans tenir compte de la relation spatiale.
- *Evaluation* : qu'ils soient entièrement ou semi-automatiques, les processus de généralisation automatique doivent s'appuyer sur des outils d'auto-évaluation pour identifier les zones à généraliser avant le processus, et les zones mal généralisées (Figure 1-8), pendant ou après le processus.

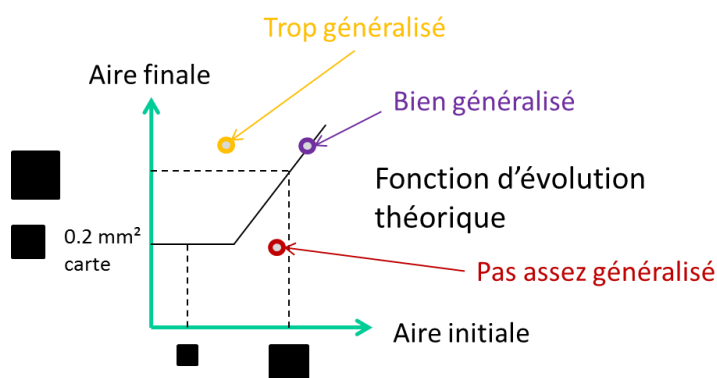


Figure 1-8 Evaluation automatique de la satisfaction d'une contrainte sur l'aire minimum d'un bâtiment sur la carte.

Etat de l'art des processus de généralisation automatique

De nombreux processus de généralisation automatique, plus ou moins génériques sur le type de données et de saut d'échelle traités, ont été proposés ces vingt-cinq dernières années. Les premiers ont suivi une approche par workflow enchainant l'application d'algorithmes. Les travaux suivants illustrent différentes manières d'approcher ce principe par workflow (Müller 1990), (Revell 2007), (Palomar-Vazquez et Pardo-Pascual 2008), (Burghardt, Petzold et Bobzien 2010), (Touya 2010), (Gould et Chaudhry 2012).

D'autres approches ont essayé de résoudre les problèmes de généralisation de manière globale, en cherchant une transformation optimale des données. Parmi ces approches, certains ont modélisé la généralisation comme un problème physique, comme par exemple des forces à équilibrer (Baeijs, Demazeau et Alvares 1996), (Hojholt 2000), (Bader, Barrault et Weibel 2005), (Joubran Abu Daoud et Doytsher 2008). D'autres ont adopté une approche par optimisation (Ware et Jones 1998), en utilisant des techniques issues d'autres domaines comme l'optimisation par moindres carrés (L. E. Harrie 1999) (Sester 2005) ou la programmation linéaire en nombres entiers (Haunert et Wolff 2010).

D'autres encore ont adopté une approche stochastique du problème, en essayant de manière itérative des algorithmes, avec plus ou moins de hasard dans le choix de l'espace de travail et de l'algorithme appliquée (Ware, Jones et Thomas 2003), (Ware, Wilson et Ware 2003), (Burghardt et Neun 2006), (Richards et Ware 2010), (M. J. Ware, I. D. Wilson, et al. 2002).

Enfin, d'autres ont adopté une approche que l'on peut qualifier de « pas à pas locale », dans laquelle s'inscrivent la plupart des travaux de l'équipe COGIT, à la suite des travaux d'Anne Ruas (Ruas et Plazanet 1996), (A. Ruas 1999). De manière itérative, on choisit l'espace de travail le plus prioritaire à traiter, et dans chaque espace de travail, on applique de manière itérative des algorithmes suivis d'une évaluation validant ou non une amélioration locale et/ou globale de la lisibilité de la carte. On peut noter qu'une grande partie des travaux suivant cette approche utilisent aussi les méthodologies issues de la recherche en système multi-agents, dont on reparlera dans le chapitre 3 de ce mémoire (Barrault, et al. 2001), (Galanda et Weibel 2002), (Duchêne 2004), (Jabeur 2006), (Gaffuri 2008), (Sabo, et al. 2008), (Zhang et Guilbert 2011), (Duchêne, Ruas et Cambier 2012), (Yan, Guilbert et Saux 2016).

Aucun de ces processus n'est toutefois assez générique pour généraliser l'ensemble d'une carte, quel que soit le paysage traité ou l'échelle finale de la carte, mais ils sont plus ou moins efficaces sur des sous-problèmes de ce problème général. C'est ainsi que dans le cadre de ma thèse, j'ai proposé un modèle, CollaGen, permettant d'interfacer plusieurs de ces processus indifféremment, pour les

enchainer et généraliser une carte dans son ensemble (Touya, Duchêne et Ruas 2010), (Touya et Duchêne 2011), (Touya 2011).

1.1.4 Autres processus de cartographie automatique

Même si c'est celui qui nous intéressera le plus dans ce mémoire, la généralisation cartographique n'est pas le seul processus nécessaire à la fabrication automatique d'une carte. Les toponymes ou de manière plus générale, les éléments de texte présents sur une carte sont des éléments fondamentaux de la compréhension de la carte car toutes les informations ne peuvent pas être abstraites par un symbole (notamment le nom des entités géographiques). Le placement de ce texte sur la carte est donc une étape clé de sa conception, et le placement doit être fait en optimisant la lisibilité du texte (à la bonne taille, bien séparés les uns des autres) et en minimisant la quantité d'information de la carte « cachée » par le texte (Imhof 1975). Tout comme la généralisation, le placement automatique de textes a très vite intéressé les chercheurs en cartographie et il existe de nombreux travaux sur le placement des toponymes droits, c'est-à-dire les textes écrits horizontalement pour décrire un symbole ponctuel comme le nom des villes (Christensen, Marks et Shieber 1995), (M. Barrault 2001) (Ribeiro et Lorena 2006) (Rylov et Reimer 2014). Il existe également des travaux, moins nombreux sur l'optimisation du placement ou de l'écriture des toponymes courbes (par exemple le nom des cours d'eau) ou des toponymes surfaciques (par exemple le nom des forêts écrit à l'intérieur, et au milieu de la forêt) (M. Barrault 2001) (Haunert, et al. 2014). D'autres travaux ont porté sur d'autres points clés de l'automatisation de ce processus, qui sont communs à la généralisation : l'évaluation d'un placement automatique (van Dijk, et al. 2002) (Rylov et Reimer 2015), ou la gestion de cartes multi-échelles (Schwartzges, et al. 2014) pour qu'un toponyme ne bouge pas trop lors d'un zoom. Enfin, les travaux en placement automatique de texte ont atteint un certain niveau de maturité ce qui a donné naissance à plusieurs logiciels de placement automatique (Lecordix, Plazanet et Chirié, et al. 1994) (Freeman 2005) (Ertz, et al. 2012).

La symbolisation ou stylisation des données est une autre étape à automatiser. Au-delà de la simple application de symboles cartographiques aux données préalablement généralisées, l'étape de symbolisation peut comporter des opérations de symbolisation avancée comme le décalage de symboles linéaires par rapport à la géométrie de l'objet, comme les représentations d'itinéraires à la manière des cartes de bus (Teulade-Denantes, Maudet et Duchêne 2016). Mais il peut également s'agir de techniques de rendu avancées (Figure 1-9) permettant de générer des textures à l'intérieur de forêts (Revell 2007), et des textures artistiques pour représenter le relief (Loi, et al. 2013) (Christophe, Dumenieu, et al. 2016), ou permettant d'imiter le style d'une carte ancienne (Christophe, Dumenieu, et al. 2016) ou d'un peintre (Christophe 2011), (Christophe et Hoarau 2013).

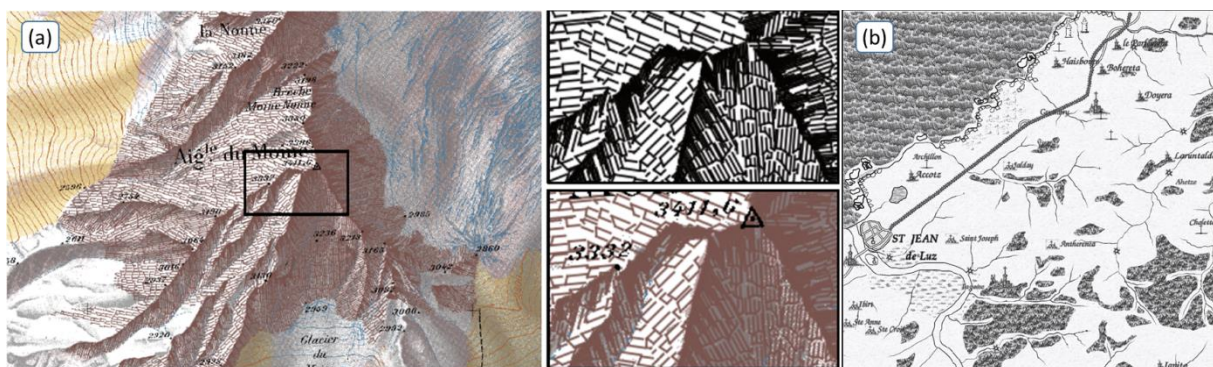


Figure 1-9 Rendus cartographiques avancés issus des travaux de (Christophe, Dumenieu, et al. 2016). (a) dessin des rochers montagneux par génération de texture (b) imitation du style cartographique des cartes de Cassini.

En plus des deux processus précédents qui sont nécessaires à la plupart des cartes, une étape de schématisation peut être nécessaire pour certaines cartes. Mackaness et Reimer (2014) définissent la schématisation de la manière suivante :

“a process that uses cartographic generalisation operators in such a way as to produce maps of a lower graphical complexity compared to maps of the same scale; the process aims to maximize task-adequacy while minimizing non-functional detail.”

Il s’agit donc d’abstraire les objets géographiques pour représenter plus leur fonction que leur forme exacte. Les plans de métro sont des exemples de cartes très schématisées (Figure 1-10), et de nombreux travaux cherchent à reproduire ce type de schématisation (Ware, Taylor, et al. 2006), (Richards et Ware 2010), (Nöllenburg et Wolff 2011) à partir de données géographiques classiques. Mais il existe d’autres types de cartes partiellement ou totalement schématisées :

- la schématisation par courbes ou rectilinéaire (Meulemans, Renssen et Speckmann 2010), (van Goethem, et al. 2013),
- les cartes avec zone de focus (van Dijk, et al. 2013),
- les cartes d’itinéraires à la manière des plans de bus (Teulade-Denantes, Maudet et Duchêne 2016),
- les cartes mentales ou schémas (Agrawala et Stolte 2001),
- les chorèmes (Brunet 1980), (Reimer et Meulemans 2011), (Mackaness et Reimer 2014)



Figure 1-10 Extrait d’une carte schématique du métro parisien réalisée manuellement ©2009 Nathan Kaufmann.

1.1.5 Une Chaîne de production automatique de cartes

L’objectif général qui guide l’ensemble des travaux présentés dans ce mémoire est la conception d’un système capable de produire automatiquement des cartes de grande qualité cartographique, c’est-à-dire des cartes dont la qualité est la plus proche possible de celle des cartographes travaillant « à la main ». Dans un premier temps, nous considérons le scénario traditionnel de production de cartes en série dans une agence de cartographie telle que l’IGN (Figure 1-11). Nous ne nous intéressons qu’aux étapes de création d’une base de données cartographique, et pas aux étapes suivantes d’édition cartographique (quadrillage, positionnement de la légende...) ou d’impression, qui sont aussi réalisées à l’IGN et peuvent prendre du temps.

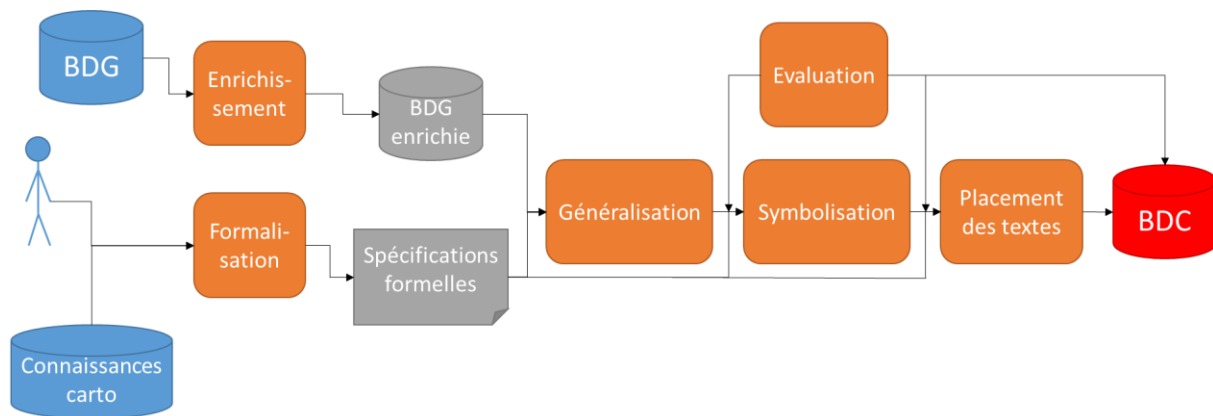


Figure 1-11 Diagramme d'un processus simplifié de cartographie automatique : obtenir une BDC à partir d'une BDG, un besoin et des connaissances cartographiques.

Nous souhaitons donc obtenir une BDC à partir : d'une BDG, de connaissances cartographiques comprenant les spécifications de la BDC, le tout piloté par un cartographe. Les premières étapes sont *l'enrichissement* de la BDG et la *formalisation* des spécifications. L'enrichissement de la BDG est une étape indispensable qui permet d'extraire et d'explicitier la géographie implicite d'une BDG pour guider les processus de cartographie automatique (Mackaness et Edwards 2002). Cela revient par exemple à identifier les bâtiments alignés le long d'une route pour mieux préserver cet alignement, ou à construire un graphe à partir des tronçons de lignes représentant les routes. La formalisation des spécifications consiste à traduire les spécifications de la carte (seuils de lisibilité, choix des thèmes présents sur la carte...) sous la forme de contraintes interprétables par les processus automatiques. De nombreux travaux en cartographie automatique ont montré que ce n'était pas une tâche aisée, même pour un cartographe chevronné (Stoter, et al. 2010). Il faut ensuite enchaîner des étapes de *généralisation*, de *symbolisation* et de *placement des textes*, éventuellement entrecoupées d'étapes d'*auto-évaluation* permettant, ou non, des retours en arrière. Chacune de ces étapes est complexe à automatiser, et pour chacune, des recherches sont encore menées pour atteindre cette automatisation, ou pour s'en rapprocher le plus possible pour minimiser les interventions correctrices de cartographes.

Nous avons ici volontairement simplifié le processus de symbolisation, car il est moins central dans nos travaux, mais les travaux de Sidonie Christophe montrent que les interdépendances avec les processus de généralisation sont plus complexes (Christophe, Dumenieu, et al. 2016).

Il existe également des travaux très intéressants qui proposent d'intégrer les opérations de généralisation et de symbolisation pour mieux gérer leur interdépendance (Brewer et Buttenfield 2007), (Roth, Brewer et Stryker 2011), (Samsonov, Podolsky et Yurova 2013), mais nous en parlerons essentiellement en perspective.

L'automatisation de la création d'une base de données cartographique de qualité est une question de recherche complexe, qui nécessite de s'attaquer à plusieurs sous-problèmes dont l'automatisation de la généralisation cartographique, au cœur de ce mémoire.

1.2 Cartes en série et carte à la carte

Nous adoptons dans ce mémoire, la plupart du temps, le point de vue de l'IGN (Institut national de l'information géographique et forestière), qui est le producteur de référence de la cartographie en France, avec pour mission de service public, entre autres, de réaliser une cartographie topographique de l'ensemble du territoire français. L'IGN produit deux types de cartes : les cartes en série qui

couvrent l'ensemble du territoire à différentes échelles (du 1 : 25000 au 1 : 1000000), et les cartes à la demande, produites à partir des données de l'IGN pour un besoin ou un utilisateur précis. Nous allons voir dans cette partie que les deux types de production cartographiques nécessitent une meilleure automatisation des processus de production, et notre recherche s'est inscrite et s'inscrit toujours dans cet objectif. Toutefois, et nous y reviendrons plus loin dans le mémoire, la portée de ces recherches ne se limite pas aux productions de l'IGN, ni même à celles d'autres agences publiques de cartographie.

1.2.1 Production de cartes en série et mise à jour

Au-delà des problèmes inhérents à l'ensemble des processus de cartographie automatique, la production de cartes en série à l'échelle d'un pays entraîne un certain nombre de problèmes spécifiques, dont on peut trouver diverses illustrations dans l'article de Duchêne et al. (2014) qui regroupe les expériences de plusieurs agences de cartographie (NMA pour *National Mapping Agency*) face à l'automatisation de leurs processus de cartographie.

Comme il n'est pas encore possible de mettre en place une chaîne de production de cartes de grande qualité tout automatique, les NMA doivent choisir où placer le curseur entre une chaîne totalement automatique, mais de qualité moindre que les cartes précédemment produites et une chaîne où l'effort est mis sur l'identification de ce qui n'a pas été bien généralisé, pour accélérer le travail des cartographes corrigeant les défauts en sortie de processus. Kadaster NL (NMA néerlandais) a, par exemple, fait le choix du tout automatique, en modifiant ses spécifications pour qu'elles correspondent aux capacités du processus de cartographie automatique, basé sur des solutions logicielles commerciales optimisées avec l'aide des cartographes locaux. A l'opposé, SwissTopo a choisi un processus plus long qui repose sur une correction minutieuse du processus de généralisation automatique par des cartographes.

La production de cartes en série à l'échelle d'un pays nécessite la gestion de grands lots de données, ce qui n'a rien d'évident car la généralisation ou le placement de textes sont des opérations très longues, et contextuelles (l'analyse du voisinage des objets est indispensable pour choisir les bonnes opérations). Donc des techniques simples de partitionnement et de parallélisation ne sont pas optimales. Par exemple, l'USGS, NMA américain réalise un partitionnement géographique de son réseau hydrographique, qui utilise les bassins versants.

Le dernier problème important rencontré dans la production de cartes en série est celui de la mise à jour. Deux stratégies s'opposent : la mise à jour par une nouvelle généralisation complète ou la mise à jour incrémentale. La première stratégie, choisie par exemple par Kadaster NL, a pour avantage la simplicité : seule la BDG initiale est mise à jour et le processus de cartographie automatique est déclenché de nouveau sur les données à jour à chaque pas de temps (tous les ans par exemple). La deuxième stratégie, choisie par exemple par l'IGN Belgique et le GST Danemark, consiste à propager la mise à jour de la BDG dans la BDC, qui est seulement modifiée « autour » de cette mise à jour : la carte est mise à jour instantanément, mais cette propagation n'est pas encore maîtrisée (comment définit-on le « autour » ?) et reste un problème de recherche.

1.2.2 La Carte à la carte ou cartographie à la demande

La carte à la carte, ou à la demande est un terme générique qui recouvre plusieurs réalités, ou plusieurs services de cartographie personnalisée. Sa version la plus simple correspond au service de l'IGN appelé « Carte à la carte » qui permet de commander une carte dont l'emprise et la couverture peuvent être personnalisées, avec également un choix entre deux échelles de fond. Un tel service ne nécessite pas d'adaptation de notre processus de cartographie automatique. Mais il existe des

formes plus avancées de cartographie à la demande dans lesquelles l'échelle, l'objectif de la carte (et donc son contenu), le style, le niveau de schématisation peuvent être choisis par l'utilisateur. Il existe même un dernier niveau de cartographie à la demande avec la création de « mashups » cartographiques mélangeant des données thématiques fournies par un utilisateur à un fond cartographique, comme c'est le cas dans la Figure 1-12 qui représente des itinéraires de randonnée fournis par un utilisateur du service « Espace Loisirs » de l'IGN, par-dessus la carte topographique au 1 : 25 000.

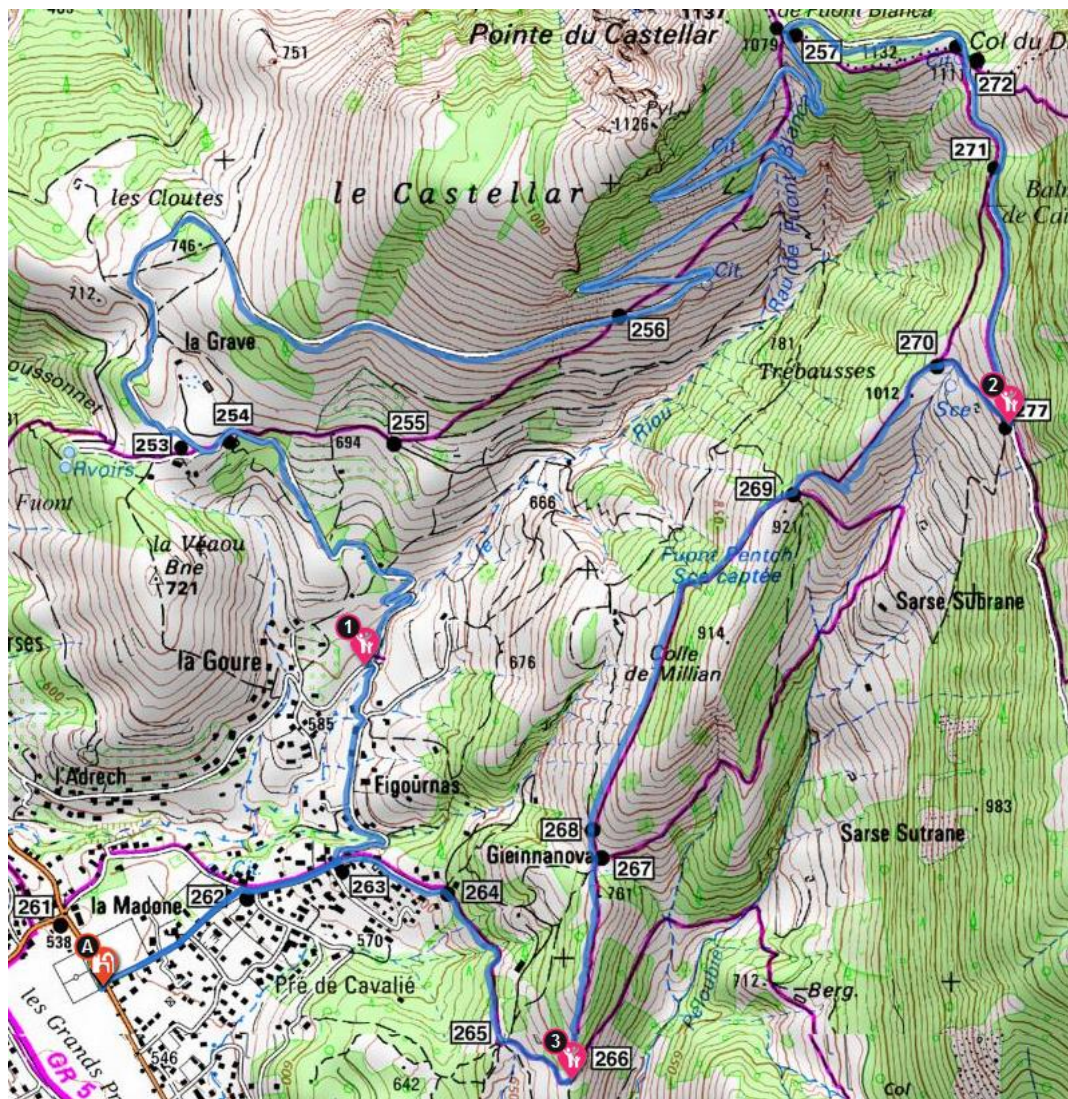


Figure 1-12 Capture d'écran issue de l'Espace Loisirs de l'IGN², site de crowdsourcing de l'IGN autour des activités de loisir en plein air : le tracé de la randonnée (issu de traces GPS) ne suit pas exactement le tracé du sentier sur la carte.

Permettre un tel niveau de personnalisation à des utilisateurs, qui ne sont pas nécessairement des experts de la cartographie, nous oblige à concevoir des méthodes de cartographie automatiques (il n'est pas possible de prendre le temps de corrections manuelles), génériques (les données en entrée peuvent varier) et adaptatives. Par exemple, on ne peut plus se contenter de processus capables de généraliser uniquement à des échelles fixes comme le 1 : 25 000 et ou le 1 : 100 000.

² <http://espaceloisirs.ign.fr>

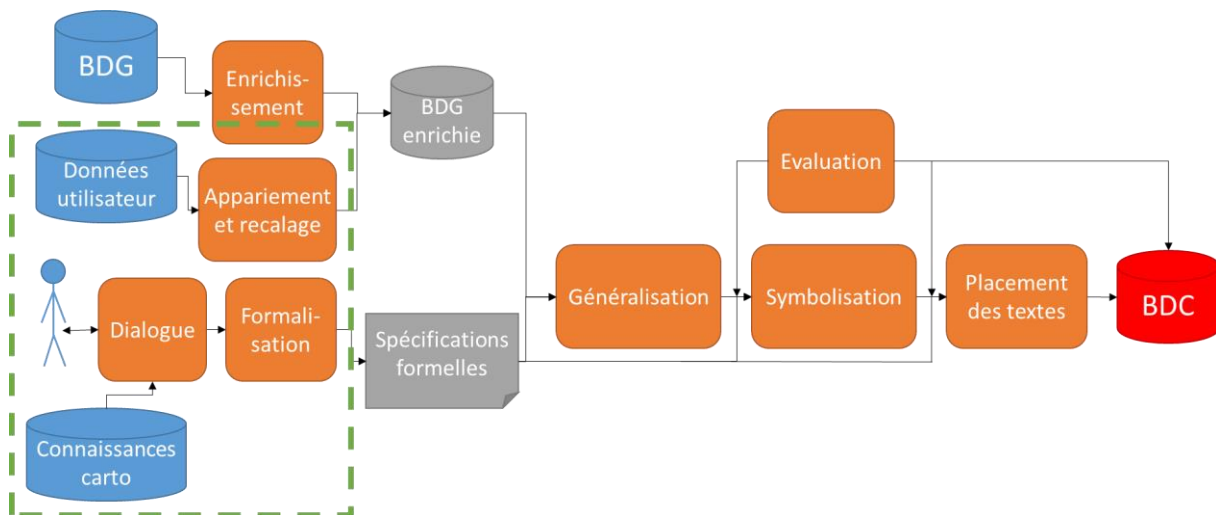


Figure 1-13 Diagramme d'un processus simplifié de cartographie automatique à la demande : il faut gérer des données utilisateur et un dialogue avec l'utilisateur pour mieux comprendre son besoin.

La Figure 1-13 montre les principales différences apportées au processus synthétique de cartographie automatique : l'ajout en entrée de données fournies par l'utilisateur, en plus de la BDG, et le fait que l'utilisateur ne soit plus un cartographe. L'ajout de données fournies par l'utilisateur a pour conséquence l'ajout d'une étape d'appariement et de recalage (voir partie 2.1) qui permet de superposer ces données au fond de carte. Même si cela n'apparaît pas dans le diagramme, les étapes suivantes sont aussi modifiées pour tenir compte de ces données reposant sur le fond de carte (Duchêne 2014). Le fait que l'utilisateur ne soit peut-être pas un spécialiste de la cartographie, l'empêche de participer à la formalisation des spécifications. Il faut dans ce cas se reposer sur des techniques d'interaction avec l'utilisateur, comme le dialogue, pour extraire et formaliser ses besoins (Balley, Baella, et al. 2014). Pour une modélisation plus précise d'un processus automatique ou semi-automatique de cartographie à la demande, voir les travaux de Balley et al. (2014).

1.3 Carte papier et application web de cartographie

Il est évident que la cartographie traditionnelle sur papier est en grande partie remplacée par les cartes sur le web et autres applications de géo-visualisation. Toutefois, une grande partie des portails cartographiques sur le web sont composés de cartes conçues pour le papier mais scannées donc il est toujours pertinent de s'intéresser à la constitution de ces cartes. Mais ce changement entraîne tout de même des différences avec la carte imprimée sur papier qui peuvent être de deux ordres. D'abord, le fait de regarder les cartes sur des écrans change sensiblement les règles traditionnelles de visibilité des symboles cartographiques, notamment quand la taille des écrans se réduit. Ensuite, la plupart des portails cartographiques permettent de zoomer/dézoomer dans les cartes, ce qui entraîne aussi des changements dans la manière de construire les cartes.

1.3.1 La Carte sur un écran

Nous avons vu dans la partie 1.1.2 que les cartes étaient traditionnellement conçues à partir de connaissances sur les seuils de perception humaine, mais ces seuils ont été déterminés empiriquement sur des cartes imprimées sur papier. Même si elle augmente continuellement, la résolution des écrans reste inférieure à celle des imprimantes utilisées en cartographie, un pixel étant donc toujours plus grand que le plus petit détail dessiné par ces imprimantes. Il faut donc s'adapter à l'écran, ce qui peut être fait en modifiant les spécifications des cartes, où en affichant les

cartes à des échelles d’affichage supérieures aux échelles de définition, solution adoptée aujourd’hui par la plupart des NMA, car elle ne nécessite pas de modifier la carte.

Nous avons vu également que l’acuité visuelle varie fortement en fonction de la distance à l’écran, mais aussi en fonction de la position par rapport à l’écran (i.e. on voit moins bien en périphérie) (C. Ware 2012). Il faut donc apprendre à concevoir des cartes optimisées pour une visualisation sur un écran.

Enfin, les cartes sont de plus en plus utilisées sur de nouveaux supports numériques, qu’ils soient petits comme les smartphones et les tablettes, ou grands comme les table tactiles ou les murs d’écrans. Si l’on trouve de plus en plus des modes de design de sites webs réactifs, qui s’adaptent le mieux possible à la taille de l’écran, ce n’est pas vraiment le cas des cartes, et la conception de cartes pour petits ou grands écrans est un champ de recherche assez vierge.

1.3.2 Les Portails cartographiques

Nous appelons portail cartographique un site, ou application, web permettant de naviguer entre plusieurs représentations cartographiques, et notamment de changer l’échelle d’affichage par des opérations de zoom. La très grande majorité fonctionne sur le même principe de niveaux de zoom fixes, principe établi par la norme WMTS de l’Open Geospatial Consortium³ : les opérations de zoom avant ou arrière change le niveau de zoom de la carte, chaque niveau de zoom correspondant à une échelle d’affichage deux fois plus grande ou petite que pour les niveaux zooms voisins. Les concepteurs de portails cartographiques possédant des cartes scannées les ont utilisées pour remplir les portails, chaque carte étant affichée à différents niveaux de zooms (qui sont plus nombreux que les cartes scannées disponibles). D’autres, comme GoogleMaps ou OpenStreetMap, ont construit une carte par niveau de zoom.

Ce nouveau mode de consultation des cartes entraîne de nouveaux problèmes de conception cartographique automatique. Tout d’abord, si les cartes utilisées ont été conçues indépendamment, les légendes et styles utilisés sont souvent très différents, ce qui entraîne des difficultés de compréhension de la carte : par exemple, une route peut être symbolisée en jaune dans une carte et en rouge au niveau de zoom suivant.



Figure 1-14 Quatre niveaux d’abstraction pour la représentation cartographique du bâti : un polygone par bâtiment individuel, agrégat de bâtiment dans un îlot, un polygone pour toute la zone bâtie, et un point par ville.

Par ailleurs, le passage d’une échelle à une autre en changeant de niveau de zoom est souvent l’occasion d’un changement de niveau d’abstraction qui peut être long à interpréter pour le lecteur de la carte. Par exemple, la Figure 1-14 montre quatre niveaux d’abstraction pour la représentation du bâti dans une carte, et le passage de l’un à l’autre peut clairement gêner la compréhension de la carte. Au-delà de la lisibilité de chacune des cartes, la fluidité de navigation entre ces cartes à niveau d’abstraction différent peut passer par des processus de cartographie adaptés ou par des interfaces

³ <http://www.opengeospatial.org/standards/wmts>

de navigation adaptées, sachant que ces deux pistes ne sont pas exclusives, mais plutôt complémentaires.

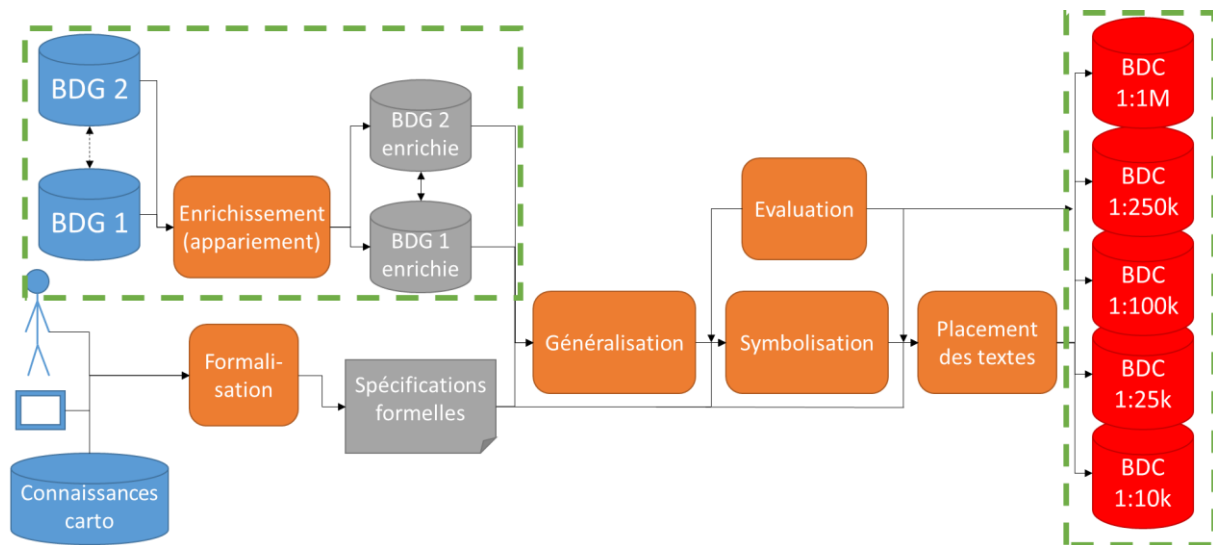


Figure 1-15 Diagramme d'un processus simplifié de cartographie automatique multi-échelles : on part de plusieurs BDG pour obtenir une pyramide de BDCs.

