

CONCEPTION DE PROJET SIG AVEC UML

François LAPLANCHE

Résumé

Avec son statut de standard, le langage UML (Unified Modelling Language) jouit d'une diffusion très importante en conception de systèmes d'information, mais qu'en est-il en conception géographique ? En fait, certaines applications utilisant le langage UML en conception de SIG ont déjà vu le jour. Après avoir succinctement présenté UML, nous présenterons deux extensions de ce langage dans le domaine de l'information géographique et réaliserons une comparaison entre UML et différents langages, mettant ainsi en évidence les attraits et les lacunes de son utilisation dans des applications géographiques. Bien que très intéressante, l'utilisation d'UML en conception de SIG ne règle pas tous les problèmes, le langage reste en effet inadapté à certaines spécificités des données géographiques. Pour l'adapter, il serait intéressant d'utiliser les capacités d'extension du langage UML. Pour notre part, nous proposons enfin une méthode de prise en compte des relations topologiques basée sur l'utilisation de deux types de matrices topologiques.

Mots-clés

conception, modélisation, SIG, UML

Abstract

With its status of standard, UML (Unified Modelling Language) relishes from a very important diffusion in the design of information systems, but what's about its use in geographic design ? Actually, some applications using UML already exist in GIS design. After introducing UML we'll present two UML extensions in the geographic information domain and we'll realize a comparison between UML and other languages showing the attractiveness and the gaps of its use in geographic applications. Although the use of UML in GIS conception is very interesting, it doesn't resolve all problems. Indeed the language stays not adapted to some specificities of geographic data. To adapt UML, it will be interesting to use its capabilities of extension. We propose in this paper an approach to consider topological relationships based on the use of two sorts of topological matrixes.

Keywords

conception, design, GIS, UML

I. INTRODUCTION

La modélisation conceptuelle de systèmes est souvent comparée au plan de l'architecte en génie civil. En effet, plus l'ouvrage est important, plus le soin qu'on porte à cette étape est primordial. D'autre part, la standardisation des méthodes de conception offre la possibilité d'une automatisation plus aisée et facilite l'échange de métadonnées. UML (Unified Modelling Language) étant un standard pour la modélisation de systèmes, l'intérêt de son utilisation en conception géographique est évidente.

Dans le paragraphe suivant, le langage UML sera brièvement présenté ainsi qu'une extension dont il fait l'objet au centre de géomatique de l'Université Laval du Québec.

Le paragraphe 3 sera consacré à une comparaison entre le formalisme UML étendu et d'autres formalismes, prin-

cipalement des langages adaptés à la conception de SIG. L'utilisation d'UML au sein du logiciel ArcGIS 8 sera envisagée au paragraphe suivant.

Nous proposerons enfin, avant de conclure, une méthode de prise en compte des relations topologiques basée sur l'utilisation de deux types de matrices topologiques.

II. UML : LE LANGAGE DE MODÉLISATION OBJET UNIFIÉ

A. Le langage UML

Ayant été reconnu comme standard, UML jouit d'une documentation importante ; en effet, de nombreux sites Internet (voir particulièrement (RATIONAL 1999)) lui sont consacrés ainsi que de nombreux ouvrages (entre autres Booch et *al.*, 2000 ; Kettani et *al.*, 1999 et Lopez et *al.*, 1998).

UML est un langage de modélisation de données orienté objet basé sur l'utilisation de neuf types de diagrammes : les diagrammes de classes, les diagrammes d'objets, les diagrammes de composants, les diagrammes de déploiement, les diagrammes de cas d'utilisation, les diagrammes de collaboration, les diagrammes de séquence, les diagrammes d'états-transitions et les diagrammes d'activités.

Les quatre premiers diagrammes permettent de visualiser les parties statiques d'un système, tandis que les cinq autres aident à décrire le comportement dynamique du même système.

transitions sont équivalents d'un point de vue sémantique. Les premiers représentent une succession d'activités tandis que les seconds représentent des objets dans des états différents et les événements entraînant ces changements d'états.

