

## Chapitre 6

# Représentation multiple et généralisation avec UML et l'outil Perceptory

### 6.1. Introduction

Les modèles de bases de données géospatiales décrivent habituellement des objets avec une représentation géométrique unique et sans opération de généralisation. Il en est ainsi principalement parce que les applications SIG sont typiquement uniéchelles (malgré les fonctions de zoom) et parce que l'objet sémantique est souvent assimilé à sa représentation géométrique, héritage d'une vision de dessin (cf. logiciel de DAO) datant de l'époque où la carte papier était le support unique de la donnée. Aujourd'hui, dans un contexte moderne d'entrepôt de données géospatiales multiéchelles et multigéométries destiné à supporter la cartographie sur demande sur le web et les nouvelles technologies décisionnelles de type SOLAP (Spatial on-line analytical processing), nous devons modéliser la représentation géométrique multiple et la généralisation cartographique. Certes celles-ci peuvent être modélisées avec les formalismes habituels tels que l'entité-relation MERISE ou l'orienté-objet UML, par contre l'utilisation d'extensions explicites à ces formalismes permet d'accroître de façon importante l'efficacité du modélisateur et la richesse du schéma (schémas plus rapides à réaliser, plus simples, plus faciles à modifier, facilitant la compréhension, etc.).

*Perceptory* est justement un outil de modélisation visuelle qui utilise le formalisme UML avec des extensions spatiales pour construire des modèles conceptuels des objets de la base de données géospatiales [BÉD 99a, BÉD 02]. Ces modèles conceptuels se composent d'un schéma et d'un dictionnaire comprenant tant des composantes UML que des composantes provenant des normes géomatiques ISO/TC211 [BRO 00]. Les fonctionnalités du dictionnaire permettent de capter, emmagasiner et visualiser la description des éléments composant le modèle orienté objet de même que des détails nécessaires à la programmation de la base de données

## Représentation multiple et généralisation avec UML et l'outil Perceptory

spatiales ou spécifiant la saisie adéquate des objets spatiaux. Comme le formalisme UML fournit des composantes formelles d'extension, appelées stéréotypes, qui permettent de définir de nouveaux éléments de modélisations, Perceptory les utilise pour introduire une approche plus efficace pour supporter les composantes spatiales, temporelles et spatio-temporelles dans les modèles conceptuels de bases de données. Ces stéréotypes sont illustrés par des pictogrammes du PVL (Plug-in for Visual Language) utilisés dans Perceptory. Essentiellement, Perceptory supporte les éléments de modélisations requis par la norme ISO/TC211 (19103, 19110) [ISO 01a, ISO 01c]: les classes, attributs, opérations, paquetages, association généralisation et spécialisation, contraintes, note et stéréotypes. *Perceptory* se compose de générateurs de rapports permettant l'impression du contenu du dictionnaire de données et de générateurs de code permettant la génération du squelette de code pour créer la base de données sous Oracle8i spatial, Microsoft SQL Server, Microsoft Access, MapInfo, ArcView et Geomedia.

L'objectif du présent chapitre est de présenter la façon d'utiliser le PVL spatial de Perceptory pour modéliser efficacement la représentation géométrique multiple et les opérations de généralisation cartographique. Ainsi nous débutons avec une introduction aux composantes des diagrammes de classes du formalisme UML qui sont utiles pour modéliser les bases de données. Par la suite, nous expliquons la philosophie symbiotique sous-tendant à la fois Perceptory, l'approche que nous utilisons pour modéliser les bases de données géospatiales ainsi que la représentation géométrique multiple et la généralisation cartographique. Suivent ensuite une section sur le PVL spatial développé dans Perceptory ainsi qu'une autre section sur la modélisation de la représentation géométrique multiple et de la généralisation. Enfin, nous terminons avec une conclusion sur ces travaux.

### 6.2. Introduction aux diagrammes de classes d'UML

Les composantes du diagramme de classes d'UML sont la base du langage supporté par Perceptory. Les composantes nécessaires à la modélisation des bases de données géospatiales ainsi que celles supportées dans notre outil sont:

Les paquetages: Les paquetages sont des thèmes contenant des classes. Ils sont utilisés pour partitionner le modèle et regrouper des éléments de modélisation.

Les classes: Les classes sont composées de trois compartiments, un pour le nom, un pour les attributs et un pour les opérations (Les opérations permettent de définir le comportement d'une classe. Elles sont déclenchées suite à une stimulation externe (un événement)). Seul le compartiment comprenant le nom est obligatoire. Les classes peuvent être réelles ou abstraites, dérivées et/ou importées.

- réelle = instanciable.

## 2 Multiéchelle et généralisation

- abstraite (que l'on différencie par son nom en italique) = non instanciable. On retrouve habituellement les classes abstraites pour les super-classes résultant d'une généralisation.

- dérivée = qui origine d'un traitement. Le nom des classes dérivées est précédé du symbole /. On retrouve dans le dictionnaire de Perceptory un champ permettant de décrire la règle de dérivation de la classe.

- importée = qui provient d'un autre paquetage. On retrouve le nom du paquetage suivi du nom de la classe.

Les relations: Les relations sont des associations, des agrégations, des compositions, des généralisations ou des spécialisations.

- l'association est de nature sémantique et exprime une relation bidirectionnelle entre les classes. Elle possède une multiplicité (cardinalité) qui exprime les quantités minimum et maximum d'occurrences possibles dans chaque direction de l'association.

- l'agrégation et la composition sont des relations qui expriment un couplage plus fort (i.e. Père-Fils) où il y a un tout et ses parties ou un composé et ses composants. Le nom d'une relation d'agrégation et de composition est AVOIR ou COMPOSER.

