

Abstraction et Changement de Langage pour Automatiser la Généralisation Cartographique

Sébastien Mustière^{1,2}

Jean-Daniel Zucker²

Lorenza Saitta³

¹Institut Géographique National – Laboratoire COGIT – France.

²Université Paris VI – LIP6 – Pôle IA - France

³Università del Piemonte Orientale – Dipartimento di Scienze e Tecnologia Avanzate - Italy

contact¹ : IGN – COGIT, 2-4 av. Pasteur - 94165 St-Mandé Cedex - France

Sebastien.Mustiere@ign.fr

Résumé

Dans cet article nous décrivons le processus de généralisation cartographique, c'est-à-dire le processus de dérivation de carte à partir de données trop détaillées, en y distinguant ses deux opérations clés : l'abstraction et le changement de langage de représentation. Nous montrons que dans ce processus l'abstraction est nécessaire pour réaliser un changement de représentation efficace et que, loin d'éliminer de l'information, l'abstraction améliore l'information perçue par le lecteur de la carte. Enfin, nous décrivons comme cette distinction entre abstraction et changement de langage guide la mise au point d'un processus d'apprentissage automatique pour acquérir les connaissances nécessaires à la généralisation cartographique.

Mots clés

Abstraction, Représentation des Connaissances, Apprentissage Automatique, Cartographie.

Abstract

In this article, we describe the cartographic generalisation process, i.e. the map derivation process from too detailed data. Two key concepts are distinguished: abstraction and change of representation language. Within the cartographic process, abstraction is necessary for an efficient change of language. We show that, far from only deleting information, abstraction increases information perceived by the map reader. Then we describe a machine learning process, guided by the previous considerations, for acquiring knowledge necessary to automate the map generalisation process.

Keywords

Knowledge Abstraction, Knowledge Representation, Machine Learning, Cartography.

1 Introduction

En cartographie, la *généralisation*¹ désigne le processus

qui permet de créer des cartes à une échelle donnée à partir de données (carte ou base de données) à plus grande échelle, c'est-à-dire plus détaillées. Le but de ce processus est donc de représenter un même espace géographique sur une surface (papier ou écran) plus petite que celle où sont représentées les données originales. Pour cela il ne suffit pas de simplifier les données de départ mais il est nécessaire de mettre en valeur les détails importants et d'éliminer les détails non significatifs, en fonction de l'utilisation finale de la carte.



Fig. 1 Extrait de carte au 1/25.000



Fig. 2 Réduction et généralisation au 1/100.000

Les extraits de cartes ci-dessus illustrent la problématique de la généralisation cartographique : on dispose de données de références, ici sous la forme d'une carte topographique au 1/25.000 (Fig. 1), et on désire réaliser une carte topographique au 1/100.000. Ce changement d'échelle est nécessaire pour des utilisations de la carte qui nécessitent soit de visualiser une zone géographique étendue, soit de changer de niveau d'analyse des données. Il s'agit donc de représenter la même zone géographique sur une surface 16 fois moins grande. Si on effectue une simple réduction de l'image cartographique (Fig. 2, à gauche), les composants de la carte (routes, maisons,

¹ Pour éviter toute confusion de langage, notons que le terme de généralisation utilisé en cartographie ne correspond pas à celui utilisé en intelligence artificielle mais se rapproche plus du terme abstraction en IA.

réseau de rues...) deviennent illisibles et la structure générale de l'espace est perdue. Une simple élimination de certains de ces composants diminuera la sensation de fouillis de la carte mais ceux restant ne seront pas plus lisibles pour autant, et la structure générale ne sera pas suffisamment mise en valeur. Il est donc nécessaire de *caricaturer* le paysage pour en faire ressortir les traits principaux. Certains éléments sont éliminés, d'autres agrandis, simplifiés, déplacés... (Fig. 2, à droite). Cela doit être fait de manière à assurer une bonne représentation de la réalité géographique et des relations entre objets, et c'est là que réside le savoir du cartographe : quoi, quand, et comment éliminer, agrandir, simplifier, déplacer... ?

Ce processus cartographique représente une phase longue et coûteuse lors de la fabrication de cartes. Des recherches sont donc menées dans le monde de la cartographie depuis une trentaine d'années pour essayer d'automatiser tout ou partie de ce processus [Lang, 69 ; Brassel & Weibel, 88 ; Nyerges, 91 ; McMaster & Shea, 92 ; Keller, 94 ; Ruas, 99]. Cet article s'inscrit dans le cadre de ces travaux, et plus précisément dans le cadre d'une étude de la possibilité d'aider, à l'aide des outils d'apprentissage automatique développés en IA, l'acquisition et la formalisation des connaissances nécessaires à la généralisation cartographique.