Enfin les diagrammes de collaboration et de séquence sont regroupés sous le terme générique de diagrammes d'interaction. Ils sont également équivalents d'un point de vue sémantique. Le diagramme de collaboration insiste sur l'organisation des objets participant à une interaction alors que le diagramme de séquence met en avant l'ordre chronologique d'apparition des messages.

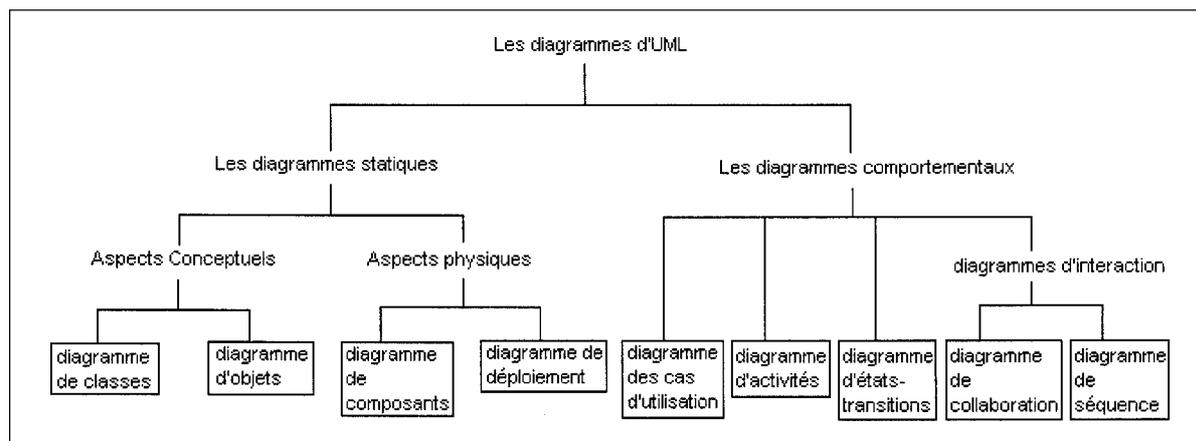


Figure 1. Les diagrammes d'UML

Au sein des diagrammes statiques, on distingue ceux décrivant les aspects conceptuels d'un système, et ceux décrivant les aspects physiques ou d'implémentation. Nous ne nous attarderons pas sur ces derniers puisqu'ils ne relèvent pas de la phase de conception.

Le diagramme de classes est le diagramme de base permettant de décrire la structure d'un système. Il est constitué de paquetages (éléments de regroupement), de classes, de relations (associations, généralisations, dépendances, agrégations ou compositions), d'interfaces (classes particulières ne contenant que des opérations), de stéréotypes (éléments permettant d'étendre le métamodèle), de notes et de contraintes.

Le diagramme d'objets, quant à lui, représente un état instantané d'une partie du système. Il comprend un ensemble d'objets et les liens entre ces objets existant à un moment donné.

La deuxième grande famille de diagrammes est celle des diagrammes comportementaux, on y retrouve d'abord les diagrammes de cas d'utilisation. Ceux-ci reposent sur les cas d'utilisation inventés par Ivar Jacobson et qui représentent une interaction particulière entre le système et les différents acteurs. Les diagrammes de cas d'utilisation décrivent les exigences auxquelles le système doit pouvoir répondre. Il constitue donc une étape d'analyse préalable à la conception d'un système.

Les diagrammes d'activités et les diagrammes d'états-

Cette volonté de représenter de deux manières différentes les mêmes événements s'explique par la conviction des auteurs que la compréhension d'un système est grandement facilitée quand on l'envisage selon des angles de vue différents.