Plusieurs changements sont apportés au processus initial de cartographie automatique lors du passage de la carte sur papier à l'application web cartographique. Tout d'abord, il n'y a plus en sortie une BDC, mais une pyramide cohérente de BDCs à plusieurs échelles. Les spécifications sont maintenant aussi guidées par le type de support de la carte, et notamment la taille et la résolution de l'écran, mais peut-être également les conditions de visualisation : les conditions d'éclairage de l'écran ne sont pas les mêmes en intérieur et en montagne un jour de beau temps. Enfin, pour pouvoir générer cette pyramide de plusieurs BDCs, plusieurs BDGs de niveaux de détail différents peuvent être utilisés en entrée. Ces BDGs en entrée devront être appariées préalablement pour permettre de préserver la cohérence des contenus des diverses BDCs produites (voir la partie 3.2.3).

1.4 IGN, données hétérogènes et collaboratives

1.4.1 Quelles différences entre toutes ces données ?

Un des changements majeurs apportés par le web 2.0 au domaine de la cartographie est la possibilité de réaliser des mashups cartographiques (Jaara, Duchêne et Ruas 2011), qui mixent des données cartographiques de référence comme celles de l'IGN, avec des données d'utilisateurs (par exemple, l'emplacement de restaurants sur une carte topographique), voire des données issues de saisies collaboratives, du *crowdsourcing*, ou de *volunteered geographic information* (VGI) pour reprendre la terminologie de Goodchild (Goodchild 2007). Au-delà de la question du lien avec le fond, déjà examinée précédemment, ces données hétérogènes sont assez différentes des données de référence traditionnelles utilisées pour réaliser des cartes. En effet, quand les données produites par des agences de cartographie peuvent être qualifiées d'homogènes, le plus fidèle possible à leurs spécifications, les données issues d'utilisateurs différents sont particulièrement hétérogènes. L'hétérogénéité la plus évidente est celle de la qualité des données, car les créateurs de ces données n'ont pas toujours les compétences de topographe nécessaires à une saisie précise et homogène. Mais ces données sont aussi hétérogènes par le niveau de détail, car non contraintes par un même jeu de spécifications. Enfin, ces données sont hétérogènes sémantiquement, car les utilisateurs n'utilisent pas les mêmes mots ou les mêmes conceptualisations pour décrire l'espace géographique.

Les problèmes posés par ces trois types d'hétérogénéité en cartographie automatique sont décrits dans les parties suivantes.

1.4.2 Difficultés liées à l'hétérogénéité sémantique

Si l'on reprend la définition de Bishr (1998) de l'hétérogénéité sémantique pour les données géographiques, elle est due à des différences de conceptualisation et de modélisation du monde réel par les différents utilisateur/producteur de données volontaires. Dans cette hétérogénéité sémantique, on peut distinguer l'hétérogénéité cognitive, c'est-à-dire que différentes personnes conceptualisent différemment des entités géographiques, et l'hétérogénéité de nommage, c'est-à-dire le fait que différentes personnes utilisent des noms différents pour désigner la même entité.

Cette hétérogénéité sémantique amène la présence de thèmes cartographiques inhabituels par rapport au contenu des cartes topographiques créées dans des agences cartographiques comme l'IGN. Par exemple, OpenStreetMap contient des taxiways (routes permettant aux avions de rallier les pistes dans les aéroports), qu'il faudra savoir généraliser et cartographier.

Cette hétérogénéité pose également problème pour la réutilisation des connaissances procédurales et paramètres très nombreux dans les processus de cartographie automatique. Par exemple, l'identification des voies à chaussée séparée est nécessaire dans certains processus de généralisation des routes car elles sont modélisées par deux lignes parallèles qu'il faut simplifier en une (Touya 2010), mais ces voies peuvent être modélisées de différentes manières dans des données VGI.

Enfin, cette hétérogénéité sémantique peut être préjudiciable dans le cas de données thématiques qu'un utilisateur souhaite attacher à des données topographiques (Duchêne 2014). Par exemple, on peut faire une carte des pistes cyclables, en les superposant au réseau routier auquel les pistes sont « attachées ». Mais une confusion sémantique entre pistes (séparées de la route) et voies cyclables peut gêner le processus d'alignement des données.

1.4.3 Difficultés liées à l'hétérogénéité de qualité

Les algorithmes de généralisation ont été conçus pour des objets avec un niveau de détail, une granularité homogène. Par exemple, quand un bâtiment a une forme rectangulaire dans une base homogène, c'est que sa forme est simple sur le terrain et les algorithmes de généralisation vont chercher à préserver au mieux cette forme. Mais avec une base de qualité hétérogène, il peut s'agir simplement d'une saisie grossière d'une forme plus complexe, et dans ce cas, il est inutile de chercher à préserver cette forme initiale qui n'est pas représentative du terrain.

Une mauvaise qualité des données peut également affecter la complétude des données : est-ce qu'une absence de données à un endroit correspond bien à un vide sur le terrain ? S'il s'agit d'un problème de complétude des données, il est inutile de chercher à préserver cet espace vide, et il vaut mieux chercher à le combler en grossissant et déplaçant les objets présents dans la carte.

Il est donc indispensable de savoir qualifier ces données hétérogènes, avant d'essayer de les généraliser/symboliser. Mais il ne s'agit pas ici nécessairement de qualité *interne*, mais plus de qualité *externe* en lien avec les processus de cartographie automatique : connaissant les processus utilisés, comment bien qualifier ces données pour que la qualification soit utile au bon fonctionnement du processus de cartographie automatique.

1.4.4 Difficultés liées à l'hétérogénéité de niveau de détail

L'hétérogénéité de niveau de détail génère aussi son lot de difficulté pour un processus de cartographie. Ces difficultés seront plus détaillées dans la partie 2.2.1 qui présente une méthode pour qualifier le niveau de détail des objets dans des données VGI, mais on peut noter deux difficultés principales :

- la cohérence inter-objets : les relations spatiales entre objets de la carte sont fondamentales à la compréhension de la carte (W. A. Mackaness 2007), et ces relations spatiales peuvent être rendues illisibles par une forte différence de niveau de détail entre les deux objets en relation.
- Si l'hétérogénéité de niveau de détail se traduit par une variabilité de la résolution géométrique ou de la granularité, il devient difficile de déterminer le paramétrage des processus cartographiques, qui est déjà un problème complexe avec des données homogènes.

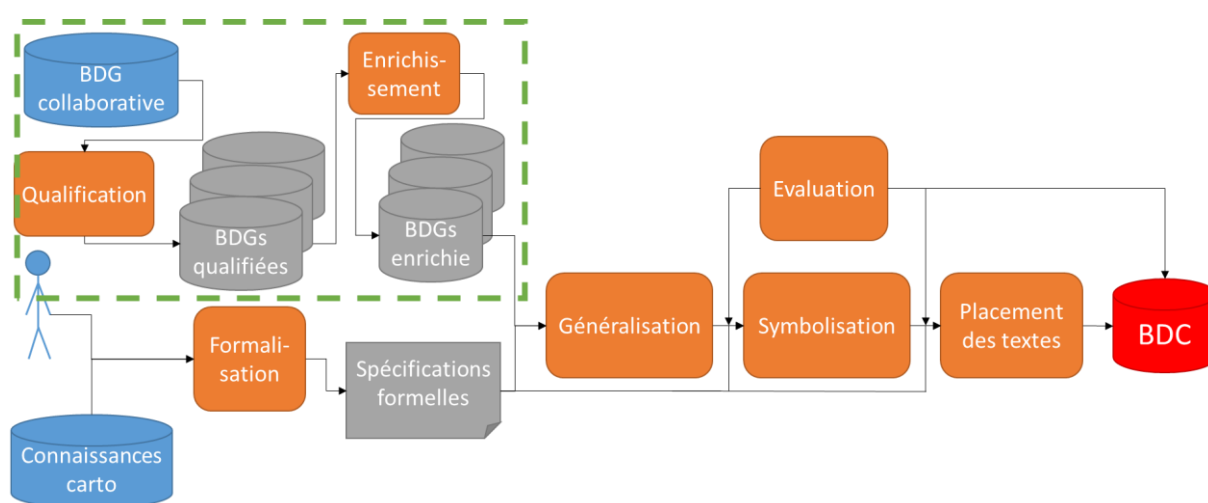


Figure 1-16 Diagramme d'un processus simplifié de cartographie automatique de données collaboratives : les objets de la BDG collaborative sont d'abord qualifiés individuellement (qualité, LoD).

Le changement principal par rapport au processus générique de cartographie automatique, est que la BDG initiale est collaborative/volontaire ou simplement hétérogène. Il est donc indispensable de qualifier ces données, pour estimer la qualité de chaque objet, et pour estimer le niveau de détail pour séparer (au moins de manière conceptuelle) la BDG hétérogène en plusieurs BDGs homogènes en niveau de détail. Par ailleurs, chacun des processus suivants (généralisation, symbolisation, etc.) doit être adapté au traitement de données hétérogènes, avec par exemple un transfert des connaissances procédurales des thèmes de la carte (par exemple, les bâtiments sont symbolisés en noir) vers les objets individuellement (par exemple, ce bâtiment avec une mauvaise qualité de positionnement doit être symbolisé en gris clair).

1.5 Synthèse et objectifs scientifiques

1.5.1 Où sont les verrous scientifiques pour l'automatisation des processus cartographiques ?

Afin de synthétiser ce premier chapitre, nous proposons de résumer les principaux verrous scientifiques liés à la cartographie automatique sous toutes ses formes. Il y a tout d'abord une série

de verrous scientifiques communs à l'ensemble des cas de dérivation automatique de cartes vus dans les parties précédentes :

- **Comprendre les besoins de l'utilisateur** : qu'il soit un cartographe ou le grand public, mieux comprendre le besoin de l'utilisateur est indispensable pour bien calibrer des processus automatiques.
- **Formaliser les spécifications de la carte**, qu'elles proviennent des besoins de l'utilisateur ou de connaissances en cartographie.
- **Développer des processus encore plus automatiques** : plutôt que de chercher le bon placement du curseur entre automatisation et qualité de la carte, visons maintenant le maximum de qualité et d'automatisation simultanément.
- **Savoir orchestrer les processus automatiques** : un processus unique n'étant pas capable de générer une carte de qualité de manière complète, il faut savoir orchestrer automatiquement différents processus pour les appliquer là où et quand ils sont efficaces.
- **Evaluer automatiquement des BDCs** : l'automatisation de la création de la carte passe par l'automatisation de son évaluation, pour permettre des orchestrations par essai/erreur, ou simplement pour identifier les zones non lisibles d'une carte.
- **Mieux comprendre la lisibilité et l'efficacité des cartes** : plus que le respect de spécifications, l'enjeu de la cartographie automatique de qualité est de concevoir des cartes plus efficaces pour les utilisateurs. Nous devons donc basculer dans une démarche scientifique systématisant le recours aux tests auprès d'utilisateurs, à l'image des communautés de recherche en visualisation d'information ou en interaction homme-machine.
- **Passer à l'échelle** : que ce soit pour la production de cartes en série à l'échelle d'un pays, ou à partir des données OpenStreetMap disponibles à l'échelle du monde, nous devons être capables de générer des cartes automatiquement sur de gros volume de données, ce qui ne pose pas que des problèmes techniques, loin s'en faut.

Il existe également des verrous plus spécifiques à la création de cartes multi-échelles, à la cartographie à la demande, ou à la cartographie à partir de données hétérogènes :

- **Concevoir des processus de cartographie adaptatifs** : les processus automatiques de généralisation, symbolisation, placement de texte doivent être capables de fonctionner avec des BDG multiples, hétérogènes, ou contenant des données inhabituelles.
- **Savoir manipuler des données thématiques et leurs données support**, à travers les échelles. Les données topographiques ont de plus en plus vocation à être le support de données thématiques de toutes sortes, et il faut à la fois savoir comment modéliser cette relation de support et comment bien la transformer à travers les échelles.
- **Savoir comment générer des cartes multi-échelles**, plutôt que des cartes avec une échelle fixe.
- **Qualifier les données hétérogènes** : même si les processus de cartographie deviennent capables de s'adapter à des données hétérogènes, il est indispensable de savoir qualifier les données hétérogènes utilisées pour construire la carte, cette qualification devant être faite au niveau de l'instance (*ce bâtiment*) plutôt que du thème (*tous les bâtiments de la carte*).

1.5.2 Plan du mémoire

Ce mémoire décrit une partie des recherches que j'ai menées ou que j'ai encadrées pour répondre à ces verrous scientifiques sur l'automatisation des processus de production cartographique, dont notamment la généralisation cartographique.

La partie 2 du mémoire est consacrée à divers travaux de recherche pouvant être vus comme des briques de base pour permettre l'automatisation des processus cartographiques. La partie 2.1 présente une méthode pour recaler géométriquement des données géographiques sur des données de références, ce qui est nécessaire pour cartographier des données hétérogènes ou les données d'un utilisateur sur une carte de référence. La partie 2.2 présente des travaux sur l'harmonisation de niveau de détail pour la conception automatique de cartes à grande échelle avec des données géographiques volontaires hétérogènes en niveau de détail (comme OpenStreetMap). Il s'agit de transformer des données peu détaillées pour les cartographier sans incohérence avec des données très détaillées : par exemple ajouter une clairière dans une forêt peu détaillée autour d'un groupe de bâtiments assez grands pour être sûr de la présence de la clairière sur le terrain. La partie 2.3 présente différents travaux réalisés en collaboration pour formaliser les connaissances nécessaires au fonctionnement des processus automatiques (ontologies notamment). Enfin, la partie 2.4 présente le post-doc de Jean-François Girres que j'ai encadré en 2012 et 2013, sur la conception d'un meta-modèle de généralisation, le ScaleMaster2.0, soit un modèle pour orchestrer des modèles de généralisation.

La partie 3 est spécifiquement dédiée à l'amélioration des processus de généralisation automatique. La partie 3.1 reprend les travaux de thèse d'Adrien Maudet (2012-2016) sur la modélisation multi-niveaux dans les modèles de généralisation par système multi-agents. La partie 3.2 décrit les travaux, encore en cours, de la thèse de Marion Dumont (démarrée en décembre 2014), sur la dérivation de cartes intermédiaires entre deux échelles préexistantes dans une pyramide d'échelles.

Chapitre 2. Comment automatiser ces processus cartographiques ?

Nous avons vu dans le chapitre précédent que la production automatique de cartes de bonne qualité nécessitait la conception de plusieurs briques de base, ces briques étant d'autant plus nombreuses qu'on s'autorise à intégrer des données hétérogènes, issues d'un utilisateur, où qu'on cherche à produire des cartes multi-échelles affichées sur écran. Avant même de chercher à intégrer toutes ces briques dans un système adaptatif capable de les enchaîner au mieux suivant le besoin, des recherches sont nécessaires pour que ces briques fonctionnent au mieux. Ce chapitre présente donc plusieurs travaux de recherche menés ces dernières années, dans des projets ou en collaboration avec d'autres chercheurs du domaine, pour concevoir des méthodes ou pour formaliser des connaissances nécessaires à la construction de ces briques de base.

Dans un premier temps, nous présentons une méthode de recalage géométrique de données géographiques. Puis, une partie plus importante est consacrée à l'harmonisation de niveau de détail dans des données hétérogènes issues de saisies collaboratives. La troisième partie est consacrée à la formalisation des connaissances pour la cartographie automatique. Enfin, la dernière partie présente des travaux sur les méta-modèles de généralisation cartographique.

2.1 Recalage géométrique de données géographiques

Lorsqu'une partie des données à cartographier a été produite de manière indépendante des données de référence de la carte, ce qui est le cas pour des données fournies par un utilisateur lors de la production d'une carte à la demande (Figure 1-13), les données peuvent être « décalées », ne pas être superposables, comme sur la Figure 2-1. Ce décalage peut avoir plusieurs raisons, comme les choix de systèmes de coordonnées, les techniques de saisie des données (GPS, restitution depuis une photo aérienne, etc.) ou des problèmes de qualité. Il est alors indispensable d'intégrer ces données extérieures et les données de référence. Cela se fait traditionnellement en deux phases :

- une phase d'*appariement* qui associe les éléments dans les deux jeux de données qui représentent la même entité du monde réel.
- une phase de *recalage géométrique* des données les moins précises (souvent les données de l'utilisateur, mais l'inverse est possible) sur les données les plus précises.

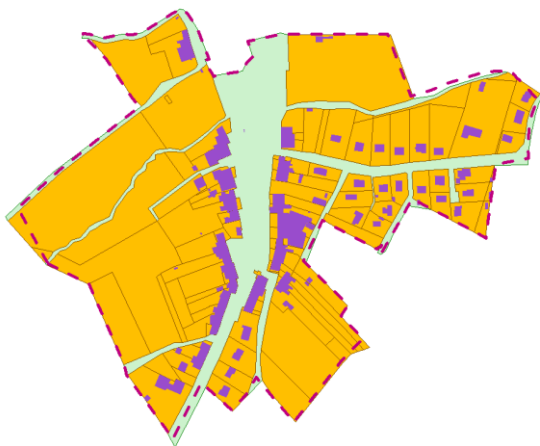


Figure 2-1 Données cadastrales et limite communale de l'IGN (tirets mauves) incohérentes : il faut déformer les données cadastrales pour les recaler sur les limites IGN plus précises.

La phase d'appariement a été traitée par de nombreux chercheurs, notamment dans l'équipe COGIT (Walter et Fritsch 1999) (Mustière et Devogele 2008) (Olteanu-Raimond, Mustière et Ruas 2015), elle n'est donc pas abordée dans ce mémoire. La phase de recalage est également complexe car elle nécessite de transformer les géométries de manière parfois importante, tout en préservant les formes générales des objets géographiques, pour qu'ils soient toujours reconnaissables. Par exemple, un bâtiment est souvent formé d'angles droits, et ceux-ci doivent être préservés dans la transformation, et il en va de même pour la courbure des cours d'eau, ou la forme des routes. Ce problème de déformation rappelle la problématique de la généralisation qui doit souvent déplacer et déformer les objets pour améliorer leur lisibilité tout en préservant leurs formes caractéristiques. Nous avons donc proposé de réutiliser les principes de la généralisation par moindres carrés développée parallèlement par (L. E. Harrie 1999) et (Sester 2005), dont on peut voir les résultats (Figure 2-2) grâce à un prototype ré-implémenté dans notre plateforme open source GeOxygene (Grosso, Perret et Brasebin 2012). Le principe est le suivant : des contraintes sur les géométries sont transformées en équations linéaires avec comme inconnues les variations en x et y de l'ensemble des points des géométries. La présence de contraintes de lisibilité et de contraintes de préservation rend le système d'équation sur-contraint, et la solution des moindres carrés donne une position optimale des points des géométries. Pour résoudre un problème de recalage, les contraintes de lisibilité sont remplacées par des contraintes forçant les points à suivre les vecteurs de déplacement dus à l'appariement de certains points dans les deux jeux de données.

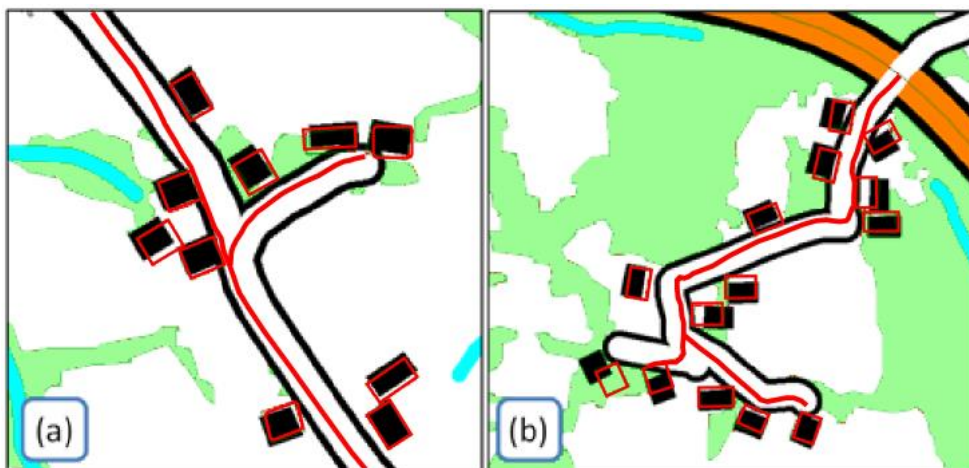


Figure 2-2 Deux résultats de généralisation obtenus pour le processus par moindres carrés issu de (L. E. Harrie, *The Constraint Method for Solving Spatial Conflicts in Cartographic Generalization* 1999) et (Sester 2005), et réutilisé dans le processus de recalage (en rouge, les géométries initiales des bâtiments et routes)

Les moindres carrés sont une méthode mathématique qui permet de trouver une solution approchée à un système d'équations linéaires, quand le nombre d'équations est supérieur au nombre d'inconnus, ce qui est le cas ici. La solution des moindres carrés minimise la somme des carrés des résidus (la différence entre la solution trouvée et la solution exacte d'une équation). Un système d'équations linéaires peut s'exprimer par des matrices, en utilisant une matrice jacobienne $\mathbf{A}$ (Equation 1).

$$\begin{cases} a_{1,1} \cdot \Delta x_1 + a_{1,2} \cdot \Delta y_1 + a_{1,3} \cdot \Delta x_2 + a_{1,4} \cdot \Delta y_2 = l_1 + v_1 \\ a_{2,1} \cdot \Delta x_1 + a_{2,2} \cdot \Delta y_1 + a_{2,3} \cdot \Delta x_2 + a_{2,4} \cdot \Delta y_2 = l_2 + v_2 \\ a_{3,1} \cdot \Delta x_1 + a_{3,2} \cdot \Delta y_1 + a_{3,3} \cdot \Delta x_2 + a_{3,4} \cdot \Delta y_2 = l_3 + v_3 \\ \dots \\ a_{n,1} \cdot \Delta x_1 + a_{n,2} \cdot \Delta y_1 + a_{n,3} \cdot \Delta x_2 + a_{n,4} \cdot \Delta y_2 = l_n + v_n \end{cases} \Leftrightarrow \mathbf{A} \cdot \mathbf{X} = \mathbf{l} + \mathbf{v} \quad (1)$$

où $\mathbf{v}$ est le vecteur des résidus

On peut ensuite ajouter une matrice de poids $\mathbf{P}$ qui traduit l'importance relative de chaque équation, c'est-à-dire qu'il est plus important d'être proche de la solution pour l'équation i que pour l'équation j . Les moindres carrés trouvent alors une solution qui minimise :

$$\mathbf{v}^T \cdot \mathbf{P} \cdot \mathbf{v} \quad (2)$$

La solution obtenue par inversion matricielle est la suivante :

$$\mathbf{x} = (\mathbf{A}^T \cdot \mathbf{P} \cdot \mathbf{A})^{-1} \cdot \mathbf{A}^T \cdot \mathbf{P} \cdot \mathbf{l} \quad (3)$$

La méthode de recalage nécessite un appariement minimal entre les deux jeux de données, que ce soit au niveau des objets (des objets homologues sont identifiés dans les deux jeux de données) ou des points (certains points des géométries des objets sont appariés dans les deux jeux de données), comme dans le schéma de la Figure 2-3. Ces vecteurs de déplacement permettent de construire un champ vectoriel contraignant tous les objets à se déformer. Plusieurs possibilités sont possibles pour calculer ce champ vectoriel, suivant le type d'appariement disponible en entrée du processus (Touya, Coupé, et al. 2013).

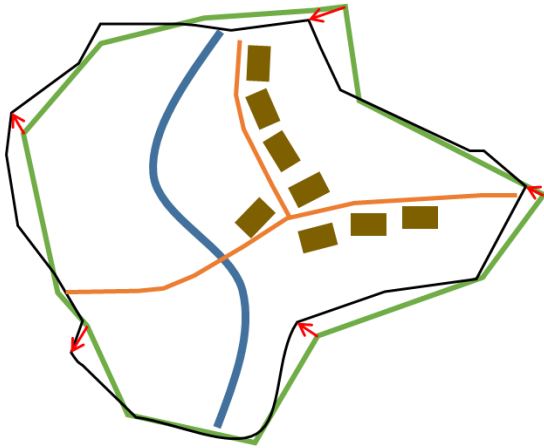


Figure 2-3 Des données à recaler (limite administrative en vert, rivière, routes, bâtiments) et une limite administrative plus précise (en noir), sur laquelle on doit tout recaler. Les flèches rouges montrent un appariement minimal de nœuds des deux limites, à partir duquel le recalage peut être calculé.

Pour contrebalancer ces contraintes de recalage et s'assurer que les objets géographiques préservent leurs caractéristiques et leurs relations spatiales, nous réutilisons les contraintes de préservation de forme introduites pour la généralisation par (L. E. Harrie 1999) et (Sester 2005), comme par exemple la contrainte de rigidité, appliquée aux bâtiments, qui tend à n'autoriser que des translations et rotations aux bâtiments :

$$\forall i \in [1, \text{nombre_de_points}], \begin{cases} \Delta x_i - \Delta x_{i+1} = 0 \\ \Delta y_i - \Delta y_{i+1} = 0 \end{cases} \quad (4)$$

D'autres contraintes de préservation sont introduites, notamment pour préserver les proximités (entre les bâtiments de la Figure 2-3 par exemple), les positionnements relatifs (un bâtiment au niveau d'un carrefour ou au bout d'une impasse par exemple) et les orientations relatives (un bâtiment parallèle à une route voisine par exemple).

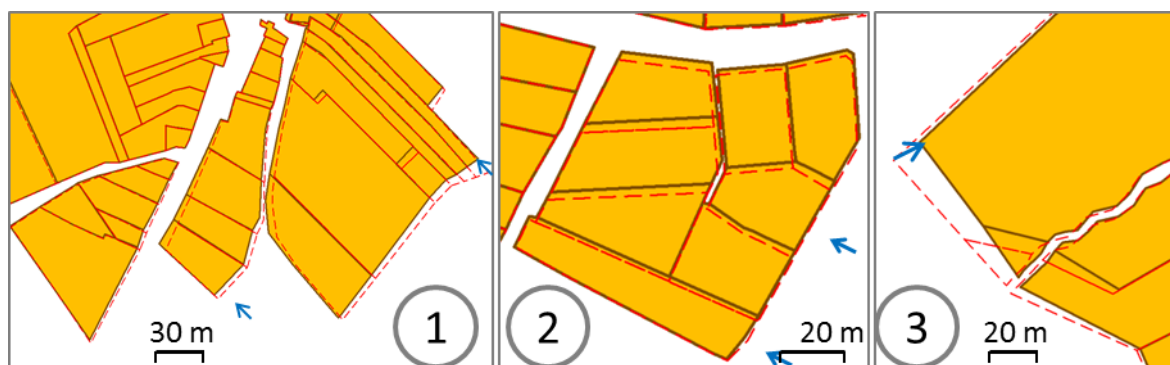


Figure 2-4 Trois extraits des résultats du recalage par moindres carrés des données de la Figure 2-1, les flèches bleues montrent les points appariés, et les tirets rouges montrent les géométries initiales

La Figure 2-4 montre des résultats obtenus pour le recalage de la zone la Figure 2-1, les flèches montrant la différence de positionnement des points (de la limite communale) appariés dans les deux jeux de données (déplacements de l'ordre de 10 m dans ce cas). Ici, ce sont les polygones de parcelles cadastrales qui ont été recalés et on voit bien que les formes des parcelles ont bien été préservées tout comme les espaces vides séparant des parcelles (grâce à la contrainte de proximité).

En résumé, ces travaux proposent une méthode pour recaler des données géographiques sur un référentiel tout en préservant les formes caractéristiques et les relations spatiales des données recalées. Cette méthode contribue à la brique de base *Appariement et recalage* (Figure 1-13) nécessaire pour automatiser la cartographie à la demande à partir de données issues de l'utilisateur.

Ces travaux de recalage ont été publiés avec plus de détails dans (Touya, Coupé, et al. 2013), et ont donné lieu à une mise en production à l'IGN pour générer des données cadastrales unifiées avec la qualité géométrique de l'IGN.

2.2 Harmonisation de niveau de détail pour les cartes à grande échelle

Nous avons vu dans la partie 1.4.4 qu'une des difficultés pour utiliser des données collaboratives, ou volontaires, dans un processus de cartographie automatique est l'hétérogénéité du niveau de détail de ces données entre les thèmes, et au sein des objets d'un même thème. La première étape est donc de pouvoir identifier le niveau de détail de chaque objet pris individuellement, pour que les processus automatiques sachent comment le cartographier à une échelle donnée. Une méthode permettant cette identification est présentée dans la sous-partie suivante, puis nous abordons (2.2.2 et 2.2.3) le problème de la cartographie automatique de ces données dans une carte à grande échelle, procédé que nous avons appelé harmonisation de niveau de détail.

2.2.1 Inférence du niveau de détail pour des données hétérogènes

Nous avons vu dans le premier chapitre (partie 1.1.2) qu'il était plus juste de parler de niveau de détail ou de niveau d'abstraction pour décrire la granularité d'une BDG. Mais cela peut aussi s'appliquer aux cartes comme le montre la Figure 2-5 avec l'agglomération de Pau extraite de trois BDC aux échelles 1 : 50 000, 1 : 250 000 et 1 : 1 000 000, et les niveaux d'abstraction et la granularité sont à chaque fois très différents. Le niveau de détail peut être défini par ses composants, que nous rappelons ici car nous allons utiliser ces composants pour inférer le niveau de détail :

- Le schéma conceptuel
- La résolution géométrique
- La résolution sémantique
- La précision géométrique
- La granularité

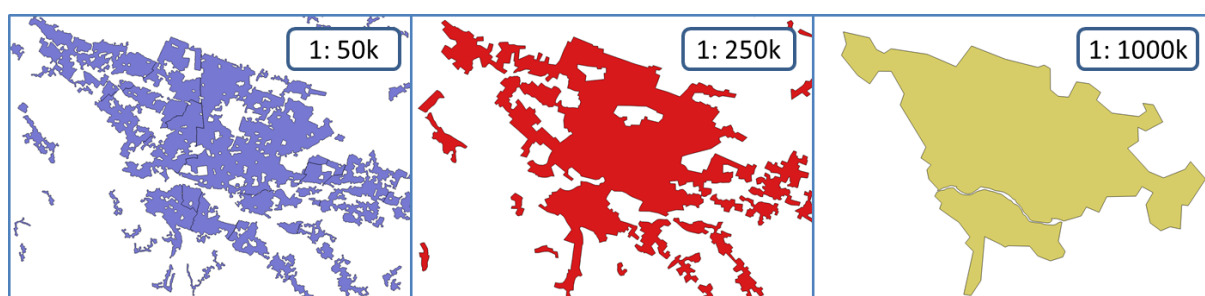


Figure 2-5 L'agglomération de Pau représentée à trois échelles dans les bases de données cartographiques de l'IGN (Scan Express™) qui correspondent également à trois niveaux de détails pour la représentation des zones bâties.

Si on s'intéresse aux données géographiques hétérogènes, particulièrement celles issues de saisies citoyennes et participatives, par opposition aux données homogènes comme celles fournies par l'IGN, on constate qu'elles présentent de fortes hétérogénéités de niveau de détail. En effet, ces données sont souvent issues de sources de données hétérogènes et de précision diverse comme des GPS, des téléphones, des images aériennes, des scans de cartes, ou même des jeux de données homogènes de niveau de détail variable. De plus, ces saisies collaboratives sont faites par des personnes dont les connaissances en géomatique sont très variables, pas besoin d'être un spécialiste pour contribuer à OpenStreetMap ou pour charger les traces GPS de sa randonnée dans l'Espace Loisirs de l'IGN. Enfin, contrairement aux spécifications d'une base de données institutionnelles qui décrivent généralement avec précision la résolution, la granularité ou la précision requise, les projets de saisie collaborative de données géographique ne donnent pas ce type de spécification pour ne pas effrayer les contributeurs potentiels et pour conserver une contribution facile et à la portée de tout citoyen. Ainsi, on trouve dans OpenStreetMap des lignes représentant les passages piétons, mais aussi des forêts dont le niveau de détail est plutôt adapté à une visualisation au 1 : 1 000 000.

Si l'on souhaite cartographier ces données automatiquement, il faut être capable de déterminer automatiquement le niveau de détail de chacun des objets, individuellement, pour savoir à quelles échelles il peut être représenté, et à quelles autres échelles il doit être transformé. La définition du niveau de détail étant floue, il nous paraît plus pertinent de définir des catégories de niveau de détail assez floues elle-même pour différencier les objets. Nous avons fait une proposition (Table 1), qui demande à être discutée, dans laquelle les gammes d'échelles sont données à titre indicatif.

Rue	Ville	Département	Région	Pays
1 : 1 à 1 : 15K	1 : 15K à 1 : 50K	1 : 50K à 1 : 200K	1 : 200K à 1 : 1M	1 : 1M et plus

Table 1. Cinq catégories proposées pour classer le niveau de détail des objets géographiques

Pour classer les objets issus de saisie collaborative dans ces catégories de niveau de détail, nous proposons de nous appuyer sur la mesure de plusieurs critères qui reprennent les propriétés du niveau de détail présentées plus haut. Ces critères ont été mis en place pour qualifier des données issues d'OpenStreetMap, mais ils sont pour la plupart directement applicables à d'autres types de données hétérogènes.

Nature de l'objet

Ce critère mesure la composante *schéma conceptuel* du niveau de détail. Pour cela, nous comparons la nature de l'objet à une ontologie des concepts géographiques (voir la partie 2.3) dans laquelle nous avons associé une gamme d'échelle à certains concepts. Par exemple, les bâtiments et les points d'intérêts sont associés aux grandes échelles, quand le concept de zone urbaine ou de forêt sont associés aux moyennes et petites échelles. Certains concepts de l'ontologie n'ont pas de gamme d'échelle, comme les routes, par exemple, qui peuvent apparaître avec une géométrie linéaire quelle que soit l'échelle de la carte.