En nous plaçant dans le cadre formel d'abstraction et de changement de langage de représentation développé par [Saitta & Zucker, 98] (décrit en partie 2), nous décrivons le processus de généralisation cartographique (partie 3). Nous décrivons en particulier comment le processus d'abstraction est nécessaire au changement de langage en cartographie. Enfin, nous décrivons comment la distinction entre abstraction et changement de langage doit être prise en compte pour améliorer l'acquisition de connaissances dans le domaine de la généralisation cartographique (partie 4).

2 Abstraction et changement de langage

La représentation et la manipulation des connaissances est l'un des principaux sujets de recherche en intelligence artificielle [Sowa, 91 ; Davis, 93]. Depuis une cinquantaine d'années de nombreux langages formels ont été utilisés en IA, et chacun d'eux est plus ou moins adapté pour représenter différents domaines de connaissance [Ginsberg, 97]. Par ailleurs, bien que de nombreux domaines de connaissances peuvent être formulés dans *un seul* langage donné, certains, comme le processus cartographique, requièrent *plusieurs* langages de représentation pour représenter les différentes données manipulées, des données brutes du monde à leur représentation finale par une carte (cf. partie 3).

En IA, la prise en compte de plusieurs niveaux de détails au sein d'un même langage de représentation et la capacité à passer d'un niveau de détail à un autre fait depuis longtemps référence à la capacité humaine dite d'*abstraction*. L'abstraction en IA a surtout été étudiée

dans les domaines de la résolution de problème [Sacerdoti, 73 ; Giunchiglia & Walsh, 92 ; Ellamn, 93] et de la reformulation de problème [Choueiry et al., 98]. De manière intuitive, l'abstraction a été définie dans ces contextes comme une transformation sur le langage de représentation qui permet de résoudre un problème (ou prouver un théorème) plus facilement, c'est-à-dire avec un coût algorithmique moindre.

[Saitta & Zucker, 98 ; Zucker, 98] ont proposé un modèle différent d'abstraction se plaçant dans un cadre de raisonnement plus large que celui de la résolution de problème au sein d'un langage particulier. Partant de l'analyse du rôle de l'abstraction dans la représentation et l'apprentissage de concept, leur étude les a conduit à différencier deux processus fondamentaux : le changement de langage de représentation et l'abstraction au sein d'un langage de représentation. Ce modèle est résumé ci-après.

2.1 Des langages de représentation pour raisonner sur le monde

A la base de toute expérience se trouve le *monde* W qui contient les entités réelles. Ce monde n'est jamais directement connu car nous n'y avons accès que par une *perception* de celui-ci $P(W)$. A ce niveau, ne sont perçus que les stimuli physiques que le monde crée sur un vecteur de notre perception. Ces stimuli n'existent que pour l'observateur et représentent une vision instantanée du monde. Afin de pouvoir utiliser ces stimuli il est nécessaire de les stocker dans une *structure* organisée S , c'est-à-dire une représentation en extension [Van Dalen, 83] de la perception du monde. Cette structure peut être par exemple organisée sous la forme de tables relationnelles, sur lesquelles une algèbre relationnelle peut être appliquée. Ensuite, pour pouvoir décrire symboliquement le monde, et pour communiquer avec d'autres agents, un *langage* L est requis. Ce langage permet de décrire le monde en intension. Enfin, une *théorie* T est nécessaire pour raisonner sur le monde. Cette théorie contient des connaissances générales ou spécifiques à un domaine qui permettent la réalisation d'inférences sur le langage.

Nous appelons par la suite un *contexte de raisonnement* l'ensemble de ces quatre langages $R = \langle P(W), S, L, T \rangle$, cf. Fig. 3.

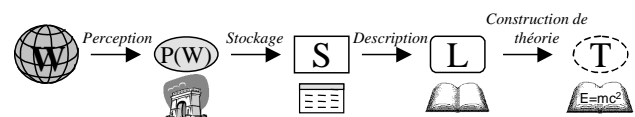


Fig. 3. Contexte de raisonnement

Nous ne détaillons pas ici la construction ni l'instanciation des différents langages de représentation du contexte de raisonnement.

2.2 Abstraction sur la perception du monde

Le monde peut être perçu d'une infinité de manières différentes par un agent intelligent en fonction des outils d'observation, du but de l'observation, de la culture de l'agent, etc. Il existe donc, pour un même monde, une infinité de perceptions possibles de celui-ci. C'est à ce niveau que le type et la quantité d'information que l'agent va stocker, décrire, et analyser sont déterminés. Plus la perception est détaillée, moins elle est abstraite. L'agent peut parfois contrôler le contenu de sa perception du monde, s'il contrôle les outils d'observation et s'il connaît à ce niveau le but de l'observation. Cependant, le plus souvent il ne peut pas contrôler cette perception : lorsqu'il collecte plus d'information qu'il ne lui est nécessaire, ou lorsqu'il a à effectuer plusieurs tâches, chacune nécessitant différentes informations qu'il est par ailleurs facile de collecter simultanément. Ceci montre qu'il est utile d'avoir des méthodes pour concrètement, ou virtuellement, transformer une perception en une autre plus abstraite. La définition suivante met en valeur ces considérations.