B. Un géo-UML : les PVL (Plug-in for Visual Language) et Perceptory

L'équipe du professeur Yvan Bédard du centre de géomatique de l'Université Laval au Québec a développé une méthode générale qui permet d'étendre n'importe quel formalisme graphique de modélisation pour la conception géographique. Cette méthode est basée sur les PVL (Plug-in for Visual Language) (Bédard, 1999 ; Brodeur et al., 2000 ; Proulx et Bédard, 2001).

Cette même équipe a également proposé une extension du formalisme UML à partir des PVL et a développé un module spécial appelé Perceptory qui est intégrable à l'atelier de génie logiciel (AGL) Visio et qui supporte cette extension (Bédard, 1999 ; Brodeur et al., 2000 ; Proulx et Bédard, 2001).

La figure 2 ci-dessus schématise la manière de représenter la spatialité des objets à l'aide des PVL.

En ce qui concerne la temporalité des objets, les auteurs proposent de distinguer, d'une part, l'existence et, d'autre part, l'évolution. Ils distinguent également une évolution descriptive et une évolution spatiale. La figure 3 présente les différents PVL spatio-temporels.

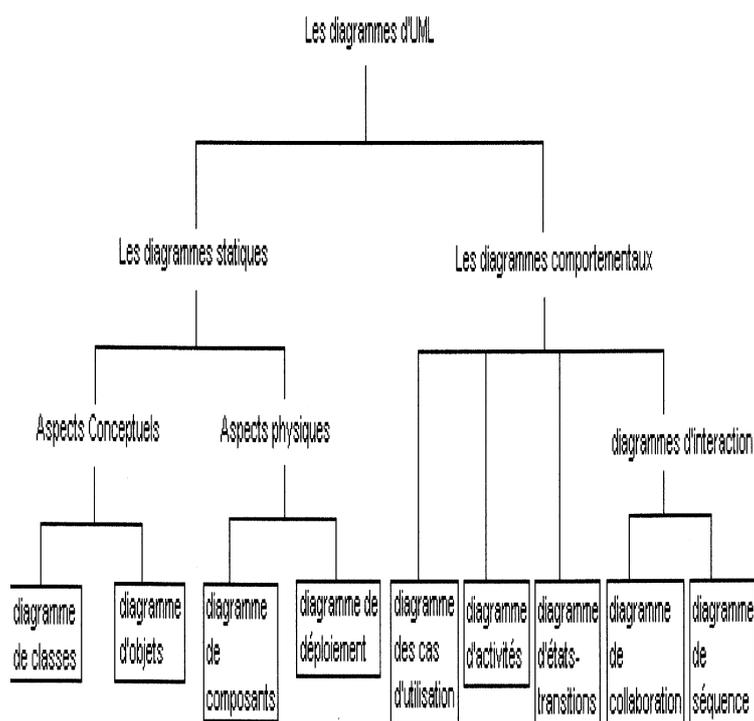


Figure 2. Les PVL spatiaux. D'après Proulx et Bédard, 2001

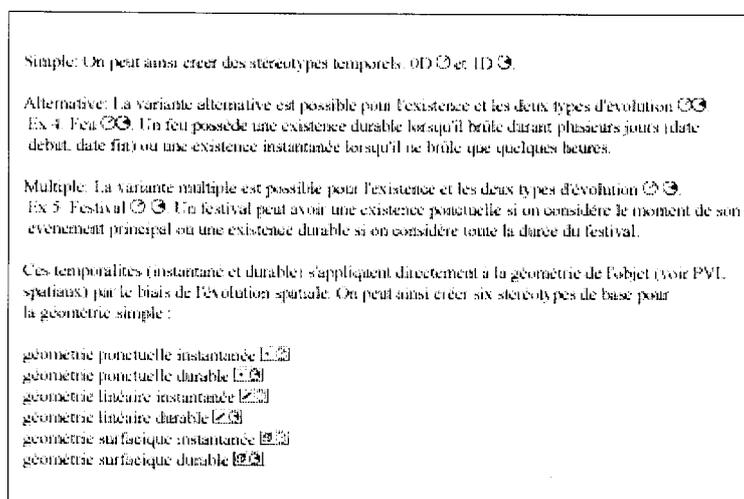


Figure 3. Les PVL temporels. D'après Proulx et Bédard, 2001

Enfin Perceptory est le module à ajouter au logiciel Visio permettant de créer des diagrammes de classes UML. Il intègre les différentes sortes de PVL pour satisfaire au besoin de la conception de base de données géographique. Il comprend également un dictionnaire précis et est capable de générer des rapports sur le contenu de celui-ci dans différents formats.

III. COMPARAISON ENTRE LE FORMALISME UML ÉTENDU ET D'AUTRES FORMALISMES

Les formalismes choisis sont parmi les plus utilisés en conception géographique, il s'agit des formalismes Entités-Associations (E/R), Modul-R, MADS et CONGOO. Le formalisme E/R a été retenu, car, bien qu'il ne soit

pas adapté à la conception géographique, il reste un standard soutenu par de nombreux logiciels.

Le formalisme Modul-R, quant à lui, a été retenu, car il fut développé au Centre de Géomatique de l'Université Laval du Québec, tout comme le géo-UML de Perceptory et parce que ce dernier est présenté par ces auteurs comme le successeur de Modul-R (Caron, 1991 et Caron et al., 1993).