- la généralisation et la spécialisation expriment une hiérarchie entre les classes par la présence d'une super-classe et de sous-classes. Les sous-classes héritent des attributs, des opérations, des relations et des pictogrammes de la super-classe. La spécialisation permet d'énoncer les particularités d'une sous-classe par rapport à la super-classe. Dans la conception d'une base de données, la super-classe d'une spécialisation est toujours réelle tandis qu'elle est toujours abstraite dans le cas d'une généralisation. Le nom d'une relation de généralisation et spécialisation est ÊTRE.

Le stéréotype: Le stéréotype est une extension prévue dans le métamodèle du formalisme UML qui permet de l'enrichir pour des domaines particuliers (comme la référence spatiale) par la création de nouveaux éléments de modélisation et éventuellement la définition d'une représentation graphique particulière. Les stéréotypes s'expriment par l'ajout dans la zone de nom de la classe, d'un mot entre guillemets «», d'un pictogramme ou encore des deux représentations.

De nombreux ouvrages décrivent le formalisme UML de façon général [ERI 98, RUM 99] ou de façon plus spécifique pour les bases de données [NAI 01] et les modèles de classes [SAT 02].

### 6.3. Philosophie symbiotique sous-tendant Perceptory

La modélisation conceptuelle d'une base de données géospatiales a pour but de décrire les besoins du client en terme de données à référence spatiale. Une représentation graphique doit être réalisée avec un formalisme simple et clair, ayant une grande richesse d'expression, afin de faciliter le processus de réflexion et de communication entre le client et l'analyste. Une description textuelle doit accompagner le schéma afin d'en expliquer les détails et d'ajouter d'autres informations complémentaires. L'outil informatique servant à la réalisation du modèle et au peuplement du dictionnaire doit lui-aussi être simple d'utilisation et le plus convivial possible.

Nos recherches en modélisation de base de données géospatiales ont considéré l'importance de la symbiose qui doit exister entre l'humain, le formalisme et l'outil informatique. Notre approche symbiotique est basée sur les résultats d'une étude empirique sur la compréhension et l'utilisation des solutions existantes, dans un contexte typique de développement de base de données spatiales. Cette étude a impliqué la réalisation de modèles dans plusieurs domaines d'application, de plusieurs variantes d'un même modèle et de grands modèles complexes. Ceci a permis d'ajuster notre approche en comparaison avec notre solution précédente (cf. MODUL-R [BED 96]).

Pour y parvenir, nous avons présenté des modèles et dictionnaires à des personnes différentes: des programmeurs, des experts de l'application et des patrons afin de voir si leur interprétation était semblable. Nous avons également observé les réactions des analystes de l'industrie ainsi que celles de quelques centaines d'étudiants dans l'utilisation de nos solutions. Des modèles produits par des analystes ont été soumis pour modification à d'autres analystes dans le but de voir si le même modèle était compris de la même façon par deux personnes différentes.

Cette étude a permis de confirmer les résultats obtenus en sciences cognitives. Les informations graphiques représentent bien les vues globales de la base de données et sont utiles pour illustrer les interdépendances entre les composantes du système. À l'opposé, l'information textuelle permet de mieux décrire les détails ponctuels et concepts abstraits. La combinaison des deux modes est indispensable pour parvenir à une bonne compréhension et offrir la richesse d'expression nécessaire pour les données spatiales tout en restant facile à lire. Ainsi, notre solution Perceptory:

- utilise de nouvelles informations graphiques au minimum ( $7 \pm 2$ );
- introduit les détails pointus dans le dictionnaire;
- introduit les contraintes d'intégrité spatiales dans le dictionnaire;
- précise les situations exceptionnelles dans le dictionnaire;

#### 4 Multiéchelle et généralisation

De plus, comme la modélisation des relations spatiales surcharge le modèle, qu'elle conduit à une implantation explicite de ces relations dans la base de données (avec les générateurs de code), et que l'on déduit principalement ces relations à partir des opérateurs spatiaux du SIG (i.e. par des opérateurs inter-objets), notre approche suggère de ne modéliser uniquement celles qui sont sémantiquement significatives ou particulièrement importantes pour le client.

Comme vue précédemment, les stéréotypes d'UML permettent d'étendre la richesse d'expression du langage de Perceptory. En utilisant un pictogramme comme stéréotype à côté du nom de la classe d'objets plutôt que de le mettre en attribut, on perçoit plus rapidement la nature des objets (i.e. spatial, temporel, spatio-temporel) dans le modèle (comme d'autres méthodes présentées dans [LAU, 01]). De plus, notre approche respecte les règles UML de construction de stéréotypes et place en évidence le caractère différent des objets.

D'abord, les pictogrammes du PVL spatial représentent le nombre de dimensions géométriques de base des objets cartographiés, soit (point 0D, ligne 1D polygone 2D), (figure 6.1, PVL spatial).

Ensuite, Perceptory offre des pictogrammes représentant la dimension temporelle de l'existence de l'objet et de son évolution descriptive (sous forme instantanée 0D ou durable 1D) (cf. figure 6.1, PVL temporel).

Enfin, Perceptory permet de combiner les pictogrammes temporels et spatiaux pour obtenir les pictogrammes d'évolution spatiale. Le pictogramme représente dans ce cas-ci la stabilité de la géométrie ou de la position de l'objet dans le temps (cf. figure 6.1, PVL spatio-temporel).















<i>PVL spatial</i>	<i>PVL temporel</i>	<i>PVL spatio-temporel</i>
 Bâtiment	Bâtiment 	 Bâtiment
 Géométrie 0D ex.: borne d'incendie	 Existence/évolution 0D ex: Accident	 Évolution spatiale ponctuelle instantanée  Évolution spatiale ponctuelle à effet durable
 Géométrie 1D ex.: route	 Existence/évolution 1D ex: Bâtiment	 Évolution spatiale linéaire instantanée  Évolution spatiale linéaire à effet durable
 Géométrie 2D ex.: lacs		 Évolution spatiale surfacique instantanée  Évolution spatiale surfacique à effet durable

Figure 6.1. :Pictogrammes du PVL spatial, temporel et spatio-temporel.