Définition 1. Etant donné un monde W , soit $R_r = \langle P_r(W), S_r, L_r, T_r \rangle$ et $R_a = \langle P_a(W), S_a, L_a, T_a \rangle$ deux contextes de raisonnement. Une abstraction est une fonction $A : R_r(W) \rightarrow R_a(W)$ tel que la perception $P_a(W)$ dite abstraite est plus simple que la perception $P_r(W)$ dite de référence.

Cette définition mérite quelques commentaires. Dans [Saitta & Zucker, 98] la relation « plus simple que » est définie formellement en termes de diminution de quantité d'information contenue dans la perception du monde.

Ensuite, la perception la moins abstraite est conventionnellement appelée de référence, mais la relation d'abstraction n'est que relative et toute perception peut être considérée de référence. En outre, le processus d'abstraction peut être réitéré plusieurs fois, la relation d'abstraction étant une relation transitive [Iwasaki, 90 ; Yoshida & Motoda, 90].

Il est également important de noter que dans cette définition, l'abstraction est définie au niveau de la perception du monde. Les abstractions au niveau de la structure, du langage, de la théorie en sont en suite dérivées, cf. Fig. 4.

Enfin, l'abstraction est un processus réversible, au sens qu'elle ne détruit pas d'information mais en cache seulement. La relation inverse (dite de *concrétion*) peut donc être définie.

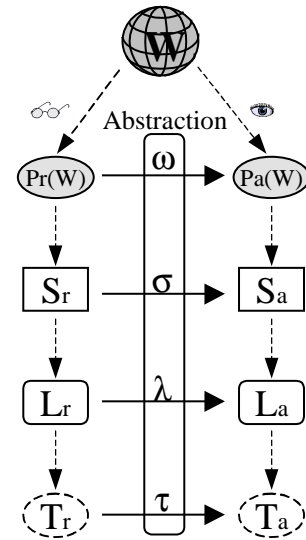


Fig. 4. L'abstraction et ses dérivées

3 Cartographie

Dans cette partie nous présentons les analogies entre le processus de création d'une carte et le modèle introduit ci-dessus. Nous présentons les différents langages de représentation utilisés en cartographie (Chap. 3.1) et le processus d'abstraction qui y est réalisé (Chap. 3.2).

3.1 Le processus cartographique

Considérons tout d'abord les différents langages de représentation utilisés en cartographie. Pour illustrer notre propos nous nous concentrons sur les cartes dites topographiques, c'est-à-dire de la surface et du relief de la terre. Les différents langages de représentation et le processus cartographique décrits ci-dessous correspondent, de manière simplifiée, à ceux utilisés dans les principaux organismes producteurs de données topographiques comme l'Institut Géographique National en France.

En topographie, le monde W étudié est bien sûr une partie du monde physique, et plus précisément la surface de la terre (son relief, l'occupation du sol, le réseau routier...).

La source principale de la perception $P(W)$ que nous en avons est constituée de photographies aériennes ou d'images satellites couvrant l'ensemble de la zone étudiée. Ces images sont le moyen le plus simple d'obtenir une grande quantité de données localisables dans l'espace.

Les éléments contenus dans ces images sont stockés dans une base de données géographique (BDG). Cette BDG contient la localisation des squelettes des objets géographiques (comme l'axe d'une route), représentés par la liste des coordonnées par où passe l'objet, et un certain nombre d'attributs (comme le nombre de voie de la route). Même si les objets de cette BDG sont nommés (routes, maison...), celle-ci représente plus un stockage de données, une structure S , qu'un langage de description a

partir duquel il est possible de communiquer. En effet celle-ci est difficilement analysable directement, une liste de coordonnées ne permettant pas de se représenter directement la forme d'un objet.

Le langage L communément utilisé pour représenter et analyser des données géographiques est la carte, qui est réalisée à partir de la BDG. Sur la carte chaque objet géographique est représenté par un symbole : une route est par exemple représentée par un trait de couleur variable suivant son importance, une rivière un trait bleu, une ville un point de taille proportionnelle au nombre d'habitants, une maison par un polygone noir.... En ce sens on peut considérer la cartographie comme un langage iconique.

Enfin, une théorie T est nécessaire pour raisonner sur la carte. Celle-ci contient les connaissances géographiques ou autres qui permettent d'analyser les différentes configurations géographiques rencontrées, que ce soit dans le cadre d'une recherche d'itinéraire, de la reconstruction mentale du paysage, de la recherche d'un modèle de développement des villes, etc. C'est cette théorie utilisée sur la carte qui guide la conception de celle-ci.

3.2 Généralisation cartographique et abstraction

Comme expliqué en introduction, le but de la généralisation cartographique est de créer des cartes à partir de données trop détaillées, et plus précisément d'une base de données géographique. C'est-à-dire créer une carte représentant une autre perception du monde que celle qui a guidée la création de la base de données géographique qui sert de référence. En effet la BDG de référence a été créée avec une certaine perception du monde (par exemple une vision très locale, à l'échelle d'un randonneur) et la généralisation cartographique a pour but de créer une carte avec une autre perception (par exemple celle moins locale d'un automobiliste).