La deuxième grande catégorie est celle des langages orientés objet. Ceux-ci se sont multipliés dans les années 80 et 90 avec l'émergence de la technologie objet. Le premier formalisme objet que nous avons envisagé est le formalisme MADS (Modélisation d'Applications à Données Spatio-temporelles). Il a été développé au Laboratoire de Bases de Données du Département d'Informatique de l'École Polytechnique Fédérale de Lausanne et il propose une méthode de modélisation originale très axée sur le graphisme (Parent et al., 1998a ; Parent et al., 1998b ; Parent et al., 1999 et Spaccapietra et al., 1998).

Le dernier formalisme envisagé est le formalisme CONGOO (CONception Géographique Orientée Objet) développé par Pantazis. Il s'inscrit dans la méthode de conception de SIG MECOSIG proposée dans le même ouvrage (Pantazis et Donnay, 1996).

Seuls les diagrammes de classes seront comparés puisque aucun des formalismes envisagés ne propose d'autres types de diagrammes. En outre, le géo-UML de Perceptory ne supporte également que ce type de diagramme.

Tableau 1. Comparaison entre UML et les formalismes basés sur E/R

| | Concepts spatiaux | Concepts temporels | Concepts objets | AGL |
|---------------|-------------------|--|-----------------|--------|
| UML + géo-UML | X | X NB : Seul le diagramme de classes est pris en compte. | XXX(X) | XXX(X) |
| E/R | | | X | XXXX |
| Modul-R | X | X | XX | XXX |

Les résultats de ces comparaisons sont synthétisés sous la forme des tableaux 1 et 2 en attribuant aux différents formalismes un certain nombre d'étoiles en fonction de leur prise en compte de différents concepts constituant la base de la comparaison. Plus spécifiquement, on peut épinglez les observations qui suivent.

Les formalismes E/R et UML bénéficient tous deux d'une large diffusion et sont supportés par de nombreux outils. Si le langage Entité-Association est encore le plus employé actuellement, il semblerait qu'il soit amené à

Tableau 2. Comparaison entre UML et les formalismes objet

| | Concepts spatiaux | Concepts temporels | Concepts objets | AGL |
|---------------|-------------------|--|-----------------|--------|
| UML + géo-UML | X | X NB : Seul le diagramme de classes est pris en compte. | XXX(X) | XXX(X) |
| MADS | XXX | XXXX | XXX | XXX |
| CONGOO | XXXX | XXX | XXX | |

céder sa place à UML dont les concepts conviennent mieux au développement de systèmes intégrant de nouvelles technologies telles la technologie réseau (concepts objets). Évidemment le formalisme E/R de base n'est pas adapté à la conception géographique.