Par le développement des PVL et de Perceptory, il devient facile de modéliser des caractéristiques spatio-temporelles complexes, le tout conformément aux plus récentes tendances des méthodes agiles ou extrêmes [AWE 02 ;COC 02]. Les PVL temporel et spatio-temporel sont décrits plus en détail dans [BED 02].

#### 6.4. Le PVL spatial (Plug-in for Visual Language)



Les trois primitives géométriques de base du PVL spatial peuvent être utilisées seules pour décrire une géométrie simple ou bien être combinées pour décrire une géométrie alternative, complexe ou multiple. On dit d'une classe d'objets qu'elle est :

- à géométrie *simple*, lorsque chaque objet est représenté par une seule occurrence géométrique dont le type de géométrie est le même pour l'ensemble de des objets de la classe. La plupart du temps, la géométrie des objets d'une classe est simple. On ajoute alors un pictogramme spatial lorsque le client désire avoir la position et la forme d'un objet sur une carte (cf. figure 6.2 1<sup>ère</sup> colonne).

- à géométrie *complexe*, lorsque chaque occurrence est composée en même temps de deux ou plusieurs géométries de type différent. Il s'agit donc d'une agrégation de géométries composée des géométries du premier type ET des géométries de l'autre type pour chaque occurrence. On compose le stéréotype à l'aide des différentes géométries introduites dans un même pictogramme (cf. figure 6.2, 2<sup>ème</sup> colonne).

- à géométrie *alternative*, lorsque ses occurrences possèdent une OU l'autre des géométries proposées (de façon mutuellement exclusive). On compose alors le stéréotype avec des pictogrammes adjacents (cf. figure 6.2, 3<sup>ième</sup> colonne).

- à géométrie *multiple*, lorsque chaque occurrence possède autant de géométries qu'il y a de pictogrammes (i.e. 2 géométries ou plus) mais dont une seule est habituellement utilisée dans une représentation cartographique ou une requête spatiale (ce qui est différent d'une géométrie complexe, laquelle est un agrégat). On compose alors le stéréotype avec des pictogrammes espacés (cf. figure 6.2., 4<sup>ième</sup> colonne).

Il est possible de combiner plus de deux pictogrammes pour illustrer la géométrie de la classe d'objets. Lorsqu'il devient trop complexe d'exprimer toutes les géométries possibles avec Perceptory (qui a une liste volontairement limitée de combinaison de géométries au niveau de son interface), on utilise le pictogramme compliqué  dans le schéma et on précise les détails de cette géométrie dans le dictionnaire. Lorsque la géométrie de l'objet peut-être toute géométrie ou toute combinaison de géométries sans restriction on utilise alors le pictogramme .

## 6 Multiéchelle et généralisation

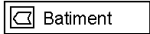
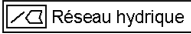
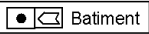

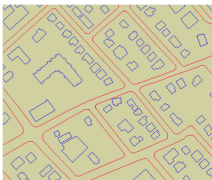
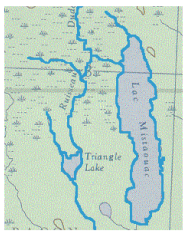

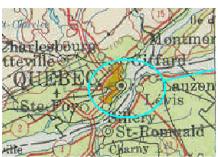
SIMPLE	COMPLEXE	ALTERNATIF	MULTIPLE
 Bâtiment	 Réseau hydrique	 Bâtiment	 Ville
 <p>Sur une carte au 1:1000, les bâtiments ont une géométrie simple de type polygone.</p>	 <p>Un réseau hydrique se compose d'éléments linéaires (rivières) et d'éléments surfaciques (lacs).</p>	 <p>Sur une carte au 1:20 000, les bâtiments ont soit une géométrie de type point ou une de type polygone.</p>	 <p>La ville de Québec a une géométrie ponctuelle pour indiquer son centre-ville et une géométrie polygonale pour représenter ses limites.</p>

Figure 6.2 Exemples des types de géométrie du PVL spatial.

### 6.4.1. L'utilisation de la cardinalité avec les pictogrammes

La cardinalité par défaut dans le cas de géométrie simple est de 1,1 et n'est pas affichée. Elle signifie qu'il y a un et un seul objet géométrique pour cartographier chaque occurrence de la classe d'objets (cf. figure 6.3, 1). Lorsque la cardinalité minimale est de 0, il s'agit d'une géométrie facultative (cf. figure 6.3, 2). Lorsque la cardinalité maximale est supérieure à 1, il s'agit de géométrie simple agrégée. Par conséquent, une géométrie avec une cardinalité 0,N est une géométrie simple agrégée facultative (cf. figure 6.3, 3).

Dans le cas de géométrie alternative, la cardinalité qui succède les pictogrammes s'applique à l'ensemble des pictogrammes et non uniquement à la dernière géométrie affichée. Par exemple, un verger peut être non cartographié, représenté par un point ou représenté par une surface (cf. figure 6.3.,5).

Pour une géométrie complexe, la cardinalité associée à chaque géométrie composant le pictogramme est sous-entendu supérieure à 1. Par exemple, un aéroport peut être composé d'un ensemble de lignes et de points (cf. figure 6.5, 6). Si une cardinalité succède le pictogramme complexe, celle-ci décrit le nombre d'agrégat géométrique de chaque occurrence sémantique et non le nombre d'objets géométriques simples composant l'agrégat. Ainsi, les aéroports peuvent être non cartographiés ou bien représentés par un agrégat de lignes et points (cf. figure 6.3., 7).