La généralisation cartographique est nécessaire pour deux raisons. Tout d'abord la phase d'acquisition des données en cartographie est longue et coûteuse, car elle nécessite à la fois la réalisation de prises de vues, un travail de terrain nécessaire à une bonne localisation des images, une saisie des données visibles sur les images, et enfin l'intégration de données non visibles sur les images (le nombre d'habitants d'une commune par exemple). Cette phase d'acquisition ne peut donc être répétée à la demande. Ensuite les informations nécessaires à une carte particulière ne correspondent jamais exactement à la quantité d'information présente sur une unique base de données de référence. Chaque nouvelle carte, créée dans un but particulier étant différente des autres.

La généralisation cartographique consiste donc tout d'abord à changer de langage de représentation : fabriquer une carte à partir d'une base de données géographique. Les objets, qui sont stockés dans la BDG comme une suite de coordonnées $\{(x,y)\}$, doivent être représentés

graphiquement par des symboles. Mais il s'agit également de changer le niveau de détail des données puisque l'on doit réaliser une carte à partir de données trop détaillées.

Illustrons ceci par deux exemples: la simplification et le déplacement de bâtiments, deux opérations courantes en cartographie. Dans ces deux exemples on choisit de symboliser les bâtiments par un polygone noir.

D'une part, si un objet est trop détaillé ses détails ne seront plus lisibles sur la carte, il faut donc le simplifier. D'autre part, si deux objets sont trop proches, on ne peut plus les distinguer sur la carte, il faut donc les éloigner l'un de l'autre.

La Fig. 5 suivante illustre ces deux transformations. On y voit l'objet tel qu'il apparaît dans les données initiales (ou plutôt tel qu'il apparaîtrait si on l'affichait sur la carte tel qu'il est dans les données initiales, sans transformation - en haut à gauche), une simple réduction de l'objet à l'échelle finale d'affichage de la carte (en haut à droite), l'objet après transformation cartographique à l'échelle finale (en bas à droite), l'objet transformé affiché à l'échelle initiale² de la BDG pour se rendre compte des transformations effectuées (en bas à gauche).

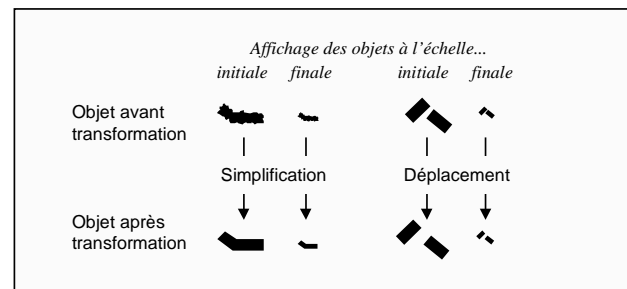


Fig. 5. Simplification et déplacement

Si on analyse ces deux opérations au niveau de la BDG, celles-ci sont sensiblement différentes. La simplification élimine des détails, elle fait donc diminuer la quantité d'information présente. Par contre, la quantité d'information présente après et avant déplacement reste constante, mais la précision de localisation des objets est dégradée, puisqu'ils sont déplacés.

Mais si on analyse ces deux opérations au niveau du langage de la carte, ces opérations sont très semblables. Ces deux transformations sont provoquées par la même cause (les objets ne sont pas lisibles si on les affiche tels quels), et ont les mêmes effets (rendre les objets lisibles en mettant en valeur un de leur caractère). Dans un cas cela est réalisé en dégradant les informations de forme dans la base de données, dans l'autre en dégradant les informations de position dans la base de données. Dans les deux cas, on améliore l'information qui aurait été

² Notons que par abus de langage on parle d'échelle d'une base de données géographique même si celle-ci n'a pas réellement d'échelle puisqu'elle n'est théoriquement pas liée à une notion d'affichage. Cette notion d'échelle est néanmoins utilisée pour évaluer rapidement le type d'information présent dans la base de données.

perçue par le lecteur de la carte si aucune transformation n'avait été effectuée. Dans le cas de la simplification, la forme complexe aurait rendu impossible ou difficile l'identification de « c'est un bâtiment globalement de cette forme là » voire seulement de « c'est un bâtiment », et dans le cas du déplacement la proximité aurait rendu impossible ou difficile l'identification de « il y a deux bâtiments proches mais disjoints ». On a donc diminué la quantité d'information stockée dans la BDG au profit d'un gain d'information perçue sur la carte.

Lors de la simplification on abstrait l'information « l'objet a globalement cette forme » car le langage de la carte n'est pas assez expressif pour représenter tous les niveaux de détails, et lors du déplacement on abstrait l'information « les objets sont proches mais disjoints » car le langage de la carte n'est pas assez expressif pour représenter à la fois la position et la séparation des objets.