Le formalisme E/R est un des premiers formalismes à avoir été proposé et la plupart des formalismes utilisés en modélisation conceptuelle de bases de données ou de systèmes d'information tel UML, même s'ils relèvent de nouvelles technologies, en restent les héritiers.

Perceptory est le successeur de Modul-R et de son AGL Orion. Bien qu'étant basé sur des formalismes différents, ils partagent bon nombre de concepts. En outre, en ce qui concerne les extensions spatiale et temporelle, elles sont tout à fait semblables. Il semble que Perceptory se soit contenté de reprendre les concepts de Modul-R et les ait intégrés au diagramme de classes d'UML sans réelle recherche complémentaire. Néanmoins Perceptory profite de la définition des concepts objets d'UML pour tout de même présenter quelques qualités supplémentaires vis-à-vis de Modul-R.

MADS est plus adapté à la conception géographique que le géo-UML. Les concepts spatiaux et temporels sont plus poussés et permettent de cerner la réalité de manière beaucoup plus fine.

C'est, à notre avis, au niveau des concepts temporels que le formalisme MADS surpasse de beaucoup le géo-UML de Perceptory ainsi que d'autres formalismes. En effet, MADS permet, comme bon nombre de formalismes, de rajouter un pictogramme aux éléments que l'on veut gérer au cours du temps, mais il propose également un ensemble de relations temporelles permettant de saisir assez précisément le comportement temporel des éléments du système.

Le formalisme CONGOO est véritablement un formalisme pour la conception géographique avec des concepts puissants (surtout au niveau des concepts spatiaux) et très efficaces dans ce domaine. Les concepts de base de la technologie objet sont, par contre, parfois définis trop succinctement dans CONGOO alors qu'ils font l'objet d'une définition plus rigoureuse dans UML (relations logiques, agrégation...).

Le géo-UML apparaît, en comparaison avec CONGOO, bien pauvre pour concevoir des SIG, du moins si la complexité de ceux-ci est quelque peu importante. En effet, l'absence de relations topologiques et la définition de certaines géométries inadaptées à la vérification d'une certaine cohérence topologique en font un langage essentiellement descriptif ne permettant pas la construction d'un modèle réellement complet. L'utilisation de ce géo-UML n'en reste pas moins beaucoup plus aisée que celle de CONGOO du fait qu'il soit supporté par un AGL. En effet, aucun AGL n'est disponible pour la modélisation en CONGOO, même si plusieurs recherches sur son intégration dans un tel outil ont été entreprises (entre autres, Dravie, 1997 et Brihaye, 1999). Pour concevoir un SIG, l'utilisation d'un langage orienté-objet est maintenant incontournable. Les langages basés sur le modèle Entité-Association sont à proscrire, car leurs concepts sont trop pauvres pour modéliser la réalité des situations actuelles. L'ensemble des concepteurs de SIG l'ont bien compris puisque les seuls langages développés actuellement sont orientés objet et les autres sont progressivement abandonnés (Modul-R). Parmi les trois formalismes objet envisagés, le choix dépend, selon nous, de la complexité des situations à décrire et du temps que l'on veut bien consacrer à la modélisation.

IV. CONCEPTION SIG AVEC UML DANS ARCGIS 8

Dans sa version 8, le logiciel ArcGIS offre la possibilité de concevoir des bases de données en utilisant le formalisme UML. Il s'agit de surcharger un modèle UML fourni par ESRI avec le modèle de sa propre base de données. Tous ces modèles sont réalisés dans le logiciel Visio (version 5 ou 2000 Entreprise).