Densité de points

Ce critère mesure la *résolution géométrique* des objets qualifiés. Il est calculé simplement en divisant le nombre de points de la géométrie par la longueur (ou le périmètre pour les objets surfaciques). Plus il y a de points pour une longueur de ligne donnée, plus la résolution géométrique est importante, et donc plus le niveau de détail général est grand. Ce critère n'est pas applicable aux objets dont la géométrie est ponctuelle.

Longueur médiane

Jean-François Girres, dans une étude pour mesurer l'impact de la généralisation sur la qualité des données (2011), montre que la densité de points n'est pas toujours le meilleur critère pour qualifier la *résolution géométrique* des objets. En effet, certaines entités géographiques sinueuses comme les routes de montagne ou les cours d'eau, nécessitent l'utilisation de plus de points pour représenter les courbes de manière réaliste. Il propose d'utiliser plutôt la médiane de la distribution des longueurs de segments entre deux points consécutifs de la géométrie, et nous ajoutons ce critère pour contrebalancer le précédent. Ce n'est pas, non plus, applicable aux objets dont la géométrie est ponctuelle.

Source de saisie

Sans analyse complémentaire, comme dans (Haklay 2010) ou (Girres et Touya 2010), il est difficile d'estimer la composante de *précision géométrique* du niveau de détail. Dans OpenStreetMap, il existe des tags de métadonnées, décrivant la source de saisie (« source »). Ce tag de métadonnées n'est pas toujours rempli, mais quand il l'est, nous en dérivant une information sur le niveau de détail. Par exemple, la valeur « GPS » induit un niveau de détail important ; la valeur « Bing », qui signifie que la saisie a été effectuée sur des images satellite mises à disposition par Bing, induit un niveau de détail moyen ; et la valeur « Corine Land Cover » induit un niveau de détail faible.

Plus petit côté

Mesurer la *granularité* d'un objet géographique peut s'avérer complexe car le « détail » à mesurer n'est pas clairement défini et peut varier d'un objet à un autre. Au contraire, la taille du plus petit côté d'un polygone, ou du plus petit segment d'une polyligne est quelque chose que l'on peut mesurer facilement sur tous les objets (linéaires ou surfaciques). Cette mesure est d'ailleurs parfois utilisée en généralisation pour contraindre la granularité des objets dans la carte (Stoter, et al. 2010).

Taille

Le critère de taille analyse la taille des objets pour compléter l'estimation de la *granularité*. En effet, la présence de très petits objets traduit une forte granularité. Nous avons utilisé les seuils de perceptibilité des objets pour relier les valeurs de taille à une échelle minimum de lisibilité et par conséquent à un niveau de détail. Pour les gros objets, le critère est neutre.

Empâtement

L'empâtement en cartographie est la superposition de portions d'un symbole appliqué à une ligne, notamment lorsque la ligne forme des virages trop serrés pour la largeur du symbole. La généralisation cartographique permet de supprimer cet empâtement et il a notamment été utilisé pour identifier les parties de lignes (routes de montagne notamment) qui nécessitent un algorithme de généralisation particulier (Mustière 1998), (Nakos, Gaffuri et Mustière 2008). Si on teste des largeurs de symbole correspondant à des échelles de carte classiques (25000^{ème}, 100000^{ème}, etc.) et qu'il n'y a pas d'empâtement, c'est que la granularité est petite pour un affichage à ces échelles. Cette idée dérive de la proposition de (Girres 2011), qui utilise l'empâtement pour vérifier si des données géographiques ont été généralisées.

Nombre d'éditions

La plupart de ces critères n'étant pas applicables aux objets à géométrie ponctuelle, nous en avons ajouté un dernier pour mieux estimer le niveau de détail de ces objets. Mais ce critère pourrait aussi être utilisé sur d'autres objets. Il dérive des travaux de Mooney et Corcoran (2012) sur le lien entre le nombre d'éditions d'un objet par différents contributeurs, et la qualité des tags associés à cet objet. Avec ce critère, plus un objet a été modifié, plus son niveau de détail est élevé.

Tous ces critères ont été calibrés à l'aide de bases de données géographiques ou cartographiques de l'IGN, à différents niveaux de détail : on analyse par exemple la distribution des densités de points dans une base de données cartographique au 1 : 100 000 pour estimer la résolution géométrique à cette échelle (Touya et Reimer 2015).

Décision multicritères et méthode ELECTRE TRI

Pour combiner efficacement ces nombreux critères, le domaine de la décision multicritères offre un panel intéressant de méthodes, et dans ce panel nous avons choisi d'utiliser la méthode ELECTRE TRI (Yu 1992), (Figueira, Mousseau et Roy 2005), déjà utilisée en généralisation par Taillandier (Taillandier et Taillandier 2012). Tout d'abord, les méthodes de décision multicritères peuvent servir à choisir le meilleur, à ordonner ou à classer (Figueira, Mousseau et Roy 2005), comme le fait ELECTRE TRI, et nous cherchons ici à classer les objets OSM dans une catégorie de LoD, donc cette méthode est adaptée. Par ailleurs, la méthode ELECTRE TRI est adaptée aux problèmes de décision multicritères dans lesquels le nombre de critère est supérieur à trois, quand les critères ne peuvent être comparés entre eux, et quand il existe une part d'incertitude sur la contribution de chaque critère à la prise de décision générale (Figueira, Mousseau et Roy 2005), (Taillandier et Taillandier 2012).

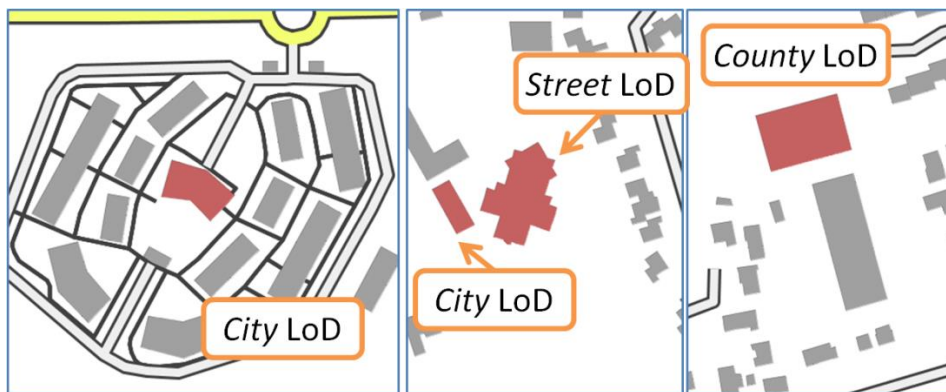


Figure 2-6 Résultats de l'inférence automatique de LoD pour des bâtiments plus ou moins détaillés (source : ©OpenStreetMap contributors).

La Figure 2-7 montre des résultats d'inférence automatique du niveau de détail pour des forêts et des zones bâties OSM en France, issues principalement d'un import automatique depuis la base de données européenne d'occupation du sol Corine LandCover, peu détaillée. On voit que les certaines forêts qui ont été éditées manuellement après import apparaissent comme détaillées quand les autres sont bien évaluées comme peu détaillés.

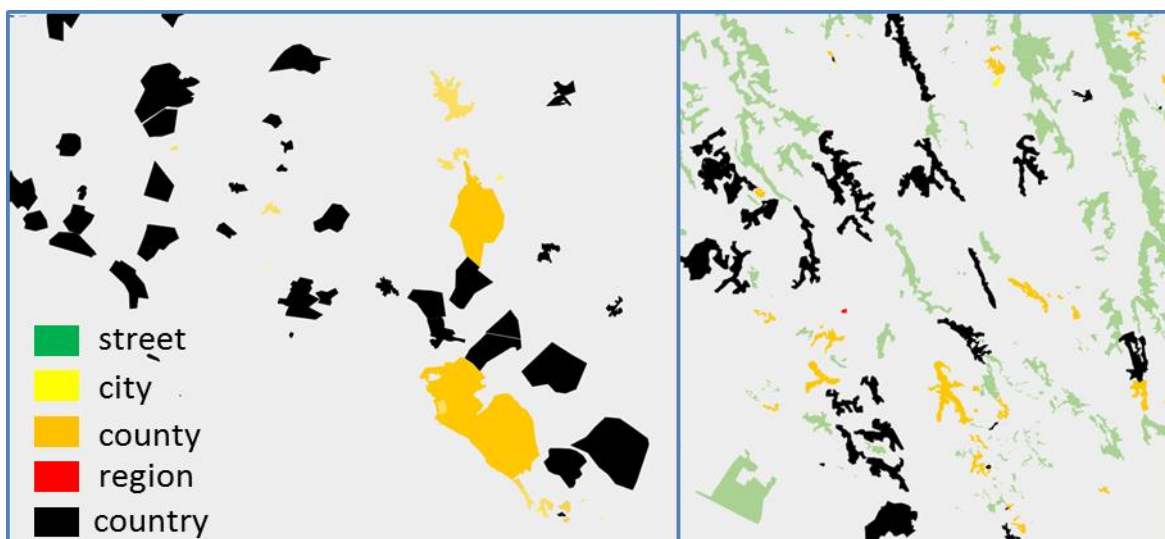


Figure 2-7 Résultat de l'inférence automatique du niveau de détail pour zones bâties à gauche et des forêts à droites (source : ©OpenStreetMap contributors).

Une fois qu'on a réussi à inférer le LoD de chaque objet individuellement, on peut identifier des incohérences dues à des différences de LoD entre objets géographiques voisins et partageant des relations spatiales. Il faut définir des relations spatiales impossibles théoriquement, comme un lotissement de bâtiments entièrement sous couvert forestier (Figure 2-8a), des bâtiments se trouvant juste à l'extérieur d'une zone bâtie (Figure 2-8b), ou un arrêt de bus situé à 50 m d'une route (Figure 2-8c) : quand des instances de ces relations spatiales impossibles sont identifiées entre objets de LoD différent, on peut en déduire que la différence de niveau de détail est très probablement responsable de cette incohérence. La Figure 2-8 montre des résultats d'identification automatique de ces incohérences de niveau de détail dans des données OSM en France (mais d'autres tests ont montré que ces incohérences sont aussi présentes dans les autres pays). Si l'on souhaite réaliser des cartes à partir de ces données, que ce soit manuellement ou par des processus

automatiques (problème qui nous intéresse ici), il faudra donc corriger ces incohérences qui peuvent perturber la lecture de la carte en induisant le lecteur en erreur.



Figure 2-8 Détection automatique de relations spatiales improbables entre objets de LoD différent : (a) ce groupe de bâtiments proches ne peut pas être dans une forêt, (b) la zone bâtie grise devrait inclure les bâtiments en rouge, (c) l'arrêt de bus marqué d'une croix rouge est trop loin de la route, (d) le chemin rouge passe dans le lac (@OpenStreetMap contributors).

Cette méthode d'inférence automatique du niveau de détail de données collaboratives contribue à la brique de base *Qualification* du processus de cartographie automatique à partir de données collaboratives (Figure 1-16).

2.2.2 Opérations d'harmonisation de niveau de détail

Avec le niveau de détail disponible pour certains objets géographiques dans OpenStreetMap, et puisque ces échelles sont disponibles dans les outils webs cartographiques, il est tentant de dériver des cartes à grande échelle. Par convention, nous considérons comme cartes à grande échelle, des cartes dont l'échelle est supérieure au 1 : 25000. Ces cartes sont susceptibles de contenir des objets présentant des incohérences de niveau de détail telles que celles présentées dans la partie précédente. En effet, les données très détaillées sont bien adaptées à une représentation à grande échelle, et certaines données non détaillées, même si elles ne sont pas adaptées à ces échelles, doivent nécessairement être présentes également dans la carte : comment se priver des lacs par exemple ? Il faut donc trouver un moyen pour éliminer ces incohérences de niveau de détail qui nuisent fortement à la qualité de la carte car elles rendent difficile l'interprétation des données. Notre proposition est de concevoir des méthodes d'harmonisation de niveau de détail qui vont transformer l'objet peu détaillé pour corriger la relation spatiale improbable ou impossible. Et puisque l'objectif est de faire une carte, et pas de corriger la base de données OSM, nous serons plus guidés par des questions de lisibilité et de compréhension, que par des questions de réalisme et de précision.

Nous avons d'abord proposé une taxonomie de ces opérations d'harmonisation, qui vient étendre une taxonomie des opérations de cartographie automatique proposée par (Roth, Brewer et Stryker 2011), et nous distinguons quatre nouvelles opérations :

- *Merge/Dissolve* : c'est un regroupement des opérateurs *Merge* et *Dissolve* qui permet d'agréger ou de désagréger des objets se trouvant en limite d'une zone et devant être inclus ou exclus de cette zone (Figure 2-9 et Figure 2-10).
- *Adjust Shape* : c'est l'ajustement de la forme d'un objet pour éviter des relations spatiales improbables comme des intersections chemin/lac (Figure 2-12).
- *Displace & enhance* : il s'agit du déplacement d'un ou plusieurs objets pour mettre en valeur une relation spatiale dans la carte (Figure 2-13).
- *Disambiguation* : parfois, le lecteur de la carte ne saura pas dire si l'incohérence est due à un problème dans la carte ou si elle existe réellement sur le terrain. Dans ce cas, la désambiguïsation supprime cette ambiguïté sans considération pour la vérité terrain, qui est inconnue (Figure 2-11).

La suite de cette partie porte sur la description d'algorithmes permettant de réaliser ces opérations pour différents exemples d'incohérences de niveau de détail. Nous avons tout d'abord développé deux algorithmes réalisant des opérations de *Merge/Dissolve*. Le premier permet de corriger le contour d'objets surfaciques décrivant des sites fonctionnels comme des écoles, hôpitaux, ou centre commerciaux (Chaudhry, Mackaness et Regnauld 2009). Les objets géographiques détaillés intersectant le bord du site sont identifiés, comme les tronçons de route entourés à la limite de l'hôpital de la Figure 2-9. Puis, on détermine leur score d'appartenance au site en fonction de critères spatiaux (il est à 70% dans le site par exemple) et de critères sémantiques (un terrain de sport ou une bibliothèque font probablement partie d'une école, ou, une autoroute ne fait sûrement pas partie d'un hôpital). Enfin, la géométrie du site est modifiée pour inclure ou exclure ces objets (Figure 2-9).

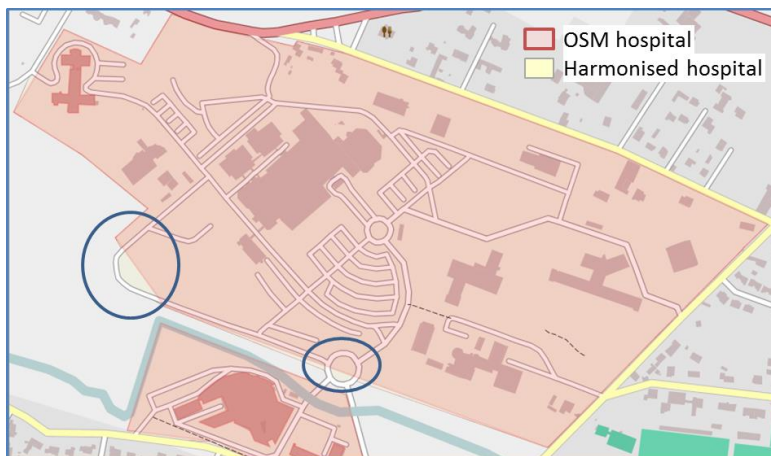


Figure 2-9 La zone de l'hôpital est harmonisée en l'étendant aux routes faisant partie de l'hôpital car permettant la circulation à l'intérieur (@OpenStreetMap contributors).

Le deuxième algorithme de *Merge/Dissolve* développé est plus simple et permet d'étendre les zones bâties aux bâtiments situés juste à l'extérieur. De manière itérative, on cherche les bâtiments proches de la limite dans un rayon de 25 m, puis on fusionne la géométrie de la zone bâtie avec une dilatation, du même seuil de 25 m, des bâtiments trouvés. Puis, on simplifie la géométrie du contour pour avoir une granularité géométrique similaire sur toute la limite. A chaque itération, le seuil est diminué pour éviter que la zone bâtie ne se propage trop loin (Figure 2-10).

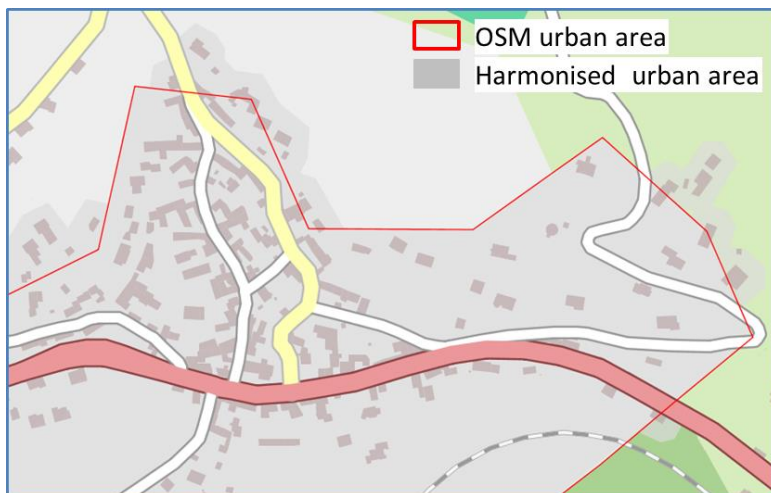


Figure 2-10 La zone urbaine peu détaillée issue d'OpenStreetMap est étendue pour intégrer les bâtiments situés directement à l'extérieur (@OpenStreetMap contributors).

Pour la désambiguïsation, nous avons développé un algorithme qui reconstruit une clairière autour d'un groupe de bâtiments dont la taille est assez grande pour rendre certaine la présence d'une clairière. La première étape est d'identifier et construire les groupes de bâtiments ce qui est fait par une technique similaire à la construction des emprises de villes à partir des bâtiments (Boffet 2000). Puis la clairière est créée en prenant une dilatation de l'enveloppe convexe du groupe, qui est coupée par les routes, qui forment souvent des barrières naturelles pour les forêts (Figure 2-11).

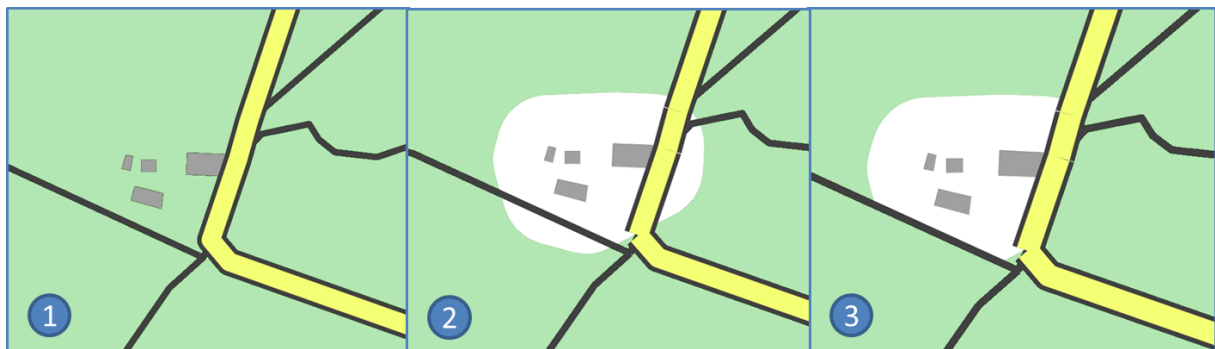


Figure 2-11 La forêt peu détaillée (1) est trouée par un polygone formé par l'union des bâtiments dilatés (2) et la clairière créée est rognée (3) par les routes et chemins proches qui forment souvent les limites d'une forêt ou clairière (@OpenStreetMap contributors).

Concernant l'opération *Displace & enhance*, nous avons développé un algorithme permettant de réaligner des arbres saisis de manière peu précise le long d'une route. On commence par identifier les arbres présents à gauche ou à droite d'une route, puis on calcule une version décalée de la géométrie de la route, la longueur du décalage étant liée à la largeur du symbole de la route dans la carte (Figure 2-12). Puis, on projette l'ensemble des arbres sur cette ligne décalée et on les écarte un peu s'ils se retrouvent trop près les uns des autres. Cela permet à la fois de corriger la superposition des symboles, et d'améliorer la relation d'alignement qui devient nette donc plus lisible pour le lecteur de la carte.



Figure 2-12 Les arbres peu détaillés sont correctement alignés le long de la route en les projetant sur une ligne décalée de la route (@OpenStreetMap contributors).

Enfin, nous avons développé un algorithme pour corriger les intersections entre un chemin, ou une route, ou une piste cyclable, et un lac (opération *Adjust Shape*). Pour chaque ligne intersectant un lac, nous calculons de quelle côté de la ligne le lac se trouve : la petite partie du lac se trouvant de l'autre côté est enlevée, alors que côté lac, nous enlevons une petite bande correspondant à la largeur du symbole plus une petite marge pour bien différencier le chemin du lac (Figure 2-13b). Mais certaines fois, une intersection chemin/lac correspond à un pont, et ne devrait donc pas être enlevée (Figure 2-13c). Pour éviter ces problèmes, nous avons développé un algorithme de détection de ces ponts par des critères géométriques, ce qui permet de ne pas ajuster la forme du lac au niveau des ponts (Figure 2-13d).

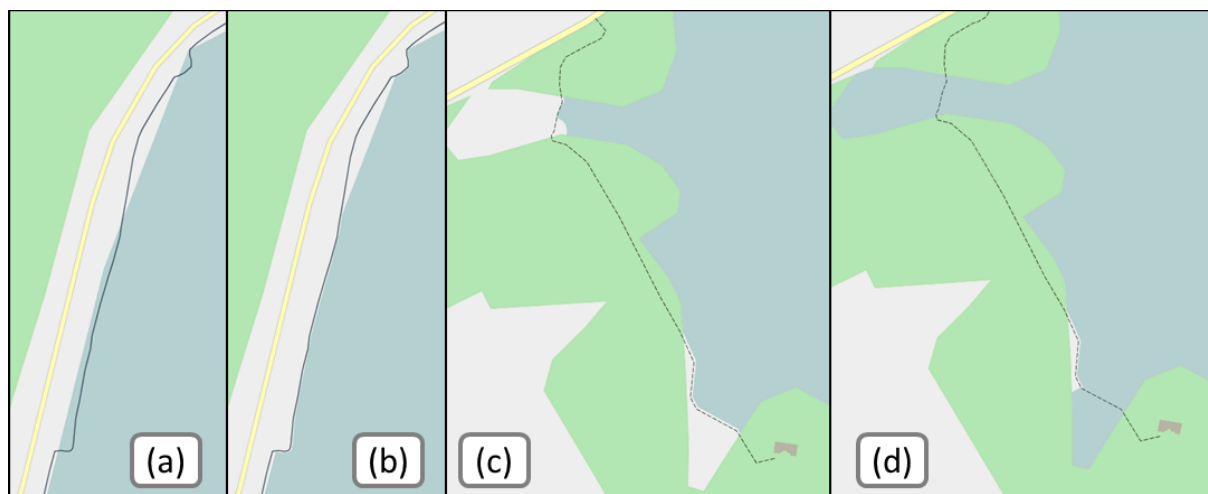


Figure 2-13 (a) un chemin détaillé intersecte un lac peu détaillé (b) harmonisation du lac (c) harmonisation d'un lac sans tenir compte d'intersections probables (les ponts) (d) harmonisation améliorée avec détection des ponts (@OpenStreetMap contributors).

2.2.3 De l'algorithme au processus, parallèle avec la généralisation

Comme c'est le cas pour la généralisation cartographique, il n'est pas possible de se contenter de lancer, l'un après l'autre, les différentes opérations d'harmonisation de niveau de détail. Il faut les orchestrer pour **choisir les valeurs de paramètres** adaptées à chaque contexte géographique (les seuils pour étendre les zones urbaines peuvent varier en fonction de la densité du tissu urbain par exemple), **déterminer le bon ordre** d'application des opérations, **gérer la propagation** des transformations générées par les opérations, pouvoir **changer de niveau d'analyse** en agrégeant par exemple des incohérences pour les traiter en une fois, et éventuellement autoriser des **fonctionnements par essai/erreur**, où une autre opération est essayée quand la première a échoué.

Nous proposons évidemment de s'inspirer de la recherche en généralisation qui a effectué un tel passage de l'algorithme au processus à la fin des années 90, à la suite du modèle proposé par Ruas et Plazanet (1996). Pour illustrer l'intérêt de s'inspirer des processus de généralisation, nous avons expérimenté l'utilisation d'un processus de déformation par moindres carrés (L. E. Harrie 1999), (Sester 2005) pour changer de niveau d'analyse et traiter plusieurs incohérences par groupes de bâtiments proches (Figure 2-14). Des bâtiments ici identifiés comme peu détaillés chevauchent le lit d'une rivière identifiée comme très détaillée. Il faut donc déplacer les bâtiments hors de la rivière tout en évitant les collisions avec les bâtiments proches. Les bâtiments sont donc regroupés par proximité (étape 2 de la Figure 2-14), puis chaque groupe est harmonisé séparément, mais les bâtiments d'un même groupe sont harmonisés ensemble par ce processus d'optimisation par moindres carrés pour éviter la propagation des incohérences (étape 3 de la Figure 2-14).



Figure 2-14 (1) détection des incohérences de LoD, ici des bâtiments peu détaillés chevauchant la rivière. (2) des groupes de bâtiments proches sont créés autour des incohérences identifiées. (3) chaque groupe est harmonisé d'un bloc en enlevant le chevauchement sans en créer de nouveaux avec les bâtiments voisins (bâtiments harmonisés en rouge) (@OpenStreetMap contributors).

Ces algorithmes et processus d'harmonisation automatique du niveau de détail de données collaboratives contribuent à la brique de base *Généralisation* du processus de cartographie automatique à partir de données collaboratives (Figure 1-16), même si dans le cas de la dérivation de cartes à grande échelle, les opérations d'harmonisation remplacent la généralisation dans le processus global.

Ces travaux sur la détermination du niveau de détail et son harmonisation ont été publiés avec plus de détails dans (Touya et Brando-Escobar 2013), (Touya et Reimer 2015) et (Touya et Baley 2016).

2.3 Formalisation des connaissances en cartographie

2.3.1 Connaissances nécessaires à un processus de généralisation automatique

Nous avons vu dans le chapitre 1 que la formalisation des connaissances cartographiques et des spécifications de la carte était une étape importante pour l'automatisation des processus cartographiques, particulièrement en cas de besoin utilisateur spécifique, mais même dans le cas d'une production en série de cartes comme le fait traditionnellement l'IGN.

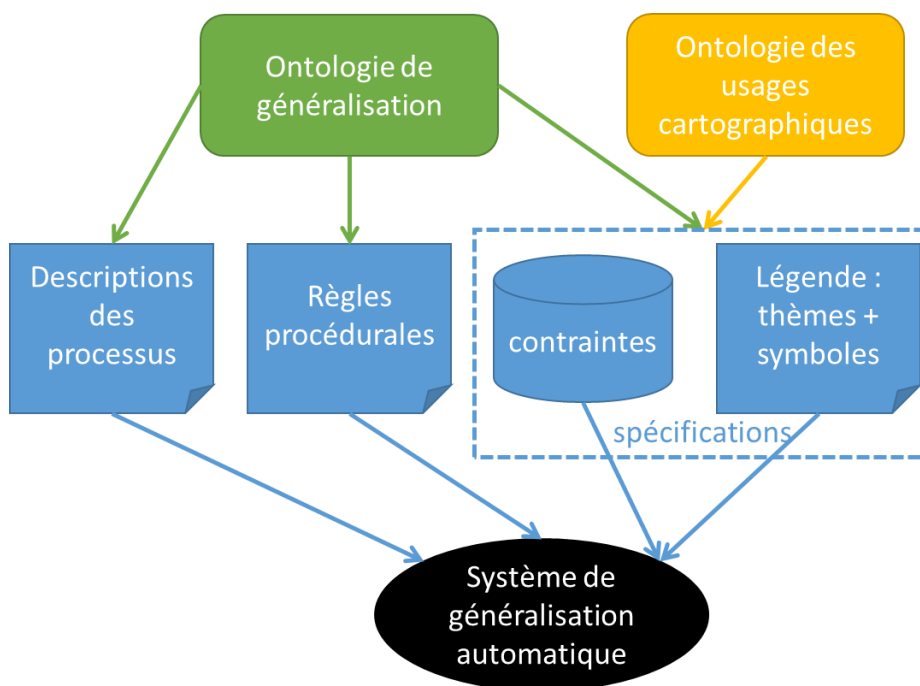


Figure 2-15 Modélisation des connaissances nécessaires à un système de généralisation automatique : des ontologies et des spécifications et règles dérivées pour une carte donnée à partir de ces ontologies.

Différents travaux passés (Regnauld 2007), (Balley et Regnauld 2011), (Touya, Duchêne et Ruas 2010), résumés par la Figure 2-15, ont listés les types de connaissances nécessaires au fonctionnement d'un processus de cartographie automatique, capable de réaliser des cartographies à la demande :

- *Spécifications* : il s'agit de définir des contraintes formelles (Touya, Duchêne et Ruas 2010) interprétables par différents processus automatiques de cartographie, mais aussi une

légende formelle qui spécifie les thèmes de la carte et leur symbolisation (Christophe, Dumenieu, et al. 2016).

- *Description des processus* : c'est un classique des problèmes d'interopérabilité, il faut décrire les capacités des processus automatiques disponibles pour pouvoir les orchestrer de manière générique.
- *Connaissances procédurales* : il s'agit des connaissances des cartographes (« on généralise les routes avant les bâtiments ») qui permettent de guider les processus automatiques (Buttenfield et McMaster 1991), ainsi que des connaissances liées au bon fonctionnement de processus particuliers (Taillandier, Duchêne et Drogoul 2011), comme par exemple le bon paramétrage de contraintes pour optimiser la généralisation d'une ville avec AGENT (A. Ruas 1999), (Barrault, et al. 2001).
- *Ontologies* : pour alimenter ces modèles formels en vocabulaire contrôlé, il faut se reposer sur des ontologies décrivant les processus cartographiques et les usages de la carte.

Différents projets de recherche et collaborations ont été initiés pour formaliser ces connaissances et les partager entre chercheurs en cartographie, et deux sont présentés dans les parties suivantes.

2.3.2 Vers une ontologie des relations spatiales

Parmi les connaissances à intégrer dans une ontologie de la généralisation telle que celle présentée en Figure 2-15, les relations spatiales sont fondamentales. En effet, elles transmettent la géographie implicite dans la carte et doivent être représentées aux mieux. Expliciter et formaliser ces relations spatiales dans une ontologie est intéressant pour plusieurs raisons :

- Faciliter leur prise en compte dans les processus automatiques, notamment en généralisation (Duchêne, Ruas et Cambier 2012).
- Permettre la formalisation des contraintes de généralisation qui portent souvent sur la préservation ou la transformation de ces relations (Steiniger et Weibel 2007).
- Permettre le référencement de données thématiques par relation spatiale (Jaara, Duchêne et Ruas 2013), (Duchêne 2014).
- Faciliter la qualification intrinsèque des données VGI (voir partie 2.2.1).

Une première étape de formalisation est de réaliser une typologie des relations spatiales dans les cartes. Il existe dans la littérature un certain nombre de typologies des relations spatiales qui varient par leur point de vue ou leur application, qui n'est pas nécessairement la cartographie, et on peut retrouver un recensement de ces typologies dans Steiniger et Weibel (2007) et Burghardt et al (2010). On peut voir un aperçu de la diversité de ces relations spatiales dans la Figure 2-16.

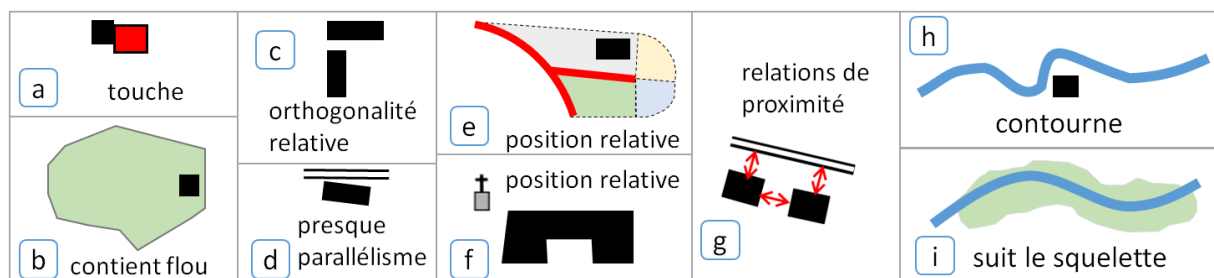


Figure 2-16 Relations spatiales de différents types, que l'on peut modéliser dans une carte.

Pour construire une ontologie des relations spatiales, il faut se situer au bon niveau de formalisation (Clementini et Laurini 2008) : la Figure 2-16 montre qu'il ne faut pas se contenter d'un niveau géométrique (inclusion d'une ligne dans un polygone par exemple) car la sémantique des objets en

relation joue un rôle dans la caractérisation de ces relations. Nous avons donc proposé les concepts clés d'une ontologie des relations spatiales, ainsi que leurs propriétés (Figure 2-20). Une relation spatiale est donc composée de plusieurs objets spatiaux (dont la nature et la primitive géométrique sont plus ou moins spécifiées). Elle est décrite par une propriété, qui peut elle-même être quantifiée par une opération : par exemple, la relation de proximité entre bâtiments de la Figure 2-16g est décrite par une propriété de distance entre les deux membres, et cette propriété peut être quantifiée par la distance euclidienne minimum entre deux polygones.