Ces deux abstractions ont été faites au niveau de la perception du monde. Dans le cas du déplacement par exemple, on est passé d'une perception du monde où la localisation précise et la séparation entre les bâtiments sont importantes à une perception où la séparation est plus importante que la localisation précise. C'est ce choix d'abstraction de perception du monde, choisie pour satisfaire les besoins de la théorie, qui permet de déduire une opération efficace de transformation (le déplacement) au niveau du langage, en fonction des possibilités d'expression de celui-ci.

Classiquement en cartographie, on classe les opérations de transformation géométrique à partir des opérations élémentaires suivantes [McMaster & Shea, 92]:

- déplacement,
- simplification,
- amélioration de la géométrie, par ex. lorsqu'un bâtiment est dessinée de manière à rendre tous les angles droits, ce qui aidera son identification en tant que bâtiment,
- élimination,
- typification, lorsqu'un groupe d'objets est représenté par un groupe d'objets moins nombreux mais ayant la même répartition,
- caricature et exagération, lorsqu'un objet ou une partie d'un objet est agrandie pour être mis en valeur,
- changement de dimension, par ex. lorsqu'une ville stockée en détail dans la BDG par la surface qu'elle couvre est représentée par un symbole ponctuel.

Sans détailler chacune de ces opérations de transformation, on peut toutes les voir, à l'instar de la simplification et du déplacement, comme la conséquence d'une abstraction définie au niveau de la perception du monde, et réalisée pendant le changement de langage pour mettre en valeur un des caractères d'un objet ou d'un groupe d'objets, au détriment d'autres caractères.

4 Apprendre à cartographier

Nos travaux se placent dans le cadre de l'étude de

l'automatisation de la généralisation cartographique. Dans cette partie nous étudions les conséquences de l'analyse du processus de généralisation cartographique faite ci-dessus pour la mise au point d'un système d'acquisition des connaissances nécessaire à cette automatisation.

Notre but est de déterminer comment un objet géographique donné peut être représenté sur une carte. Cet objet est décrit dans une base de données géographique par le jeu des coordonnées d'un certain nombre de ses points (les coordonnées de chaque angle d'une maison par exemple). Le but de la généralisation cartographique est de transformer ce jeu de coordonnées afin que cet objet soit lisible sur une carte une fois symbolisé.

Le chapitre 4.1 explique les difficultés de représentation liées à ce problème. Le chapitre 4.2 montre une reformulation de ce problème qui le rend plus facile. Enfin le chapitre 4.3 montre les conséquences de la différenciation entre abstraction et changement de langage sur l'acquisition des connaissances dans le domaine de la généralisation cartographique.

4.1 Apprendre à transformer

De nombreuses études ont été réalisées en cartographie pour mettre au point des algorithmes génériques de transformation d'un jeu de coordonnées d'un objet en un jeu de coordonnées simplifié respectant les règles cartographiques [Lang, 69 ; Fritsch, 97]. Cependant, tous ces algorithmes ne se sont avérés efficaces que pour un type de transformation sur un type particulier d'objets.

Cette transformation générique peut être vue comme une fonction de l'espace des jeu de coordonnées vers ce même espace. La mise au point d'une telle fonction étant difficile, des études ont été menées pour essayer d'apprendre cette fonction à partir d'exemples donnés par un cartographe. (cf. Fig. 6, « h ? » représentant la phase d'apprentissage).

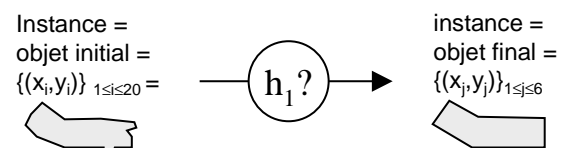


Fig. 6 Apprendre de coordonnées à coordonnées

[Werschlein & Weibel, 94 ; Weibel et al, 95] proposent de représenter cette fonction sous la forme d'un réseau de neurones. Confrontés au problème de la représentation des entrées et sorties du réseau de neurones, ils proposent de changer le mode de représentation des objets. Une grille de taille constante est appliquée sur les objets, et les objets initiaux (resp. finaux) représentés à l'origine par la liste de leur coordonnées sont représentés par l'ensemble des cases de cette grille, la valeur 1 est affectée à chaque

case si l'objet passe par cette case, 0 sinon³. Chacune de ces cases est alors un des neurones d'entrée (resp. de sortie) du réseau de neurones. Tous les exemples ont de cette manière le même nombre d'attributs (le nombre de cases). Sans rentrer en détail dans cette étude, retenons en qu'elle a mis en évidence les principales difficultés de ce type d'approche :

- les espaces de départ et d'arrivée (espace des jeux de coordonnées) sont grands et complexes, et donc le nombre d'exemples nécessaire à l'apprentissage est trop grand pour être recueilli en pratique.
- les descriptions des objets, même de types identiques, sont trop différentes les unes des autres, le nombre de coordonnées des objets étant variable, d'où des difficultés de représentation des exemples et la nécessité de changer de représentation.
- de plus ces descriptions « vecteur » et « raster » permettent de bien stocker les objets mais ne sont pas pertinentes du point de vue de l'analyse des objets. Par exemple une simple rotation d'un objet change complètement le jeu de ses coordonnées alors que le traitement cartographique est indépendant de l'orientation des objets.