Lorsque le modèle de la base à réaliser est construit avec toutes les spécifications imposées par ESRI (nombreux emplois de « tagged value », éléments d'extension d'UML défini comme la définition explicite d'une propriété par une paire nom-valeur (RATIONAL, 1999)), il s'agit de l'exporter vers un « repository ». Le « repository » est une base de données au format MS Access reprenant toute l'information contenue dans le modèle. On peut ensuite vérifier la cohérence de son modèle en utilisant une macro développée par ESRI et utilisable dans Visio.

Une fois le modèle exempt d'incohérences, le « repository » est utilisable au sein du logiciel ArcCatalog (un des trois logiciels composant ArcGIS) pour créer une base de données personnelle ou mettre à jour une base de données multi-utilisateurs en réseau gérée par ArcSDE et un système de gestion de bases de données d'entreprise.

Nous nous sommes concentrés sur la réalisation d'une base de données personnelle puisque nous ne disposons pas du matériel adéquat pour l'implémentation d'une base de données multi-utilisateurs.

Cette utilisation d'un langage de modélisation offre de nombreux attraits vis-à-vis de la méthode standard puisqu'il permet une vue globale du projet permettant d'envisager toutes les relations pouvant exister entre les différentes données. En outre, le modèle est réutilisable pour la mise à jour de la base de données et il constitue un document de métadonnées très utile.

Néanmoins nous reprochons au modèle UML destiné à être importé dans ArcGIS, une importante complexification des concepts du langage pour la prise en compte des données géographiques. Cette complexité pourrait être très fortement réduite par l'emploi d'un formalisme adapté à la conception géographique.

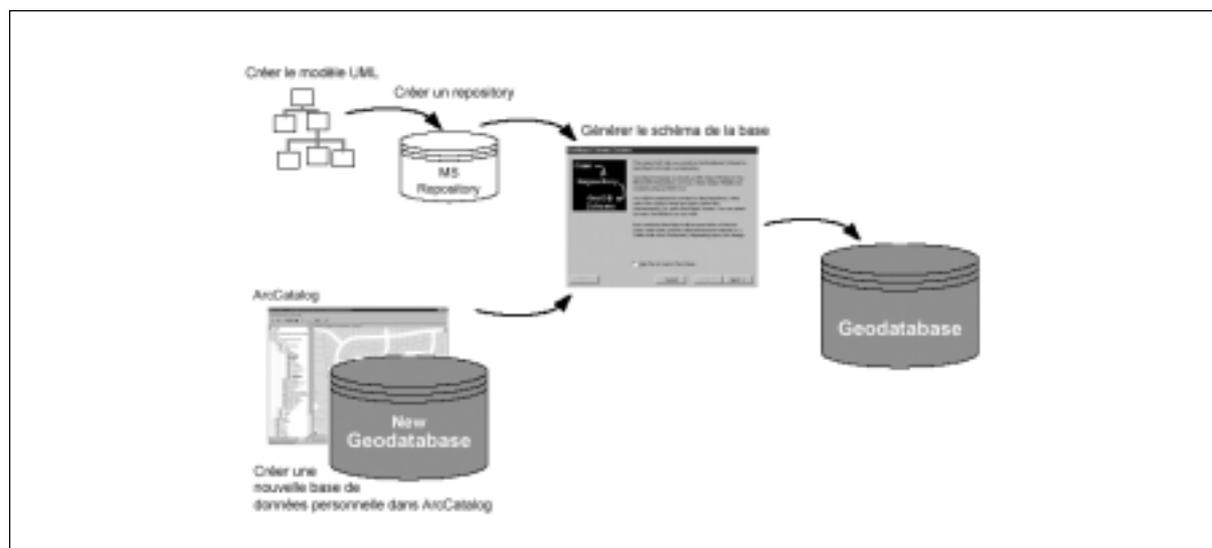


Figure 4. Création d'une base de données géographiques dans ArcGIS 8 avec UML

| | Région | Commune |
|---------|--------|--------------------------|
| Région | | -St,-Vt +Sn,p ; +Vp,r |
| Commune | | |

Matrice topologique originak

La matrice signifie que, si on envisage toutes les régions et toutes les Communes :

- une région ne peut pas être superposée totalement à une commune.
- une région ne peut pas être voisine totalement d'une commune.
- une région peut être superposée partiellement à une commune.
- une région peut ne pas être superposée à une commune.
- une région peut être voisine partiellement d'une commune.
- une région peut ne pas être voisine d'une commune.