## Représentation multiple et généralisation avec UML et l'outil Perceptory

La géométrie multiple étant composée d'une séquence de pictogrammes pouvant être simple, complexe ou alternatif, l'application de la cardinalité dépendra donc des types de géométrie utilisés. Ainsi, l'exemple 8 de la figure 6.3. illustre une géométrie multiple dont l'une des géométries simples est facultative et l'exemple 9 montre une géométrie multiple ou l'une des géométries simples est agrégée.

Finalement, si l'expression de la cardinalité à l'aide des pictogrammes est trop complexe, le pictogramme *compliqué* est utilisé et il est décrit comment seront cartographiés les objets de cette classe dans le dictionnaire (cf. figure 6.3, 10).

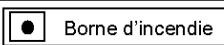
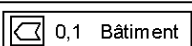
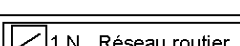
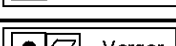
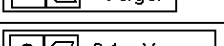
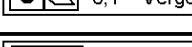
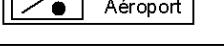
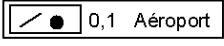
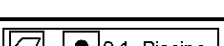
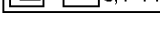
1	 Borne d'incendie	<b>Géométrie simple:</b> L'ensemble des bornes d'incendie sont cartographiées par un seul objet géométrique ponctuel.
2	 0,1 Bâtiment	<b>Géométrie simple facultative:</b> Certaines de occurrences de bâtiment ne sont pas cartographiées.
3	 1,N Réseau routier	<b>Géométrie simple agrégée:</b> un réseau routier est composé de 1,N segments linéaires.
4	 Verger	<b>Géométrie alternative:</b> les vergers sont soit représentés par un point ou bien par une surface.
5	 0,1 Verger	<b>Géométrie alternative facultative:</b> les vergers sont soit: non représentés, représentés par un point ou bien par une surface.
6	 Aéroport	<b>Géométrie complexe:</b> les aéroports sont représentés par une agrégation de lignes et points.
7	 0,1 Aéroport	<b>Géométrie complexe facultative:</b> les aéroports sont soit: non cartographiés ou cartographiés par un agrégat de lignes et de points.
8	 0,1 Piscine	<b>Géométrie multiple:</b> l'ensemble des piscines publiques sont cartographiées à grande échelle par un polygone et seulement les principales sont représentées par une point à petite échelle.
9	 1,N Verger	<b>Géométrie multiple:</b> chaque verger est représenté par un ensemble de points à grande échelle et par un polygone à petite échelle.
10	 Verger	<b>Géométrie compliquée:</b> les vergers sont soit: non représentés, représentés par plusieurs points ou bien par une surface.

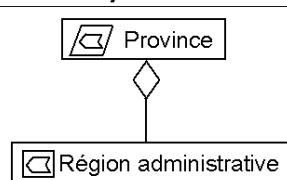
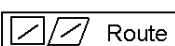

Figure 6.3. Exemples d'utilisation de la cardinalité avec les pictogrammes spatiaux.

### 6.4.2. La géométrie dérivée

Tout comme les attributs d'une classe, la géométrie peut être dérivée. Les pictogrammes sont alors écrits en *italique* pour rappeler visuellement la barre oblique / de UML qui est utilisée devant les attributs dérivés. Le traitement doit être complètement automatique, i.e. sans aucune manipulation manuelle, sinon nous considérons que la géométrie n'est pas dérivée. Dans le cas des géométries



multiples et alternatives, il est possible que seulement l'une des géométries soit dérivée (cf. figure 6.4).

<b>Géométrie simple dérivée</b>	<b>Géométrie alternative dérivée</b>	<b>Géométrie multiple dérivée</b>
		
<p>La géométrie de la province est déduite de la géométrie du contour de l'ensemble des régions administratives qui la composent.</p>	<p>Chaque occurrence de route a soit: sa géométrie saisie ou bien sa géométrie dérivée.</p>	<p>Chaque occurrence de route a une géométrie linéaire saisie à grande échelle et une géométrie linéaire dérivée à petite échelle.</p>

**Figure 6.4.** Exemples de géométrie dérivée.

#### 6.4.3. Pictogrammes aux attributs

Il est possible pour un attribut d'avoir une géométrie que l'on représente avec un ou plusieurs pictogrammes placés à droite du nom de l'attribut. Cette géométrie est alors dépendante de la géométrie de la classe d'objets à laquelle elle est liée. Il n'y a pas de restriction quant aux types de géométrie (simple, complexe, alternatif ou multiple) que l'on peut utiliser. Évidemment, il s'agit d'exprimer ici la perception que le client a de sa réalité et non ce qui sera implanté puisque c'est au niveau conceptuel que nous modélisons avec Perceptory.

Le pictogramme d'un attribut peut être dérivé, i.e. que la géométrie de l'attribut est obtenue par traitement. La dérivation de la valeur de l'attribut et la dérivation de sa géométrie ne dépendent pas nécessairement l'une de l'autre.

#### 6.4.4. Détails de la géométrie

Il est possible dans le dictionnaire de l'outil Perceptory de détailler la géométrie plus qu'elle ne l'est dans le diagramme. Il est d'ailleurs souhaitable et reconnu comme bonne pratique de ne pas surmodéliser dans le schéma mais plutôt de décrire les détails en langage naturel (ex. : français) ou en langage formel (ex. :Z, OCL).

Une section du dictionnaire permet de définir précisément la géométrie de l'objet, i.e. les spécifications de numérisation (le « quoi » et non le « comment »).

