

**Modélisation des systèmes d'information à référence spatiale
et de leurs domaines d'utilisation spécialisés:
aspects méthodologiques, organisationnels et technologiques**

THESE NO 1080 (1992)

Présentée au département de génie rural

ECOLE POLYTECHNIQUE FEDERALE DE LAUSANNE

pour l'obtention du grade de docteur ès sciences techniques

par

François GOLAY
ingénieur du génie rural et géomètre diplômé de l'EPFL
de nationalité suisse

acceptée sur proposition du jury :

Prof. P. Howald, rapporteur
Prof. A. Carosio, corapporteur
Prof. J.-J. Chevallier, corapporteur
Prof. G. Leclerc, corapporteur
Prof. S. Spaccapietra, corapporteur

Lausanne, EPFL
1992

“Die neuere Zeit schätzt sich selbst zu hoch, wegen der großen Masse Stoffes, den sie umfaßt. Der Hauptvorzug des Menschen beruht aber nur darauf, inwiefern er den Stoff zu behandeln und zu beherrschen weiß.”

Goethe, “zur Geschichte der Farbenlehre”, um 1800.

“Nous sommes aujourd’hui trop imbus de nous-mêmes, du fait du grand volume de connaissances dont nous disposons. Nous oublions cependant que l’homme ne peut en tirer profit que dans la mesure où il sait les traiter et les dominer.”

Goethe, “de l’histoire de la théorie des couleurs”, vers 1800

Table des matières

Résumé

Abstract

1.....	Introduction	1
1.1.....	Contexte	1
1.2.....	Objectifs de l'étude	2
1.3.....	Contexte méthodologique	3
1.4.....	Structure du rapport	4
2.....	Bases scientifiques et étude de littérature	7
2.1.....	Systemique	7
2.2.....	Bases de données	7
2.3.....	Méthodes et outils de conception de SIRS	8
2.4.....	Conclusion du chapitre	10
3.....	Présentation de quelques domaines spécialisés	11
3.1.....	Assainissement urbain	11
3.2.....	Aménagement du territoire	15
3.3.....	Gestion de l'entretien routier	16
3.4.....	Analyse statistique géographique	17
3.5.....	Cartographie	18
3.6.....	Conclusion du chapitre	19
4.....	Aspects de la théorie de la modélisation	21
4.1.....	Les modèles et leur évolution dans le temps	21
4.2.....	Modèles et méta-modèles	25
4.3.....	Les modèles spatiaux	27
4.4.....	Mécanismes d'abstraction	36
4.5.....	Conclusion du chapitre	40
5.....	Problèmes et solutions relatifs aux traitements	41
5.1.....	Systemes de prise de décision: de l'agrégation des informations à l'intégration des décisions	41
5.2.....	Typologies des traitements	45
5.3.....	Classification de quelques fonctions essentielles des systemes d'agrégation d'informations	47
5.4.....	Modélisation des traitements	49
5.5.....	Approche du niveau organisationnel	51
5.6.....	Conclusion du chapitre	52
6.....	Problèmes et solutions relatifs aux données	53
6.1.....	Les formalismes géosémantiques	53
6.2.....	Cahier des charges des formalismes géosémantiques	53
6.3.....	Apports et lacunes des formalismes actuels	54
6.4.....	Approche du niveau conceptuel	55
6.5.....	Nouvelle composante sémantique d'un modèle de données: l'agrégat	57
6.6.....	Approche du niveau organisationnel	64
6.7.....	Conclusion du chapitre	66

7.	Problèmes et solutions relatifs aux organisations	67
7.1.	Découpage d'une entreprise en domaines de gestion	67
7.2.	Systemes d'information à référence spatiale: des sous-systemes du systeme d'information d'entreprise	68
7.3.	Le SIT: un concept fédérateur des SIRS sur un territoire donné	70
7.4.	Systemes de prise de décision à référence spatiale: .. des systemes spécialisés exploitant et produisant des informations localisées	72
7.5.	Synthèse et conclusion du chapitre: un modèle organisationnel global	73
8.	Application aux domaines d'utilisation considérés	77
8.1.	Assainissement urbain	77
8.2.	Aménagement du territoire	83
8.3.	Gestion de l'entretien routier	84
8.4.	Analyse statistique géographique	85
8.5.	Cartographie	86
8.6.	Conclusion du chapitre	87
9.	Implications de l'étude sur les méthodes de conception et les outils informatiques	89
9.1.	Implications de l'étude sur les méthodes de conception	89
9.2.	Implications de l'étude sur l'exploitation des systèmes d'information	91
9.3.	Implications de l'étude sur les outils technologiques destinés	92
10.	Conclusion	93
10.1.	Résultats de l'étude	93
10.2.	Perspectives de recherches et d'applications	95
Bibliographie		
Index		
Remerciements		
Aperçu biographique		

Résumé

Le domaine des systèmes d'information à référence spatiale est en pleine expansion de nos jours. Les problèmes toujours plus aigus rencontrés par l'homme pour gérer son environnement et la disponibilité croissante de systèmes informatiques performants en constituent sans doute les principales raisons. Les gestionnaires du territoire – ingénieurs, aménagistes, géographes, etc. – sont parmi les principaux intéressés à disposer de sources d'informations abondantes et fiables relatives au territoire. Des méthodes et des outils leur permettant d'exploiter au mieux de leurs besoins ces sources d'informations leur sont aussi nécessaires. Cette thèse a précisément pour objectif de définir un cadre méthodologique, organisationnel et technologique permettant aux gestionnaires du territoire d'intégrer de manière efficace et cohérente les SIRS à leurs propres réalisations.

Trois thèmes constituent le coeur de l'étude:

- la modélisation des traitements appliqués aux données localisées dans les domaines spécialisés d'utilisation des SIRS;
- la modélisation des données localisées dans les domaines spécialisés d'utilisation des SIRS;
- la définition d'un cadre organisationnel des systèmes traitant des informations relatives à un territoire donné.

Les solutions apportées à ces trois thèmes se réfèrent toutes à des éléments de théorie de la modélisation, qui fait l'objet d'un chapitre préliminaire.

Les principaux résultats de cette thèse sont les suivants:

- Un modèle des *systèmes de prise de décision* est proposé, pertinent pour décrire des systèmes qui ont pour objectif de produire des décisions.
- Une classification des traitements appliqués aux données localisées est proposée, qui fait appel à de nouvelles notions telles que:
 - les référentiels spatiaux, au sein desquels la mesure de la localisation des objets à référence spatiale est de nature et de qualité homogènes;
 - les modèles spatiaux;
 - la topologie, dont la perception traditionnelle dans le domaine des SIRS est remise en cause.
- Quelques enrichissements sémantiques sont apportés aux formalismes utilisés pour décrire les structures de données à référence spatiale.
- Un modèle d'organisation des systèmes gérant et traitant des informations relatives à un territoire donné est proposé. Ce modèle est basé sur une organisation fédérée des SIRS mis en oeuvre par les partenaires impliqués dans la gestion du territoire.