3A

| | Région | Commune |
|---------|--------|---------|
| Région | | +Sp N |
| Commune | | |

Matrice topologique forte

La matrice signifie qu'une région doit obligatoirement être en superposition partielle avec N communes

3B

Tableau 3. Matrice topologique et matrice topologique forte

V. TENTATIVE DE PRISE EN COMPTE DE LA TOPOLOGIE DANS UN MODÈLE RÉALISÉ AVEC PERCEPTORY

Pour prendre en compte la topologie dans un modèle réalisé avec Perceptory, nous pouvions soit surcharger le diagramme de classe avec les relations topologiques, soit utiliser une autre méthode de représentation de celles-ci. Nous avons choisi, dans un souci de clarté, de ne pas surcharger le modèle.

Le concept de matrice topologique et la définition des relations topologiques de Pantazis (Pantazis et Donnay, 1996) ont été réutilisés pour leurs qualités de concision et de complétude (voir tableau 3). Il semble, en effet, que les travaux de Pantazis soient parmi les plus complets dans la prise en compte de la topologie dans un modèle conceptuel de données.

Comme l'utilisation de la matrice topologique telle que proposée par Pantazis (voir tableau 3A) ne permettait pas d'exprimer précisément les contraintes topologiques (ou relations topologiques fortes), celles-ci étant à la base présentes dans le MCD CONGOO, nous avons proposé l'utilisation d'un nouveau type de matrice topologique appelée matrice topologique forte (voir tableau 3B). Celle-ci permet de définir les contraintes topologiques entre classes d'objets géographiques en utilisant les cardinalités 0 (aucun), 1 (un), N (plusieurs), T (tous). En fait, ces cardinalités permettent d'exprimer le nombre d'objets de la classe présente en colonne qui participent à une relation topologique donnée avec un seul objet (= une seule entité) de la classe située en ligne.

Nous avons également proposé une méthode de vérification des contraintes topologiques. Comme la topologie n'est généralement pas conservée dans les bases de

données géographiques, cette vérification a pris la forme de contraintes sur le résultat de requêtes spatiales. Ces requêtes spatiales ont été effectuées avec le logiciel ArcMap, un des composants d'ArcGIS 8. Le résultat des vérifications était satisfaisant, néanmoins certaines situations ne peuvent être discernées l'une de l'autre par le seul emploi de ces opérateurs spatiaux.

Enfin l'automatisation de la saisie et de la vérification des relations topologiques a été envisagée et une petite macro permettant une saisie rapide des relations topologiques a été réalisée.

CONCLUSIONS ET PERSPECTIVES

De nombreux concepts véhiculés par le formalisme UML apparaissent comme très intéressants pour la conception de systèmes d'information géographique. Ces concepts sont, entre autres, ceux relevant de la technologie objet et ceux permettant une prise en compte des aspects dynamiques.

Néanmoins il y a une nécessité d'adapter UML pour la modélisation de données spatiales. Ces adaptations doivent permettre de prendre en compte la représentation, la topologie et, en général, l'ensemble des spécificités des données spatiales. Certaines de ces adaptations ont déjà été partiellement réalisées.

Enfin, comme perspectives, on pourrait souhaiter qu'un effort de standardisation et d'uniformisation, tel celui qui a présidé à l'élaboration d'UML, naisse au sein des foyers de développement de l'information géographique où un trop grand cloisonnement persiste encore. Un tel effort amènerait, selon nous, au développement de méthodes et de logiciels qui jouiraient d'une large diffusion, car réellement adaptés à la conception géographique.