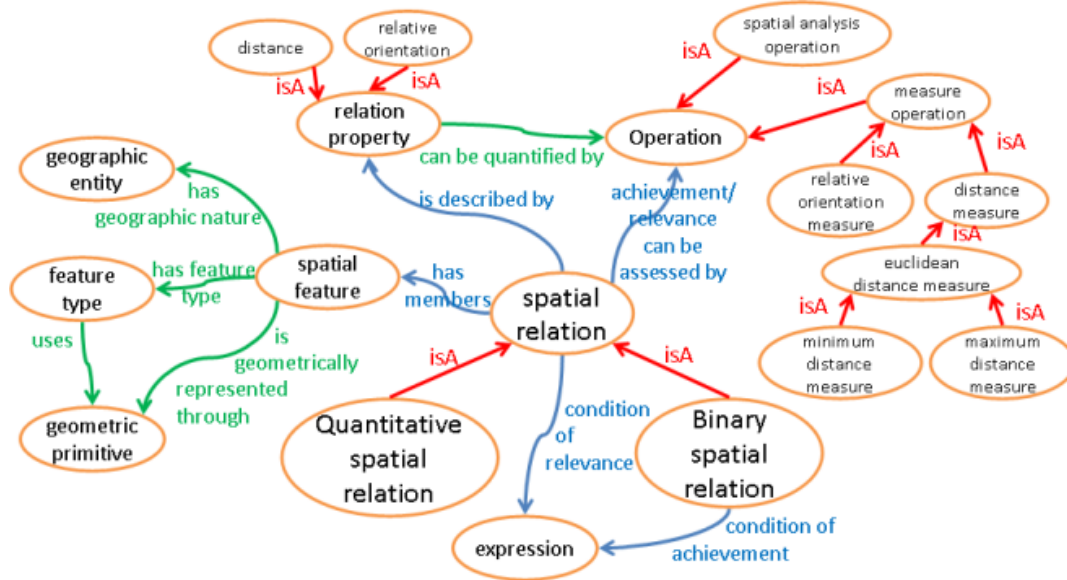


Figure 2-17 Modèle proposé pour la constitution d'une ontologie des relations spatiales dans les cartes avec quelques exemples de spécialisation pour certains des concepts.

A partir de cette ontologie des relations spatiales, nous avons également proposé une ontologie des contraintes de généralisation portant sur des relations spatiales, qui synthétise les différentes opérations que peuvent subir ces relations pendant un processus de généralisation : préservation de la relation, caricature de la relation, transformation d'une relation en une autre, et non création d'une relation qui n'existe pas en réalité (Figure 2-18).

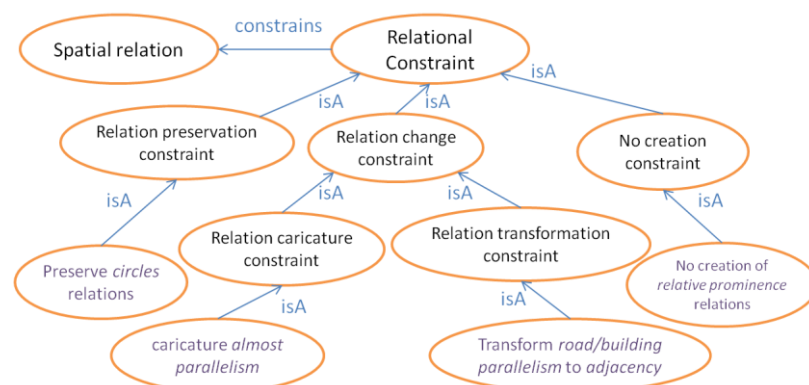


Figure 2-18 Modèle proposé pour la constitution d'une ontologie des contraintes « relationnelles » à partir de l'ontologie des relations spatiales.

Ces travaux de modélisation des relations spatiales sont issus d'une collaboration avec des chercheurs l'Ordnance Survey, l'IGN britannique, et ont été publiés avec plus de détails dans (Touya, Balley, et al. 2012) et (Touya, Bucher, et al. 2014).

2.3.3 Des ontologies partagées par les chercheurs en généralisation ?

Nous avons, comme d'autres (Gould et Chaudhry 2012), (Gould, et al. 2014), démontré que des ontologies pouvaient être utilisées avec succès pour automatiser des processus de cartographie (Touya 2011), (Touya et Girres 2013). Mais pour que ces ontologies soient réellement utiles au-delà du simple prototype de recherche, il faut qu'elles soient partagées par une grande partie des acteurs du domaine. Pour cela, nous avons participé au montage d'un atelier dont l'objectif était de construire de manière collaborative une ou des ontologies pour la généralisation et plus généralement la cartographie à la demande.

L'atelier a regroupé des chercheurs en généralisation cartographique de la commission de l'ACI⁴ sur la généralisation et la représentation multiple, mais aussi des chercheurs plutôt intéressés par la cartographie à la demande, et des chercheurs intéressés par l'utilisation d'ontologies en sciences de l'information géographique. Pour animer l'atelier, nous avons invité deux chercheurs de l'université de Manchester, spécialisés dans l'animation de réunion pour la conception d'ontologies. Les participants ont exprimé des objectifs différents pour l'utilisation d'ontologies :

- définir un vocabulaire partagé permettant de décrire les algorithmes, processus, contraintes de manière partagée, facilitant l'interopérabilité de ces processus et contraintes,
- formaliser les connaissances des cartographes pour permettre leur réutilisation alors que ces compétences se perdent,
- permettre des raisonnements logiques pour orchestrer des processus de cartographie.

L'atelier a été une succession de sessions de discussions en groupes ou avec tous les participants autour de différents utilisateurs possibles de ces ontologies imaginés en début d'atelier (Mackaness, Gould, et al. 2015). Les conclusions de l'atelier les besoins déjà identifiés et résumés dans la Figure 2-15, avec quatre types d'ontologies nécessaires au fonctionnement d'un système de cartographie à la demande :

- une ontologie du domaine modélisant les entités géographiques représentées dans la carte,
- une ontologie de la cartographie qui contient les connaissances nécessaires au choix de représentation des données,
- une ontologie des relations spatiales, à mi-chemin entre l'ontologie du domaine et l'ontologie de la cartographie,
- une ontologie des tâches d'utilisation d'une carte, qui modélise les besoins de l'utilisateur.

A la suite de cet atelier, nous avons construit une version initiale d'une ontologie pour la généralisation sur la plateforme webprotégé qui permet à plusieurs personnes de collaborer sur la construction d'une ontologie. Cette ontologie essaie de synthétiser les efforts précédents, essentiellement (Gould, et al. 2014) et (Touya 2011), avec les premières conclusions de l'atelier. Cette ontologie, appelée *GeneProcessOnto*, décrit les processus, les opérations et les algorithmes de généralisation, mais s'appuie pour cela sur des ontologies des entités géographiques pouvant apparaître de manière explicite ou implicite dans une carte, ce qui inclue l'ontologie des relations spatiales présentée dans la partie précédente. L'ontologie s'appuie aussi sur des ontologies de plus haut niveau (géométries, systèmes de coordonnées par exemple) utilisées dans le domaine de l'information géographique.

⁴ Association de Cartographie Internationale

Ces travaux de conception d'ontologies partagées ne peuvent se faire qu'en collaboration avec d'autres chercheurs du domaine et j'ai collaboré ici principalement avec Cécile Duchêne du COGIT, William Mackaness, Nick Gould, Nicolas Regnauld et Dirk Burghardt. Les ontologies partagées sont disponibles sur la plateforme webprotégé⁵, qui permet la construction collaborative d'ontologies.

En résumé, les travaux présentés dans cette partie 2.3 proposent des modèles formels pour exprimer les besoins de l'utilisateur et les spécifications de la carte de manière compréhensible pour un processus automatique. Ils contribuent donc à la brique de base *Formalisation* du processus global de cartographie automatique. Cette brique est d'autant plus importante qu'elle est présente dans l'ensemble des processus de cartographie.

2.4 Meta-modèles de généralisation

2.4.1 Opérateur, algorithme, modèle et meta-modèle de généralisation

Nous avons vu dans le premier chapitre que la recherche en généralisation a d'abord travaillé sur la conception d'algorithme permettant la généralisation d'un objet ou d'un groupe d'objets de la carte, comme par exemple la simplification du contour d'un bâtiment ou l'élimination des routes les moins importantes. Puis, les chercheurs se sont concentrés sur les moyens d'enchaîner automatiquement l'application de ces algorithmes, ou d'en appliquer plusieurs en même temps, comme le font par exemple le modèle AGENT (A. Ruas 1999) ou la généralisation par moindres carrés (L. E. Harrie 1999) (Sester 2005). Puis, j'ai essayé dans ma thèse d'enchaîner plusieurs de ces processus pour généraliser automatiquement l'ensemble d'une carte (Touya et Duchêne 2011). Cette partie reprend ces différents niveaux pour attaquer le problème de la généralisation cartographique pour mieux les définir.

Opération (ou opérateur) de généralisation : un type de transformation que l'on applique à un objet géographique pour réaliser sa généralisation. L'agrégation et la squelettisation sont deux exemples d'opérations de généralisation. Il existe différentes propositions de taxonomies de ces opérations, voir la partie 2.3.3.

Algorithme de généralisation : nous appelons algorithme de généralisation un processus qui réalise une, ou plusieurs opérations de généralisation, pour un type de géométrie (polyligne par exemple), et/ou pour un type d'objet géographique (les bâtiments par exemple), grâce à un programme informatique automatique. Par exemple, le classique algorithme de Douglas et Peucker (Douglas et Peucker 1973) réalise une opération de filtrage sur les polygones.

Modèle de généralisation : nous appelons modèle de généralisation un modèle conceptuel (éventuellement implémenté) qui permet d'enchaîner et/ou combiner différentes opérations de généralisations. La mise en œuvre d'un modèle de généralisation pour un cas d'application particulier est considérée ici comme un processus de généralisation. AGENT (A. Ruas 1999), CartACom (Duchêne, Ruas et Cambier 2012) ou les moindres carrés (L. E. Harrie 1999) (Sester 2005) sont des modèles de généralisation. AGENT permet l'enchaînement par essai/erreur de divers algorithmes de généralisation, quand les moindres carrés permettent de combiner des opérations de simplification, caricature et déplacement dans une optimisation globale sur l'ensemble des objets.

⁵ Plus de renseignements sur l'accès à cette ontologie sur le site web de la commission en Généralisation de l'ACI : <http://generalisation.icaci.org/index.php/generalisation-ontologies>

Processus de généralisation : nous appelons processus de généralisation un processus qui enchaîne et/ou combine différentes opérations de généralisation de manière automatique. Un processus peut instancier un modèle de généralisation, comme les processus issu du modèle AGENT, ou non.

Méta-modèle de généralisation : nous appelons un méta-modèle de généralisation un modèle de généralisation capable d'enchaîner/combiner plusieurs processus de généralisation, comme le modèle CollaGen développé dans ma thèse (Touya et Duchêne 2011).

2.4.2 Le ScaleMaster2.0

Cette partie décrit des travaux réalisés en collaboration avec Jean-François Girres lors de son post-doc que j'ai encadré de début 2012 à fin 2013. Nous nous plaçons dans le contexte de la dérivation automatique de plusieurs cartes/BDC à plusieurs échelles, à partir d'une ou plusieurs BDG, comme dans la Figure 1-15.

Le ScaleMaster de Brewer et Buttenfield (Brewer et Buttenfield 2007) (Brewer et Buttenfield 2009), est un modèle de représentation des connaissances pour la création de cartes à plusieurs échelles (Figure 2-19). Le ScaleMaster fonctionne comme une frise pour laquelle le temps serait remplacé par les échelles, de la plus grande à la plus petite. Chaque thème de la carte est représenté par une telle frise dans laquelle on définit les opérations à effectuer sur les données pour que ce thème soit bien représenté dans la carte finale. Par exemple, dans la Figure 2-19, pour représenter les zones d'hydrographie surfacique au 1 : 70K, il faut prendre les objets du DLM d'échelle de référence 1 : 50K et réduire la taille de leur toponyme. Ce ScaleMaster a été conçu pour réaliser des cartes manuellement et les opérations qu'il contient sont donc des opérations classiques pour un logiciel SIG.

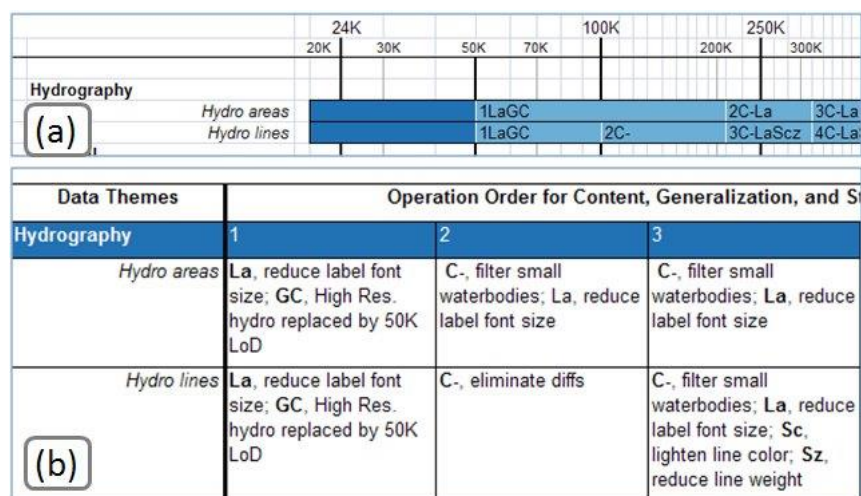


Figure 2-19 (a) extrait d'un ScaleMaster pour les thèmes hydrographiques, le bleu représente un des DLMs disponibles pour construire la carte. (b) signification des codes de ce ScaleMaster pour les opérations de filtrage, généralisation et symbolisation.

Si l'on souhaite utiliser le ScaleMaster pour représenter des connaissances permettant de guider une dérivation automatique de cartes, plusieurs problèmes se posent. Tout d'abord, on doit pouvoir décrire dans le ScaleMaster des algorithmes de généralisation automatique, et même des processus de généralisation pour certains thèmes complexes, et il faut pouvoir spécifier leur paramètres pour une échelle donnée. De plus, si plusieurs algorithmes/processus doivent être appliqués à un thème, ce qui est très courant, on doit pouvoir spécifier l'ordre d'application des opérations. Enfin, il est souvent nécessaire d'enrichir les données initiales par des objets géographiques implicites (par exemple les ronds-points dans un ensemble de routes linéaires) qui permettent d'améliorer les

résultats de la généralisation, ou d'appliquer la généralisation sur ces objets (Mackaness et Edwards 2002). Automatiser le ScaleMaster nécessite donc la prise en compte de ces enrichissements et leur exécution avec des paramètres donnés.

Afin d'automatiser le ScaleMaster, le ScaleMaster2.0 a été modélisé en reprenant la structure en frise du ScaleMaster. La Figure 2-20 montre ce modèle dans un diagramme de classes UML. La classe centrale est le *ScaleMaster*. Un *ScaleMaster* est composé de lignes associées à un thème de la carte, appelées *ScaleLines* par analogie avec les frises (« timeline » en anglais). Une ligne est composée d'intervalles d'échelles, eux-mêmes d'un (ou plusieurs) *ScaleMasterElement*, élément contenant les connaissances sur la généralisation du thème pour cette intervalle d'échelles. Un élément peut nécessiter des enrichissements (par exemple, la construction de la topologie et la détection des ronds-points pour la généralisation des routes), avoir une sélection attributaire (on ne garde que les routes départementales ou nationales par exemple) et surtout possède une liste de processus ou algorithmes de généralisation à appliquer, avec les paramètres valables pour cette échelle.

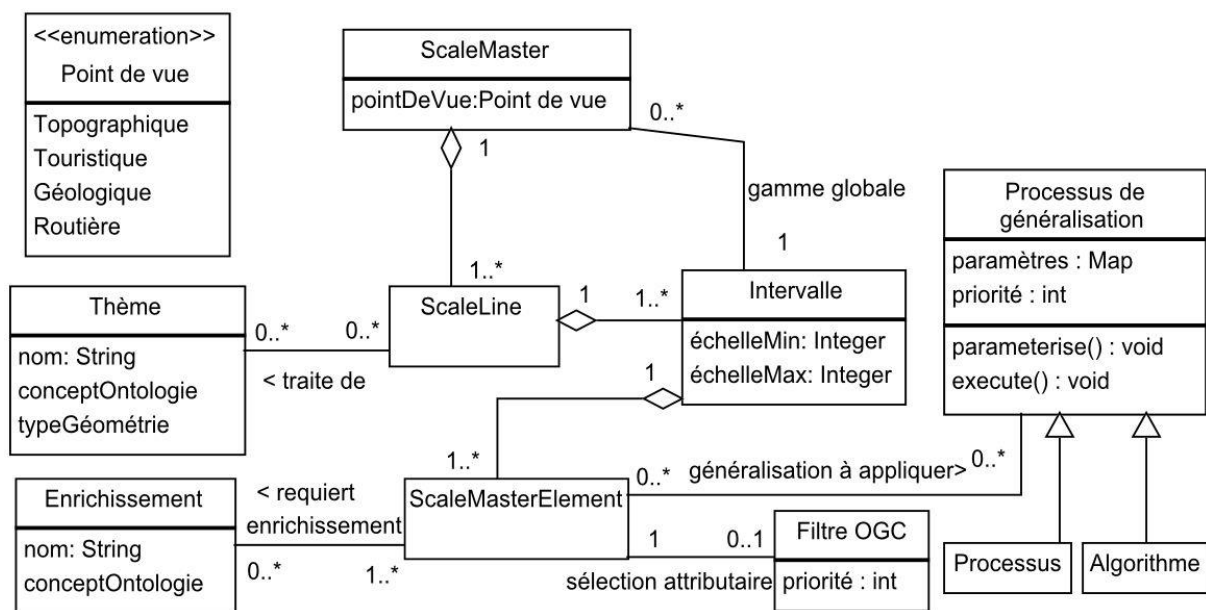


Figure 2-20 Diagramme de classe UML du ScaleMaster2.0.

Le fonctionnement du ScaleMaster2.0 est assez simple, et le déroulement de l'exécution pour une échelle donnée est décrit dans la Figure 2-21. Les thèmes sont généralisés ligne par ligne, de haut en bas. Pour une échelle donnée, on récupère le *ScaleMasterElement* correspondant et on en extrait les processus à appliquer (enrichissement, filtrage attributaire, processus de généralisation) et on les applique sur les données choisies (également décrites dans le *ScaleMasterElement*).

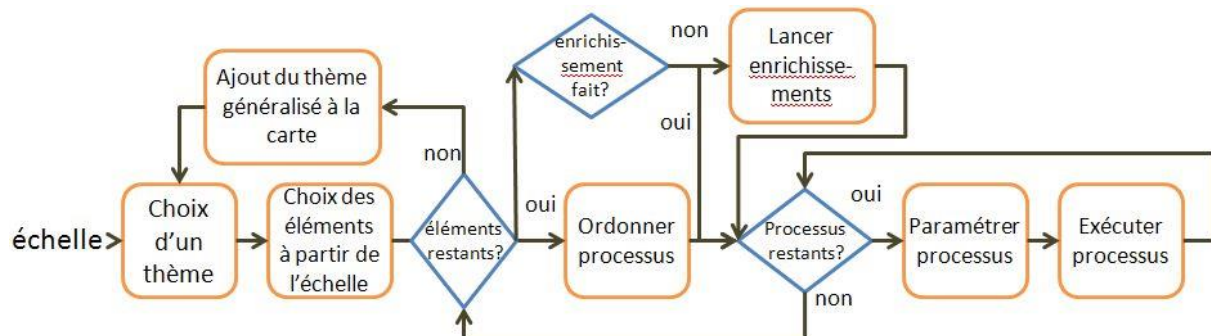


Figure 2-21 Diagramme des étapes de l'exécution automatique du ScaleMaster2.0 pour une échelle donnée.

Le système est également capable de s'appliquer avec un partitionnement préalable des données puis chaque partition est généralisée séparément (et éventuellement parallèlement), puis les données sont fusionnées après généralisation. Un partitionnement par mailles carrés a été testé dans notre prototype.

Le ScaleMaster2.0 a été implémenté dans un prototype sur la plateforme open source du COGIT, GeOxygene (Grosso, Perret et Brasebin 2012), et des tests ont été réalisés à différentes échelles à partir de trois DLMs dont les échelles de références correspondent à peu près au 1 : 50K, 1 : 250K et 1 : 1M. La Figure 2-22 montre des résultats à différentes échelles pour des cours d'eau, alors que la Figure 2-23 montre la généralisation de divers thèmes sur Dakar au 1 : 100K.

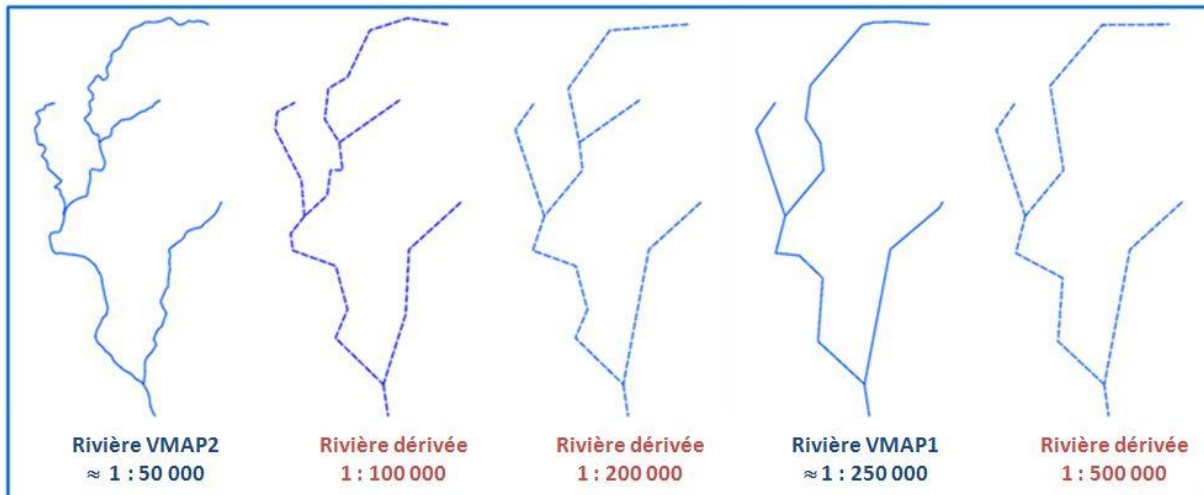


Figure 2-22 Résultats à de la généralisation de cours d'eau à différentes échelles en partant de deux DLMs (VMAP2 et VMAP1).

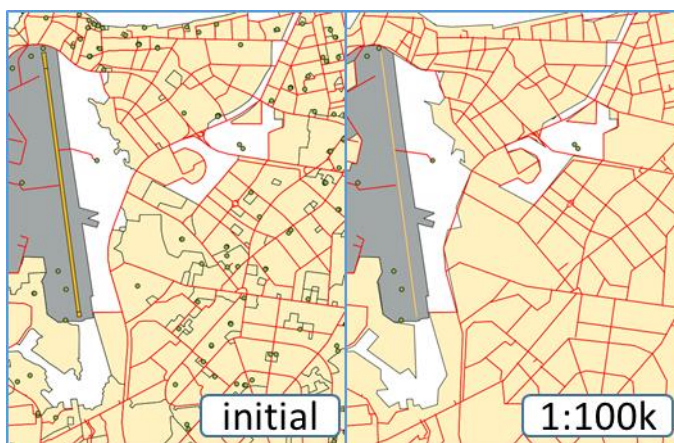


Figure 2-23 Résultat de généralisation par ScaleMaster2.0 au 1 : 100000 sur Dakar avec des routes, points d'intérêt, zones bâties, un aéroport et une piste d'aéroport.

Si les résultats sont très prometteurs, la mécanique du ScaleMaster2.0 ne peut autoriser des processus plus complexes, fonctionnant par itérations et essai/erreur, comme nous le faisons généralement pour obtenir de très bons résultats de généralisation, notamment pour les cartes à grande échelle. Ce méta-modèle est donc a priori réservé à des dérivations de cartes à moyenne et petite échelle (le cas d'étude démarrant au 1 : 50K convient bien). Par ailleurs, le paramétrage de tous ces processus est très complexe, et le ScaleMaster2.0 serait plus pratique à utiliser s'il pouvait être contrôlé par des contraintes cartographiques, plus proches de ce qu'un concepteur de carte

connait de son résultat attendu : il est plus facile de dire qu'on veut une résolution géométrique supérieure à 0,1 mm sur la carte, que de dire qu'on veut utiliser l'algorithme de Visvalingam-Whyatt avec un paramètre de 250 m².

En résumé, le ScaleMaster2.0 est un modèle permettant à la fois de formaliser des connaissances sur la généralisation de cartes multi-échelles, mais aussi d'orchestrer des processus de généralisation à l'aide de ces connaissances. Il contribue ainsi aux trois briques de base *Formalisation*, *Généralisation*, et *Symbolisation* (même si cette dernière n'est pas détaillée dans ce mémoire) pour un processus de création automatique de cartes multi-échelles (Figure 1-15).

Ces travaux de conception du ScaleMaster2.0 et d'amélioration du modèle ont été publiés avec plus de détails dans (Touya et Girres 2013) et (Girres et Touya 2014).

Chapitre 3. Amélioration des processus de généralisation automatique

Même si de nombreuses recherches ont été fructueuses dans le domaine ces dernières années, menant notamment à la mise en production de processus automatiques dans divers IGNs européens (Duchêne, Baella, et al. 2014), il existe encore de nombreux verrous pour améliorer les processus de généralisation automatique. Ce chapitre présente différents travaux de recherche s'inscrivant dans cet objectif. La première partie décrit des travaux sur la modélisation multi-niveaux pour améliorer les modèles de généralisation multi-agents précédemment développés dans notre équipe. La deuxième partie décrit la dérivation de cartes à des échelles intermédiaires dans une pyramide de cartes entre différentes échelles préexistantes.

3.1 Modélisation multi-niveaux dans la généralisation par agents

Cette partie reprend les travaux de thèse d'Adrien Maudet, thèse que j'ai encadrée, et qui a été co-dirigée par Sébastien Picault du Laboratoire d'Informatique Fondamentale de Lille et Cécile Duchêne du COGIT.

3.1.1 Généralisation cartographique et modélisation multi-niveaux

Nous avons vu dans le premier chapitre que parmi les approches pour automatiser la généralisation cartographique les chercheurs de l'équipe COGIT de l'IGN, comme d'autres, ont historiquement privilégié l'approche pas à pas locale. Cette approche cherche itérativement le bon espace de travail (objet de la carte ou groupe d'objet) et cherche une solution locale à ce sous-problème par une séquence optimale d'algorithmes de généralisation. Cette approche est particulièrement propice à l'utilisation de systèmes multi-agents (SMA). Les SMA sont un des sous-domaines de l'intelligence artificielle. En vulgarisant les différentes définitions, nous pouvons dire qu'un agent est une entité informatique autonome capable de percevoir son environnement, de communiquer, de réfléchir et d'agir pour atteindre ses objectifs ; un système multi-agents est un système informatique composé de plusieurs agents. Dans le cadre de la généralisation cartographique, un agent est une entité de la carte (un bâtiment ou un tronçon de route par exemple) dont l'objectif est d'être lisible et bien généralisé à une échelle donnée ; il est capable de percevoir ses entités voisines et ainsi d'évaluer sa lisibilité (est-ce qu'un bâtiment empiète sur le symbole de la route voisine par exemple), de communiquer avec ces voisins, de choisir la meilleure séquence d'algorithme pour être lisible et de lancer l'exécution de ces algorithmes qui vont transformer l'agent et ses voisins.

Plusieurs modèles suivant ces principes ont été proposés de manière complémentaire au COGIT : AGENT (A. Ruas 1999), (Barrault, et al. 2001) dédié à la généralisation des villes et des routes de montagne (Figure 3-1), CartACom (Duchêne 2004), (Duchêne, Ruas et Cambier 2012) dédié aux zones rurales (Figure 3-1), et GAEL (Gaffuri 2008) dédié à la gestion des champs déformables comme le relief. Ailleurs, d'autres modèles de généralisation par SMA ont été proposés (Baeijs, Demazeau et Alvares 1996), (Galanda et Weibel 2002), (Li, Zhou et Jones 2002), (Jabeur 2006), (Sabo, et al. 2008), (Zhang et Guilbert 2011), (Guilbert 2016), (Yan, Guilbert et Saux 2016).



Figure 3-1 Résultats automatiques au 1 : 50 000 des processus AGENT et CartACom obtenus sur la plateforme GeOxygene. Les données « avant » sont issues de la BD Topo, BDG de l'IGN, symbolisées pour le 1 : 50 000.

La diversité des données représentées dans les cartes topographiques amène une diversité dans les relations multi-niveaux qu'il est nécessaire de modéliser dans de tels SMA pour la généralisation :

- **Relations hiérarchiques arborescentes** : il s'agit d'agents se regroupant en agents de niveau supérieur pour une prise de décision plus centralisée (les bâtiments se regroupent en îlots dans AGENT), ou d'agents qui se décomposent en plusieurs agents de niveau inférieur (le relief triangulé se décompose en sous-agents points dans GAEL).
- **Relations transversales** : il s'agit d'agents voisins d'autres agents de même niveau, et partageant par exemple une relation spatiale (comme un bâtiment au bout d'une impasse). Ces relations sont déjà gérées dans CartACom.
- **Relations de situation** : il s'agit d'agents situés sur ou dans un autre agent sans qu'il y ait de relation hiérarchique entre eux, comme des arrêts de bus situés sur les routes (Figure 3-2b).
- **Relations non arborescentes** : il n'est pas toujours possible de modéliser les relations hiérarchiques de manière arborescente car des agents peuvent appartenir à plusieurs groupes de même niveau, comme un bâtiment appartenant à deux alignements de bâtiments (Figure 3-2c).
- **Relations diagonales** : dans le cas où plusieurs hiérarchies non arborescentes cohabitent, il peut y avoir des relations entre agents haut niveau mais appartenant à des hiérarchies différentes, relations que nous appelons diagonales. Par exemple, deux bâtiments collés forment un groupe autonome chargé du maintien de leur adjacence durant la généralisation des bâtiments, mais ce groupe appartient aussi à un îlot s'il est situé en ville (Figure 3-2a).
- **Relations de voisinage** : des agents peuvent se trouver dans le voisinage d'autres agents et être asservis par ces agents. Par exemple, un bâtiment le long d'une route devrait pouvoir suivre le mouvement si cette route est déplacée ou déformée, pour qu'il se situe toujours le long de la route.

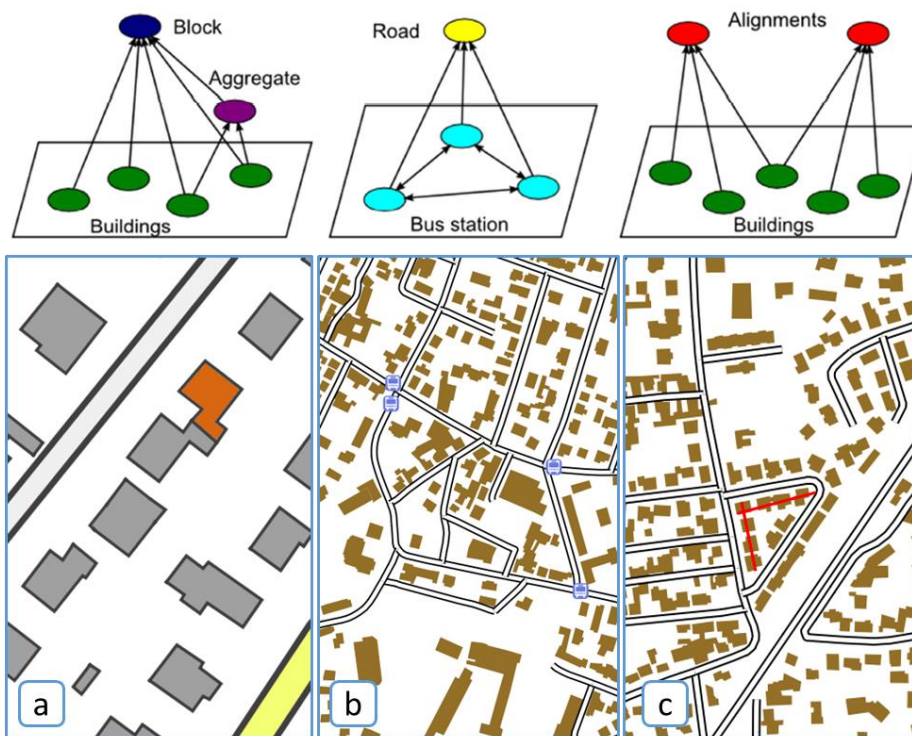


Figure 3-2 Exemples de relations multi-niveaux complexes. (a) relation diagonale entre un îlot et un groupe de bâtiments collés ; (b) des arrêts de bus « situés » sur les routes ; (c) un bâtiment appartenant à deux alignements (Maudet, Touya, et al. 2013).

Parmi toutes ces relations multi-niveaux possibles entre agents pour la généralisation, seules les deux premières, qui sont aussi les plus courantes, sont gérés par les trois modèles développés au COGIT : AGENT, CartACom et GAEL. Il est donc nécessaire de trouver un moyen pour intégrer les autres types de relation multi-niveaux dans un processus de généralisation.

Par ailleurs, ces trois modèles SMA du COGIT ne sont pas complètement intégrés car ils ont été développés de manière relativement indépendante. Ainsi, un agent bâtiment en ville ne peut pas avoir à la fois les comportements d'un agent hiérarchique d'AGENT et d'un agent CartACom capable d'interactions transversales. Si le modèle CollaGen (Touya 2011), développé pendant ma thèse permet d'utiliser ces modèles de manière conjointe sur des zones différentes, une fusion de ces modèles permettrait d'avoir des agents capables d'interactions hiérarchiques et transversales, et capables de se décomposer en agents points. La fusion permettrait de gagner en expressivité et en généricité pour les connaissances de guidage de ces processus, ce qui faciliterait leur réutilisation. La fusion de ces trois modèles est donc un autre objectif des travaux de thèse d'Adrien Maudet.