La pertinence du cadre théorique proposé est finalement démontrée pour quelques domaines spécialisés d'utilisation des SIRS:

- le génie urbain, et plus particulièrement l'assainissement urbain, qui a servi de fil conducteur à cette étude;
- l'aménagement du territoire;
- la gestion de l'entretien routier;
- l'analyse statistique géographique;
- la cartographie.

Abstract

The field of spatially referenced information systems (SRIS) is currently in full expansion. The principal reasons are undoubtedly the increasingly acute problems encountered by man to manage his environment and the growing availability of high-performance information systems. Territorial managers - engineers, planners, geographers, etc. - are among the main people interested in the availability of abundant, reliable information sources concerning territory. Methods and tools permitting the best use of these information sources according to their needs are also necessary. The goal of this thesis is to define a methodological, organizational and technological framework which allows territorial managers to integrate SRIS for their respective applications in an efficient and coherent manner.

Three themes make up the core of the study:

- modeling of treatments applied to localized data in the specialized fields of SRIS utilization;
- modeling of localized data in the specialized fields of SRIS utilization;
- definition of an organizational context of systems which process information relative to a given territory.

The solutions brought to these three themes all refer to elements of modeling theory, which is the subject of a preliminary chapter.

The principal results of this thesis are the following:

- A model of decision making system is proposed, relevant for the description of systems with the goal of producing decisions.
- A classification of treatments applied to localized data is proposed, which calls upon new notions such as:
 - spatial references, within which the measure of the location of spatially referenced objects is of homogeneous nature and quality;
 - spatial models;
 - topology, for which the traditional perception in the SRIS field is re-examined.
- A few semantic enrichments are brought to the formalisms used to describe the structure of spatially referenced data.
- An organizational model of systems which manage and process information relative to a given territory is proposed. This model is based on a federated organization of SRIS established by partners involved in territorial management.

The relevance of the proposed theoretical framework is finally demonstrated for a few fields specialized in SRIS utilization:

- urban engineering, particularly urban drainage, which served as the guiding thread for this study;
- land planning and improvements;
- road maintenance management;
- geographic statistical analysis;
- cartography.

1. Introduction

1.1. Contexte

Un système d'information à référence spatiale ¹(SIRS) constitue, selon la définition de la Fédération internationale des géomètres de 1981, *un instrument de décision dans les domaines juridique, administratif et économique, et une aide pour la planification et le développement; il comprend, d'une part, une base de données se rapportant au sol sur un territoire donné, et, d'autre part, les procédures et les techniques nécessaires à la mise à jour systématique, au traitement et à la diffusion des données. Les systèmes d'information à référence spatiale se distinguent des systèmes d'information classiques par leur aptitude à gérer et à exploiter des informations localisées* ².

Le domaine des SIRS est en pleine expansion de nos jours. L'origine de ce phénomène nous semble devoir être rattachée à deux raisons principales:

- Les relations entre l'homme et son environnement naturel, économique et social sont toujours plus critiques; les décisions prises en matière d'aménagement du territoire, d'urbanisme, de gestion d'infrastructures, d'élimination et de recyclage des déchets, etc. sont donc toujours plus complexes et plus interdépendantes. Dans ce contexte, la qualité et l'adéquation des informations relatives au territoire revêtent une importance croissante pour l'évaluation de l'impact des décisions envisagées et la justification des solutions retenues.
- Les solutions technologiques sont en constante progression. Les performances du matériel informatique croissent selon une courbe quasi exponentielle, la technologie des bases de données garantit une flexibilité et une sécurité essentielles aux applications de gestion, et les réseaux de télécommunication permettent le partage et l'échange de données entre différentes organisations. Une activité de recherche très intense a permis de mettre en oeuvre ces nouvelles technologies pour la gestion des informations localisées, et de développer des systèmes informatiques susceptibles de répondre à des cahiers de charges toujours plus exigeants.

Ces deux facteurs ont permis l'éclosion d'un nombre considérable de réalisations durant cette dernière décennie, destinées en particulier à la gestion des infrastructures dans les grandes agglomérations. Cette première génération de réalisations a permis la constitution de bases de données numériques, qui suscitent maintenant de nouvelles réalisations. En outre, la prolifération des réalisations et des (bonnes et mauvaises) expériences conduit à la création d'associations d'utilisateurs (AM/FM, URISA). Ces associations développent à leur tour une intense activité prospective pour promouvoir des réalisations répondant toujours mieux aux besoins des utilisateurs.

Mais d'autres facteurs peuvent encore servir de catalyseurs au déclenchement d'un tel processus. En Suisse par exemple, la réforme de la mensuration officielle (REMO) et ses mesures cantonales d'exécution (digitalisation rapide en particulier) devraient fournir l'impulsion nécessaire à la mise sur pied de SIRS sur l'ensemble du territoire.

¹ Dans le cadre de notre étude, nous considérerons les termes de *systèmes d'information à référence spatiale* (SIRS), *systèmes d'information du territoire* (SIT) et *systèmes d'information géographiques* (SIG) comme synonymes. Le lecteur restera néanmoins attentif à ce que certains auteurs actuels donnent des sens différents à ces termes.

² Dans la suite de cette étude, nous utiliserons l'expression *informations localisées* de préférence aux autres expressions traditionnellement utilisées dans un sens identique (*informations à référence spatiale*, *informations sur le territoire*, ...).

Les ingénieurs, tout comme les aménagistes, les géographes et tous les gestionnaires du territoire, ont été de tout temps utilisateurs des SIRS. Que ce soit pour la construction et l'entretien des routes, des réseaux de canalisation, l'implantation et la construction de bâtiments et d'ouvrages d'arts, l'ingénieur civil a toujours dû intégrer à ses projets la forme et la nature du territoire environnant. L'aménagiste doit concevoir et mettre en oeuvre un développement de la ville et des campagnes qui soit conforme aux objectifs politiques de l'aménagement. L'expansion actuelle des SIRS offre donc un intérêt considérable pour l'ingénieur et les autres gestionnaires du territoire. Mais en quoi peuvent-ils concrètement bénéficier de l'apport des SIRS ? S'agit-il avant tout d'un apport *technologique*, comme une grande partie de la littérature et des réalisations actuelles le laisse penser ? Ou s'agit-il plutôt d'un apport *organisationnel* ou *méthodologique*, les SIRS permettant aux gestionnaires de mettre en oeuvre les outils et les méthodes propres à leur art dans un contexte plus cohérent et plus efficace ?

Ce sont là les questions auxquelles notre étude apporte des éléments de réponse.

1.2 Objectifs de l'étude

1.2.1. Objectif général

Cette étude a pour objectif la définition d'un cadre méthodologique, organisationnel et technologique permettant aux ingénieurs d'intégrer de manière efficace et cohérente les SIRS à leurs propres applications de l'informatique.

Il s'agit plus particulièrement d'améliorer le potentiel d'utilisation des informations localisées dans des champs d'activité spécialisés, dans la perspective de processus de prise de décisions.