## Représentation multiple et généralisation avec UML et l'outil Perceptory

On y retrouve l'information supplémentaire requise suivant le type de géométrie de l'élément (0D, 1D, 2D) ainsi que les dimensions minimales requises pour numériser l'objet et choisir son type de géométrie lors de la numérisation. On y inscrit la largeur, la longueur, la hauteur, l'étendu et une règle.

La règle permet de décrire toute particularité ou information complémentaire sur la définition géométrique d'un élément. Si nécessaire, on explique les détails de la forme de la géométrie de l'élément.

Une autre section permet de décrire plusieurs métadonnées dont:

- **Les sources initiales** utilisées pour numériser ou importer l'élément (ex. documents cartographiques, descriptifs, des photographies aériennes, des images satellites, des relevés terrain). On indique de plus la date de mise à jour (document) ou la date de prise de données (relevé) pour cette classe d'objets.

- **La mise en garde** permet à l'utilisateur de mieux évaluer la validité de l'information géométrique incluse dans la base de données (ex. normes d'acquisition, précision).

- **Le responsable** ou les personnes responsables des informations relatives à la géométrie de l'élément.

Toutefois, si ces métadonnées deviennent plus importantes et que l'on désire les gérer de façon plus efficace, il est alors conseillé de plutôt favoriser un lien vers un géorépertoire [PRO 97] ou un outil de gestion de métadonnées conforme aux normes ISO/TC 211 (19115) [ISO 01d]

Finalement, le champ **Derivation Rule** permet de décrire la règle de dérivation de la géométrie lorsque celle-ci provient d'un autre élément du modèle (i.e. est dérivée).

Il est possible de déterminer les types de données spatiales pour cette classe pour chacune des plates-forme d'implantation et dans la norme ISO TC/211 (19107) [ISO 01b]. Quoiqu'il ne s'agisse pas ici d'information de niveau conceptuel, il est possible de la spécifier immédiatement dans le modèle et ainsi faciliter la génération automatique de code (une telle possibilité de modélisation incrémentale est typique des approches orientées-objets).

### 6.5 Modélisation de la représentation multiple et de la généralisation dans Perceptory.

Il existe plusieurs façon de modéliser la généralisation. Sur le plan conceptuel, toutes ces manipulations d'objets se traduisent par des opérations sur la géométrie des objets. Le résultat de ce processus peut ensuite être stocké ou non résultant ou

non en représentation multiple. Voici les différents types de généralisation et de représentation multiple supportés dans Perceptory:

- On peut généraliser à la volée, i.e. sans stockage des résultats obtenus et ainsi ne pas nécessiter l'utilisation de représentations multiples.
- On peut généraliser automatiquement les objets et stocker le résultat sous forme de représentations multiples (pour des besoins d'optimisation);
- On peut généraliser semi-automatiquement les objets et stocker les résultats<sup>1</sup>;
- On peut généraliser manuellement les objets et stocker les résultats sous forme de représentations multiples ;
- On peut saisir les différentes représentations multiples des objets et les stocker telles quelles;

### **6.5.1 Modélisation- règles générales**

Dans notre approche, les opérations de généralisation sont modélisées de la même façon que n'importe quelle autre opération dans un modèle orienté-objets. Selon les règles typiques de modélisation, seules les données et les opérations à programmer seraient modélisées et non les opérations ayant servi à peupler cette base de données. Par conséquent, les généralisations donnant un résultat à la volée (i.e. entièrement automatisées) seraient modélisées tandis que les opérations de généralisation qui servent à peupler les représentations multiples de la base de données ne le seraient pas. En appliquant ces principes de modélisation, on obtient un modèle de base de données qui correspond au système tel qu'implanté.

Cependant, dans un contexte d'incertitude technologique actuel, il est souhaitable d'utiliser une approche différente. Puisque la technologie tend à évoluer rapidement entre la phase de conception et la réalisation de tels systèmes, certaines décisions prises au début du processus de modélisation devront être reconsidérées au moment de l'implantation du système. Par conséquent, comme on ne peut prévoir initialement ce qui peut être programmé pour obtenir un traitement entièrement automatisé versus ce qui doit être peuplé à priori (ex. résultant de généralisation manuelle), il est profitable de procéder par itération. Dans un contexte technologique idéal, il sera possible de dériver automatiquement et rapidement l'ensemble des géométries à l'aide de traitements automatisés permettant ainsi d'éviter de stocker ces représentations dans le système final. Il en résultera un modèle final sans représentation multiple.

Une approche itérative permet de produire des modèles intermédiaires illustrant l'existant (i.e. les traitements requis pour peupler la base de données) et de faire plus

---

1. Il est à ce jour préférable de stocker le résultat de généralisation semi-automatisée afin d'assurer un résultat identique d'une généralisation à l'autre.

## Représentation multiple et généralisation avec UML et l'outil Perceptory

facilement évoluer les modèles en fonction de l'évolution des logiciels de généralisation ou du développement interne de routines spécialisées. Ceci adhère totalement au principe de modélisation itératif et incrémental suggéré par les méthodes les plus récentes en développement de système (ex. RUP, AGILES) [COC 02, AWE 02].

### 6.5.2 Modélisation règles spécifiques

Les règles de modélisation permettant d'exprimer la généralisation et la représentation multiple selon notre approche sont décrites ci-dessous:

Règle #1: Toutes les opérations de généralisation sont définies par une opération dans la classe d'objet (section opération) que ces processus soient manuels, semi-automatisés ou entièrement automatisés.

Règle #2: Toutes les représentations multiples sont définies par un pictogramme de géométrie multiple dans la classe d'objet.

Règle #3: Si la géométrie résultante d'une opération de généralisation doit être stockée dans le système car trop complexe à mettre en œuvre ou à exploiter à la volée, cette géométrie de l'objet sera ajoutée en représentation multiple dans la classe d'objet à l'aide des pictogrammes de géométrie multiple (cf. règle # 2).