3.1.2 Le modèle DIOGEN

L'hypothèse, ou l'intuition de départ, est que le modèle de simulation multi-niveaux PADAWAN (Picault, Mathieu et Kubera 2010) peut servir de modèle pour unifier AGENT, CartACom et GAEL, et surtout pour intégrer les autres relations multi-niveaux entre agents pour la généralisation. PADAWAN repose sur les principes de IODA (Kubera, Mathieu et Picault 2011), c'est-à-dire que tous les éléments du système sont considérés comme des agents dont le comportement est simplement décrit par les interactions possibles entre chacun d'eux. A ces principes, PADAWAN ajoute la possibilité pour certains agents d'encapsuler des environnements dans lesquels les agents peuvent évoluer, et leur comportement peut s'adapter en fonction des environnements dans lesquels ils se situent. En SMA, un environnement décrit l'espace dans lequel les agents évoluent et comment ils

peuvent percevoir leurs voisins (Mathieu, Picault et Secq 2014). Dans notre cas, l'environnement principal est simplement la carte et sa métrique. Mais d'autres espaces géométriques présents dans la carte peuvent aussi devenir des environnements. Ainsi, un îlot urbain encapsule un environnement, soit le polygone délimité par les routes entourant l'îlot, et les agents situés géométriquement dans l'îlot, comme des bâtiments, appartiennent à cet environnement. Si on considère les bâtiments alignés de la Figure 3-2c, l'alignement est un agent encapsulant un environnement dans lequel sont inclus les bâtiments alignés, qui sont également inclus dans l'environnement de l'îlot.

Dans les trois modèles multi-agents développés au COGIT, le principe est relativement similaire : les agents ont des contraintes de lisibilité et de préservation de leurs caractéristiques, qu'ils doivent satisfaire au mieux. Les actions sont pilotées par les contraintes non satisfaites qui sont capables de proposer une liste d'algorithmes de généralisation pour les satisfaire. Comme on le voit dans la Figure 3-3, PADAWAN ne fonctionne pas de la même manière, car il n'y a pas de contrainte pour piloter le choix des actions à effectuer : un agent dispose d'une liste d'interactions possibles avec les autres agents, qui se déclenchent dans certaines conditions. Par exemple, un loup déclenche l'action « manger » s'il a faim et s'il se retrouve à proximité d'un agent mouton.

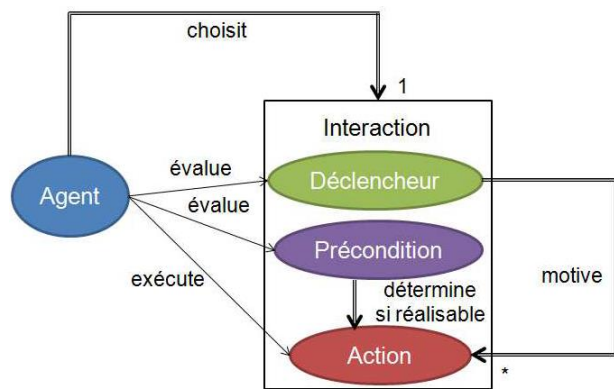


Figure 3-3 Modélisation des interactions entre agents dans PADAWAN (Maudet 2016).

Pour déclencher des actions de généralisation en suivant les principes de PADAWAN, le modèle DIOGEN d'A. Maudet propose une inversion de la perspective de décision des agents de généralisation (Figure 3-4). Ce n'est plus la contrainte qui déclenche et détermine l'action, mais l'agent lui-même. Comme le choix de la bonne action dépend du contexte, principalement traduit par la satisfaction des contraintes, la prise de décision de l'agent s'appuie sur l'opinion que peut avoir chaque contrainte sur la bonne action à effectuer.

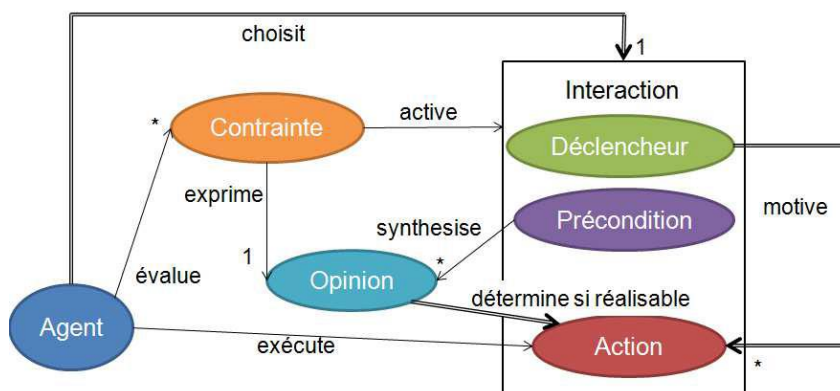


Figure 3-4 Modélisation des interactions entre agents et de leurs contraintes dans DIOGEN (Maudet 2016).

Si ces changements laissent entrevoir de fortes possibilités d'extension des capacités de généralisation de nos modèles, la première étape est d'abord de s'assurer que DIOGEN est capable de reproduire les résultats de ces modèles au moins aussi bien. DIOGEN a ainsi été configuré pour fonctionner à la manière d'AGENT dans les zones urbaines, et nous avons pu vérifier que DIOGEN est bien capable de répliquer le modèle AGENT avec des résultats principalement identiques (Figure 3-5), et même de temps en temps meilleurs : meilleure satisfaction globale des contraintes, ou résultat similaire obtenu plus rapidement. Le même test de réplcation a été réalisé avec CartACom et donne les mêmes résultats, mais n'a pas été réalisé avec GAEL, car nous ne disposions pas d'une implémentation fonctionnelle sur la plateforme de recherche CartAGEN (Renard, Gaffuri et Duchêne 2010) sur laquelle a été réalisé ce travail.



Figure 3-5 Réplication par DIOGEN du fonctionnement d'AGENT (Maudet, Touya, et al. 2013).

Une fois que l'on s'est assuré de la capacité du modèle DIOGEN à faire aussi bien que les modèles précédents pris individuellement, il s'agit de vérifier qu'il permet bien d'aller au-delà, notamment en couplant des capacités issues de ces trois modèles. Par exemple, nous souhaitons intégrer les interactions entre deux agents de CartACom, dans la généralisation urbaine d'AGENT. DIOGEN nous permet bien de le faire comme le montrent les exemples ci-dessous. Dans le premier exemple, nous avons intégré des relations entre deux impasses trop proches à l'intérieur d'un îlot urbain, et l'interaction possible entre ces deux impasses est une action qui fait « glisser » les impasses sur la route qui les porte, pour les éloigner l'une de l'autre (Figure 3-6). Une deuxième relation spatiale est introduite ici, celle entre une impasse et les bâtiments desservis par cette impasse : cela permet de faire suivre les bâtiments quand une impasse glisse pour garantir la préservation de la relation de desserte.



Figure 3-6 Intégration d'une relation de voisinage dans la généralisation AGENT des villes (Maudet, Touya, et al. 2013).

Le deuxième exemple est l'introduction d'une relation de parallélisme parfait entre les murs d'un bâtiment et la route la plus proche, quand il y a un « presque parallélisme » (Figure 3-7). L'action ici est une simple rotation du bâtiment, mais ce comportement typique de CartACom n'était pas intégrable dans AGENT jusque-là, et laisse espérer une forte amélioration des résultats de généralisation automatique des cartes topographiques grâce à DIOGEN.



Figure 3-7 Intégration de capacités d'AGENT et CartACom dans DIOGEN : les bâtiments de l'îlot peuvent aussi communiquer avec les routes voisines pour s'aligner avec elles (Maudet 2016).

3.1.3 Application aux cartes de randonnées à la manière des plans de bus

En plus des classiques cartes topographiques, nous avons considéré que DIOGEN était très adapté à dérivation automatique de cartes dans lesquelles des données thématiques sont plaquées sur un fond topographiques, car ces cartes recèlent de nombreuses relations multi-niveaux entre les données thématiques et les données du fond de carte. Nous nous sommes particulièrement intéressés aux cartes de randonnées sur lesquelles des itinéraires sont ajoutés sur les routes et chemins de la carte topographique. La Figure 3-8 montrent deux manières de réaliser ces cartes d'itinéraires, celle des cartes de randonnée actuelles de l'IGN où il y a un seul symbole magenta, et la différenciation d'itinéraires empruntant les mêmes voies se fait par la texte ; et la manière plus schématique, inspirée des plans de réseaux de transport comme les bus ou les métro, dans lesquelles une couleur est associée à chaque itinéraire, et les itinéraires sont décalés pour être tous visibles quand ils empruntent les mêmes voies. C'est ce dernier cas d'étude que nous avons décidé d'étudier. Il pose deux problèmes principaux :

- Comment optimiser le placement des itinéraires le long des routes ?
- Comment modifier la carte topographique pour ne pas perdre des informations sous les symboles d'itinéraire ?

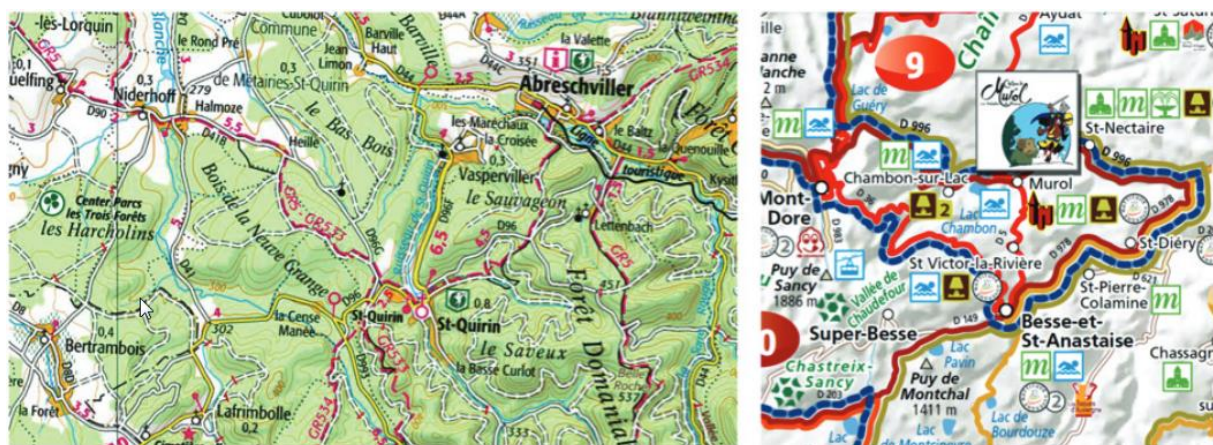


Figure 3-8 Deux manières de représenter plusieurs itinéraires sur une carte : un symbole mais plusieurs toponymes à gauche, et plusieurs symboles adjacents à droite (Teulade-Denantes, Maudet et Duchêne 2016).

Pour le premier problème, nous avons décidé de le modéliser sous la forme d'un problème de satisfaction de contraintes (Bellicha, et al. 1994), les contraintes étant par exemple la minimisation des croisements de symboles lorsqu'un itinéraire prend une voie différente, la minimisation des croisements avec les routes ou chemins, ou encore la maximisation de la continuité visuelle de chaque itinéraire (Teulade-Denantes, Maudet et Duchêne 2016). La Figure 3-9 montre les bons résultats de placement obtenus avec cette méthode. Elle est intégrée à DIOGEN sous la forme d'une action déclenchée au démarrage de la généralisation, et le modèle adapte ensuite le reste de la carte à ce placement optimal initial, se laissant la possibilité de modifier ponctuellement le positionnement d'un itinéraire si les objets de la carte situés en-dessous ne peuvent pas être déplacés sans générer trop de problèmes de lisibilité dans la carte.

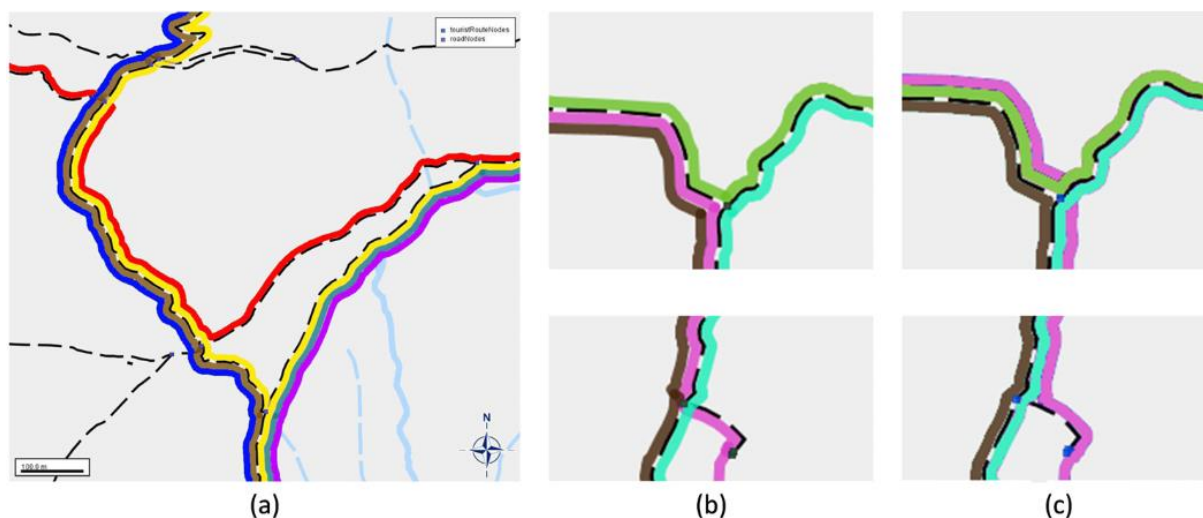


Figure 3-9 Résultats du placement automatique des itinéraires de randonnée. (b) et (c) montrent deux bonnes solutions différentes obtenus avec des poids différents sur la contrainte de minimisation des croisements (Teulade-Denantes, Maudet et Duchêne 2016).

Pour notre deuxième problème, nous faisons donc appel aux capacités de DIOGEN à modéliser des interactions multi-niveaux entre les agents de la carte. Ici, ces interactions reposent sur deux types de relations multi-niveaux :

- La relation entre un itinéraire et les différentes voies (routes ou chemins) qui « supportent » cet itinéraire.
- La relation entre chacun des objets ou groupes d'objets situés au bord de la route et cette route.

Ces relations, lorsqu'elles se retrouvent insatisfaites suite au placement initial des itinéraires, ou à la généralisation des données topographique du fond de carte, entraînent des actions de généralisation correctives, comme le déplacement groupé des bâtiments qui se retrouvent sous le symbole d'itinéraire dans la Figure 3-10. Dans ce cas, le déplacement est possible car il y a de la place au-dessus, mais si cela n'avait pas été possible, la décision aurait pu être de mettre l'itinéraire de l'autre côté sur cette portion de route.



Figure 3-10 Les bâtiments en relation de voisinage avec la route sont déplacés quand l'itinéraire change de côté (Maudet 2016).

Dans certains cas, les problèmes posés par le placement des itinéraires concernent des parties d'un objet et non un objet en entier, ou un groupe d'objets. C'est le cas du cours d'eau partiellement recouvert par l'itinéraire dans la Figure 3-11. Dans ce cas, DIOGEN permet des interactions entre l'itinéraire et les agents points (réplication du fonctionnement de GAEL) composant le cours d'eau, et ceux-ci se déplacent pour résoudre le conflit tout en préservant la forme du cours d'eau, et d'éventuelles relations de l'autre côté du cours d'eau (Figure 3-11).

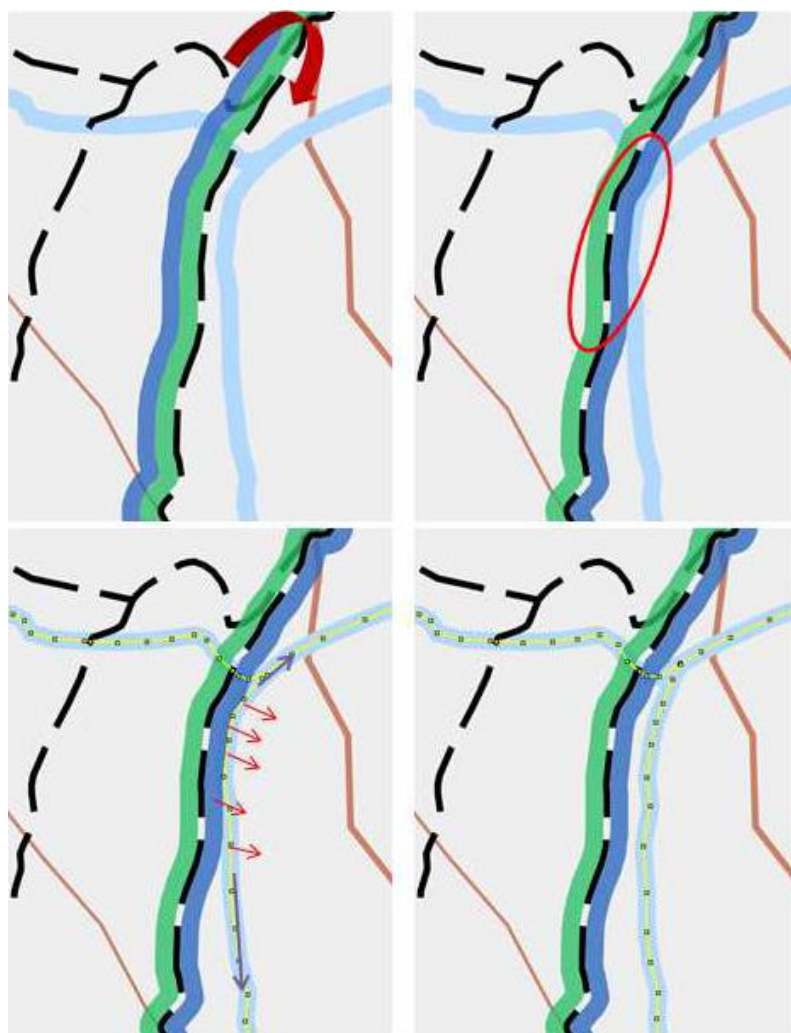


Figure 3-11 Quand l'itinéraire change de côté, les agents points du cours d'eau en relation avec le chemin sont activés et se déplacent pour préserver leur lisibilité et la forme générale du cours d'eau (Maudet 2016).

En résumé, grâce à la modélisation d'interactions multi-niveaux variées, DIOGEN nous permet de généraliser automatiquement des cartes topographiques classiques par la

mutualisation des capacités de nos précédents modèles SMA, et de nouvelles formes de cartes couplant des données thématiques et des données topographiques de support. Toutes les questions ne sont pas encore résolues, notamment s'agissant de l'orchestration des agents dans DIOGEN, c'est-à-dire l'optimisation de l'ordre d'activation des agents. Pour l'instant, cette orchestration est réalisée ad hoc dans les prototypes que nous avons développés, l'utilisation de systèmes de vote, ou de décision multicritères doivent être envisagés pour rendre générique cette orchestration.

Ces travaux sur la construction de cartes de randonnées à la manière des plans de bus sont décrits avec plus de précision dans (Teulade-Denantes, Maudet et Duchêne 2016) et (Maudet 2016).

3.1.4 Patterns d'analyse pour les SMA multi-niveaux

Les interactions multi-niveaux spécifiques mises en œuvre à l'aide du modèle DIOGEN nous permettent également d'exhiber des interactions plus génériques entre agents de niveaux différents qui pourraient se retrouver dans d'autres systèmes multi-agents multi-niveaux. Nous appelons ces interactions génériques des *patterns* d'analyse pour les SMA multi-niveaux. Une formalisation est proposée pour chacun de ces patterns (Maudet, Touya, et al. 2015), et nous reprenons ici deux exemples de ces formalisations pour les discuter et en montrer l'intérêt.

Pattern décomposition

Le pattern décomposition est issu de l'utilisation de DIOGEN pour reproduire les comportements du modèle GAEL (Gaffuri 2008), dans lequel des agents peuvent se décomposer en agents points contrôlant les points de la géométrie de niveau supérieur. Cela permet notamment de déformer localement une géométrie tout en préservant la forme générale, et est utilisé principalement pour la déformation du relief dans les cartes topographiques. Dans le pattern décomposition (Figure 3-12), on part donc d'un état initial avec un agent quelconque qui se décompose en plusieurs agents de niveau inférieur dépendants de lui et hébergés dans sa géométrie. Ces agents sont dépendants car ils n'ont de sens que par l'existence de leur hôte, contrairement aux agents bâtiments qui ont un sens même hors d'un éventuel agent hôte de type îlot.

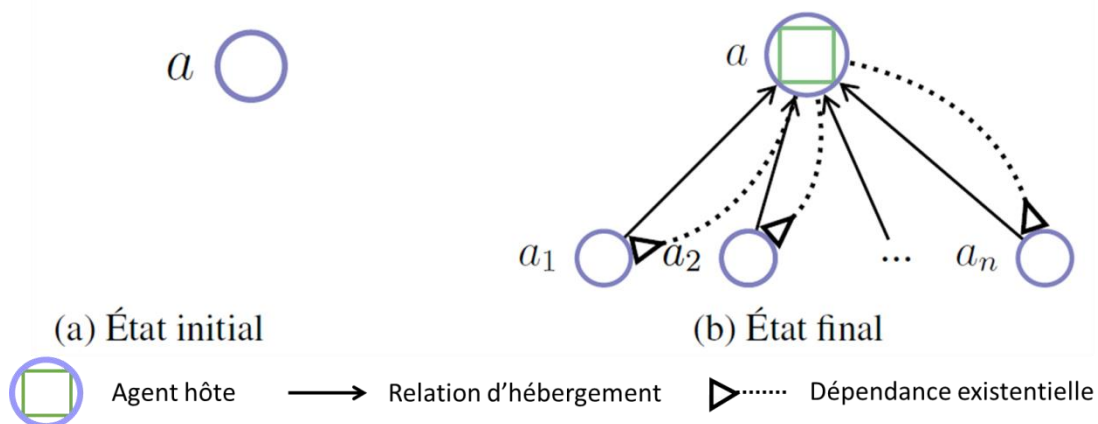


Figure 3-12 Formalisation du pattern décomposition : l'agent devient hôte et héberge de nouveaux agents dépendants (ils n'existent que parce que *a* existe).

Ces travaux sur la formalisation de patterns multi-niveaux sont décrits avec plus de précision dans (Maudet, Touya, et al. 2015) et (Maudet 2016).

3.2 Génération de niveau intermédiaires dans une pyramide d'échelles

Cette partie reprend les travaux de thèse, encore en cours, de Marion Dumont. Dans ce travail, nous nous intéressons à la dérivation de cartes multi-échelles pour des applications web de cartographie (voir partie 1.3) comme le géoportail⁶ de l'IGN ou GoogleMaps⁷ (Figure 3-14).

Les travaux de thèse de Marion Dumont se situent dans le contexte du projet MapMuxing (ANR-14-CE24-0011-01) démarré fin 2014 et dont je suis le porteur. Le projet MapMuxing a pour objectif d'améliorer l'efficacité de la navigation dans un portail web cartographique, lors des changements fréquents de représentation cartographique, que ce soit par un zoom, ou par un passage de la carte à une ortho-photographie. Pour cela, nous proposons une collaboration inédite avec des chercheurs en interactions homme-machine, de l'équipe de recherche ILDA de l'INRIA Saclay, pour combiner des techniques d'interactions avec des techniques de cartographie automatique. Enfin, le dernier membre du consortium du projet est l'UMR GRED de Montpellier qui associe au projet des géographes spécialisés dans la gestion des risques naturels. L'application clé du projet est d'améliorer la navigation dans ces portails dans le contexte de la gestion des risques de tsunamis aux Antilles. Dans ce projet, les chercheurs de l'IGN sont particulièrement chargés de la conception de méthodes permettant la construction de continuums de représentations cartographiques, soit des séries de représentations cartographiques intermédiaires entre deux extrêmes de représentation cartographique, qui peuvent être une carte topographique et une ortho-photographie, deux cartes avec des légendes et styles différents, ou deux cartes à des échelles différentes. Ces continuums peuvent alors être exploités par des interfaces innovantes sur une même représentation (principe des loupes ou lentilles par exemple) ou animant la transition (Pietriga, Bau et Appert 2010) (Pindat, et al. 2012) (Alvina, et al. 2014).



Figure 3-14 Pyramide multi-échelles de l'IGN appelée ScanExpress, avec une certaine continuité dans les légendes et le style.

La thèse de Marion Dumont s'inscrit donc dans l'objectif de constitution de continuums cartographiques entre cartes d'échelle différente d'une pyramide multi-échelles, ce qui passe par la génération de niveaux intermédiaires entre les niveaux existant dans cette pyramide. Il s'agit

⁶ <http://geoportail.gouv.fr>

⁷ <http://maps.google.com>

réellement d'une nouvelle manière de construire des cartes, et les connaissances que nous avons accumulées sur la généralisation automatique de cartes topographiques de l'IGN, imprimées sur papier, ne sont plus forcément valides, et le premier verrou de la thèse est de réapprendre comment construire une bonne carte dans ce contexte, notamment en testant nos propositions auprès d'utilisateurs. Dans ce contexte, la cohérence des niveaux entre eux devient aussi importante que la lisibilité des niveaux pris individuellement. L'autre verrou de la thèse est donc de déterminer quelles sont les meilleures stratégies de généralisation pour assurer cette cohérence ou continuité entre les échelles, car rien ne dit qu'il s'agisse des mêmes stratégies de généralisation que pour une carte statique.

3.2.1 Etude des portails cartographiques multi-échelles existants

La construction de ces continuums cartographiques entre deux échelles pose beaucoup de questions car nous n'avons pas encore de connaissances sur les bonnes pratiques pour faire un bon continuum d'échelles, c'est-à-dire une bonne pyramide multi-échelles. Pour enrichir ces connaissances, et faire émerger des bonnes pratiques, nous avons étudié l'existant, c'est-à-dire les applications actuelles de cartographie permettant la navigation dans une pyramide multi-échelles. Nous avons identifié 16 applications : principalement des portails de diffusion des cartes de NMA, mais aussi des portails comme GoogleMaps ou OpenStreetMap (Dumont, Touya et Duchêne 2015). Nous avons étudié ces applications selon divers points de vue :

- A quel point les informations d'échelle fournies par ces applications sont-elles dépendantes de la résolution et de la taille de l'écran de l'utilisateur ?
- Puisque la plupart de ces applications remplissent les niveaux de zoom avec des cartes papier scannées, à quels niveaux de zoom (et donc échelle d'affichage) sont utilisés les cartes avec une échelle de symbolisation donnée ?
- Quel contenu est affiché à quelle échelle dans ces cartes ?

Pour répondre à la première question, chaque application a été testée avec trois écrans différents pour vérifier que l'échelle affichée était bien changée quand la résolution ou la taille de l'écran changeait, à une ou deux exceptions près. Cela signifie pour nous que nous pouvons comparer les échelles d'affichage de chaque application si on le fait avec la même configuration d'écran, ce que nous avons mis en place pour la suite de cette étude. Par ailleurs, l'étude a permis de vérifier que toutes les applications suivaient la norme WMTS de l'Open Geospatial Consortium qui fixe l'échelle d'affichage associée à chacun des 21 niveaux de zoom de la norme (zoom 0 = 1 : 500 000 000 et zoom 20 = 1 : 100), avec quelques approximations dues aux projections cartographiques différentes dans chaque pays, ou à l'emprise de certains pays.

Si on exclue temporairement les applications qui ne sont pas construites à partir de cartes papier numérisées, comme OpenStreetMap ou GoogleMaps, on peut étudier pour chaque application et chaque niveau de zoom (et son échelle d'affichage associée), la carte papier utilisée et son échelle de symbolisation (« definition scale » dans la Figure 3-15). Dans cette figure, on remarque une tendance claire qui est l'utilisation d'une carte sur une plage d'échelle d'affichage allant de l'échelle de symbolisation à 1/3 de cette échelle de symbolisation. Du fait des limitations de la perception sur écran, qui réduit l'acuité visuelle par rapport à la lecture sur papier (C. Ware 2012), il est logique de ne pas afficher les cartes à des échelles d'affichage supérieures à l'échelle de symbolisation. Les quelques exceptions du graphique se situent à des échelles très petites, continentales, où peu d'information est affichée sur la carte en dehors des frontières, ce qui rend ces cartes relativement lisibles malgré tout. Les trois exceptions de l'autre côté sont des exemples de mauvaise utilisation d'une pyramide multi-échelles, avec des cartes papier numérisées extrêmement grossières et donc très

pixélisées. De manière générale, la Figure 3-15 montre des cas d'utilisations d'une même échelle de symbolisation à plusieurs échelles d'affichage et inversement de plusieurs échelles d'affichage pour une même échelle de symbolisation dans des pays différents (donc pas vraiment de consensus sur le contenu d'une carte à une échelle d'affichage donnée).

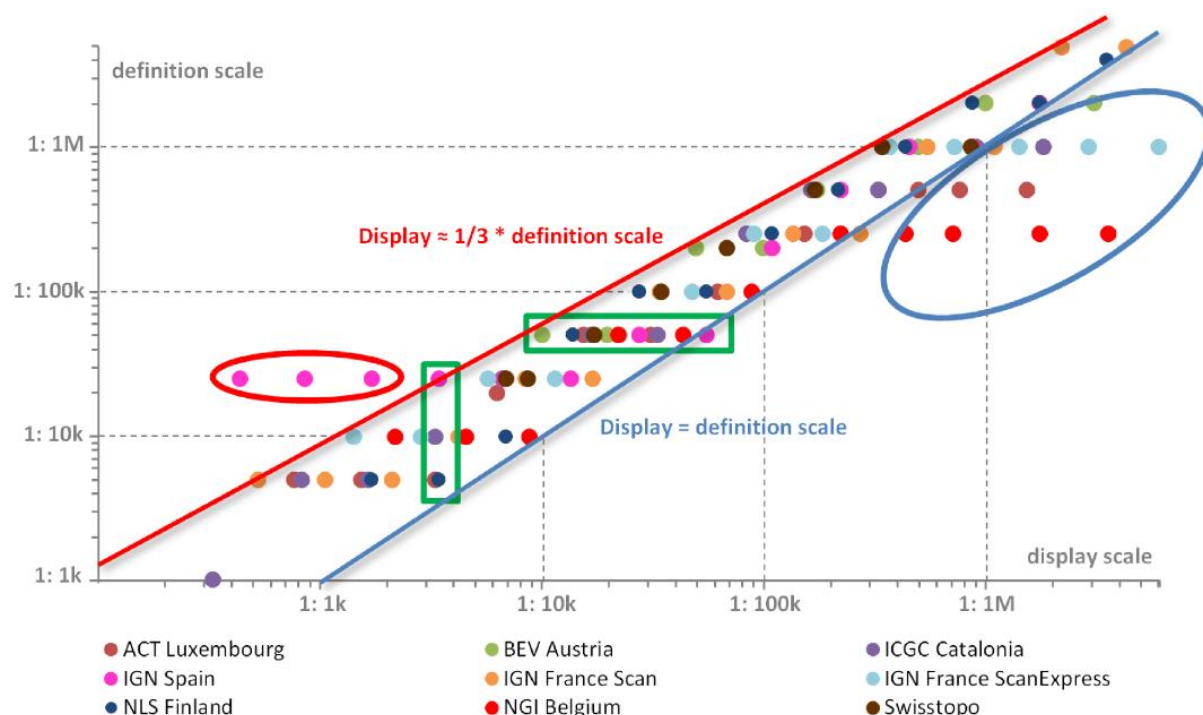


Figure 3-15 Distribution échelle d'affichage/échelle de symbolisation pour chaque zoom par application (une couleur par application).

L'étude a porté ensuite sur le contenu concret des cartes à chaque échelle d'affichage. Nous nous sommes concentrés sur deux des thèmes principaux de ces cartes, les thèmes « bâti » et « réseau routier ». Seuls les résultats sur le thème « bâti » sont présentés succinctement dans ce mémoire. Nous avons d'abord identifié les différents niveaux d'abstraction de représentation de ce thème dans les cartes à travers les échelles : représentation du bâti individuel aux grandes échelles, représentations des zones bâties dans chaque îlot urbain, représentation de la zone bâtie ou urbaine par un polygone, et enfin, représentation des villes par un symbole ponctuel (Figure 3-16). Chacune des applications utilise plusieurs de ces niveaux d'abstraction à travers ses échelles, sans nécessairement les utiliser toutes : par exemple, les cartes du géoportail de l'IGN n'utilisent pas le niveau îlot urbain.

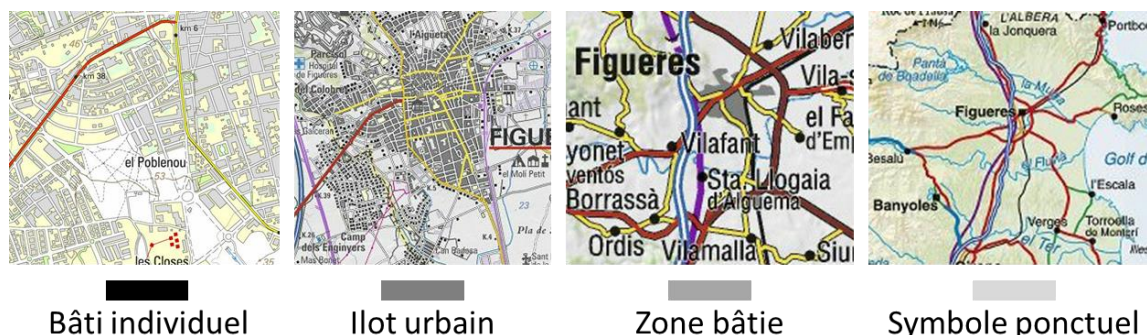


Figure 3-16 Les quatre niveaux d'abstraction identifiés pour la représentation du thème « bâti » dans les cartes multi-échelles étudiées.