1.2.2. Objectifs spécifiques

Le cadre général défini ci-dessus est très large. Nous le précisons donc en énonçant un ensemble d'objectifs plus spécifiques, qui aideront à canaliser notre démarche:

- a. définir les objectifs et les caractéristiques générales des domaines spécialisés d'utilisation d'informations localisées;
- b. identifier, classifier et représenter les traitements mis en oeuvre dans ces domaines spécialisés;
- c. enrichir et adapter les concepts et les formalismes de représentation des structures de données à référence spatiale;
- d. valider un cadre organisationnel de mise en oeuvre des SIRS dans des domaines spécialisés;
- e. déduire des contraintes sur les outils méthodologiques ³ et technologiques à mettre en oeuvre pour la conception et l'exploitation des systèmes spécialisés.

³ Un outil méthodologique constitue une aide à la conception ou à la réalisation d'un système. Il sous-tend une ou plusieurs étapes d'une démarche méthodologique.

1.3. Contexte méthodologique

Dans de nombreux travaux de recherche en cours actuellement, on applique une démarche basée sur le triangle de la figure 1.1a⁴. Ce triangle définit la technologie, les applications et la recherche comme les trois piliers des SIRS.

Ce triangle nous paraît cependant peu satisfaisant, car il ne permet pas de situer les SIRS de manière univoque dans le référentiel ainsi défini. La recherche ne constitue en particulier pas à nos yeux un critère de classification en lui-même, mais est un facteur de développement et d'intégration de tous les autres piliers. En outre, le facteur *technologie* y est surévalué, dans la mesure où ni les *contraintes humaines* (organisationnelles, psychologiques, etc.) ni les *méthodes* (de mise en oeuvre et d'exploitation) ne sont explicitement mentionnées.

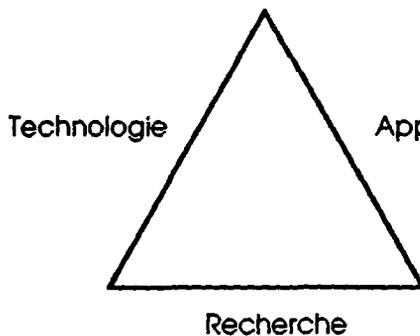


Figure 1.1a:

Triangle des SIRS selon Dangermond

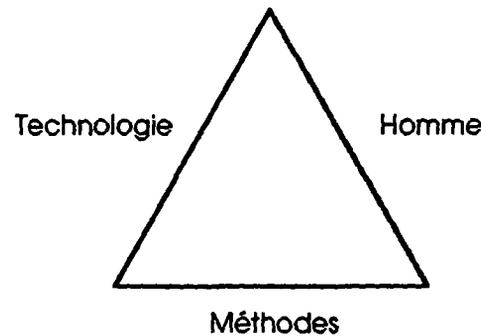


Figure 1.1b:

Triangle des SIRS à la base de cette étude

Les approches basées sur le *triangle de Dangermond* confortent ainsi l'idée à nos yeux fortement réductrice que les SIRS seraient avant tout une *technologie* adaptée à la gestion du territoire. Vus sous cet angle, les problèmes méthodologiques liés à la mise en oeuvre et à l'exploitation des SIRS deviennent alors des problèmes d'*interface utilisateurs*, et conduisent à des recherches visant à améliorer le dialogue homme-machine [Brassel 1990]!

Ces considérations nous conduisent à baser notre étude sur une approche illustrée par le triangle de la figure 1.1b. Ce triangle définit la *technologie*, l'*homme* et les *méthodes* comme les trois piliers des SIRS (et de leurs applications):

- Sans la *technologie* et ses constants développements, l'intégration en un système cohérent de l'ensemble des informations sur un territoire donné resterait une vue de l'esprit.
- L'*homme* reste cependant un acteur intervenant dans tout processus de décision sur le territoire. C'est d'abord pour l'homme et pour son mieux-être que des SIRS sont mis en oeuvre. Ce sont ensuite des hommes qui utilisent les SIRS pour étayer leurs décisions. Pour atteindre ses objectifs, un SIRS doit donc se greffer harmonieusement sur une organisation existante, et qu'il soit correctement accepté et assimilé par ses utilisateurs.
- Enfin, il faut des *méthodes* intégrant de manière cohérente l'homme et la technologie. Ces méthodes doivent permettre de spécifier comment la technologie peut être mise de manière cohérente et efficace au service de l'homme, et réciproquement comment l'homme peut appliquer la technologie existante pour résoudre les problèmes qui se posent à lui.

⁴ Ce triangle a été présenté par Jack Dangermond, président de ESRI, au cours d'un exposé non publié lors du congrès URISA 1990 à Edmonton.

Nous pouvons constater que ces trois côtés du triangle correspondent précisément aux trois volets du cadre que nous avons pour objectif d'élaborer dans cette étude: volet méthodologique, volet organisationnel (qui précise en particulier le rôle de l'homme dans les SIRS) et volet technologique.

1.4. Structure du rapport

La structure de ce rapport reflète la démarche adoptée pour réaliser notre étude. Elle constitue une trame modulaire de chapitres, reflet de nos divers objectifs spécifiques. Chaque chapitre constitue une étape clairement jalonnée par son introduction et sa conclusion.