Lors de l'élaboration du modèle, il peut être nécessaire de procéder par itération afin de statuer sur l'ensemble des processus de toutes les classes d'objets de la base de données. Éventuellement, il ne restera dans le modèle du système que les opérations qui nécessitent de la programmation.





### 6.5.3 Notation graphique:



Voici la notation graphique utilisant les opérations et les PVL spatiaux de Perceptory permettant de modéliser les règles de représentation multiple et de généralisation décrites à la section précédente.

#### Opération de généralisation:



- Le nom de l'opérateur de généralisation sert à identifier l'opération s'il est unique (ex. raffinement, réduction de dimension, simplification, resymbolisation, caractérisation, exagération, déplacement, déformation, agrégation et lissage [ MAR 97]). S'il s'agit d'un ensemble complexe d'opérateurs, l'opération générale "GEN" est utilisée en consignnant la séquence des opérateurs dans le dictionnaire de données de Perceptory.

## 12 Multiéchelle et généralisation

- La géométrie initiale et la géométrie résultante de la généralisation sont indiquées dans l'opération à l'aide des pictogrammes spatiaux. Le pictogramme de géométrie résultante est en italique s'il est en partie automatisable (i.e. dérivable). (ex. Gen  -> **). Par conséquent, on peut mettre l'emphase sur la complexité des traitements entièrement manuels d'une opération de généralisation en enlevant l'italique du pictogramme de la géométrie résultante (i.e. Gen  -> ).

- L'échelle ou le niveau géographique des géométries est identifié dans l'opération de généralisation. Pour ce faire une codification prédéfinie dans Perceptory ou l'échelle cartographique abrégée est utilisée (ex. GEN  20K -> ** 100K).

### Représentation multiple:

- Les géométries sont définies par un pictogramme de géométrie multiple dans la classe d'objet (coin supérieur gauche) (ex.  ).

- Les géométries sont ordonnées du niveau géographique détaillé au niveau plus grossier (ex. de la plus grande à la plus petite échelle).

- Le pictogramme de géométrie résultante est indiqué sans italique puisque la géométrie provient directement d'une saisie sans être dérivée.


## 6.6. Exemples d'applications



Cette nouvelle approche a été expérimentée pour la modélisation de l'entrepôt de données géospatiales du Ministère des Ressources Naturelles du Québec (MRNQ) [BÉD 01a]. Des exemples utilisant des objets cartographiés aux échelles 20 000, 100 000 et 250 000 (cf. 20K, 100K et 250K) sont tirés directement de cette application et expliquent les différents concepts présentés jusqu'à maintenant. Le premier exemple de modélisation illustre la généralisation d'un objet où toutes les occurrences sont généralisables. Le second exemple présente la généralisation d'un objet où certaines occurrences ne peuvent pas être généralisées à petite l'échelle et, par conséquent, doivent être saisies en représentation multiple. Finalement, le dernier exemple de modélisation illustre un objet où la généralisation à plus petite échelle implique un remplacement d'objet sémantique en plus du changement de géométrie.

### **6.6.1. Modélisation d'un objet où toutes les occurrences sont généralisables.**

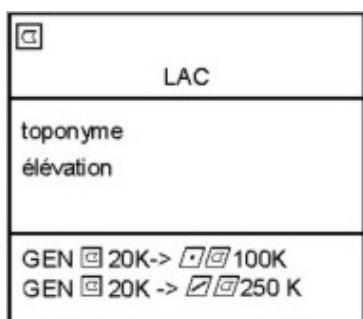
Dans le modèle, toutes les géométries saisies ainsi que toutes les généralisations potentielles sont décrites dans la classe d'objets. Par exemple, un lac est saisi à

## Représentation multiple et généralisation avec UML et l'outil Perceptory

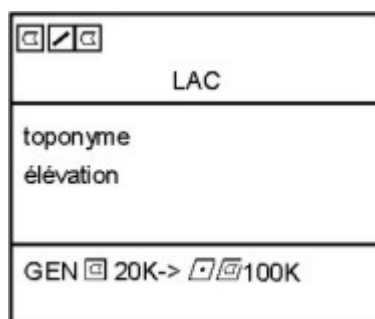
l'échelle du 20K comme une surface si superficie  $> 600\text{m}^2$ . Dans la figure 6.5, un pictogramme de géométrie surfacique (ex. ) est indiqué dans le coin gauche de la classe. La règle géométrique indiquant la superficie minimale de l'objet sera consignée dans le dictionnaire de données .

Le lac est ensuite généralisé du 20K vers le 100K comme une surface si superficie  $> 15000\text{m}^2$  avec une largeur minimale  $> 100\text{m}$  et comme un point si superficie  $< 15000\text{m}^2$ . Dans la figure 6.5, une opération de généralisation est inscrite dans la classe d'objet. Cette opération est complexe car elles se composent de plusieurs opérateurs de généralisation exécutés dans un ordre précis (une sélection d'objet selon des dimensions précises, la simplification de vecteurs, des déplacements et un changement de dimension pour passer du polygone au point (i.e.  à )). L'ensemble des détails de ces manipulations est consigné dans le dictionnaire de données de Perceptory.

Finalement, de la même manière que pour les opérations de généralisation précédentes, nous indiquerons que le lac est généralisé du 20K au 250K comme une surface si superficie  $> 50000\text{m}^2$  avec une largeur minimale  $> 1000\text{m}$  et comme une ligne si sa largeur minimale  $< 1000\text{m}$ . Il en résulte dans la classe d'objet LAC illustrée dans la figure 6.5, le stockage de la représentation au 20K et 2 opérations de généralisation pour passer aux échelles 20K -> 100K et 20K -> 250K.