Dans chaque niveau d'abstraction, nous avons également cherché s'ils étaient plus ou moins généralisés par élimination de certains objets, grossissement ou simplification des géométries. Nous avons donc collecté pour chaque application et chaque niveau de zoom les niveaux d'abstraction et de généralisation, et nous avons placé ces données sur une frise à la manière du ScaleMaster pour les visualiser (Figure 3-17). Cela permet de visualiser rapidement l'utilisation courante de plusieurs niveaux d'abstraction en même temps dans la même carte, ou encore les plages d'utilisation assez différentes de chaque niveau d'abstraction à travers les applications étudiées.

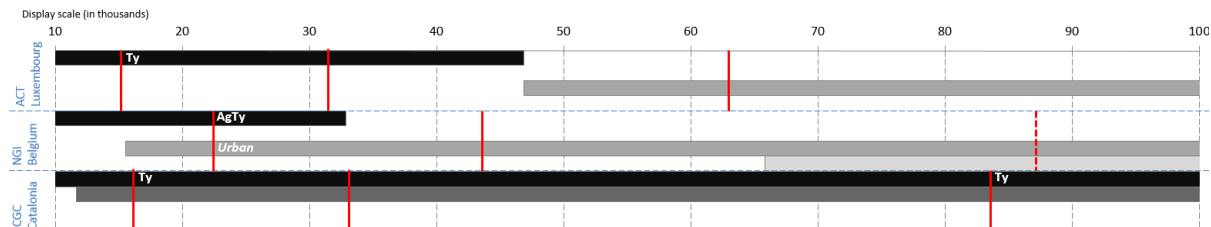


Figure 3-17 Extrait sur trois cartes multi-échelles de la frise réalisée pour comparer la répartition des niveaux d'abstraction sur la représentation du bâti. Les traits verticaux rouges représentent les échelles d'affichage de ces cartes multi-échelles.

Cette frise nous a permis de remarquer des informations intéressantes. Tout d'abord, il y a bien un ordre d'enchaînement entre les niveaux d'abstraction du thème bâti, et on peut définir pour chacun des plages d'utilisation courante. Nous avons également remarqué que les applications que nous avons jugées comme les plus efficaces utilisent l'affichage combiné de deux niveaux d'abstraction pour faciliter la transition de l'un à l'autre. La Figure 3-18 montre deux méthodes de combinaison identifiées : la coexistence où deux niveaux d'abstraction sont utilisés pour des espaces différents de la carte (par exemple l'îlot urbain en centre-ville et le bâti individuel en zone périurbaine), et la superposition où deux niveaux d'abstraction sont affichés l'un sur l'autre (par exemple le bâti individuel affiché sur la zone urbaine). Cette dernière remarque nous invite à utiliser les représentations combinées de niveaux d'abstraction pour construire nos niveaux intermédiaires censés faciliter la navigation dans la pyramide.

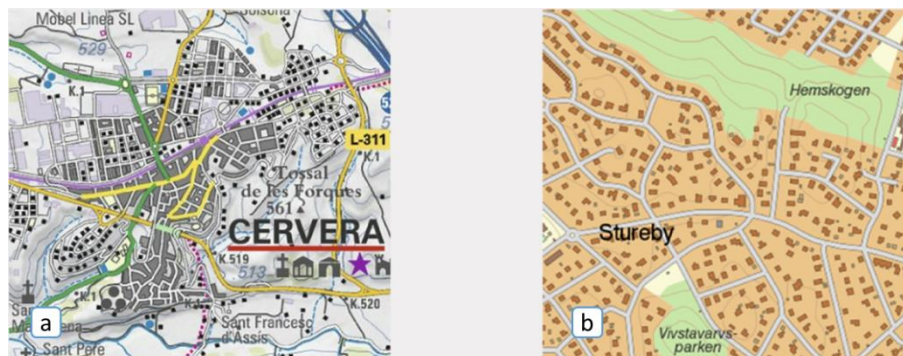


Figure 3-18 Deux types de combinaison de niveaux d'abstraction : (a) la coexistence où les niveaux d'abstraction sont utilisés à des endroits différents (b) la superposition où deux niveaux d'abstraction sont utilisés l'un sur l'autre.

Une des manières de d'assurer une transition fluide entre deux échelles est d'avoir des cartes dont le niveau de complexité est similaire. La complexité d'une carte représente la difficulté de lecture et de compréhension d'une carte (Jégou et Deblonde 2012) et recouvre des questions de complexité graphique (quantité d'information) et de complexité intellectuelle (difficulté à comprendre le sens des symboles utilisés). En vision par ordinateur, on utilise la notion de *clutter*, qui n'a pas vraiment de traduction française, mais se rapproche de cette notion de complexité de la carte. L'effet de *clutter* se manifeste dans une image lorsque la quantité d'information dans cette image est trop importante et/ou quand cette information est mal organisée, hiérarchisée (Rosenholtz, Li et Nakano

2007). Il existe un très grand nombre de mesures du clutter issues des recherches en vision par ordinateur (Moacdieh et Sarter 2014), qui pourraient être exploitées pour mesurer la complexité des cartes. Des travaux précédents ont d'ailleurs montré le potentiel de ces mesures, mais aussi leurs limites, lorsqu'elles sont appliquées à des cartes à échelles différentes, pour évaluer le niveau de généralisation de ces cartes (Touya, Decherf, et al. 2015) (Brychtová, Çöltekin et Paszto 2016).

Nous avons donc enregistré une capture d'écran de chaque niveau de zoom, pour chaque application, et nous avons appliqué les mesures de clutter déjà utilisées dans (Touya, Decherf, et al. 2015). La Figure 3-19 montre le premier enseignement prévisible de cette mesure du clutter : utiliser une même carte numérisée à deux niveaux de zoom consécutifs entraîne des variations du niveau de complexité, qui baisse quand on passe à une version grossie d'une carte. Pour avoir une complexité invariable, ou en tout cas plus monotone, il est donc préférable d'avoir plus de cartes que dans les applications actuelles.



Figure 3-19 Evolution du clutter à travers les niveaux de zooms : la complexité est nettement plus faible aux niveaux où on s'est contenté de grossir une carte adaptée à un autre niveau (Dumont, Touya et Duchêne 2016).

L'IGN produit deux pyramides multi-échelles différentes, avec des processus différents : l'une contient les cartes en papier numérisées (appelée plus loin « Scan IGN ») et l'autre est produite par des processus presque entièrement automatiques (appelée plus loin « Scan Express »). Le Scan Express contient des cartes à plus de niveaux de le Scan IGN, notamment au niveau 13 qui correspond au 1 : 50 000, car l'IGN ne produit plus de carte en papier à cette échelle. Nous avons donc comparé les valeurs de clutter obtenues sur ces deux pyramides, notamment autour du niveau 13 (Figure 3-20). On remarque clairement une meilleure continuité de la complexité pour la pyramide Scan Express, celle qui contient une carte spécifique au niveau 13 et pas un simple grossissement du niveau 12, ce qui tend à démontrer que des niveaux intermédiaires bien construits aident à préserver une complexité plus stable à travers les niveaux, ce qui aiderait par conséquent à fluidifier la navigation entre ces échelles, même si ce dernier point est encore une hypothèse qu'il faudra démontrer en approfondissant la question.

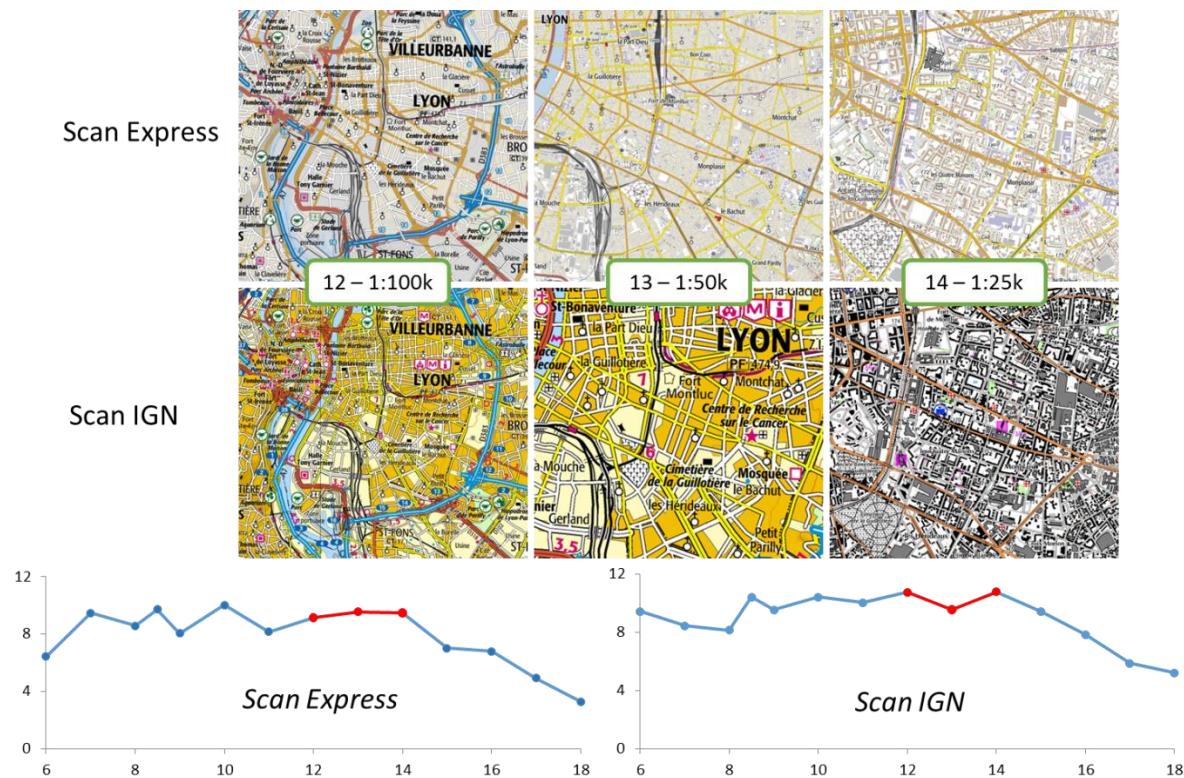


Figure 3-20 Deux pyramides produites par l'IGN qui diffèrent au niveau de zoom 13 avec une carte grossière pour la pyramide "Scan IGN" et une carte spécifique pour la pyramide "Scan Express" : la variation de clutter est plus faible pour le Scan Express.

En résumé, on peut d'abord remarquer qu'il n'existe pas encore de méthode idéale pour construire une carte multi-échelles, toutes celles consultées lors de cette étude présentent des défauts, soit par manque de cartes à certains niveaux de zooms, soit par l'utilisation de cartes peu adaptées à ce contexte, ou de mauvaise qualité, soit parce qu'il n'y pas assez de niveaux dans la pyramide. Ces résultats ne confirment pas directement notre intuition que de nouveaux niveaux intermédiaires amélioreraient la navigation de l'utilisateur dans ces cartes, mais il commence à y avoir un faisceau de présomption. L'utilisation conjointe de deux niveaux d'abstraction pour assurer une transition plus fluide entre ces deux niveaux est une idée prometteuse qui sera utilisée dans la suite de ces travaux de recherche.

Cette étude comparative des applications cartographiques multi-échelles a été publiée avec plus de détails dans (Dumont, Touya et Duchêne 2015) et (Dumont, Touya et Duchêne 2016) et plus de détails devraient être également disponibles dans une future publication de Marion Dumont.

3.2.2 Dérivation de niveaux intermédiaires entre deux niveaux préexistants

L'état des connaissances sur la génération de niveaux intermédiaires est très faible et nous souhaitons mieux connaître ce qui est efficace et ce qui ne l'est pas, pour améliorer la navigation d'un utilisateur. Pour cela, nous proposons de mettre en place un test auprès d'utilisateurs qui nous permettrait de confirmer ou infirmer nos hypothèses et de nous donner des premières réponses à nos questions :

- L'ajout de niveaux intermédiaires améliore bien l'expérience de navigation dans une pyramide multi-échelles, et améliore surtout l'efficacité de cette navigation.
- La transition entre niveaux d'abstraction est facilitée par l'ajout de niveaux intermédiaires contenant des représentations combinant ces niveaux d'abstraction.
- Comment mesurer la ressemblance ou la différence entre deux cartes avec des niveaux d'abstraction différents, dans l'objectif de maximiser cette ressemblance pour plus de fluidité ?
- Y a-t-il un nombre optimal de niveaux intermédiaires pour une pyramide multi-échelles pour l'efficacité de la navigation ?
- Quelle est la bonne quantité d'information pour échelle de définition/d'affichage donnée ? Et quelle est la meilleure méthode pour mesurer la quantité d'information dans la carte ?
- Quelle est la meilleure stratégie de généralisation pour améliorer la navigation dans une pyramide multi-échelles ?

Pour mettre en place un test permettant de valider ces hypothèses et de répondre, au moins partiellement, à ces questions, nous devons disposer d'un certain nombre de pyramides de cartes multi-échelles alternatives. Cela permet de comparer leur efficacité dans l'accomplissement de tâches similaires. Nous avons décidé de nous concentrer sur des stratégies alternatives sur le thème « bâti » de la carte, faisant suite aux travaux précédents de comparaison, portant déjà sur ce thème. La Figure 3-21 montre une vue simplifiée de la méthode établie pour constituer ces pyramides alternatives. L'idée est de construire un certain nombre de BDC ne traitant que certains des thèmes de la carte, à des échelles intermédiaires et de piocher dans ces BDC pour constituer nos pyramides avec des BDC contenant tous les thèmes de la carte. Nous avons volontairement réduit le problème au saut d'échelle entre le 1 : 25 000 et le 1 : 100 000, c'est-à-dire entre les niveaux de zoom 12 et 14. Il n'y a donc qu'un niveau intermédiaire dans la norme WMTS, mais nous nous autorisons à faire varier ce nombre de niveaux intermédiaires de 1 à 3.

Concernant la généralisation du thème bâti, nous souhaitons comparer trois stratégies :

- La stratégie dite « AGENT », dans laquelle nous utilisons les méthodes classiques du COGIT pour la généralisation des bâtiments dans une carte topographique, soit une combinaison des modèles AGENT (Barrault, et al. 2001) et CartACom (Duchêne 2004).
- La stratégie par typification, dans laquelle on cherche plus à préserver la structure et la répartition des bâtiments, plutôt que les bâtiments pris individuellement. Nous utilisons une version adaptée de l'algorithme de typification de Burghardt et Cecconi (2007).
- La stratégie par point de repère, dans laquelle les bâtiments sont remplacés petit à petit par des îlots « grisés » en laissant tout de même apparaître par-dessus les bâtiments servant de point de repère dans la carte (Elias 2003).

Nous avons vu dans la comparaison des géoportails décrite dans la partie précédente que des stratégies de combinaison des niveaux d'abstraction pouvaient faciliter la transition, ici entre le niveau d'abstraction « bâtiment individuel » et le niveau « zone bâtie ». Nous panachons donc ces stratégies avec les deux formes de combinaison de niveaux d'abstraction identifiées : dans la Figure 3-21, la pyramide 1 combine bâtiment individuel typifié et zone bâtie dans ses deux niveaux intermédiaires. Les BDC de nos pyramides sont complétées par des versions généralisées des autres thèmes de la carte (végétation, hydrographie, routes, etc.).

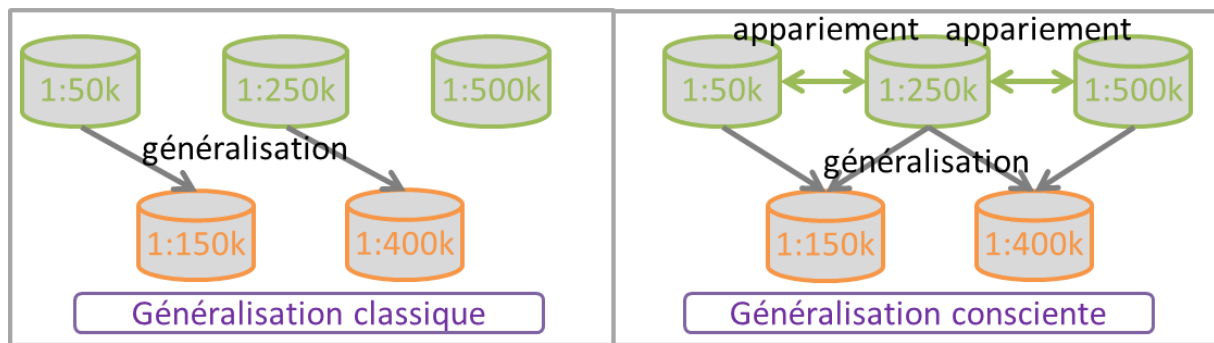


Figure 3-22 Différence entre la généralisation classique et la généralisation consciente des niveaux multiples, dans le cadre de la dérivation de niveaux intermédiaires.

Deux stratégies sont possibles pour utiliser tous les niveaux existants dans une généralisation consciente des niveaux multiples : utiliser les niveaux en entrée des processus de généralisation (stratégie de *pré-traitement*), et utiliser les niveaux en sortie du processus de généralisation pour le corriger (stratégie de *post-traitement*).

Il y a deux types de techniques de *pré-traitement* :

- Exclure de la généralisation les objets présents dans les autres niveaux, les représentations de ces objets dans les niveaux intermédiaires étant dérivées des autres niveaux.
- Modifier le fonctionnement des processus de généralisation pour les rendre eux aussi conscients des niveaux multiples. Ce type de technique est a priori le plus prometteur en termes de qualité des résultats, mais aussi le plus complexe à mettre en œuvre car il faudrait modifier beaucoup de processus déjà fonctionnels.

Il y a également deux types de techniques de *post-traitement* :

- Simplement ajouter dans le niveau intermédiaire, les objets présents dans d'autres niveaux, mais qui ont été éliminés par le processus de généralisation (Figure 3-23).
- Comme l'ajout d'objets dans le niveau intermédiaire peut augmenter trop fortement le nombre d'objets de la carte, cet ajout peut être compensé par un posttraitement qui élimine les objets les moins importants (non présents dans les autres niveaux) pour équilibrer la quantité d'objets.

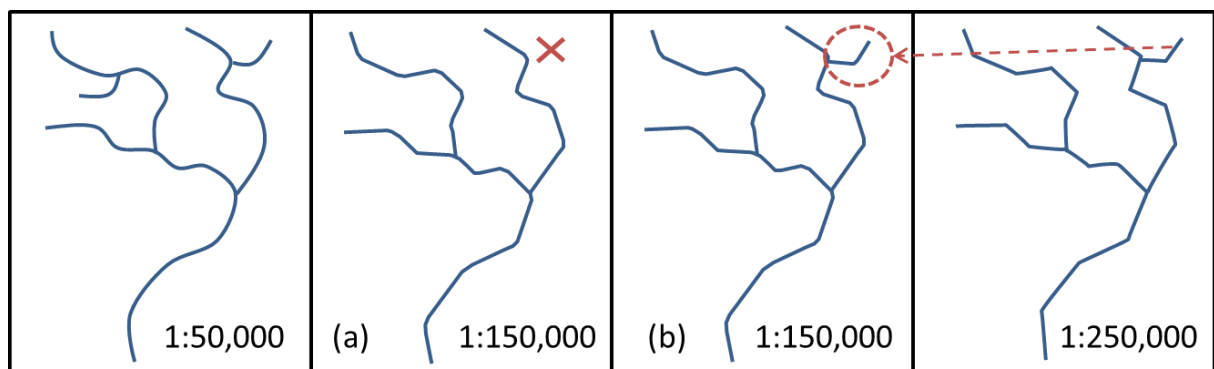


Figure 3-23 Stratégie de *post-traitement* : la rivière oubliée (a) est ajoutée après généralisation au 1 : 150k de la BDC au 1 : 250k (b).

La meilleure stratégie dépend de l'endroit où l'on souhaite placer le curseur entre complexité des développements et qualité des résultats. Pour aider à faire ce choix, nous avons réalisé divers tests

sur des données et des processus différents, avec les versions simples des deux stratégies. Par exemple, la Figure 3-24 montre un processus de sélection des routes au 1 : 100k à partir du 1 : 50k et du 1 : 250k, avec une stratégie simple de post-traitement. Les différents tests effectués montrent à chaque fois que la généralisation consciente est nécessaire et que même les stratégies simples améliorent fortement les résultats par rapport à une généralisation classique. Les tests comparatifs entre les deux stratégies, dans leur version simple, montrent aussi que la stratégie de post-traitement est meilleure que celle de prétraitement car cette dernière enlève une partie du contexte géographique nécessaire au bon fonctionnement des processus de généralisation (Girres et Touya 2014).

Intuitivement, il nous semble nécessaire de combiner les stratégies en fonction des thèmes généralisés et des processus utilisés, et de nouveaux tests plus poussés doivent être menés pour bien optimiser ce principe de généralisation consciente des niveaux multiples.

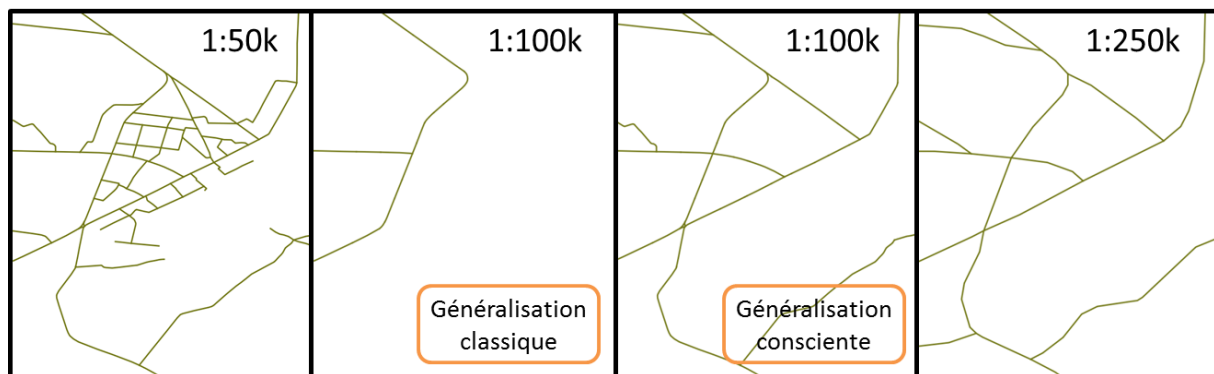


Figure 3-24 Stratégie de post-processing appliquée à un processus de généralisation du réseau routier au 1 : 100k.

Ces travaux sur la généralisation consciente des niveaux multiples ont été publiés avec plus de détails dans (Girres et Touya 2014) et seront étendus dans la fin du projet MapMuxing.

Chapitre 4. Conclusions et perspectives de recherche

Résumé des contributions pour l'automatisation des processus cartographiques

L'objectif général qui guide l'ensemble des travaux présentés dans ce mémoire est de concevoir des méthodes nous permettant à court ou moyen terme d'automatiser les processus de création d'une carte. Les contributions principales sont celles des deux doctorants que j'ai encadrés. Adrien Maudet a proposé le modèle DIOGEN qui permet de modéliser facilement des interactions multi-niveaux dans les SMA pour la généralisation cartographique. Nous l'avons appliqué à la fusion de nos trois précédents modèles SMA, mais aussi à un cas d'étude nouveau de cartographie à la demande : des cartes d'itinéraires de randonnée offrant à la fois une vue topographique du territoire et une vue schématique des itinéraires.

Marion Dumont, même si sa thèse est encore en cours, a étudié le problème de la généralisation de niveaux intermédiaires dans une pyramide multi-échelles pour améliorer la fluidité de navigation dans ces pyramides. Elle a extrait des premières bonnes pratiques pour la conception de tels niveaux intermédiaires et elle met en place un test de ses hypothèses par des utilisateurs en construisant ces niveaux intermédiaires de différentes manières pour les comparer. Plus globalement, le projet MapMuxing que je pilote s'éloigne de la production automatique de cartes à une échelle donnée pour s'intéresser à l'efficacité de cette carte dans un ensemble de représentations dans lesquelles un utilisateur peut naviguer. Si les premiers résultats du projet sont encourageants, nous espérons que les prochains mois de collaborations amèneront une synergie encore meilleure entre cartographie et interactions homme-machine.

Nous avons vu dans le premier chapitre de ce mémoire que la production de carte avait maintenant de nombreuses facettes, et nous avons présenté des contributions sur l'automatisation de la cartographie à la demande, grâce à des collaborations internationales sur la formalisation des connaissances en généralisation par constitution d'ontologies partagées. Enfin, nous nous sommes intéressés à la cartographie automatique à partir de données hétérogènes, notamment issues de saisies collaboratives, comme celles d'OpenStreetMap, en proposant des techniques de qualification du niveau de détail de ces données, et en proposant des techniques de transformation automatique de ces données en tenant compte de leur hétérogénéité (harmonisation et généralisation).

Besoins futurs en cartographie automatique

Il ne fait aucun doute que les tendances actuelles, décrites dans le premier chapitre, vont encore s'accroître :

- Les données géographiques vont continuer à être plus massives et mises à jour plus rapidement, et encore plus centrées sur les besoins de chaque utilisateur/fournisseur de données. La cartographie sera donc nécessairement de plus en plus automatique et sur mesure pour chaque utilisateur.
- Cette accentuation de la masse de données va s'accompagner d'une prédominance des données hétérogènes par rapport aux données classiques issues des NMAs. Il faudra donc être capable de cartographier ces données hétérogènes, les intégrer et les combiner dans une même carte, notamment en combinant des données collaboratives à des données de référence comme celles de l'IGN.

- Si nous sommes encore dans une période où il existe deux modes de production de cartes, sur papier et sur écran, nous allons vers une cartographie résolument multi-supports, ou « responsive », le mot à la mode en développement web, qui désigne des interfaces qui s’adaptent au support de diffusion (écrans de tailles différentes notamment). Plutôt que de concevoir des cartes pour chaque nouveau support qui apparaît, il faudra concevoir des cartes adaptatives.
- Enfin, le mode de diffusion actuel des cartes sur le web est le tuilage d’images de cartes. Ce mode de diffusion est très rapide et a permis la forte diffusion de cartes dans des applications web. Mais ce mode de diffusion limite fortement les interactions avec la carte, et on voit actuellement se développer l’alternative, le tuilage vectoriel (Figure 4-1). Avec le tuilage vectoriel, chaque objet de la carte est dessiné indépendamment permettant des changements faciles dans la légende, ou d’interroger la sémantique attaché à l’objet. Il faudra donc proposer des techniques de cartographie automatique adaptées à ce nouveau support de diffusion.

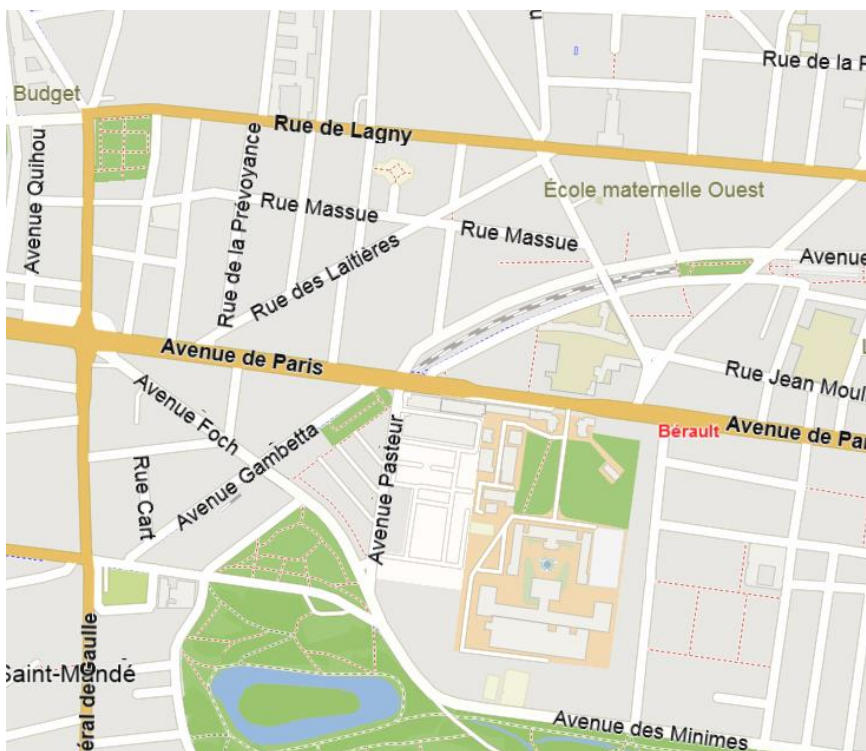


Figure 4-1 Extrait d'OpenScienceMap, plateforme de diffusion en tuilage vectoriel des données OpenStreetMap (@OpenStreetMap contributors).

Perspectives en généralisation cartographique

Il reste encore beaucoup de travail pour les chercheurs en généralisation cartographique pour s’adapter à cette évolution des besoins en cartographie. Tout d’abord, nos propositions pour améliorer la navigation dans les cartes multi-échelles, fondées sur la construction de niveaux intermédiaires, ne sont pas les seules. D’autres considèrent que la généralisation continue, soit la dérivation d’une représentation de la carte à n’importe quelle échelle vue comme une quantité continue, permettrait d’obtenir de meilleurs résultats (van Oosterom, Meijers, et al. 2014). Nous considérons qu’il n’est pas possible de « rendre continu » chaque changement de niveau d’abstraction, mais les techniques de généralisation continue pourraient sûrement améliorer nos résultats pour certains thèmes et certains sauts d’échelles, en les couplant avec nos méthodes de

génération de niveaux intermédiaires. Une collaboration avec Alexander Wolff de l'Université de Würzburg a été initiée pour creuser cette complémentarité, avec la visite au COGIT de son doctorant Dongliang Peng. Durant sa visite, il a travaillé avec nous sur la conception d'une méthode de transformation continue des bâtiments présents dans une carte au 1 : 25 000 vers un polygone de zone bâtie. Cette stratégie continue pourra être comparée aux stratégies plus « discrètes » proposées par Marion Dumont dans le cadre de sa thèse sur les mêmes données et le même saut d'échelle.

Une des évolutions possibles de la gestion des bases de données géographiques, qui sera peut-être précipitée par l'intégration de données hétérogènes dans ces bases, est la mise en place de base de données multi-représentations. Il s'agit de base de données géographiques, contenant plusieurs représentations à des niveaux de détail différents, liées entre-elles, de la même entité du monde réel. Le fait de produire des cartes à partir d'une base de données multi-représentation pose deux problèmes spécifiques. Le premier est celui de la mise à jour incrémentale, c'est-à-dire la propagation d'une mise à jour intégrée au niveau le plus détaillée, et propagée dans les autres représentations sans avoir à tout recalculer. Le deuxième problème est celui de la généralisation consciente, présentée dans la partie 3.2.3. Nous avons présenté des premières réflexions et expérimentations, qui devraient être complétées dans la fin du projet MapMuxing. Si l'utilisation des bases de données multi-représentations se développe, le besoin de telles techniques devra aussi se développer, et il faut donc dès maintenant travailler sur ce sujet.

L'équipe de recherche en généralisation du laboratoire à laquelle j'appartiens depuis douze ans, mais dont les premiers travaux remontent au milieu des années 90, a toujours eu comme objectif d'automatiser la production des cartes topographiques de l'IGN. Si l'évolution des besoins en cartographie a élargi la palette de nos objectifs, nous n'avons jamais perdu de vue cet objectif initial. Le modèle CollaGen issu de mes travaux de thèse, et le modèle DIOGEN issu de la thèse d'Adrien Maudet, et présenté dans le troisième chapitre de ce mémoire, ont posé des bases théoriques qui devraient nous permettre de viser la généralisation tout automatique des bases de données de l'IGN pour obtenir des cartes de qualité aux échelles du 1 : 25 000 et du 1 : 50 000. Le principal travail est ici d'exploiter les capacités de ces modèles en peaufinant leur paramétrage. Le paramétrage des processus de cartographie automatique est d'ailleurs lui-même un problème de recherche, et des techniques de dialogues basées sur des échantillons seraient certainement efficaces (Christophe 2011) pour aider les utilisateurs de nos systèmes à bien les paramétrer.

Puisque les données géographiques vont devenir de plus en plus hétérogènes, avec une part importante issue de saisies collaboratives, il est important que les techniques de généralisation s'adaptent à ces nouvelles données source. Nous avons vu dans ce mémoire que la caractéristique importante de ces données quand il s'agit de les cartographier est l'hétérogénéité de leur niveau de détail. Nous avons proposé des méthodes pour caractériser cette hétérogénéité de niveau de détail, et l'étape suivante est maintenant de concevoir des méthodes de généralisation qui s'adaptent au niveau de détail de chacun des objets dans la carte. En plus du contexte géographique et des propriétés géométriques de l'objet (forme à préserver par exemple), le choix du bon algorithme de généralisation devra aussi tenir compte du niveau de détail. Quand le niveau de détail est insuffisant pour la cartographie précise à une échelle, il faudra développer les premières méthodes d'harmonisation présentées dans le mémoire.

Enfin, si on s'intéresse à la généralisation à la demande, amenée à se développer dans les années à venir, le point clé pour nous est la généralisation d'une carte avec un fond topographique, vu comme le support des données, et des données thématiques fournies par un utilisateur par-dessus. Pour que les cartes soient compréhensibles, il faut dans un premier temps modéliser les relations spatiales qui

relient les données thématiques et les données support, comme le recommande Cécile Duchêne (2014). Puis, à la manière de la proposition d'Adrien Maudet dans les cartes d'itinéraires de randonnée, il faut que le processus de généralisation prenne en compte ces relations particulières pour que les données thématiques restent au centre de la carte, mais en cohérence avec le fond, et sans endommager sa lisibilité. Il reste beaucoup à faire sur ce sujet, notamment du fait de la forte diversité des cartes thématiques.