BIBLIOGRAPHIE

- BÉDARD Y., 1999. Visual modelling of spatial databases : towards spatial PVL and UML, *Géomatica*, 53(2), pp.169-186.
- BOOCH G., RUMBAUGH J. et JACOBSON I., 2000. *Le guide de l'utilisateur UML*, collection Technologies objet/Référence, Paris, Eyrolles, 534 p.
- BRIHAYE M., 1999. *Contribution au développement d'un Outil CASE pour le soutien du formalisme CONGOO*, travail de fin d'étude, Université de Liège, Faculté des Sciences appliquées, inédit.
- BRODEUR J., BÉDARD Y. et PROULX M.-J., 2000. Modelling geospatial application databases using UML-based repositories aligned with international standards in geomatics, *ACMGIS 2000*, november 10-11, Washington DC, USA.
- CARON C., 1991. *Nouveau formalisme de modélisation conceptuelle adapté aux SIRS*, mémoire de maîtrise, Université Laval, Québec, inédit, 247 p.
- CARON C., BÉDARD Y. et GAGNON P., 1993. MODUL-R : un formalisme adapté pour les SIRS, *Revue internationale de Géomatique*, 3(3), pp. 283-306.
- DRAVIE S., 1997. *Contribution au développement d'un ensemble d'Outils CASE pour le formalisme de conception géographique Orienté Objet (CONGOO)*, travail de fin d'étude, Université de Liège, Faculté des Sciences appliquées, inédit.
- KETTANI N., MIGNET D., PARÉ P. et ROSENTHAL-SABROUX C., 1999. *De Merise à UML*, Paris, Eyrolles, 434 p.
- LOPEZ N., MIGUEIS J. et PICHON E., 1998. *Intégrer UML dans vos projets*, collection Informatiques Magazine, Paris, Eyrolles, 256 p.
- PANTAZIS D. et DONNAY J.-P., 1996. *La conception de SIG : méthode et formalisme*, collection Géomatique, Paris, Hermes, 343 p.
- PARENT C. et al., 1998a. *MADS ou l'information spatio-temporelle à portée de ses utilisateurs*, Site de l'École Polytechnique Fédérale de Lausanne - Laboratoire de Bases de Données (<http://lbdwww.epfl.ch>), consultation le 20 novembre 2000.
- PARENT C. et al., 1998b. *Modeling Spatial Data in the MADS Conceptual Model*, Site de l'École Polytechnique Fédérale de Lausanne - Laboratoire de Bases de Données (<http://lbdwww.epfl.ch>), consultation le 20 novembre 2000.
- PARENT C. et al., 1999. *Spatio-Temporal Conceptual Models : Data Structures + Space + Time*, Site de l'École Polytechnique Fédérale de Lausanne - Laboratoire de Bases de Données (<http://lbdwww.epfl.ch>), consultation le 20 novembre 2000.
- PROULX M.-J. et BÉDARD Y., 2001. *Perceptory 2000, guide à l'usager du logiciel*, (<http://sirs.scg.ulaval.ca/perceptory/>), consultation le 8 janvier 2001.
- RATIONAL SOFTWARE CORPORATION, 1999. *OMG Unified Modeling Language Specification*, (<http://www.rational.com/uml/>), consultation le 18 novembre 2000.
- SPACCAPIETRA S. et al., 1998. *Modeling Time from a Conceptual Perspective*, Site de l'École Polytechnique Fédérale de Lausanne - Laboratoire de Bases de Données (<http://lbdwww.epfl.ch>), consultation le 20 novembre 2000.

Adresse de l'auteur :

François LAPLANCHE
 Laboratoire SURFACES, Unité de géomatique
 Université de Liège
 Allée du 6 Août, 17 – Bât. B5
 B-4000 Sart Tilman, Liège