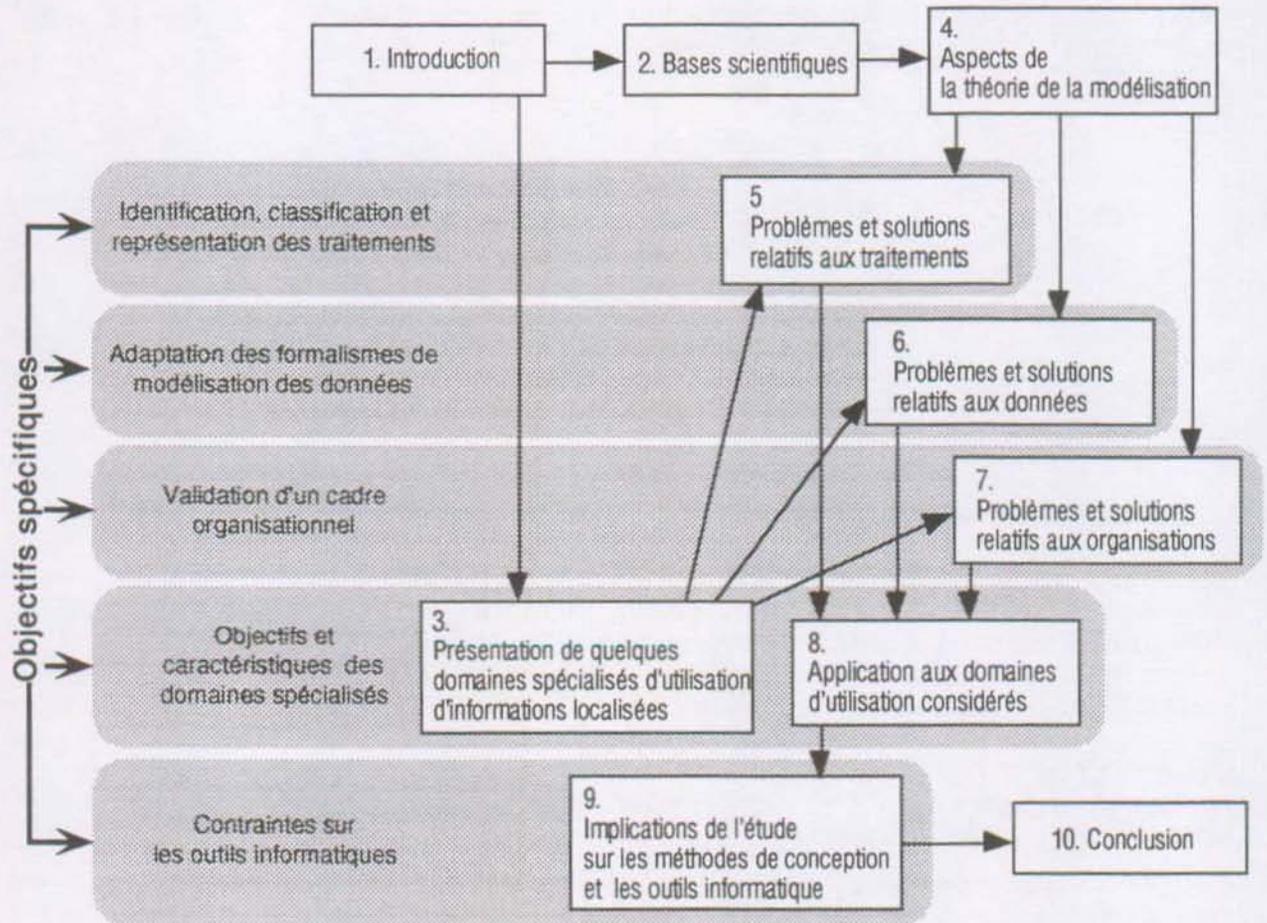


Figure 1.2: Trame de l'étude et structure du rapport

Ce rapport peut donc être parcouru par le lecteur en suivant différents enchaînements de chapitres. La figure 1.2 présente l'ordonnancement des chapitres et constitue une "aide à la navigation" pour le lecteur. On y distingue les objectifs spécifiques et leurs implications sur les différents chapitres, ainsi que les interdépendances entre les chapitres. Le lecteur intéressé à un objectif particulier peut ainsi localiser les chapitres pertinents et se constituer un canevas de lecture respectant ces interdépendances.

Les principales articulations de notre démarche peuvent être explicitées comme suit:

- Les domaines spécialisés d'utilisation d'informations localisées, décrits au chapitre 3, sont les "catalyseurs" de l'étude. L'assainissement urbain a été l'objet d'une analyse prioritaire, alors que les autres domaines ont simplement contribué à valider et étayer les apports de ce domaine prioritaire.

- **Les problèmes identifiés lors de l'analyse des domaines spécialisés ont été rassemblés en trois classes:**
 1. **l'identification et la classification de traitements informationnels nécessaires aux domaines spécialisés;**
 2. **la nature des informations localisées utilisées dans les domaines spécialisés et la modélisation de leur structure;**
 3. **les problèmes organisationnels liés aux interdépendances étroites entre les organisations gestionnaires du territoire.**

Chacune de ces trois classes de problèmes est l'objet d'un chapitre de notre étude (chapitres 5, 6 et 7 respectivement).

- **Les solutions à chacune de ces trois classes de problèmes sont toutes dépendantes d'une théorie de la modélisation. Les aspects de celles-ci nécessaires à la résolution des problèmes posés sont étudiés au chapitre 4.**
- **Les conséquences de notre étude sur les domaines spécialisés sont présentées au chapitre 8, sur un canevas analogue à celui du chapitre 3.**
- **Les conséquences de notre étude sur les méthodes de conception et les outils informatiques de conception et de mise en oeuvre de systèmes à référence spatiale, indépendantes des domaines spécialisés considérés, sont explicitées au chapitre 9.**

2. Bases scientifiques et étude de littérature

Cette étude repose sur un ensemble de domaines scientifiques avec lesquels le lecteur doit être préalablement familiarisé. Ces domaines sont passés en revue dans ce paragraphe, et des références bibliographiques proposées à l'intention de ceux qui souhaiteraient se documenter sur l'un ou l'autre de ces thèmes.

Les domaines suivants sont présentés dans ce chapitre:

1. systémique (brève introduction);
2. bases de données (brève introduction);
3. méthodes et outils de conception de SIRS (état de l'art)

Quelques domaines spécialisés utilisant des informations localisées seront présentés au chapitre 3.

2.1. Systémique

La première théorie à laquelle nous nous référerons est la *systémique*. Postulant que "le tout vaut plus que la somme de ses parties", la systémique axe son analyse sur les finalités et l'environnement du système qu'elle décrit, contrairement aux méthodes analytiques traditionnelles qui étudient la structure d'un système indépendamment de tout objectif et de tout contexte.

Le lecteur peu averti de systémique pourra se référer à une littérature abondante, en particulier [Le Moigne 1977], ouvrage de référence, et [De Rosnay 1975], ouvrage de vulgarisation.

Si la systémique est une théorie qui peut être appliquée formellement à la modélisation de la réalité, elle est aussi un mode de pensée, qui préfère une vision d'ensemble à une connaissance exhaustive et détaillée, une connaissance fonctionnelle à une connaissance organique, un exemple à une description théorique, un schéma à une description textuelle [De Rosnay 1975]. Une application de la systémique aux SIRS a été proposée en 1983 déjà dans [Chevallier 1983].

2.2. Bases de données

Une *base de données* (BD) est un ensemble de données numériques structurées de manière à en permettre la consultation et la gestion à l'aide de logiciels adéquats, les *systèmes de gestion de bases de données* (SGBD). Une base de données et le SGBD correspondant constituent ensemble une *banque de données* [Zehnder 1985]. Nous appellerons en outre *système de base de données* un ensemble de logiciels permettant de concevoir, de créer, de gérer et d'exploiter des bases de données.

Dans ce contexte, les informaticiens ont développé au cours des dernières décennies plusieurs théories et modèles, qui ont joué un rôle important dans le développement des SIRS (modèles hiérarchique, en réseau et relationnel, langages de description et de manipulation de données) [Zehnder 1985], [Date 1987]. Aujourd'hui encore, les développements propres aux systèmes de base de données ouvrent de nouvelles perspectives pour les SIRS (bases de données orientées objets, bases de données multi-média, bases de connaissances, etc.). Pour permettre la mise en oeuvre de telles bases de données, les informaticiens ont aussi développé des formalismes permettant de passer d'une description aussi générale que possible de la réalité à la spécification de modèles informatiques [Brodie, Mylopoulos et al. 1984], donnant ainsi naissance à l'*ingénierie des connaissances* [Laurini 1989]. La définition du modèle entité-relation par P. Chen à la fin des années 70, objet aujourd'hui encore de nombreuses recherches, en est un exemple

symptomatique [Chen 1976]. La complexité et la multiplicité des modèles nécessaires à la description de systèmes complexes ont conduit à la définition de nombreuses méthodes d'analyse et de conception des systèmes d'information [Olle, Hagelstein et al. 1988], dont la méthode française MERISE [Tardieu, Rochfeld et al. 1984], très utilisée sur le marché européen.