Démontrer l'utilité de la généralisation cartographique

Jusqu'à aujourd'hui, l'objectif des chercheurs en généralisation, en particulier ceux travaillant avec les NMA, était de faire aussi bien que les cartographes donc on ne se posait pas la question de l'utilité de la généralisation, c'était une évidence. Avec le foisonnement de cartes mal généralisées utilisées aujourd'hui, sans que les utilisateurs soient conscients d'une gêne, nous devons nous poser la question de l'utilité de nos processus. Quelle est notre valeur ajoutée ?

Des premières réponses à cette question vont être apportées par la fin de la thèse de Marion Dumont et du projet MapMuxing, où plusieurs tests seront menés de front pour identifier les meilleures stratégies pour créer des continuums cartographiques, et pour démontrer que ces continuums, couplées à des interfaces adaptées, améliorent sensiblement l'expérience de navigation de l'utilisateur dans les cartes multi-échelles.

Mais il faudra aller plus loin, et mettre en place des tests démontrant l'intérêt de simplifier les données, de caricaturer parfois pour mieux préserver les relations spatiales clés pour la compréhension de la carte. Dans certaines communautés de recherche, comme la visualisation d'information ou les interactions homme-machine, la validation auprès des utilisateurs est devenue systématique et c'est dans cette direction que nous devons aller pour légitimer notre savoir-faire et notre valeur ajoutée.

Vers une meilleure intégration de tous les processus cartographiques

Nous avons vu dans le premier chapitre que les différents processus pour la création d'une carte automatiquement sont interdépendants. La phase de stylisation est ainsi placée après la généralisation, mais dans le même temps, les processus de généralisation ont besoin de connaître les symboles appliqués sur les objets de la carte pour éviter des superpositions ou occlusions de symboles. Malgré nos efforts de collaborations, ces deux processus sont encore et toujours vus et réalisés séparément. Quelques travaux ont déjà proposé des modèles pour une meilleure intégration des différents travaux (Brewer et Buttenfield 2009), (Roth, Brewer et Stryker 2011), (Samsonov, Podolsky et Yurova 2013), et ils proposent d'utiliser des opérations de stylisation au cours du processus de généralisation, à la place de certaines opérations de symbolisation : on peut par exemple diminuer la largeur d'un symbole pour résoudre des problèmes de superposition plutôt que de déplacer les objets et leurs symboles.

Une deuxième manière d'intégrer ces deux processus est d'adapter le niveau de détail cible de la généralisation aux besoins d'une stylisation avancée. Les processus de stylisation artistique comme l'aquarelle ou l'encre de Chine ne peuvent pas donner un résultat lisible si les données sont aussi détaillées que pour un simple style topographique (Christophe, Dumenieu, et al. 2016), et c'est le même problème quand on imite un style de cartes anciennes comme les cartes de Cassini. Il faut donc configurer les processus de généralisation de manière spécifique pour chacun de ces styles.

Pour aller plus loin dans l'intégration, il faut clairement imaginer un unique processus de création cartographique automatique, incluant également le placement de textes, qui fonctionnerait par

itérations de généralisation et stylisation, avec des retours en arrière possible, jusqu'à obtenir la carte la plus efficace et la plus belle possible. Cela passera également par une fusion des spécifications et paramètres de tous les processus, et l'extension de nos propositions dans le cadre du modèle ScaleMaster2.0 nous pourrait très prometteuse pour unifier la norme SLD/SE de description du style et la description des paramètres de généralisation.

Cartographier de gros volumes de données

Maintenant que nous sommes capables de démontrer l'efficacité de processus automatiques de cartographie sur des zones choisies volontairement petites pour nos expérimentations, il faut être capable de passer à l'échelle pour qu'ils soient plus que de simples prototypes. Mais ce passage à l'échelle n'est pas simplement une affaire d'ingénierie, il pose plusieurs questions de recherche que nous allons devoir résoudre. Le premier de ces problèmes est de savoir comment bien partitionner les données pour découper des lots de données les plus indépendants possibles entre eux, pour éviter au maximum l'absence de contexte géographique lors du traitement d'un lot. Des stratégies de découpage régulier par grille ou quad tree ont été proposées (Briat, Monnot et Punt 2011) (Thiemann, et al. 2013), ainsi que des techniques de découpage géographique (Chaudhry et Mackaness 2010) (Touya 2010), et il paraît évident que suivant les opérations, les découpages ne peuvent pas être les mêmes : par exemple, le contexte nécessaire à la généralisation des routes est plus étendu que celui nécessaire à la généralisation d'un bâtiment.

Le deuxième problème est celui de la parallélisation des traitements une fois les bons découpages identifiés et effectués. Même si le partitionnement est optimisé, il paraît peu probable qu'on arrive à encapsuler tout le contexte géographique dans une seule partition, et il faudra sûrement trouver une technique pour accéder aux résultats des partitions voisines, en gelant par exemple leur exécution le temps du traitement de la partition. Il faudra aussi établir des stratégies de réconciliation pour les objets se trouvant dans deux partitions, et donc potentiellement traités deux fois de manière différente.

Le dernier problème est celui de la généralisation à la volée et il est indirectement lié à la cartographie de gros volumes de données. En effet, sur de gros volumes de données mis à jour fréquemment, on ne souhaite pas relancer de lourds calculs à chaque mise à jour, et il faut donc prévoir que certains traitements soient lancés à la volée, c'est-à-dire en temps réel quand un utilisateur veut visualiser la version à jour d'une zone qui n'a pas encore été recalculée. Plusieurs stratégies ont déjà été proposées pour des mécanismes couplant traitements de fond plus lourds et traitements plus cosmétiques à la volée (Lehto et Kilpeläinen 2000), (Cecconi, Weibel et Barrault 2002), (Harrie et Sarjakoski 2002), (van Oosterom 2005). Ces méthodes doivent être évaluées pour voir celles qui sont les plus à même de passer à l'échelle pour de gros volumes de données.

Une Cartographie multi-échelles vectorielle d'OpenStreetMap ?

A l'interface de toutes ces perspectives se trouve le projet qui va structurer nos recherches pendant les prochaines années : est-il possible, à moyen terme, de passer des deux modes de conception de cartes évoqués en introduction du mémoire, « vite fait mal fait », et « lentement fait bien fait », au seul mode viable avec les évolutions de la cartographie, le « vite fait bien fait » à partir des données d'OpenStreetMap ? En effet, OpenStreetMap étant présent et actif partout dans le monde, construire des cartes de qualité à partir de ces données donnerait accès à une cartographie de qualité à de très nombreux citoyens du monde qui n'y ont pas encore accès pour des raisons politiques ou de développement économique. Ces cartes pourraient être complétées avec d'autres

données disponibles suivant les zones : des données provenant de NMA ou d'autres données géographiques libres. Cet objectif recouvre plusieurs défis importants :

- Atteindre réellement la cartographie tout automatique. Cela passe par une poursuite des recherches sur l'amélioration des processus de généralisation et de leur orchestration. Mais cela passe aussi par une diminution des exigences de qualité définies jusqu'à présent par les cartographes dont on souhaitait automatiser le travail. L'objectif tout automatique sera certainement atteignable si on remplace une partie de ces exigences, par des spécifications guidées par l'efficacité des cartes, mesurée par des tests de performance sur des utilisateurs de ces cartes. On ne cherchera plus à faire « aussi bien et beau qu'avant », mais plutôt à faire des cartes « aussi efficaces qu'avant et aussi belles que possible ».
- Etre capable de traiter le monde entier, avec toutes les spécificités géographique que cela implique en plus de la masse de données (voir les défis sur la cartographie de gros volumes de données juste au-dessus).
- Savoir qualifier automatiquement et sans données de référence (qui ne sont pas disponibles sur le monde entier) les données OpenStreetMap pour savoir comment bien les cartographier.
- Dans les territoires où d'autres données qu'OpenStreetMap sont disponibles, comme en France, il faudrait être capable de tirer profit de l'ensemble des données disponibles, de les combiner dans les différentes échelles pour produire des cartes encore meilleures que celles produites avec un seul de ces jeux de données.

Bibliographie

- Agrawala, Maneesh, et Chris Stolte. «Rendering effective route maps: improving usability through generalization.» *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*. New York, NY, USA: ACM, 2001. 241-249.
- Alvina, Jessalyn, Caroline Appert, Olivier Chapuis, et Emmanuel Pietriga. «RouteLens: Easy Route Following for Map Applications.» *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces*. New York, NY, USA: ACM, 2014. 125-128.
- Bader, Matthias, Mathieu Barrault, et Robert Weibel. «Building displacement over a ductile truss.» *International Journal of Geographical Information Science* (Taylor & Francis) 19, n° 8 (2005): 915-936.
- Baeijs, Christof, Yves Demazeau, et L. Alvares. *Sigma : Application of multi-agent systems to cartographic generalization*. Vol. 1038, chez *Agents Break Away*. *Proceedings of the European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, édité par W. Van de Velde et J. W. Perram, 163-176. Springer, 1996.
- Balley, Sandrine, Blanca Baella, Sidonie Christophe, Maria Pla, Nicolas Regnaud, et Jantien Stoter. «Map Specifications and User Requirements.» Dans *Abstracting Geographic Information in a Data Rich World*, édité par Dirk Burghardt, Cécile Duchêne et William Mackaness, 17-52. Springer International Publishing, 2014.
- Balley, Sandrine, et Nicolas Regnaud. «Models and Standards for On-Demand Mapping.» *Proceedings of 25th International Cartographic Conference ICC'11*. Paris, France, 2011.
- Balley, Sandrine, Kusay Jaara, et Nicolas Regnaud. «Towards a Prototype for Deriving Custom Maps from Multisource Data.» *Proceedings of 15th ICA Workshop on Generalisation and Multiple Representation*. Istanbul, Turkey, 2012.
- Barrault, M. «A methodology for placement and evaluation of area map labels.» *Computers, Environment and Urban Systems* 25, n° 1 (jan 2001): 33-52.
- Barrault, Mathieu, et al. «Integrating multi-agent, object-oriented, and algorithmic techniques for improved automated map generalisation.» *20th International Cartographic Conference*. 2001. 2110-2116.
- Beard, Kate M. «Constraints on rule formation.» Dans *Map Generalization*, édité par Barbara Buttenfield et Robert McMaster, 121-135. Longman Pages, 1991.
- Bellicha, Amit, et al. «Autour Du Problème De Satisfaction De Contraintes.» *In Actes des 5èmes journées nationales du PRC GDR Intelligence Artificielle*. 1994. 159-178.
- Biljecki, Filip, Hugo Ledoux, Jantien Stoter, et Junqiao Zhao. «Formalisation of the level of detail in 3D city modelling.» *Computers, Environment and Urban Systems* 48 (nov 2014): 1-15.
- Bishr, Yaser. «Overcoming the semantic and other barriers to GIS interoperability.» *International Journal of Geographical Information Science* (Taylor & Francis) 12, n° 4 (jun 1998): 299-314.
- Boffet, Annabelle. «Creating urban information for cartographic generalisation.» *International Symposium on Spatial Data Handling (SDH)*. 2000.

- Brassel, Kurt E., et Robert Weibel. «A review and conceptual framework of automated map generalization.» *International Journal of Geographical Information Systems* (Taylor & Francis) 2, n° 3 (jan 1988): 229-244.
- Brewer, Cynthia A., et Barbara Buttenfield. «Framing Guidelines for Multi-Scale Map Design Using Databases at Multiple Resolutions.» *Cartography and Geographic Information Science* 34, n° 1 (jan 2007): 3-15.
- Brewer, Cynthia A., et Barbara Buttenfield. «Mastering map scale: balancing workloads using display and geometry change in multi-scale mapping.» *Geoinformatica* (Springer Netherlands) 14, n° 2 (apr 2009): 221-239.
- Briat, Marc-Olivier, Jean-Luc Monnot, et Edith M. Punt. «Scalability of Contextual Generalization Processing Using Partitioning and Parallelization.» *Proceedings of 14th ICA Workshop on Generalisation and Multiple Representation*. Paris, France, 2011.
- Brunet, Roger. «La composition des modèles dans l'analyse spatiale.» *L'Espace Géographique* 8 (1980): 253-265.
- Brychtová, Alžbeta, Arzu Çöltekin, et Vit Paszto. «Do the visual complexity algorithms match the generalization process in geographical displays?» *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Commission II, TeS: WG II/6 Geovisualization and Virtual Reality*. Prague, Czech Republic, 2016.
- Burghardt, D., et A. Cecconi. «Mesh simplification for building typification.» *International Journal Geographic Information Science* (Taylor & Francis, Inc.) 21, n° 3 (mar 2007): 283-298.
- Burghardt, Dirk, Cécile Duchêne, et William Mackaness, . *Abstracting Geographic Information in a Data Rich World*. Berlin Heidelberg: Springer, 2014.
- Burghardt, Dirk, et Moritz Neun. «Automated Sequencing of Generalisation Services Based on Collaborative Filtering.» Dans *Geographic Information Science - 4th International Conference GIScience*, édité par Martin Raubal, Harvey J. Miller, Andrew U. Frank et Michael F. Goodchild, 41-46. Münster, Germany: IFGI prints, 2006.
- Burghardt, Dirk, Ingo Petzold, et Matthias Bobzien. «Relation Modelling within Multiple Representation Databases and Generalisation Services.» *Cartographic Journal, The* (Maney Publishing) 47, n° 3 (aug 2010): 238-249.
- Burghardt, Dirk, Stefan Schmid, et Jantien Stoter. «Investigations on cartographic constraint formalisation.» *10th ICA Workshop on Generalisation and Multiple Representation*. 2007.
- Buttenfield, Barbara, et David M. Mark. «Expert systems in cartographic design.» Dans *Geographic Information Systems: The Microcomputer and Modern Cartography*, édité par D. R. F. Taylor, 129-150. Oxford: Pergamon Press, 1991.
- Buttenfield, Barbara, et Robert McMaster, . *Map Generalization. Making Rules for Knowledge Representation*. London: Longman, 1991.
- Cecconi, Alessandro, Robert Weibel, et Mathieu Barrault. «Improving Automated Generalisation for On-Demand Web Mapping by Multiscale Databases.» Dans *Advances in Spatial Data Handling*, édité par Dianne E Richardson et Peter van Oosterom, 515-531. Springer Berlin Heidelberg, 2002.

- Chaudhry, Omair Z., et William A. Mackaness. «DTM Generalisation: Handling Large Volumes of Data for Multi-Scale Mapping.» *Cartographic Journal, The* (Maney Publishing) 47, n° 4 (nov 2010): 360-370.
- Chaudhry, Omair Z., William A. Mackaness, et Nicolas Regnauld. «A functional perspective on map generalisation.» *Computers, Environment and Urban Systems* 33, n° 5 (sep 2009): 349-362.
- Christensen, Jon, Joe Marks, et Stuart Shieber. «An Empirical Study of Algorithms for Point-feature Label Placement.» *ACM Trans. Graph. (ACM)* 14, n° 3 (jul 1995): 203-232.
- Christophe, Sidonie. «Creative Colours Specification Based on Knowledge (COLorLEGend system).» *Cartographic Journal, The* 48, n° 2 (2011): 138-145.
- Christophe, Sidonie, et al. «Map Style Formalization: Rendering Techniques Extension for Cartography.» Édité par David Mould et Pierre Bénard. *Proceedings of Non-Photorealistic Animation and Rendering*. Lisbon, Portugal, 2016.
- Christophe, Sidonie, et Charlotte Hoarau. «Expressive Map Design Based on Pop Art: Revisit of Semiology of Graphics?» *Cartographic Perspectives* (North American Cartographic Information Society) 73 (2013).
- Clementini, Eliseo, et Robert Laurini. «Un cadre conceptuel pour modéliser les relations spatiales.» *Revue des Nouvelles Technologies de l'Information RNTI-E-14* (nov 2008).
- Cuenin, R. *Cartographie générale*. Paris: Collections scientifiques de l'Institut Géographique National, 1972.
- Douglas, David H., et Thomas K. Peucker. «Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or its Caricature.» *Cartographica: The International Journal for Geographic Information and Geovisualization* 10, n° 2 (oct 1973): 112-122.
- Duchêne, Cécile. «Généralisation Cartographique par Agents Communicants : le modèle CartACom. Application aux données topographiques en zone rurale.» Ph.D. dissertation, Université Paris 6, 2004.
- . «Making a map from "thematically multi-sourced data": the potential of making inter-layers spatial relations explicit.» *Proceedings of 17th ICA Workshop on Generalisation and Multiple Representation*. 2014.
- Duchêne, Cécile, Anne Ruas, et Christophe Cambier. «The CartACom model: transforming cartographic features into communicating agents for cartographic generalisation.» *International Journal of Geographical Information Science* 26, n° 9 (apr 2012): 1533-1562.
- Duchêne, Cécile, et al. «Generalisation in Practice Within National Mapping Agencies.» Dans *Abstracting Geographic Information in a Data Rich World*, édité par Dirk Burghardt, Cécile Duchêne et William Mackaness, 329-391. Springer International Publishing, 2014.
- Dumont, Marion, Guillaume Touya, et Cécile Duchêne. «Assessing the Variation of Visual Complexity in Multi-Scale Maps with Clutter Measures.» *Proceedings of 19th ICA Workshop on Generalisation and Multiple Representation*. Helsinki, Finland, 2016.
- . «Automated Generalisation of Intermediate Levels in a Multi-Scale Pyramid.» *Proceedings of 18th ICA Workshop on Generalisation and Multiple Representation*. Rio de Janeiro, Brazil, 2015.

- Elias, Birgit. *Extracting Landmarks with Data Mining Methods*. Vol. 2825, chez *Spatial Information Theory. Foundations of Geographic Information Science*, édité par Walter Kuhn, Michael F. Worboys et Sabine Timpf, 375-389. Springer Berlin Heidelberg, 2003.
- Ertz, Olivier, Maxence Laurent, Daniel Rappo, et Eric Taillard. «Towards an Automated Label Placement Service.» Dans *Geospatial Free and Open Source Software in the 21st Century*, édité par Erwan Bocher et Markus Neteler, 157-171. Springer Berlin Heidelberg, 2012.
- Figueira, José, V. Mousseau, et B. Roy. «ELECTRE Methods.» Dans *Multiple Criteria Decision Analysis: State of the Art Surveys*, édité par José Figueira, S. Greco et M. Ehrogott, 133-162. Springer-Verlag, 2005.
- Foerster, Theodor, Jantien Stoter, et Barend Köbben. «Towards a Formal Classification of Generalization Operators.» *23rd International Cartographic Conference*. 2007.
- Freeman, Herbert. «Automated cartographic text placement.» *Pattern Recognition Letters* 26, n° 3 (feb 2005): 287-297.
- Gaffuri, Julien. «Généralisation automatique pour la prise en compte des thèmes champ : le modèle GAEL.» Ph.D. dissertation, Université Marne-la-Vallée, 2008.
- Gaffuri, Julien. «Improving Web Mapping with Generalization.» *Cartographica: The International Journal for Geographic Information and Geovisualization* 46, n° 2 (jan 2011): 83-91.
- Galanda, Martin, et Robert Weibel. «An agent-based framework for polygonal subdivision generalization.» Dans *Advances in spatial data handling, 10th international symposium on spatial data handling*, édité par Diane Richardson et Peter van Oosterom, 121-136. Springer, 2002.
- Galanda, Martin, et Robert Weibel. «An agent-based framework for polygonal subdivision generalization.» Dans *Advances in spatial data handling, 10th international symposium on spatial data handling*, édité par Diane Richardson et Peter van Oosterom, 121-136. Springer, 2002.
- Girres, Jean-François. «An evaluation of the impact of cartographic generalisation on length measurement computed from linear vector databases.» *Proceedings of 25th International Cartographic Conference (ICC'11)*. Paris, France, 2011.
- Girres, Jean-François, et Guillaume Touya. «Cartographic Generalisation Aware of Multiple Representations.» Édité par Matt Duckham, Kathleen Stewart et Edzer Pebesma. *Proceedings of GIScience 2014 - Poster session*. Vienna, Austria, 2014.
- Girres, Jean-François, et Guillaume Touya. «Quality Assessment of the French OpenStreetMap Dataset.» *Transactions in GIS* (Blackwell Publishing Ltd) 14, n° 4 (2010): 435-459.
- Goodchild, Michael F. «Citizens as voluntary sensors: spatial data infrastructure in the world of Web 2.0.» *International Journal of Spatial Data Infrastructures Research*. 2007. 24-32.
- Gould, Nicholas M., William A. Mackaness, Guillaume Touya, et Glen Hart. «Collaboration on an Ontology for Generalisation.» *Proceedings of 17th ICA Workshop on Generalisation and Multiple Representation*. Vienna, Austria, 2014.
- Gould, Nicholas, et Omair Chaudhry. «Generation and Validation of Workflows for On-Demand Mapping.» *GeoProcessing 2012*. Valencia, Spain: IARIA, 2012. 238-243.

- Grosso, Eric, Julien Perret, et Mickaël Brasebin. «GEOXYGENE : une plate-forme interoperable pour le développement d'applications géographiques.» Chap. 3 dans *Développements logiciels en géomatique : innovations et mutualisation*, édité par Bénédicte B. Le Ber F. Hermès - Lavoisier, 2012.
- Grünreich, D. «Ein Vorschlag zum Aufbau einer grossmassstäbigen topographischkartographischenDatenbank unter besonderer Berücksichtigung der Grundrissdateides ALK-Systems.» *Nachrichten aus dem Karten- und Vermessungswesen, Series I* 95 (1985): 55+.
- Guilbert, Eric. «Feature-Driven Generalization of Isobaths on Nautical Charts: A Multi-Agent System Approach.» *Transactions in GIS* 20, n° 1 (2016): 126-143.
- Haklay, M. «How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets.» *Environment and Planning B: Planning and Design* (Pion Ltd) 37, n° 4 (2010): 682-703.
- Harrie, Lars E. «The Constraint Method for Solving Spatial Conflicts in Cartographic Generalization.» *Cartography and Geographic Information Science* (Cartography and Geographic Information Society) 26, n° 1 (jan 1999): 55-69.
- Harrie, Lars E., et Tapani Sarjakoski. «Simultaneous Graphic Generalization of Vector Data Sets.» *Geoinformatica* (Kluwer Academic Publishers) 6, n° 3 (2002): 233-261.
- Harrie, Lars, L. Tiina Sarjakoski, et Lassi Lehto. «A Mapping Function for Variable-Scale Maps in Small-Display Cartography.» *Journal of Geospatial Engineering* 2, n° 3 (2002): 111-123.
- Haunert, J. H., H. Haverkort, B. Niedermann, A. Nocaj, A. Slingsby, et J. Wood. «Labeling Curves with Curved Labels.» *Abstracts of the Schematic Mapping Workshop 2014*. 2014.
- Haunert, Jan-Henrik, et Alexander Wolff. «Area aggregation in map generalisation by mixed-integer programming.» *International Journal of Geographical Information Science* (Taylor & Francis) 24, n° 12 (2010): 1871-1897.
- Hojholt, Peter. «Solving Space Conflicts in Map Generalization: Using a Finite Element Method.» *Cartography and Geographic Information Science* (Cartography and Geographic Information Society), jan 2000: 65-74.
- Imhof, Eduard. «Positioning Names on Maps.» *American Cartographer* 2, n° 2 (1975): 129-144.
- Jaara, Kusay, Cécile Duchêne, et Anne Ruas. «Preservation and Modification of Relations Between Thematic and Topographic Data Throughout Thematic Data Migration Process.» Dans *Cartography From Pole To Pole, Selected Contributions to the 26th International Conference of the ICA*, édité par M. Buchroithner, 103-118. Springer, 2013.
- . «Toward the generalisation of cartographic mashups: Taking into account the dependency between the thematic data and the reference data throughout the process of automatic generalisation.» *Proceedings of 14th ICA Workshop on Generalisation and Multiple Representation*. Paris, France, 2011.
- Jabeur, Nafaâ. «A Multi-Agent System for On-the-fly Web Map Generation and Spatial Conflict Resolution.» Ph.D. dissertation, Université Laval, Quebec, 2006.

- Jégou, Laurent, et Jean-Philippe Deblonde. «Vers une visualisation de la complexité de l'image cartographique.» *Cybergeo : European Journal of Geography*, n° article 600 (avr 2012).
- Joubran Abu Daoud, J., et Y. Doytsher. «An Automated Cartographic Generalization Process: A Pseudo-Physical Model.» *Proceedings of ISPRS Commission II, WG II/3*. 2008.
- Kubera, Yoann, Philippe Mathieu, et Sébastien Picault. «IODA: an interaction-oriented approach for multi-agent based simulations.» *Autonomous Agents and Multi-Agent Systems* (Springer Netherlands), jan 2011: 1-41.
- Lecordix, François, Corinne Plazanet, et Jean-Philippe Lagrange. «A Platform for Research in Generalization: Application to Caricature.» *Geoinformatica* 1, n° 2 (1997): 161-182.
- Lecordix, François, Corinne Plazanet, François Chirié, Jean-Philippe Lagrange, T. Banel, et Y. Cras. «Placement automatique des écritures d'une carte avec une qualité cartographique.» *6ème rendez-vous européen des acteurs de l'information géographique numérique MARI*. 1994. 22-32.
- Lehto, L., et T. Kilpeläinen. «Real-Time Generalisation of Geodata in the Web.» *International Archives of Photogrammetry and Remote Sensing XXXIII* (2000): 559-566.
- Li, Maozhen, Sheng Zhou, et Christopher Jones. «Multi-agent Systems for Web-Based Map Information Retrieval.» Dans *Proceedings of the 2nd International Conference on GIScience*, 161-180. 2002.
- Loi, Hugo, Thomas Hurtut, Romain Vergne, et Joëlle Thollot. «Discrete Texture Design Using a Programmable Approach.» *ACM SIGGRAPH 2013 Talks*. New York, NY, USA: ACM, 2013.
- Mackaness, William A. «Understanding Geographic Space.» Chap. 1 dans *The Generalisation of Geographic Information: Models and Applications*, édité par William A. Mackaness, A. Ruas et T. Sarjakoski, 1-10. Amsterdam, Netherlands: Elsevier, 2007.
- Mackaness, William A., Anne Ruas, et L. Tiina Sarjakoski, . *Generalisation of Geographic Information: Models and Applications*. Amsterdam, The Netherlands: Elsevier, 2007.
- Mackaness, William A., et al. «Thematic workshop on building an ontology of generalisation for on-demand mapping.» *Proceedings of 4th UK Ontology Network Workshop*. Leeds, UK, 2015.
- Mackaness, William A., et Geoffrey Edwards. «The Importance of Modelling Pattern and Structure in Automated Map Generalisation.» *Proceedings of the Joint ISPRS/ICA Workshop on Multi-Scale Representations of Spatial Data*. 2002. 7-8.
- Mackaness, William A., et Mike Steven. «An Algorithm for Localised Contour Removal over Steep Terrain.» *Cartographic Journal, The* (Maney Publishing) 43, n° 2 (jul 2006): 144-156.
- Mackaness, William A., et Peter F. Fisher. «Automatic Recognition and Resolution of Spatial Conflicts in Cartographic Symbolisation.» *Proceedings of Auto-Carto 8*. 1987. 709-718.
- Mackaness, William, et Andreas Reimer. «Generalisation in the Context of Schematised Maps.» Dans *Abstracting Geographic Information in a Data Rich World*, édité par Dirk Burghardt, Cécile Duchêne et William Mackaness, 299-328. Springer International Publishing, 2014.
- Mathieu, Philippe, Sébastien Picault, et Yann Secq. «Les environnements : en avoir ou pas ? Formalisation du concept et patterns d'implémentation.» Édité par Rémy Courdier et Jean-

- Paul Jamont. *22e Journées Francophones sur les Systèmes Multi-Agents*. Lorient-sur-Drôme, France: Cépaduès, 2014. 55-64.
- Maudet, Adrien. «Interactions entre niveaux dans un modèle orienté agent de généralisation cartographique. Le modèle DIOGEN.» Ph.D. dissertation, Université Paris-Est, 2016.
- Maudet, Adrien. «Interactions multi-niveaux dans un modèle de généralisation cartographique. Le modèle DIOGEN.» Ph.D. dissertation, Université Paris-Est, 2016.
- Maudet, Adrien, Guillaume Touya, Cécile Duchêne, et Sébastien Picault. «Improving multi-level interactions modelling in a multi-agent generalisation model: first thoughts.» *Proceedings of 16th ICA Generalisation Workshop*. Dresden, Germany, 2013.
- . «Multi-Agent Multi-Level Cartographic Generalisation in CartAGen.» *Advances in Practical Applications of Heterogeneous Multi-Agent Systems: Proceedings of the 12th International Conference, PAAMS 2014*. Springer, 2014. 355-358.
- . «Patterns multi-niveaux pour les SMA.» Édité par Laurent Vercoeur et Gauthier Picard. *23es Journées Francophones sur les Systèmes Multi-Agents (JFSMA'15)*. Rennes, France: Cépaduès, 2015. 19-28.
- Maudet, Adrien, Guillaume Touya, Cécile Duchêne, et Sébastien Picault. *Representation of Interactions in a Multi-Level Multi-Agent Model for Cartography Constraint Solving*. Vol. 8473, chez *Advances in Practical Applications of Heterogeneous Multi-Agent Systems. The PAAMS Collection*, édité par Yves Demazeau, Franco Zambonelli, JuanM Corchado et Javier Bajo, 183-194. Springer International Publishing, 2014.
- McMaster, Robert B., et K. S. Shea. «Cartographic Generalization in Digital Environment: A Framework for implementation in a GIS.» *GIS/LIS'88*. 1988. 240-249.
- McMaster, Robert, et K. S. Shea. *Generalization in Digital Cartography*. Association of American Geographers Press, 1992.
- Meulemans, Wouter, André Renssen, et Bettina Speckmann. *Area-Preserving Subdivision Schematization*. Vol. 6292, chez *Geographic Information Science*, édité par Sara Irina Fabrikant, Tamasch Reichenbacher, Marc Kreveld et Christoph Schlieder, 160-174. Springer Berlin Heidelberg, 2010.
- Moacdieh, Nadine, et Nadine Sarter. «Display Clutter: A Review of Definitions and Measurement Techniques.» *Human Factors: The Journal of the Human Factors and Ergonomics Society* (SAGE Publications) 3 (jul 2014).
- Mooney, Peter, et Pádraig Corcoran. «The Annotation Process in OpenStreetMap.» *Transactions in GIS* (Blackwell Publishing Ltd) 16, n° 4 (2012): 561-579.
- Müller, Jean-Claude. «Rule based generalization: Potentials and impediments.» *Proceedings of 4th International Symposium on Spatial Data Handling*. Zurich, Switzerland, 1990. 317-334.
- Mustière, Sébastien. «Algorithm for Legibility Evaluation of Symbolised Lines.» *Proceedings of 4th Intercarto conference*. 1998. 43-47.
- Mustière, Sébastien. «Apprentissage supervisé pour la généralisation cartographique.» Ph.D. dissertation, Université Pierre et Marie Curie, 2001.

- Mustière, Sébastien, et Thomas Devogele. «Matching Networks with Different Levels of Detail.» *Geoinformatica* (Springer Netherlands) 12, n° 4 (dec 2008): 435-453.
- Nakos, Byron, Julien Gaffuri, et Sébastien Mustière. «A Transition from Simplification to Generalisation of Natural Occuring Lines.» *Proceedings of 11th ICA Workshop on Generalisation and Multiple Representation*. Montpellier, France, 2008.
- Nöllenburg, M., et A. Wolff. «Drawing and Labeling High-Quality Metro Maps by Mixed-Integer Programming.» *Visualization and Computer Graphics, IEEE Transactions on* (IEEE) 17, n° 5 (may 2011): 626-641.
- Nöllenburg, Martin, Damian Merrick, Alexander Wolff, et Marc Benkert. «Morphing polylines: A step towards continuous generalization.» *Computers, Environment and Urban Systems* 32, n° 4 (jul 2008): 248-260.
- Olteanu-Raimond, Ana-Maria, Sébastien Mustière, et Anne Ruas. «Knowledge Formalization for Vector Data Matching Using Belief Theory.» *Journal of Spatial Information Science* 10 (2015): 21-46.
- Ormeling, Ferjan. «Cartography and Geoinformation in the 20th and 21st Centuries.» *Meta-carto-semiotics, Journal for Theoretical Cartography* 4 (2011).
- Palomar-Vazquez, Jesus, et Josep E. Pardo-Pascual. «Automated spot heights generalisation in trail maps.» *International Journal of Geographical Information Science* (Taylor & Francis) 22, n° 1 (2008): 91-110.
- Peng, Dongliang, Jan-Henrik Haunert, Alexander Wolff, et Christophe Hurter. «Morphing Polylines Based on Least-Squares Adjustment.» *Proceedings of 16th ICA Workshop on Generalisation and Multiple Representation*. Dresden, Germany, 2013.
- Picault, Sébastien, Philippe Mathieu, et Yoann Kubera. «PADAWAN, un modèle multi-echelles pour la simulation orientée interactions.» Édité par Michel Ocello et Lilia Rejeb. *Systèmes Multi-Agents, Défis Sociétaux, JFSMA'10*. Mahdia, Tunisie: Cépaduès, 2010. 193-202.
- Pietriga, Emmanuel, Olivier Bau, et Caroline Appert. «Representation-Independent In-Place Magnification with Sigma Lenses.» *IEEE Transactions on Visualization and Computer Graphics* (IEEE Computer Society) 16, n° 03 (2010): 455-467.
- Pindat, Cyprien, Emmanuel Pietriga, Olivier Chapuis, et Claude Puech. «JellyLens: content-aware adaptive lenses.» *Proceedings of the 25th annual ACM symposium on User interface software and technology*. New York, NY, USA: ACM, 2012. 261-270.
- Regnauld, Nicolas. «Evolving from automating existing map production systems to producing maps on demand automatically.» *10th ICA Workshop on Generalisation and Multiple Representation*. 2007.
- Regnauld, Nicolas, et Robert C. McMaster. «A synoptic View of Generalisation Operators.» Dans *Generalisation of Geographic Information*, édité par William A. Mackaness, Anne Ruas et L. Tina Sarjakoski, 37-66. Elsevier, 2007.
- Reimer, Andreas, et Wouter Meulemans. «Parallelity in Chorematic Territorial Outlines.» *Proceedings of 14th ICA Workshop on Generalisation and Multiple Representation*. Paris, France, 2011.