4.2 Apprendre quelle type de transformation appliquer

Pour contourner ces problèmes de représentation, nous nous plaçons dans un autre type d'approche. Notre but n'est plus de déterminer directement un algorithme générique de transformation d'un jeu de coordonnées (comme un réseau de neurones), mais de déterminer, à partir de la description d'un objet, quelle transformation lui appliquer. En cela nous nous plaçons dans une approche plus adaptative [Lecordix et al., 96 ; Ruas 99].

Les recherches en généralisation cartographique automatique ont mis au point de nombreux algorithmes de transformation géométrique dédiés à un type de transformation particulier (simplification, exagération, déplacement...). Un des problèmes non encore résolu est de déterminer quelle opération appliquer sur un objet donné.

Des expériences de mise au point « manuelle » de bases de règles pour répondre à cette question ont été réalisées [Ruas, 98 ; Mustière, 98 ; Ruas, 99]. Ces expériences encourageant cette approche par la qualité cartographique des résultats obtenus. Cependant ces expériences mettent aussi en évidence les difficultés de cette approche : ces règles sont difficiles à formaliser tout d'abord parce que les règles cartographiques sont peu formalisées, ensuite parce qu'elles sont nombreuses, et enfin parce qu'elles sont contradictoires. Par exemple, deux règles sont "d'écarter les objets trop proches pour les rendre bien distincts" mais aussi "de conserver au mieux la précision de localisation des objets". Formaliser ces règles nécessite de définir ce que signifie « trop proche », « distincts » et

de formaliser le compromis entre précision de localisation et écartement suffisant.

Dans la suite de cette article, nous étudions la possibilité d'utiliser les outils d'apprentissage automatique pour formaliser ces règles.

Reformulons donc notre problème dans cette approche (cf. Fig. 7). Tout d'abord, l'espace des observables est modifié : nos objets ne sont plus représentés par leur jeu de coordonnées mais par un ensemble de mesures descriptives (taille, allongement, nombre de points...) issues des travaux réalisés en analyse spatiale [McMaster, 95 ; Plazanet, 96]. Ensuite, l'espace des classes est aussi modifié : notre but n'est plus d'apprendre une description de l'objet cartographié mais quelle transformation lui appliquer pour le cartographier. Par transformation nous entendons à la fois le type de transformation (simplification, exagération...), l'algorithme de transformation adapté, et enfin les valeurs paramétriques de cet algorithme. Ceci nous permet d'avoir une description plus pertinente de l'objet initial et de réduire considérablement l'espace des classes de notre apprentissage par rapport à l'approche précédente décrite au chapitre 4.1.

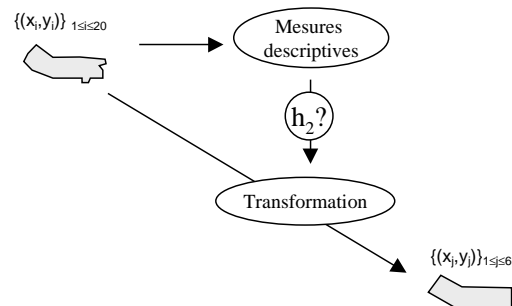


Fig. 7 Apprendre quelle transformation est efficace

4.3 Différencier abstraction et changement de langage dans l'apprentissage

Nous avons expliqué (cf. chapitre 3.2) que nos transformations cartographiques sont le fruit d'une abstraction pour réaliser un changement de langage efficace.

Ces considérations conduisent à définir notre apprentissage plus précisément. Il est nécessaire de distinguer lors de l'apprentissage l'acquisition des connaissances d'abstraction (quel caractère doit on faire ressortir d'un objet par rapport aux autres caractères) et les connaissances sur le changement de langage (quelles transformations seront efficaces pour mettre en valeur ce caractère).

Ce découpage en plusieurs phases de l'apprentissage est nécessaire à plusieurs niveaux.

Tout d'abord les hypothèses apprises, sous forme de règles par exemple, seront beaucoup plus faciles à analyser et à

³ cette représentation des objets est dite « raster » à opposer la représentation par la liste des coordonnées dite « vecteur »

valider par les experts cartographes.

Ensuite, ces règles ont pour but d'être utilisées au sein d'un système expert. Comme [Clancey, 83] l'a montré en étudiant le système expert MYCIN, si les règles du système contiennent différents types de connaissances sans que ceux-ci soient différenciés, il est impossible pour le système d'expliquer clairement les inférences qu'il effectue. Or cette capacité d'explication est une des premières qualités demandée à un système expert. Pour que ces règles soient utilisées en pratique efficacement dans un système expert il est donc nécessaire de différencier les différentes connaissances introduites dans ces règles dès la phase de leur acquisition [Thomas, 96].