2.3. Méthodes et outils de conception de SIRS

Les technologies, concepts et méthodes issus de la recherche sur les bases de données ont donné lieu à de nombreuses adaptations aux caractéristiques de la référence spatiale. Depuis plusieurs années déjà, l'Université Laval à Québec s'intéresse aux méthodes de conception des systèmes d'information à référence spatiale et à leur modélisation [Bédard et Paquette 1988], [Paquette 1990], [Caron 1991], Cette étude intègre les résultats des nombreux travaux qui y ont été réalisés, et les complète là où cela s'est révélé nécessaire pour les domaines d'utilisation spécialisés des SIRS. Ce paragraphe décrit l'état de l'art et de la recherche en la matière.

La conception et la mise en oeuvre d'un système d'information impliquent la structuration du modèle de la réalité. Pour ce faire, nous devons disposer d'un ou de plusieurs langages composés chacun d'un vocabulaire, répertoire des concepts fondamentaux, d'un ensemble de signes pour communiquer, et d'une syntaxe définissant les règles d'assemblage des concepts fondamentaux. Conformément à [Caron 1991], nous appelons *formalismes*⁵ les langages formels auxquels nous ferons appel; les *composantes sémantiques* en constituent le vocabulaire, les *règles structurelles* la syntaxe, et la *notation* est l'ensemble des signes et des règles de représentation. A titre d'exemple, le formalisme individuel [Tardieu, Rochfeld et al. 1984] a pour composantes sémantiques principales les types d'objets (dont la notation est un rectangle) et les types de relations (dont la notation est un ovale); un type de relation associe par au moins deux liens un ou plusieurs types d'objets (règle structurelle).

Plusieurs auteurs ont déjà contribué au développement de formalismes intégrant la référence spatiale:

- a. André Frank a été parmi les premiers à étudier systématiquement les structures des données à référence spatiale, avec l'aide d'outils développés dans le cadre des sciences de l'information; il a adopté les *diagrammes de blocs d'entités*, étendus pour prendre en compte le mécanisme de la généralisation [Frank 1982].
- b. François Paquette, dans le cadre de sa maîtrise à l'Université Laval, s'est basé sur le formalisme individuel, dérivé du formalisme entité-relation [Chen 1976] dans le cadre de la méthode de conception MERISE [Tardieu, Rochfeld et al. 1984]. Il a ajouté à ce formalisme plusieurs nouvelles composantes sémantiques, dont la *généralisation* mentionnée ci-dessus, ainsi que les *sous-modèles* spatiaux [Paquette 1990]. Une technique de *substitution de sous-modèles* (SSM) lui a permis de remplacer la description détaillée des structures spatiales par ces sous-modèles, dont la notation se résume à une icône accompagnant chaque type d'objet à référence spatiale; ces sous-modèles permettent d'éviter la description formelle complète des structures de données spatiales, qui sont particulièrement complexes et surtout qui n'appartiennent pas au niveau conceptuel⁶ de la modélisation d'un SIRS. Paquette a en outre choisi de modéliser les traitements s'appliquant aux données d'un SIRS à l'aide de *diagrammes de flux de données* (DFD) [Gane et Sarson 1979], outils classiques du génie logiciel, fréquemment appliqués à la conception de systèmes.

⁵ On sera attentif au fait que dans la littérature, le terme de *modèle* est très souvent employé dans le sens de *formalisme*, créant ainsi une confusion dangereuse entre les notions de *modèle d'information* et de *formalisme*. Par exemple, selon notre définition, le *modèle entité-relation* [Chen] est un formalisme, et non un modèle !

- c. Max Egenhofer [Egenhofer 1989; Egenhofer et Franzosa 1991], Pullar [Pullar et Egenhofer 1989], Bruegger [Bruegger et Kuhn 1990] et al. mènent depuis plusieurs années à l'Université du Maine à Orono des recherches sur la nature des relations spatiales, et plus particulièrement des relations topologiques

En utilisant la théorie mathématique des sets de points, Egenhofer a pu identifier dans un référentiel rigoureux 16 types de relations spatiales binaires définies par les intersections des frontières et des intérieurs de deux objets spatiaux.

Ce cadre très rigoureux est malheureusement encore limité aux relations entre objets de même dimension que l'espace considéré (i.e.: polygones définis dans le plan), ce qui est encore trop restrictif pour la conception de SIRS qui intègrent aussi des objets à 0, 1 voire 3 dimensions.

- d. Pierrette Champoux a conduit à l'Université Laval une recherche ayant pour objectif de répertorier et de structurer les fonctions d'analyse spatiale utilisées dans les SIRS [Champoux 1991]; cette recherche a débouché sur la définition d'un nouveau formalisme, DENOT, adapté à la description des traitements sur les données. Un résultat particulièrement intéressant pour les besoins de cette étude en a été la réduction de toutes les relations topologiques (voisinage, incidence, adjacence, etc.) à la composition de deux relations de base:

- l'*inclusion* d'un objet spatial dans un autre, qui est définie par le partage des intérieurs des deux objets mis en relation; cette relation peut prendre les valeurs "nulle", "partielle" ou "totale";
- l'*adjacence* d'un objet spatial à un autre, qui est définie par le partage des frontières des deux objets mis en relation; cette relation peut aussi prendre les valeurs "nulle", "partielle" ou "totale".

Pierrette Champoux s'est placée dans une optique plus globale et mieux orientée vers la modélisation conceptuelle que Egenhofer, dont l'approche est plus mathématique, mais aussi plus restrictive. Nous nous baserons donc pour cette étude sur les propositions de Pierrette Champoux, tout en étant conscients que des recherches sont encore nécessaires sur la nature des relations et des contraintes spatiales.

- e. Claude Caron [Caron 1991] a poursuivi à l'Université Laval les travaux de Paquette, en leur donnant cependant une envergure plus vaste et plus solide, et en intégrant les réflexions de [Bédard et Paquette 1990]:

- Certaines lacunes du travail de Paquette ont été corrigées, en particulier dans le catalogue des sous-modèles proposé.
- Seule la dimension topologique (0, 1, 2 ou 3) doit être décrite dans le modèle conceptuel des données; le reste des spécifications spatiales est décrit dans le modèle conceptuel cartographique (MCC), clairement distinct du modèle conceptuel des données (MCD).
- Une approche analogue à la référence spatiale a été adoptée pour la gestion de la dimension temporelle; le modèle conceptuel des données contient les spécifications essentielles relatives à la référence temporelle des types d'objets et des types de propriétés: *existence* (continue ou discontinue) dans le passé, le présent et/ou le futur et *évolution* graduelle ou instantanée; les spécifications détaillées sont quant à elles reléguées dans un *modèle conceptuel temporel*.
- L'importance des *contraintes* pouvant s'exercer sur des relations spatiales et sur des groupes de relations spatiales a été identifiée, et a conduit à la définition de contraintes inter-relationnelles de type "et", "ou" ou "et/ou".