**Figure 6.5.** Classe d'objet LAC telle que modélisée.



**Figure 6.6.** Classe d'objet LAC telle que modélisée après une itération.

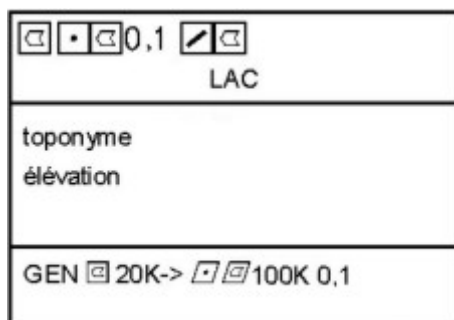
En cours de modélisation, il est possible d'identifier les opérations trop complexes à automatiser qui devront être stockées en représentations multiples dans le système. Par exemple, l'opération de généralisation permettant de produire l'objet au 250K est si complexe et demande tant d'interventions humaines qu'il est souhaité de stocker ces nouvelles géométries en représentation multiple une fois que la généralisation est complétée. Par conséquent, l'opération de généralisation sera

détruite de la classe et sera remplacée par les pictogrammes correspondant aux représentations multiples stockées. La figure 6.6 illustre la classe d'objet LAC modélisée après une itération. D'autres versions du modèle pourront être réalisées tout au long des expérimentations techniques avant d'obtenir le modèle final.

### 6.6.2. Modélisation d'un objet où certaines occurrences sont généralisables et d'autres doivent être saisies en représentation multiple.

Dans l'exemple suivant, il est possible d'envisager que la couverture spatiale d'un niveau géographique source soit différente de la couverture du niveau résultant. Par exemple, il est possible que les lacs saisis à l'échelle du 20K soient ceux localisés au sud du 52<sup>ème</sup> parallèle, mais qu'on désire tout de même couvrir l'ensemble du territoire québécois avec les objets à l'échelle du 100K. Alors, seulement un sous-ensemble des lacs du 100K pourront être généralisés à partir de l'échelle du 20K et les autres lacs complétant la couverture devront être saisis au 100K. Par conséquent, il est utile d'indiquer une telle caractéristique dans notre modèle.


Pour y parvenir, les cardinalités appliquées aux éléments de modélisation permettront d'indiquer que les géométries et l'opération sont facultatives. Ainsi, les cardinalités associées aux géométries et à l'opération du niveau 100K indiquent que l'objet LAC peut être saisi (*cf.* figure 6.7 pictogramme de représentation multiple facultatif 0,1) ou généralisé (*cf.* figure 6.7, opérations de généralisation -> 100K facultative 0,1). Les détails et les règles d'application de l'opération sont consignés dans le dictionnaire d'objets.



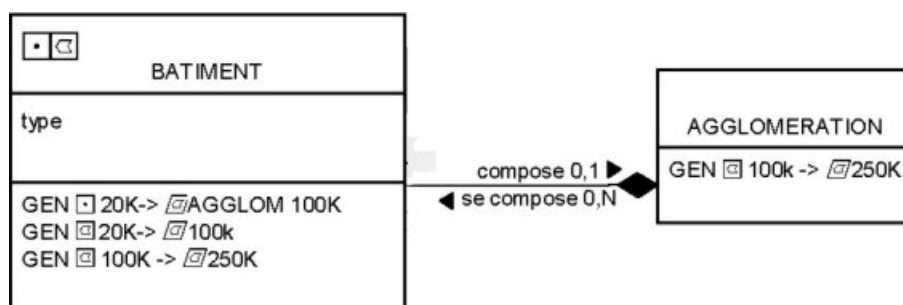
**Figure 6.7.** Classe d'objet LAC exprimant qu'un sous-ensemble des objets d'un niveau cartographique est saisi ou généralisé.

**6.6.3. Modélisation d'un objet où la généralisation implique un remplacement d'objet sémantique en plus du changement de géométrie.**

Dans certains cas de généralisation, des objets d'un niveau cartographique détaillé changent sémantiquement d'appellation. Par exemple, des bâtiments ponctuels qui se trouvent à une certaine distance les uns des autres deviennent une agglomération urbaine polygonale lors de la généralisation à plus petite échelle et les bâtiments isolés (i.e plus de la distance) ne sont pas conservés.

Dans le modèle, il doit exister au départ une relation sémantique de composition entre les deux objets. Lors du processus de généralisation, l'objet initial (bâtiment) cèdera la place à un autre objet (agglomération) s'il répond aux critères (moins d'une certaine distance). Il est possible à l'aide de notre notation graphique d'exprimer précisément un tel remplacement en indiquant directement dans l'opération de généralisation le nom de l'objet de remplacement (ex.  AGGLOM 100K).

L'exemple de la figure 6.8 illustre le cas de bâtiments ponctuels et polygonaux à l'échelle du 20K. Lors de la généralisation à l'échelle du 100K, les bâtiments ponctuels sont remplacés par une agglomération urbaine si ces bâtiments sont situés à moins de 200m les uns des autres. Les bâtiments ponctuels isolés ne sont pas conservés à plus petite échelle (cf. absence de géométrie ponctuelle résultante au 100K, figure 6.8). Par contre, les bâtiments polygonaux sont généralisés à l'échelle du 100K (cf. opération sur les polygones 20K, figure 6.8).



**Figure 6.8** Classe d'objet *BATIMENT* modélisée pour indiquer le changement sémantique vers l'*AGGLOMÉRATION*.

Évidemment, dans le modèle final, il est fort probable que l'analyste décide qu'il est plus approprié pour des raisons d'efficacité de stocker la géométrie polygonale de l'agglomération à l'échelle du 100K au lieu de la dériver de la géométrie initiale puisque cette géométrie sert à la généralisation de d'autres géométries à l'échelle du



250K. Par conséquent, dans le modèle final l'agglomération aura une représentation polygonale stockée au 100K.