Notre conception du problème d'apprentissage est donc mieux représentée dans la Fig. 8.

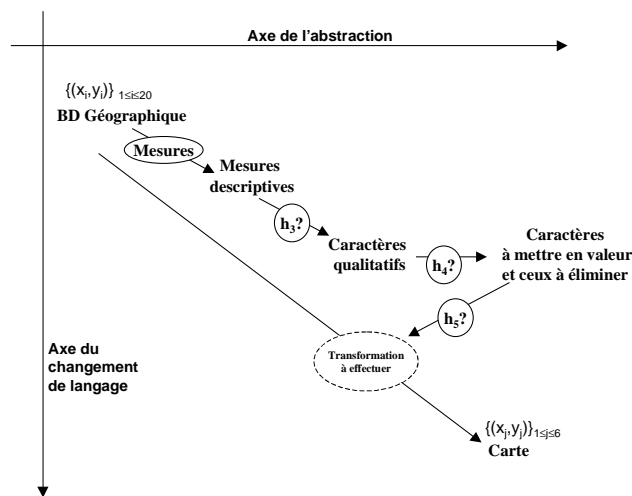


Fig. 8 Apprendre en différenciant abstraction et changement de langage.

L'apprentissage doit donc comporter plusieurs phases:

- tout d'abord relier les attributs descriptifs mesurés d'un objet à une abstraction et une reformulation de cette description, constituée de descripteurs plus abstraits et formulés qualitativement (c'est-à-dire des hypothèses du type « surface $< x \Rightarrow$ taille = petit ».
- ensuite à relier cette description abstraite au choix des caractères à mettre en valeur (taille = petit \wedge forme = très détaillé \Rightarrow petite taille à mettre en valeur, forme très détaillée à simplifier).
- et enfin de relier cette description à l'opération à effectuer sur cet objet pour bien le représenter. (petite taille à mettre en valeur \wedge forme très détaillée à simplifier \Rightarrow effectuer une simplification).

Ce processus d'apprentissage doit permettre de créer des règles plus faciles à analyser par un expert du domaine, comme celles par exemple décrites dans la Fig. 9.

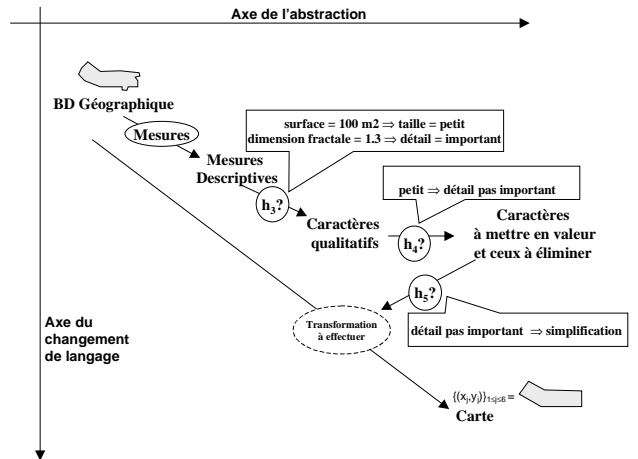


Fig. 9 Exemple de règles apprises

Un tel apprentissage nécessite une phase d'acquisition des exemples plus complexe qu'un apprentissage direct (cf. Fig. 7). En effet il devient nécessaire de demander à l'expert de décrire non seulement l'opération cartographique qu'il effectuerait sur un objet donné, mais également de décrire qualitativement cet objet et quel caractère il désire en abstraire pour le représenter sur la carte.

Cette complexité d'acquisition des exemples est néanmoins nécessaire à une analyse des hypothèses produites par les experts du domaine et à une utilisation en pratique de celles-ci.

5 Conclusion

Nous avons expliqué, de manière générale et en particulier en cartographie, la différence entre changer de langage de représentation et abstraire de l'information au sein d'un langage de représentation donné, et les connexions entre ces deux processus. Nous avons défini l'abstraction comme une relation entre deux contextes de raisonnement où la perception du monde tient une place primordiale. Ceci nous a guidé pour la mise au point d'un processus d'apprentissage de connaissances difficiles à formaliser en cartographie. Nous sommes actuellement en train de recueillir auprès de cartographes des jeux d'exemples pour valider l'intérêt de cette approche. Nous désirons comparer empiriquement les différentes approches présentées dans cet article, et en particulier l'intérêt d'un découpage du processus d'apprentissage en plusieurs phases. Cette évaluation doit être faite à deux niveaux, celui de l'expressivité des règles et celui de leur pouvoir prédictif.