6 Dans de nombreuses méthodes de conception et de réalisation de systèmes d'information, on appelle *niveau conceptuel* le niveau de description le plus abstrait d'un système. Il regroupe les éléments les plus stables de l'organisation considérée.

- Les activités de l'organisation et les événements qui les déclenchent peuvent être intégrés (de manière rudimentaire) au modèle conceptuel des données, facilitant en particulier la mise en oeuvre d'une base de données orientée objets.
 - Conformément à une approche systémique de la conception du système d'information, conduisant du général au particulier, différents niveaux de détail du modèle conceptuel sont définis. Ils intègrent progressivement des spécifications toujours plus détaillées tout en se focalisant successivement sur les différents thèmes recensés dans le système d'information (vues).
 - Toutes les techniques proposées sont intégrées dans un formalisme modulaire cohérent (MODUL-R).
- f. Bédard dirige depuis plusieurs années les recherches de l'Université Laval en matière de conception de systèmes d'information. L'ensemble des recherches actuelles a pour objectif la mise sur pied d'un ensemble cohérent de formalismes, mis en oeuvre dans des outils informatiques CASE⁷ intégrés autour d'un dictionnaire commun. Les recherches citées aux points b, d et e ci-dessus s'intègrent précisément dans ce contexte de recherche, soutenu d'ailleurs par l'industrie (firme INTERGRAPH).

24. Conclusion du chapitre

Ce chapitre nous a permis de siter notre étude par rapport aux principales bases scientifiques et technologiques retenues: la systémique d'une part, et les bases de données d'autre part.

Il nous a en outre permis de faire le point sur l'état de l'art en matière de méthodes et d'outils de conception de SIRS, auxquels nous ferons fréquemment référence dans la suite de notre étude.

Les domaines d'utilisation spécialisés des SIRS retenus pour notre étude n'ont cependant pas encore été présentés. Ce sera précisément le thème de notre prochain chapitre.

⁷ Dans ce contexte: Computer Aided System Engineering.

On distingue les méthodes et outils "UPPER-CASE", liés à phase d'analyse et de conception d'un système, des méthodes et outils "LOWER-CASE", liés à la phase de réalisation d'un système. L'expression française AGL (atelier de génie logiciel) n'offre pas les mêmes nuances.

3. Présentation de quelques domaines spécialisés d'utilisation d'informations localisées

Conformément aux objectifs de notre étude, nous voulons définir un cadre conceptuel, organisationnel et technologique pertinent pour des domaines spécialisés d'utilisation d'informations localisées.

Pour éviter l'écueil d'une étude théorique faisant fi des problèmes concrets, nous nous proposons de fonder notre réflexion sur un domaine spécialisé particulier: l'*assainissement urbain*. Au sein de ce vaste domaine, nous focaliserons plus particulièrement notre attention sur la problématique du diagnostic de fonctionnement à l'aide d'un *modèle de simulation*.

L'assainissement urbain paraît représentatif de l'ensemble des problèmes liés à l'utilisation d'informations localisées dans des domaines spécialisés. Nous examinerons cependant aussi certains aspects de l'utilisation d'informations localisées spécifiques à d'autres domaines spécialisés. Il s'agit de:

- l'aménagement du territoire;
- la gestion de l'entretien routier;
- l'analyse statistique géographique;
- la cartographie.

Dans ce contexte, ce chapitre a pour objectif de dégager les principaux problèmes d'utilisation d'informations localisées propres à chacun de ces domaines. Les contributions de notre étude à la résolution de ces problèmes seront mises en exergue au chapitre 8.

3.1. Assainissement urbain

Les installations d'assainissement constituent une infrastructure destinée à l'élimination des eaux de pluie et des déchets liquides. Leclerc [Leclerc et Chevallier 1986] distingue les réseaux (égouts) des ouvrages majeurs (stations d'épuration par exemple).

3.1.1. Activités principales de l'ingénieur urbain

L'ingénieur urbain planifie et dimensionne les composantes des infrastructures urbaines, veille à leur construction, établit des stratégies d'opération et d'entretien et participe à la mise en oeuvre de ces stratégies d'intervention [Leclerc, Golay et al. 1989]. Ces activités sont précisées dans [Leclerc et Chevallier 1986] par l'inventaire suivant:

1. planification et dimensionnement de nouveaux éléments d'un réseau;
2. élaboration de plans et devis;
3. gestion opérationnelle d'un réseau;
4. entretien du réseau;
5. contrôle et suivi de projets;
6. coordination des projets;
7. entretien d'urgence.

Leclerc propose encore de rattacher ces activités à deux fonctions principales, différentes par les responsabilités auxquelles elles font appel et par les informations qu'elles exploitent:

- a) *L'étude de projets* a pour finalité l'analyse de variantes et l'élaboration de stratégies d'intervention. Cette fonction est caractérisée par des besoins en informations sur le territoire et son mobilier, sur les phénomènes mis en jeu et sur les ressources disponibles pour réaliser les stratégies recommandées.
- b) *La mise en oeuvre de projets* a pour finalité la coordination des projets, l'allocation, le suivi et le contrôle des ressources et l'évaluation de l'efficacité des stratégies d'intervention. Des renseignements variés sur l'évolution de la situation sont indispensables, et une base cartographique permettant de situer d'un simple regard ces informations sur le territoire est particulièrement utile.

On constate que la fonction d'études de projets sert de base à des décisions aux niveaux stratégiques et tactiques de la pyramide de décision [Davis et Olson 1985], alors que la fonction de mise en oeuvre de projets est caractérisée par des décisions de niveau plus opérationnel, assorties de courts délais de décision. En outre, au sein d'une organisation, ces deux fonctions se complètent dans le temps: l'étude d'un projet aboutit généralement à sa mise en oeuvre.

Leclerc, complétant [Chocat 1981], rattache enfin les informations nécessaires à l'ingénieur urbain pour accomplir ces activités à quatre classes génériques:

- a. les *informations structurelles*, localisant et décrivant des objets réels (tels que routes, bâtiments, canalisations) ou virtuels (tels que parcelles, zones d'aménagement, limites administratives) sur le territoire.
- b. les *informations phénoménologiques (de comportement)*, décrivant le comportement d'un système ou d'une infrastructure sous diverses conditions d'opération (courbe hauteur-débit d'un déversoir par exemple).
- c. les *informations thématiques*, ciblées sur les différents thèmes qui intéressent l'ingénieur dans son activité. Lorsque ces informations sont localisées, elles peuvent être présentées sous forme de cartes thématiques. Ces informations sont souvent le résultat d'agrégations et de traitements des informations originales, telles que la constitution de zones à partir d'informations localisées ponctuellement sur le territoire.
- d. les *informations de gestion de projet*, nécessaires à l'ingénieur pour suivre l'avancement du projet (disponibilité et utilisation des ressources humaines, matérielles, durée et coût de chaque tâche, etc.). Ces informations ne sont généralement pas localisées.