- Renard, Jérémy, Julien Gaffuri, et Cécile Duchêne. «Capitalisation problem in research - example of a new platform for generalisation : CartAGen.» *11th ICA Workshop on Generalisation and Multiple Representation*. Zurich, Switzerland, 2010.
- Revell, Patrick. «Automated Generalisation and Representation of Ordnance Survey Polygonal Landcover Data at 1:10 000 Scale.» *Proceedings of ICC2007, the International Cartography Association Conference*. 2007.
- Ribeiro, Glaydston M., et Luiz A. Lorena. «Heuristics for cartographic label placement problems.» *Computers & Geosciences* 32, n° 6 (jul 2006): 739-748.
- Richards, Nigel, et Mark J. Ware. «Ant Colony Optimization Applied to Map Generalization.» *13th Workshop of the ICA commission on Generalisation and Multiple Representation*. 2010.
- Rosenholtz, Ruth, Yuanzhen Li, et Lisa Nakano. «Measuring visual clutter.» *Journal of Vision* (Association for Research in Vision and Ophthalmology) 7, n° 2 (aug 2007).
- Roth, Robert E., Cynthia A. Brewer, et Michael S. Stryker. «A typology of operators for maintaining legible map designs at multiple scales.» *Cartographic Perspectives* 68 (2011).
- Ruas, A., et C. Plazanet. «Strategies for Automated Generalization.» *7th International Symposium on Spatial Data Handling*. Delft, Netherlands, 1996. 319-336.
- Ruas, Anne. *Généralisation et représentation multiple*. Lavoisier, 2002.
- . «Le changement de niveau de détail dans la représentation de l'information géographique.» *Le changement de niveau de détail dans la représentation de l'information géographique*. 2004.
- Ruas, Anne. «Modèle de généralisation de données géographiques à base de contraintes et d'autonomie.» Ph.D. dissertation, Université de Marne-la-Vallée, 1999.
- Ruas, Anne, et A. Bianchin. «Echelle et Niveau de détail.» Chap. 1 dans *Généralisation et représentation multiple*, édité par Anne Ruas. Hermès Lavoisier, 2002.
- Rylov, Maxim A., et Andreas W. Reimer. «A Comprehensive Multi-criteria Model for High Cartographic Quality Point-Feature Label Placement.» *Cartographica* 49, n° 1 (2014): 52-68.
- Rylov, Maxim A., et Andreas W. Reimer. «Improving label placement quality by considering basemap detail with a raster-based approach.» *Geoinformatica* (Springer US) 19, n° 3 (jul 2015): 463-486.
- Sabo, Mamane N., Yvan Bédard, Bernard Moulin, et Eveline Bernier. «Toward Self-Generalizing Objects and On-the-Fly Map Generalization.» *Cartographica: The International Journal for Geographic Information and Geovisualization* 43, n° 3 (jan 2008): 155-173.
- Salichtchev, K. A. *Cartographic communication: A theoretical survey*. Vol. 2, chez *Graphic Communication and Design in Contemporary Cartography*, édité par D. R. F. Taylor, 11-36. New-York: Wiley, 1983.
- Samsonov, Timofey, Anton Podolsky, et Nina Yurova. «Multimapper A Prototype System for Designing Multi-Scale Maps.» *Proceedings of 26th International Cartographic Conference*. Dresden, Germany, 2013.
- Schwartzges, Nadine, Jan-Henrik Haunert, Alexander Wolff, et Dennis Zwiebler. «Point Labeling with Sliding Labels in Interactive Maps.» Dans *Connecting a Digital Europe Through Location and*

- Place*, édité par Joaquin Huerta, Sven Schade et Carlos Granell, 295-310. Springer International Publishing, 2014.
- Sester, Monika. «Optimization approaches for generalization and data abstraction.» Édité par Peter F. Fisher, M. Gahegan et B. Lees. *International Journal of Geographical Information Science* (Taylor & Francis) 19, n° 8 (2005): 871-897.
- Shea, K. Stuart, et Robert B. McMaster. «Cartographic Generalization in a Digital Environment: When and How To Generalize.» *Proceedings of Auto-Carto 9*. 1989. 56-67.
- Steiniger, Stefan, et Robert Weibel. «Relations among Map Objects in Cartographic Generalization.» *Cartography and Geographic Information Science* (Taylor & Francis) 34, n° 3 (jan 2007): 175-197.
- Stoter, Jantien E., et al. *State-of-the-art of automated generalisation in commercial software*. Tech. rep., EuroSDR series on research projects, Amsterdam: Gopher, 2010.
- Taillandier, Patrick, Cécile Duchêne, et Alexis Drogoul. «Automatic revision of the control knowledge used by trial and error methods: Application to cartographic generalisation.» *Applied Soft Computing* 11, n° 2 (mar 2011): 2818-2832.
- Taillandier, Patrick, et Franck Taillandier. «Multi-criteria diagnosis of control knowledge for cartographic generalisation.» *European Journal of Operational Research* 217, n° 3 (#mar# 2012): 633-642.
- Teulade-Denantes, Jules, Adrien Maudet, et Cécile Duchêne. «Routes visualization: Automated placement of multiple route symbols along a physical network infrastructure.» *Journal of Spatial Information Science* 11 (2016): 53-79.
- Thiemann, Frank, Stefan Werder, Thomas Globig, et Monika Sester. «Investigations into partitioning of generalization processes in a distributed processing framework.» Édité par Manfred F. Buchroithner. *Proceedings of the 26th International Cartographic Conference*. Dresden, Germany, 2013.
- Touya, Guillaume. «A Road Network Selection Process Based on Data Enrichment and Structure Detection.» *Transactions in GIS* 14, n° 5 (2010): 595-614.
- . «Relevant Space Partitioning for Collaborative Generalisation.» *Proceedings of 12th ICA Workshop on Generalisation and Multiple Representation*. Zurich, 2010.
- . «Un système multi-agent pour l'orchestration de processus de généralisation cartographique.» Édité par Emmanuel Adam et Jean-Paul Sansonnet. *Systemes Multi-Agents, Initiatives Mixtes - JFSMA 11*. Cepadues Editions, 2011. 73-82.
- Touya, Guillaume, Adeline Coupé, Jérémie Jollec, Olivier Dorie, et Frank Fuchs. «Conflation Optimized by Least Squares to Maintain Geographic Shapes.» *ISPRS International Journal of Geo-Information* 2, n° 3 (jul 2013): 621-644.
- Touya, Guillaume, Bénédicte Bucher, Gilles Falquet, Kusay Jaara, et Stefan Steiniger. «Modelling Geographic Relationships in Automated Environments.» Chap. 3 dans *Abstracting Geographic Information in a Data Rich World*, édité par Dirk Burghardt, Cécile Duchêne et William Mackaness, 53-82. Berlin Heidelberg: Springer, 2014.

- Touya, Guillaume, Blandine Decherf, Mayeul Lalanne, et Marion Dumont. «Comparing Image-Based Methods for Assessing Visual Clutter in Generalized Maps.» *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* II-3/W5 (sep 2015).
- Touya, Guillaume, Cécile Duchêne, et Anne Ruas. *Collaborative Generalisation: Formalisation of Generalisation Knowledge to Orchestrate Different Cartographic Generalisation Processes*. Vol. 6292, chap. 19 dans *Geographic Information Science*, édité par Sara Fabrikant, Tumasch Reichenbacher, Marc van Kreveld et Christoph Schlieder, 264-278. Berlin, Heidelberg: Springer Berlin / Heidelberg, 2010.
- Touya, Guillaume, Cécile Duchêne, et Anne Ruas. *Collaborative Generalisation: Formalisation of Generalisation Knowledge to Orchestrate Different Cartographic Generalisation Processes*. Vol. 6292, chap. 19 dans *Geographic Information Science*, édité par Sara Fabrikant, Tumasch Reichenbacher, Marc van Kreveld et Christoph Schlieder, 264-278. Berlin, Heidelberg: Springer Berlin / Heidelberg, 2010.
- Touya, Guillaume, et Andreas Reimer. «Inferring the Scale of OpenStreetMap Features.» Dans *OpenStreetMap in GIScience*, édité par Jamal Jokar Arsanjani, Alexander Zipf, Peter Mooney et Marco Helbich, 81-99. Springer International Publishing, 2015.
- Touya, Guillaume, et Carmen Brando-Escobar. «Detecting Level-of-Detail Inconsistencies in Volunteered Geographic Information Data Sets.» *Cartographica: The International Journal for Geographic Information and Geovisualization* 48, n° 2 (jan 2013): 134-143.
- Touya, Guillaume, et Cécile Duchêne. *CollaGen: Collaboration between automatic cartographic Generalisation Processes*. Vol. 1, chap. 30 dans *Advances in Cartography and GIScience*, édité par Anne Ruas, 541-558. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011.
- Touya, Guillaume, et Jean-François Girres. «ScaleMaster 2.0: a ScaleMaster extension to monitor automatic multi-scales generalizations.» *Cartography and Geographic Information Science* (Taylor & Francis) 40, n° 3 (jun 2013): 192-200.
- Touya, Guillaume, Sandrine Balley, Cécile Duchêne, K. Jaara, Nicolas Regnauld, et Nick Gould. «Towards an Ontology of Generalisation Constraints and Spatial Relations.» *Proceedings of 15th ICA workshop on generalisation and multiple representation*. Istanbul, Turkey, 2012.
- van Dijk, Steven, Marc V. Kreveld, Tycho Strijk, et Alexander Wolff. «Towards an evaluation of quality for names placement methods.» *International Journal of Geographical Information Science* (Taylor & Francis) 16, n° 7 (nov 2002): 641-661.
- van Dijk, Thomas C., et Jan-Henrik Haurert. «Interactive focus maps using least-squares optimization.» *International Journal of Geographical Information Science* (Taylor & Francis), mar 2014: 1-24.
- van Dijk, Thomas, Arthur van Goethem, Jan H. Haurert, Wouter Meulemans, et Bettina Speckmann. «Accentuating Focus Maps via Partial Schematization.» *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. New York, NY, USA: ACM, 2013. 428-431.
- van Goethem, Arthur, Wouter Meulemans, Andreas Reimer, Herman Haverkort, et Bettina Speckmann. «Topologically Safe Curved Schematisation.» *Cartographic Journal, The* (Maney Publishing) 50, n° 3 (aug 2013): 276-285.

- van Oosterom, Peter. «Variable-scale Topological Data Structures Suitable for Progressive Data Transfer: The GAP-face Tree and GAP-edge Forest.» *Cartography and Geographic Information Science* (Cartography and Geographic Information Society) 32, n° 4 (oct 2005): 331-346.
- van Oosterom, Peter, Martijn Meijers, Jantien Stoter, et Radan Suba. «Data Structures for Continuous Generalisation: tGAP and SSC.» Dans *Abstracting Geographic Information in a Data Rich World*, édité par Dirk Burghardt, Cécile Duchêne et William Mackaness, 83-117. Springer International Publishing, 2014.
- Visvalingam, M., et J. D. Wyatt. «Line Generalization by Repeated Elimination of Points.» *The Cartographic Journal* 30, n° 1 (1993): 46-51.
- Vo, Duc-An, A. Drogoul, et J. Zucker. «An Operational Meta-Model for Handling Multiple Scales in Agent-Based Simulations.» *Computing and Communication Technologies, Research, Innovation, and Vision for the Future (RIVF), 2012 IEEE RIVF International Conference on*. IEEE, 2012. 1-6.
- Walter, V., et D. Fritsch. «Matching spatial data sets: a statistical approach.» *International Journal of Geographical Information Science* (Taylor & Francis) 13, n° 5 (jul 1999): 445-473.
- Ware, Colin. *Information Visualization, Third Edition: Perception for Design Interactive Technologies*. 3. Morgan Kaufmann, 2012.
- Ware, Mark J., Christopher B. Jones, et Nathan Thomas. «Automated map generalization with multiple operators: a simulated annealing approach.» *International Journal of Geographical Information Science* (Taylor & Francis) 17, n° 8 (2003): 743-769.
- Ware, Mark J., et Christopher B. Jones. «Conflict Reduction in Map Generalization Using Iterative Improvement.» *Geoinformatica* (Kluwer Academic Publishers) 2, n° 4 (dec 1998): 383-407.
- Ware, Mark J., George E. Taylor, Suchith Anand, et Nathan Thomas. «Automated Production of Schematic Maps for Mobile Applications.» *Transactions in GIS* (Blackwell Publishing) 10, n° 1 (jan 2006): 25-42.
- Ware, Mark J., I. D. Wilson, et J. A. Ware. «A knowledge based genetic algorithm approach to automating cartographic generalisation.» *Knowledge-Based Systems* 16, n° 5-6 (jul 2003): 295-303.
- Ware, Mark J., Ian D. Wilson, Andrew J. Ware, et Christopher B. Jones. «A tabu search approach to automated map generalisation.» *GIS '02: Proceedings of the 10th ACM international symposium on Advances in geographic information systems*. New York, NY, USA: ACM, 2002. 101-106.
- Weibel, Robert, et G. Dutton. «Generalising Spatial Data and Dealing with Multiple Representations.» Dans *Geographical Information Systems, Vol. 1, Principles and Technical Issues*, édité par P. Longley, M. F. Goodchild, D. J. Maguire et D. W. Rhind, 125-150. Wiley, 1999.
- Werder, Stefan. «Formalization of Spatial Constraints.» *12th AGILE International Conference on Geographic Information Science*. 2009.
- Yan, Jingya, Eric Guilbert, et Eric Saux. «An ontology-driven multi-agent system for nautical chart generalization.» *Cartography and Geographic Information Science* (Taylor & Francis), feb 2016: 1-15.

- Yu, Wei. «Aide multicritère à la décision dans le cadre de la problématique du tri : concepts, méthodes et applications.» Ph.D. dissertation, Université Paris Dauphine, 1992.
- Zhang, Xiang, Jantien Stoter, et Tinghua Ai. «Formalization and Automatic Interpretation of Map Requirements.» *11th ICA Workshop on Generalisation and Multiple Representation*. 2008.
- Zhang, Xunruo, et Eric Guilbert. *A multi-agent System Approach for Feature-driven Generalization of isobathymetric Line*. Vol. 1, chap. 27 dans *Advances in Cartography and GIScience*, édité par Anne Ruas, 477-495. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011.
- Zhou, Sheng. «Towards a Multi-Representation Database Approach to On-Demand Mapping and Continuous Zooming.» *Proceedings of 17th ICA Workshop on Generalisation and Multiple Representation*. Vienna, Austria, 2014.

Eléments de CV

Parcours professionnel

- **2004-2005**: Chargé d'étude au laboratoire COGIT de l'IGN:
Etude sur l'intégration de la BD Carto et de la BD Topo par généralisation automatique : jusqu'à quel point peut-on généraliser automatiquement un équivalent de la BD Carto à partir de la BD Topo.
- **2006-2007** : Recherche sur la sélection des réseaux géographiques en généralisation de base de données (mise au point de modèles et algorithmes).
- **2008-2011** : Thèse en sciences de l'information géographique de l'Université Paris-Est.
- **2011-2013** : Responsable COGIT du projet de R&D pour la DGA "PEA REP".
- **2013-2014** : Responsable de l'action de recherche en cartographie
- **depuis 2014** : Porteur du projet ANR MapMuxing

Activités de recherche

Projets collaboratifs :

Nom du projet	Dates du projet	Financement	Rôle
State-of-the-art of automated generalisation in commercial software	2006-2009	EuroSDR	participant
PEA REP	2012-2013	PEA - DGA	porteur COGIT
MapMuxing	2014-2018	ANR	porteur
MobiTourGeo	2016-2017	PEPS UPE	participant
BigGeoQUP	2016-2017	CNRS Mastodons	porteur COGIT

Organisation d'évènements, société savantes :

- Co-organisateur local du Workshop ACI de généralisation en 2011
- Comité de programme et d'organisation de la conférence ISSDQ 2015
- Comité de programme du 1^{er} International Workshop on Big Geo Data Quality and Privacy (BIGQP 2017), à Venise en Italie, en parallèle de la conférence EDBT.
- Organisation de plusieurs workshops (COST TD1202 en 2014, Ontologies et généralisation en 2015 et COST ENERGIC 2015 et 2016)
- Représentant français au COST ENERGIC (2014-2016)
- Porteur avec Cyril De Runz de l'Action Prospective VGI/Crowdsourcing du GDR MAGIS (2016-)

Autres :

- Administration de plusieurs modules (généralisation, VGI, cartographie...) du logiciel open source GeOxygene.
- Enseignements depuis 2005 essentiellement à l'Ecole Nationale des Sciences Géographiques (ENSG) en M2 et cycle ingénieurs (15 à 20 h par an).
- Evaluator pour la plupart des revues du domaine (IJGIS, CaGIS, Transactions in GIS, IJGI...), 5 à 15 articles par an depuis 2011.
- Evaluator pour des conférences (ICC, ISPRS, Workshop annuel de généralisation) depuis 2010.

- Evalueur ANR en 2015.

Travaux encadrés

Thèse dirigée par dérogation

- Sébastien Biniek, 2016-

Thèses encadrées

- Quy Thy Truong, 2016-
- Adrien Maudet, 2012-2016
- Marion Dumont, 2014-

Post-doctorants/ingénieurs d'étude encadrés

- Jean-François Girres, 2012-2013, post-doctorant
- Imran Lokhat, 2014-, ingénieur d'étude
- Thibault Chassin 2015-2016, ingénieur d'étude
- Dieudonné Tsatcha, 2015, post-doctorant

Stages encadrés (ou co-encadrés)

- Justin Berli (co-encadré avec Imran Lokhat et Nicolas Regnauld) 2016
- Blandine Decherf et Mayeul Lalanne, 2014-2015
- Matthieu Baley, 2013
- Camille Delahaye, 2010
- Aurélien Velten, 2007
- Nicolas Culson, 2006

Publications

Revue à comité de lecture

Lokhat I. and **G. Touya** (2016). Enhancing Building Footprints with Squaring Operations, *Journal of Spatial Information Science* 13, in press.

Touya, G., S. Christophe and C. Hoarau (2016). Clutter and Map Legibility in Automated Cartography: A Research Agenda, *Cartographica: The International Journal for Geographic Information and Geovisualization*, in press.

Olteanu-Raimond, A.-M., G. Hart, **G. Touya**, T. Kellenberger, G. Foody and D. Demetriou (2016). The scale of VGI in map production: A perspective of European National Mapping Agencies, *Transactions in GIS* 20, in press.

Touya, G. and J.-F. Girres (2013). ScaleMaster 2.0: a ScaleMaster extension to monitor automatic multi-scales generalizations. *Cartography and Geographic Information Science* 40 (3), 192-200.

Touya, G. and C. Brando-Escobar (2013). Detecting Level-of-Detail inconsistencies in volunteered geographic information data sets. *Cartographica: The International Journal for Geographic Information and Geovisualization* 48 (2), 134-143.

Touya, G., A. Coupé, J. Jollec, O. Dorie, and F. Fuchs (2013). Conflation optimized by least squares to maintain geographic shapes. *ISPRS International Journal of Geo-Information* 2 (3), 621-644.

Touya, G. (2010). Enrichissement automatique de données par analyse spatiale pour la généralisation de réseaux. *Revue Internationale de Géomatique* 20 (2), 175-200.

Touya, G., C. Duchêne, and S. Mustière (2010). Généralisation et intégration pour un fond vert commun entre l'IFN et l'IGN. *Revue Internationale de Géomatique* 20 (1), 65-86.

Touya, G. (2010). A road network selection process based on data enrichment and structure detection. *Transactions in GIS* 14 (5), 595-614.

Girres, J.-F. and **G. Touya** (2010). Quality assessment of the french OpenStreetMap dataset. *Transactions in GIS* 14 (4), 435-459.

Stoter, J., D. Burghardt, C. Duchêne, B. Baella, N. Bakker, C. Blok, M. Pla, N. Regnauld, **G. Touya**, and S. Schmid (2009). Methodology for evaluating automated map generalization in commercial software. *Computers, Environment and Urban Systems* 33 (5), 311-324.

Actes de conférence (papier long)

Maudet, A., **G. Touya**, C. Duchêne, and S. Picault (2015, June). Patterns multi-niveaux pour les SMA. In L. Vercoeur and G. Picard (Eds.), *23es Journées Francophones sur les Systèmes Multi-Agents (JFSMA'15)*, Rennes, France, pp. 19-28. Cépaduès.

Touya, G., B. Decherf, M. Lalanne and M. Dumont (2015). Comparing Image-Based Methods for Assessing Visual Clutter in Generalized Maps. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. II-3/W5, Proceedings of ISSDQ, ISPRS

Savino, S. and **G. Touya** (2015, August). Automatic structure detection and generalization of railway networks. In Proceedings of 27th International Cartographic Conference, Rio de Janeiro, Brazil. ICA.

Maudet, A., **G. Touya**, C. Duchêne, and S. Picault (2014, June). Multi-Agent Multi-Level cartographic generalisation in CartAGen. In Advances in Practical Applications of Heterogeneous Multi-Agent Systems: Proceedings of the 12th International Conference, PAAMS 2014, Lecture Notes in Computer Science, pp. 355-358. PAAMS: Springer.

Maudet, A., **G. Touya**, C. Duchêne, and S. Picault (2014). Representation of interactions in a Multi-Level Multi-Agent model for cartography constraint solving. In Y. Demazeau, F. Zambonelli, J. Corchado, and J. Bajo (Eds.), Advances in Practical Applications of Heterogeneous Multi-Agent Systems. The PAAMS Collection, Volume 8473 of Lecture Notes in Computer Science, pp. 183-194. Springer International Publishing.

Touya, G. (2012). Social welfare to assess the global legibility of a generalized map. In N. Xiao, M.-P. Kwan, M. F. Goodchild, and S. Shekhar (Eds.), *Geographic Information Science 7th International Conference, GIScience 2012*, Volume 7478 of Lecture Notes in Computer Science, Chapter 15, pp. 198-211. Berlin, Heidelberg: Springer Berlin / Heidelberg.

Touya, G. (2011). Un système multi-agent pour l'orchestration de processus de généralisation cartographique. In E. Adam and J.-P. Sansonnet (Eds.), *Systèmes Multi-Agents, Initiatives Mixtes - JFSMA 11*, pp. 73-82. Cepadues Editions.

Touya, G. and C. Duchêne (2011). CollaGen: Collaboration between automatic cartographic generalisation processes. In A. Ruas (Ed.), *Advances in Cartography and GIScience*, Volume 1 of

Lecture Notes in Geoinformation and Cartography, Chapter 30, pp. 541-558. Berlin, Heidelberg: Springer Berlin Heidelberg.

Duchêne, C. and **G. Touya** (2010, October). Emergence de zones conflits dans deux modèles de généralisation cartographique multi-agents. In M. Ocelllo and L. Rejeb (Eds.), *Systèmes Multi-Agents, Défis Sociétaux*, Toulouse, pp. 33-42. Cépaduès.

Touya, G., C. Duchêne, and A. Ruas (2010). Collaborative generalisation: Formalisation of generalisation knowledge to orchestrate different cartographic generalisation processes. In S. Fabrikant, T. Reichenbacher, M. van Kreveld, and C. Schlieder (Eds.), *Geographic Information Science*, Volume 6292 of Lecture Notes in Computer Science, Chapter 19, pp. 264-278. Berlin, Heidelberg: Springer Berlin / Heidelberg.

Touya, G. (2008). Enrichissement de données par analyse spatiale pour la généralisation des réseaux. In *Colloque International de Géomatique et d'Analyse Spatiale SAGEO'08*. Montpellier, France.

Chapitres de livre

Antoniou, V., **G. Touya**, and A.-M. Olteanu-Raimond (2016). Quality analysis of the parisian OSM toponyms evolution. In C. Capineri, M. Haklay, H. Huang, V. Antoniou, J. Kettunen, F. Ostermann, and R. Purves (Eds.), *European Handbook of Crowdsourced Geographic Information*, Chapter 8, pp. 97-112. London: Ubiquity Press.

Touya, G. and M. Baley (2016). Level of Details Harmonization Operations in OpenStreetMap Based Large Scale Maps. In M. Leitner and J. Jokar Arsanjani (Eds.), *Citizen Empowered Mapping, Geotechnologies and the Environment*. Springer International Publishing. In press.

Touya, G. and A. Reimer (2015). Inferring the scale of OpenStreetMap features. In J. Jokar Arsanjani, A. Zipf, P. Mooney, and M. Helbich (Eds.), *OpenStreetMap in GIScience*, Lecture Notes in Geoinformation and Cartography, pp. 81-99. Springer International Publishing.

Regnauld, N., **G. Touya**, N. Gould, and T. Foerster (2014, March). Process modelling, web services and geoprocessing. In D. Burghardt, C. Duchêne, and W. Mackaness (Eds.), *Abstracting Geographic Information in a Data Rich World*, Chapter 7, pp. 198-225. Berlin Heidelberg: Springer.

Touya, G., B. Bucher, G. Falquet, K. Jaara, and S. Steiniger (2014, March). Modelling geographic relationships in automated environments. In D. Burghardt, C. Duchêne, and W. Mackaness (Eds.), *Abstracting Geographic Information in a Data Rich World*, Chapter 3, pp. 53-82. Berlin Heidelberg: Springer.

Ateliers ou conférence sur papier court

Dumont, M., **G. Touya**, and C. Duchêne (2016, June). Assessing the variation of visual complexity in Multi-Scale maps with clutter measures. In *Proceedings of 19th ICA Workshop on Generalisation and Multiple Representation*, Helsinki, Finland.

Dumont, M., **G. Touya** and C. Duchêne (2015) Automated Generalisation of Intermediate Levels in a Multi-Scale Pyramid. In *Proceedings of 18th ICA Workshop on Generalisation and Multiple Representation*, Rio de Janeiro, Brazil.

Touya, G. (2015). Lessons Learned From Research on Multimedia Summarization. In *Proceedings of 18th ICA Workshop on Generalisation and Multiple Representation*, Rio de Janeiro, Brazil.

Mackaness, W. A., N. Gould, S. Bechhofer, D. Burghardt, C. Duchene, R. Stevens, and **G. Touya** (2015, April). Thematic workshop on building an ontology of generalisation for on-demand mapping. In *Proceedings of 4th UK Ontology Network Workshop*, Leeds, UK.

Girres, J.-F. and **G. Touya** (2015, August). A framework to formalize multi-scales land-use generalization in the ScaleMaster 2.0. In *Proceedings of 27th International Cartographic Conference*, Rio de Janeiro, Brazil. ICA.

Touya, G. and J.-F. Girres (2014, September). Generalising unusual map themes from OpenStreetMap. In *Proceedings of 17th ICA Workshop on Generalisation and Multiple Representation*, Vienna, Austria. ICA.

Gould, N. M., W. A. Mackaness, **G. Touya**, and G. Hart (2014, September). Collaboration on an ontology for generalisation. In *Proceedings of 17th ICA Workshop on Generalisation and Multiple Representation*, Vienna, Austria. ICA.

Maudet, A., **G. Touya**, C. Duchêne, and S. Picault (2013, August). Improving multi-level interactions modelling in a multi-agent generalisation model: first thoughts. In *Proceedings of 16th ICA Generalisation Workshop*, Dresden, Germany. ICA.

Touya, G. (2012). What is the level of detail of OpenStreetMap? In Workshop on Role of Volunteered Geographic Information in Advancing Science: Quality and Credibility.

Touya, G., S. Balley, C. Duchêne, K. Jaara, N. Regnauld, and N. Gould (2012). Towards an ontology of generalisation constraints and spatial relations. In *Proceedings of 15th ICA workshop on generalisation and multiple representation*, Istanbul, Turkey. ICA.

Touya, G. (2010). Relevant space partitioning for collaborative generalisation. In *Proceedings of 12th ICA Workshop on Generalisation and Multiple Representation*, Zurich, Switzerland.

Stoter, J., B. Baella, C. Blok, D. Burghardt, F. Dávila, C. Duchêne, M. Pla, N. Regnauld, and **G. Touya** (2010, September). EuroSDR research on state-of-the-art of automated generalisation in commercial software: main findings and conclusions. In *13th Workshop of the ICA commission on Generalisation and Multiple Representation*. Zurich, Switzerland.

Touya, G. (2006, September). A method for generalization of river networks based on "strokes" and database enrichment. In M. Raubal, H. Miller, A. U. Frank, and M. F. Goodchild (Eds.), *4th International Conference on Geographic Information Science (GIScience'2006) – Extended Abstracts*, Volume 28 of IFGI Prints, pp. 191-194.

Stoter, J., K.-H. Anders, B. Baella, D. Burghardt, F. Davila, C. Duchêne, M. Pla, N. Regnauld, P. Rosenstand, S. Schmid, **G. Touya**, and H. Uitermark (2008). A study on the state-of-the-art in automated map generalisation implemented in commercial out-of-the-box software. In *11th ICA Workshop on Generalisation and Multiple Representation*, Montpellier, France.

Touya, G. (2007). River network selection based on structure and pattern recognition. In *Proceedings of International Cartographic Conference*, Moscow, Russia.

Touya, G. (2007). A road network selection process based on data enrichment and structure detection. In *workshop on generalisation and multiple representation*, Moscow, Russia. ICA.

Touya, G. (2008, June). First thoughts for the orchestration of generalisation methods on heterogeneous landscapes. In *workshop on generalisation and multiple representation*, Montpellier, France.

Autres (posters, revues sans comité de lecture, rapports techniques...)

Dumont, M., **G. Touya**, and C. Duchêne (2016) A comparative study of existing multi-scale maps: what content at which scale? In *Proceedings of GIScience 2016*, poster session, Montreal, Canada.

Touya, G., S. Christophe, and C. Hoarau (2015, December). Problématiques de recherche autour de la notion de clutter en cartographie automatique. *Cartes & Géomatique - Revue du Comité Français de Cartographie* 226, 91-100.

Girres, J.-F. and **G. Touya** (2015, December). Un cadre formel pour la généralisation multi-échelle de l'occupation du sol au sein du ScaleMaster2.0. *Cartes & Géomatique - Revue du Comité Français de Cartographie* 226, 61-70.

Touya, G. and S. Savino (2015, December). Enrichissement automatique et généralisation de réseaux ferrés. *Cartes & Géomatique - Revue du Comité Français de Cartographie* 226, 71-80.

Girres, J.-F. and **G. Touya** (2014, September). Cartographic generalisation aware of multiple representations. In M. Duckham, K. Stewart, and E. Pebesma (Eds.), *Proceedings of GIScience 2014 - Poster session*, Vienna, Austria.

Touya, G. and M. Baley (2014, September). Harmonizing level of detail in OpenStreetMap based maps. In M. Duckham, K. Stewart, and E. Pebesma (Eds.), *Proceedings of GIScience 2014 - Poster session*, Vienna, Austria.

Baley, M. and **G. Touya** (2014). Intégration et correction automatique de données OpenStreetMap. *Géomatique Expert* 96, 32-43.

Touya, G. and J.-F. Girres (2013). Le ScaleMaster 2.0 : une extension du modèle ScaleMaster pour gérer des généralisations multi-échelles automatiques. *Cartes & Géomatique - Revue du Comité Français de Cartographie* 217, 113-128.

Touya, G. and C. Brando (2013). Détection d'incohérences de niveau de détail dans des données collaboratives. *Cartes & Géomatique* 217, 59-72.

Touya, G. (2012). CollaGen: Collaboration entre des processus de généralisation cartographique. *Cartes & Géomatique - Revue du Comité Français de Cartographie* 211, 33-46.

Renard, J., J. Gaffuri, C. Duchêne, and **G. Touya** (2011). Automated generalisation results using the agent-based platform CartAGen. In *Proceedings of 25th International Cartographic Conference (ICC'11) - poster session*, Paris, France. ICA.

Stoter, J. E., B. Baella, C. Blok, D. Burghardt, C. Duchêne, M. Pla, N. Regnauld, and **G. Touya** (2010). State-of-the-art of automated generalisation in commercial software. Technical report, EuroSDR series on research projects, Amsterdam.

Touya, G. and J.-F. Girres (2009). Première évaluation de la qualité des données libres d'OpenStreetMap en France. In *Colloque International de Géomatique et d'Analyse Spatiale SAGEO'09*, Saint-Mandé, France.

Touya, G., E. Grosso, S. Mustière, and C. Duchêne (2006). Automatisation de l'appariement de la BDCarto avec la BDTopo pays et de la généralisation de la BDCarto à partir de la BDTopo pays, en vue de l'intégration de ces bases - « la BDCARTO fait peau neuve ». Technical report, IGN France.

Touya, G., C. Duchêne, and S. Mustière (2006). Généralisation et intégration du fond vert. Technical report, IGN France.

Touya, G., E. Grosso, S. Mustière, and C. Duchêne (2006, September). Generalization or data matching for integration? insights from experiments on databases at different levels of details. In M. Raubal, H. Miller, A. U. Frank, and M. F. Goodchild (Eds.), *4th International Conference on Geographic Information Science (GIScience'2006) – Extended Abstracts*, IFGI Prints, pp. 373-376.

Touya, G. (2007). Un processus de sélection du réseau hydrographique, basé sur la détection des structures. *Le Monde des Cartes* 194, 46-56.