Bibliographie

- Brassel K., Weibel R. 1988. A review and conceptual framework of automated map generalization. International Journal Geographical Information Systems, 1988, Vol. 2, no. 3 229-244
- Choueiry B., McIlraith S., Iwasaki Y., Loeser T., Neller T., Englemore R., Fikes R., 1998. Thoughts on a Practical

- Theory of Reformulation for Reasoning about Physical Systems. Actes de SARA'98. Pacific Grove, Californie.
- Clancey W. 1983. The Epistemology of a rule-based expert system - A framework for Explanation. Artificial Intelligence n. 20
- Davis R., Shrobe H., Szolovits P. 1993. What is in a Knowledge Representation. AI Magazine, Printemps 1993.
- Ellman T. 1993. Synthesis of Abstraction Hierarchies for Constraint Satisfaction by Clustering Approximately Equivalent Objects. Actes de International Conference on Machine Learning (Amherst, MA) pp. 104-111
- Fritsch E. 1997 *Représentations de la Géométrie et des Contraintes Cartographiques pour la Généralisation du Linéaire Routier*. Thèse de Doctorat, Université Marne-la-Vallée. Décembre 97.
- Ginsberg, 1997. Essentials of Artificial Intelligence, Morgan Kaufmann, 1997.
- Giunchiglia, F. and Walsh, T. 1992. A theory of Abstraction. Artificial Intelligence, 56, pp. 323-390.
- Iwasaki Y. 1990. Reasoning with Multiple Abstraction Models. Proceedings AAAI, Workshop on Automatic Generation of Abstraction for Planning. Artificial Intelligence, 68 (2).
- Keller S., 1994. On the use of case-based reasoning in generalization. Proceedings of SDH'94, Vol. 2, pp. 1118-1132
- Lang T. 1969 *Rules for the Robot Droughtsmen*. The Geographical Magazine Vol. 42(1) pp 50-51
- Lecordix F., Plazanet C., Lagrange J.-P., 1997. *A Platform for Research in Generalization: Application to Caricature*. GeoInformatica 1:2, pp 161-182 (1997).
- McMaster R. 1995 Knowledge Acquisition for Cartographic Generalization : experimental methods in Müller, Weibel, Lagrange Eds, GIS and Generalization, Ed. Taylor and Francis, London, pp 161-180
- McMaster R. and Shea K. 1992. Generalization in Digital Cartography. ISBN 8-89291-209-X. Association of American Geographers, Washington
- Mustière S. 1998. *GALBE: Adaptive Generalisation. The need for an Adaptive Process for Automated Generalisation, an Example on Roads*. GIS'Planet 98 proceedings. Lisbon September 98.
- Nyerges 1991 Representing Geographic Meaning in Map Generalization Ed. Longman Scientific & Technical pp 59-85
- Plazanet C. 1996 *Enrichissement des bases de données géographiques : analyse de la géométrie des objets linéaires pour la généralisation cartographique (application au routes)*. Rapport de Thèse, Univ. Marne-la-Vallée / IGN
- Ruas A. 1998. OO-Constraint modelling to automate urban generalisation process. Actes de 8th Spatial Data Handling. pp 225-235
- Ruas, A. 1999. *Modèles de généralisation de données géographiques à base de contraintes et d'autonomie*. Thèse de Doctorat. Université Marne-la-Vallée.
- Sacerdoti. E. 1973. Planning in a hierarchy of abstraction spaces. In Proceedings of the 3rd International Joint Conference in Artificial Intelligence (IJCAI), pp. 412-422
- Saitta, L. and J.-D. Zucker (1998). Semantic Abstraction for Concept Representation and Learning. Symposium on Abstraction, Reformulation and Approximation (SARA98), Asilomar Conference Center, Pacific Grove, California
- Sowa, J. F. 1984. Conceptual Structures: Information Processing in Mind and Machine. Addison-Wesley Publishing Company.
- Thomas J., *Vers l'intégration de l'apprentissage symbolique et l'acquisition de connaissances basée sur les modèles: le système ENIGME*. Thèse d'université, Univ P. & M. Curie Décembre 96
- Van Dalen, D. 1983. Logic and Structure. Springer-Verlag, Berlin, Germany.
- Weibel R., Keller S. & Reichenbacher T. 1995 Overcoming the Knowledge Acquisition Bottleneck in Map Generalization : the Role of Interactive Systems and Computational Intelligence. COSIT
- Werschlein T., Weibel R., 1994. Use of Neural Networks in Line Generalisation. EGIS/MARI'94 proceedings, Vol. 1, pp 77-85
- Yoshida K. and Motoda H. (1990). Towards Automatic Generation of Hierarchical Knowledge Bases. Proceedings AAAI, Workshop on Automatic Generation of Approximations and Abstractions (Boston, MA). pp. 98-109
- Zucker, J.-D. 1998. Abstraction for Concept Representation. The Fourth International Workshop on Multistrategy Learning (MSL'98), Lorenza Saitta ed., Desenzano del Garda (Brescia, Italy).