3.1.2. Planification et dimensionnement de réseaux d'assainissement

Si nous examinons plus en détails les études techniques liées à la planification et au dimensionnement de nouveaux éléments d'un réseau ⁸, nous constatons qu'elles impliquent essentiellement les actions suivantes:

- a. l'évaluation de débits de projet, qui dépend en particulier des aménagements planifiés et du temps de retour toléré des surcharges [Leclerc et Schaake 1973];
- b. le diagnostic relatif à l'adéquation d'une conduite ou d'un réseau de conduites aux débits de projet;
- c. la planification et le dimensionnement de nouvelles conduites, qui nécessitent l'élaboration et le diagnostic de variantes de réseaux;

⁸ Nous nous limitons volontairement à quelques aspects techniques de la planification d'un réseau, négligeant en particulier les interactions importantes avec l'aménagement du territoire.

- d. le choix et le calage des modèles de simulation permettant de répondre aux points précédents.

Chacune de ces actions implique le recours à des simulations du comportement du réseau d'assainissement considéré sous diverses conditions, selon divers scénarios. De nombreux modèles mathématiques ont été développés à cet effet, qui se distinguent en particulier par les phénomènes modélisés (ruissellement, infiltration, transport, etc.), l'unité d'évaluation des phénomènes (ensemble d'un bassin versant, sous-bassins, mailles, etc.), la nature de la modélisation (déterministe, stochastique, etc.). Une étude détaillée de ces modèles serait hors de notre sujet; le lecteur intéressé pourra se référer à un ouvrage général d'hydrologie urbaine, tel que [Walesh 1989]. Pour illustrer notre propos, nous avons retenu un modèle déjà ancien, URBAN [Leclerc et Schaake 1973], mais représentatif des besoins en informations de modèles déterministes recourant à un découpage du territoire en sous-bassins.

La démarche de simulation à l'aide du modèle URBAN comporte les étapes suivantes [Leclerc, Golay et al. 1989]:

1. définition du bassin versant considéré pour le projet;
2. décomposition du bassin en sous-bassins;
3. schématisation des sous-bassins versants en segments de ruissellement;
4. schématisation du réseau d'assainissement en segments de transport;
5. paramétrisation des segments en fonction des caractéristiques physiques des sous-bassins et du réseau d'assainissement (pente, rugosité, etc.);
6. schématisation spatiale de la pluie (modèle des polygones de Thiessen, par exemple);
7. paramétrisation de la fonction d'infiltration (paramètres de l'équation de Horton);
8. discrétisation des observations de la pluie (identification, choix et modélisation des épisodes pluvieux);
9. simulation du comportement du réseau à l'aide du modèle URBAN, selon des paramètres de calcul donnés (pas de temps, durée de la simulation, etc.);
10. visualisation et interprétation des résultats (hydrogrammes en divers points du réseau).

Ces traitements d'informations se basent sur une connaissance générale du territoire, en particulier de l'altimétrie, de l'hydrographie, de la couverture du sol et des réseaux de canalisations existants. Des informations plus spécialisées sont aussi nécessaires, telles que les bassins versants, à différents niveaux de détail (bassin principal, sous-bassins, etc.), ainsi que les précipitations et leur répartition spatiale.

Les traitements appliqués à ces informations de base nécessitent l'expertise de l'hydrologue. La schématisation et la paramétrisation des bassins versants et du réseau d'assainissement ainsi que la discrétisation des pluies constituent des agrégations d'informations faites dans une perspective clairement finalisée. Ces agrégations nécessitent une bonne expertise de l'assainissement urbain, et exigent donc l'intervention d'un ingénieur compétent.

Une simulation du comportement du réseau fournit des résultats sous la forme d'un ensemble d'hydrogrammes rattachés à différents points du réseau. Ces hydrogrammes doivent être conservés pour permettre la comparaison de plusieurs scénarios et variantes. L'accumulation des variantes peut rendre difficile la gestion et l'exploitation de ces informations; la gestion des données de projet constituent donc une difficulté particulière des études de projets.

Sur le plan organisationnel, on sera amené à distinguer les informations "institutionnelles", qui ont une valeur à long terme, des informations dont la durée de vie correspond à celle du projet. Chocat [Chocat 1981] propose de distinguer deux types de bases de données différentes pour gérer l'ensemble de ces informations:

- les bases de données permanentes (BDP), dans lesquelles sont stockées les informations institutionnelles. Les BDP contiennent en particulier les informations nécessaires à la gestion courante des infrastructures, ainsi que les informations de base nécessaires au démarrage de nouveaux projets⁹.
- les bases de données de travail (BDT), associées à un projet particulier. Ces bases sont constituées en début de projet à partir des bases de données permanentes. En cours de projet, la BDT peut fonctionner de manière autonome. A l'issue du projet, les informations pertinentes pour d'autres projets ou pour la gestion courante des installations sont intégrées à une BDP.

Chocat propose un ensemble de transformateurs pour constituer les BDT à partir des BDP, et réciproquement pour enrichir les BDP à partir des BDT. Nous constatons que ces transformations peuvent être purement *formelles*, lorsqu'elles n'impliquent qu'un changement de format. Elles peuvent être *informationnelles* lorsqu'elles exigent l'apport de nouvelles informations ou de compléments d'informations en cours de processus [Leclerc, Golay et al. 1989].