Le même exemple aurait pu être illustré par des pistes d'atterrissage linéaires qui céderaient la place à plus petite échelle à un aéroport polygonal.

### **6.7. Conclusion**

Il a été démontré dans le présent chapitre qu'il est possible d'utiliser le PVL spatial de Perceptory pour modéliser efficacement la représentation géométrique multiple et les opérations de généralisation cartographique. Par sa philosophie symbiotique, Perceptory est un outil adéquat pour faciliter le travail de l'analyste tout en supportant la richesse d'expression pour modéliser les spécificités des bases de données géospatiales multiéchelles. Un sous-ensemble seulement de notre solution a été présenté, mais elle supporte également les représentations multiples et généralisations sémantiques et temporelles (ex. représentation multiple de plusieurs attributs pour une classe d'objet, généralisation de deux attributs en un seul, raffinement d'attribut (i.e. absence d'attribut à une échelle)).

Cette nouvelle approche a été expérimentée pour la modélisation de l'entrepôt de données géospatiales du Ministère des Ressources Naturelles du Québec (MRNQ) [BÉD 01a]. Il a été ainsi possible de représenter adéquatement les différents niveaux d'agrégation spatiale de l'entrepôt, selon 5 échelles cartographiques (20K, 100K, 250K, 1M et 8M) pour les objets de 10 produits sources (ex. Système de découpage administratif 20K et 250K, Base de données topographique du Québec 20K, Base de données topographiques et administratives 250K). La construction du modèle du MRNQ permettra à l'organisation de répartir adéquatement les opérations de représentation multiéchelle ou de généralisation cartographique sur les objets sources pour concevoir un entrepôt de données géospatiales optimal. D'autres projets sont en cours et permettront de tester davantage notre solution sur des domaines d'applications connexes afin d'apporter des améliorations dans le futur.

### **6.8. Remerciements**

Les auteurs tiennent à remercier pour leur participation financière le Réseau de Centres d'Excellence GEOIDE par le biais du projet DEC#9: "Développement de techniques automatisées pour l'extraction et la généralisation de l'information géospatiale à partir de données de télédétection hyperspatiale", le Conseil de Recherche en Sciences Naturelles et en Génie du Canada (CRSNG) et le Ministère des Ressources Naturelles du Québec.

## 6.9 Bibliographies

- [AWE 02] AWERK, K & MILLER R., « Extreme Programming Applied », Addison-Wesley, 326 p.
- [BÉD 99a] BÉDARD Y., « Visual Modeling of Spatial Databases Towards Spatial Extensions and UML », *Geomatica*, Vol 53, No2, pp. 169-186, 1999.
- [BÉD 01a ]Y. BÉDARD, PROULX, M.-J. LARRIVÉE, S., NADEAU, M., FRENETTE, P., RIVEST S., « Mise en place d'une infrastructure d'accès aux données de la DGIG- Modélisation des données de l'entrepôt, Avis sur les orientations d'architecture technologique pouvant influencer la modélisation. » Version préliminaire du 21 mai, Rapport de recherche produit pour le Ministère des ressources naturelles du Québec, Centre de recherche en géomatique, Université Laval, 65 pages.
- [BÉD 02] BÉDARD Y., « Site Web de Perceptory ». <http://sirs.scg.ulaval.ca/perceptory>, 27 mars.
- [BRO 00] BRODEUR J., BÉDARD Y., PROULX M.-J., « Modelling Geospatial Application Databases using UML-bases Repositories Aligned with International standards in Geomatics », *ACMGIS 2000*, November 10-11, Washington DC, USA
- [COC 02] COCKBURN A, « Agile Software Development », Addison-Wesley, 278p.
- [ERI 98] ERIKSSON H-E, PENKER M, « UML Toolkit », Willey & son, 397 pages.
- [ISO 01a] INTERNATIONAL ORGANISATION FOR STANDARDIZATION, ISO/TC 211 19103 « Geographic Information- Conceptual schema language », Draft for International Standard, 75 pages
- [ISO 01b] INTERNATIONAL ORGANISATION FOR STANDARDIZATION, ISO/TC 211 19107 « Geographic Information- Spatial Schema », Draft for International Standard, 185 pages
- [ISO 01c] INTERNATIONAL ORGANISATION FOR STANDARDIZATION, ISO/TC 211 19110 « Geographic Information- Methodology for feature cataloguing », Draft for International Standard, 60 pages
- [ISO 01d] INTERNATIONAL ORGANISATION FOR STANDARDIZATION, ISO/TC 211 19115 « Geographic Information- Metadata », Draft for International Standard, 151 pages
- [LAU, 01] LAURINI, R., « Information Systems for Urban Planning », Taylor & Francis, 349 pages.
- [NAI, 01] NAIBURG E.J & R.A. MAKSIMCHUK, « UML for Database Design », Addison-Wesley, 300 pages
- [PRO 97] PROULX M.-J., BÉDARD Y., LETOURNEAU F., MARTEL C., « Catalogage des données spatiales sur le world wide web: concepts, analyses des sites et présentation du géorépertoire personnalisable GEOREP », *Revue Internationale de Géomatique*, volume 7, no.1, Éditions Hermès, Paris, p. 7-32, mars 1997
- [RUM 99] RUMBAUGH, J., JACOBSON, I, BOOCH G., « The Unified Modeling Language Reference Manual », Addison-Wesley, 550 pages.
- [SAT 02] STARR, L. Executable UML: How To Build Class Models. Prentice Hall, 418 p.