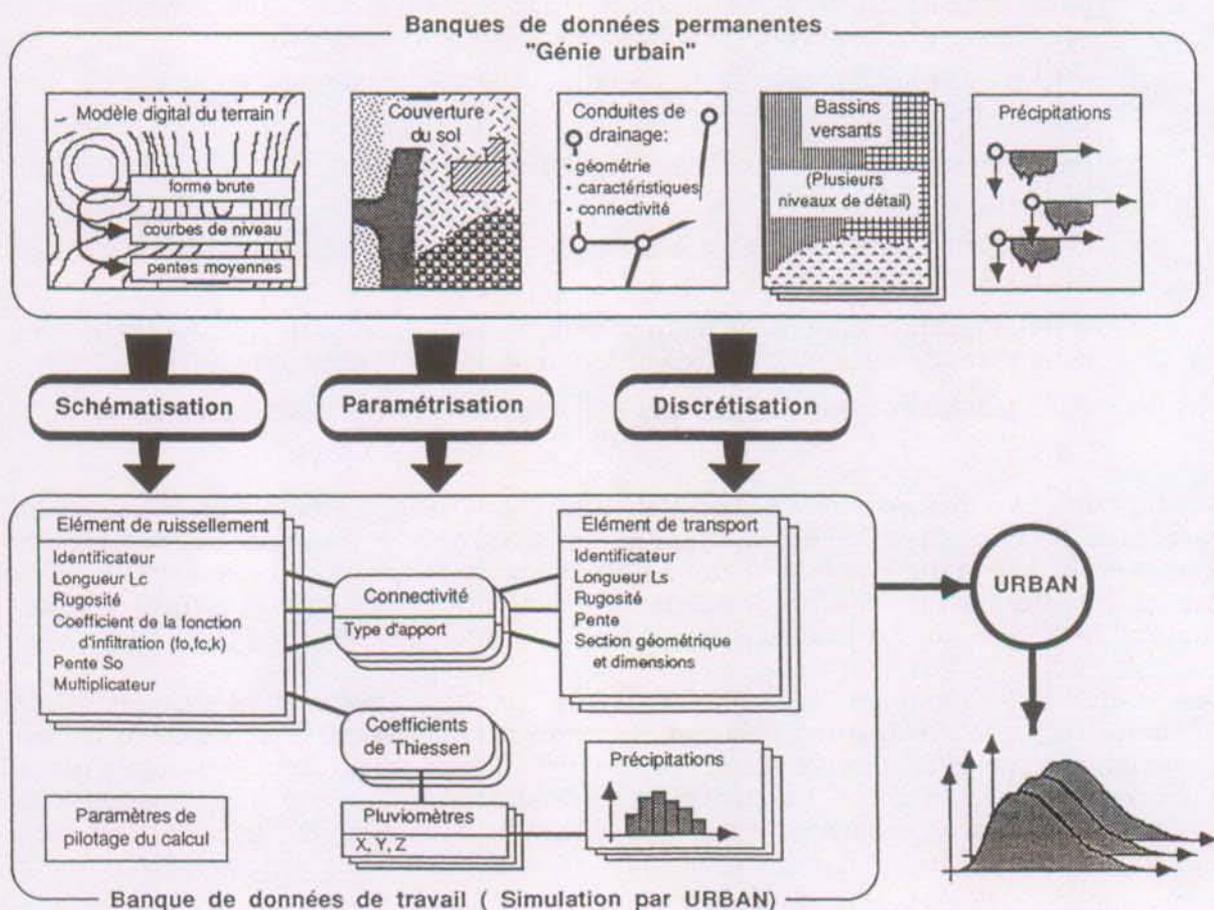


Figure 3.1: Agrégations d'informations pour le modèle URBAN

⁹ Chocat évoque en outre la périodicité de l'activité de prévision, qui implique de disposer en permanence des entrées nécessaires.

La figure 3.1, tirée de [Leclerc, Golay et al. 1989], illustre la démarche, les informations, les transformateurs et l'organisation présentées dans ce paragraphe.

Pour faciliter la tâche de l'ingénieur, les solutions suivantes peuvent être esquissées:

- Un accès efficace aux informations institutionnelles pertinentes pour son domaine doit lui être offert.
- Un ensemble modulaire de transformateurs doit être mis à sa disposition, pour lui permettre d'agréger à partir des informations disponibles des informations pertinentes pour son domaine d'activité.
- Ces transformateurs seront constitués d'outils informatiques conviviaux, permettant à l'ingénieur d'appliquer au mieux son expertise tout en confiant les travaux "de routine" à l'ordinateur (calculs, dessins, etc.).
- L'abondance des informations disponibles (scénarios et variantes diverses) en rend nécessaire une classification rigoureuse et une documentation complète et aisément accessible: la *méta-connaissance*, ensemble des informations relatives aux informations disponibles, joue donc un rôle essentiel dans ce domaine.

Dans le contexte de cette étude, Buogo [Buogo 1990] a conçu un outil informatique conforme à un tel cahier des charges et en a élaboré une maquette informatique. Les principaux résultats en seront repris dans la suite de ce document.

3.2 Aménagement du territoire

La "raréfaction" du sol et l'imbrication croissante entre les activités humaines et leur impact sur notre milieu de vie confèrent une importance croissante à l'aménagement du territoire.

L'aménagement du territoire a donc pour objectif de permettre, voire de susciter, les initiatives individuelles des propriétaires fonciers, tout en préservant à long terme l'intérêt public par un développement du territoire harmonieux et conforme aux priorités politiques.

Les activités liées à l'aménagement du territoire couvrent par conséquent chacun des trois niveaux de la pyramide de décision [Davis et Olson 1985]. Nous les illustrons par des exemples issus de l'aménagement du territoire en Suisse:

1. Au niveau stratégique correspond la planification à long terme de l'aménagement, qui se traduit par les *plans directeurs*. Simultanément instruments de synthèse de l'état existant et outils prospectifs permettant de cerner des solutions d'aménagement par la négociation et le dialogue politique, ces plans sont basés sur des informations synthétiques très générales.
2. Au niveau tactique correspond la planification normative détaillée, traduite dans les *plans d'aménagement*¹⁰. Ces plans sont basés sur des informations plus détaillées relatives au territoire et à son équipement.
3. Au niveau opérationnel correspond l'étude des *demandes d'autorisation ou de permis de construire*. Il s'agit d'une activité de routine, qui exige néanmoins l'interprétation à grande échelle des plans d'aménagement, selon une jurisprudence rigoureuse. Des informations très détaillées sur l'état et l'équipement du territoire au niveau local sont donc nécessaires.

¹⁰ Suivant le pays ou l'entité politique considérés, les plans d'aménagement ont des dénominations et des contenus variables. On peut citer "en vrac" les plans d'affectation, les plans d'affectation partiels, les plans d'occupation du sol, les plans de quartier, les plans de sites, ...

A l'instar des activités d'ingénierie urbaine (§ 3.1), les activités de l'aménagiste peuvent aussi être classifiées en deux fonctions principales:

- a. Les *études de projets*, qui ont pour objectif l'élaboration des documents d'aménagement (plans directeurs, plans d'aménagement, etc.).
- b. La *mise en oeuvre* de projets, qui implique le traitement des demandes d'autorisation, la surveillance générale de l'aménagement conformément aux différentes lois qui le régissent, la coordination entre toutes les parties et les intérêts en présence, ainsi que la tenue à jour des bases de contrôle et de décision.

Les besoins en informations et les contraintes de mise en oeuvre pour ces deux fonctions principales sont similaires à celles décrites pour l'ingénierie des réseaux d'assainissement. La complexité et l'imbrication thématique, spatiale et temporelle des dispositions d'aménagement rendent le problème de la méta-connaissance particulièrement aigu dans ce domaine.

3.3. Gestion de l'entretien routier

Pour la délimitation du domaine public, pour l'aménagement détaillé d'un carrefour, pour la construction des canalisations d'évacuation des eaux de ruissellement, une route peut être considérée comme une surface.

Pour la gestion de l'entretien routier par contre, le détail des équipements routiers n'est généralement pas déterminant. La direction de la route constitue un axe de référence privilégié, par rapport auquel les informations sont localisées. Le repérage routier est donc basé sur une *abscisse curviligne* longeant la route (axe de maintenance). Toutes les mesures qui se rapportent à l'entretien routier sont rapportées à cet axe, et éventuellement assorties d'un écart à l'axe. La figure 3.2, extraite de [Petersen, Marschal et al. 1991], illustre un système de repérage adapté à la gestion de l'entretien routier.

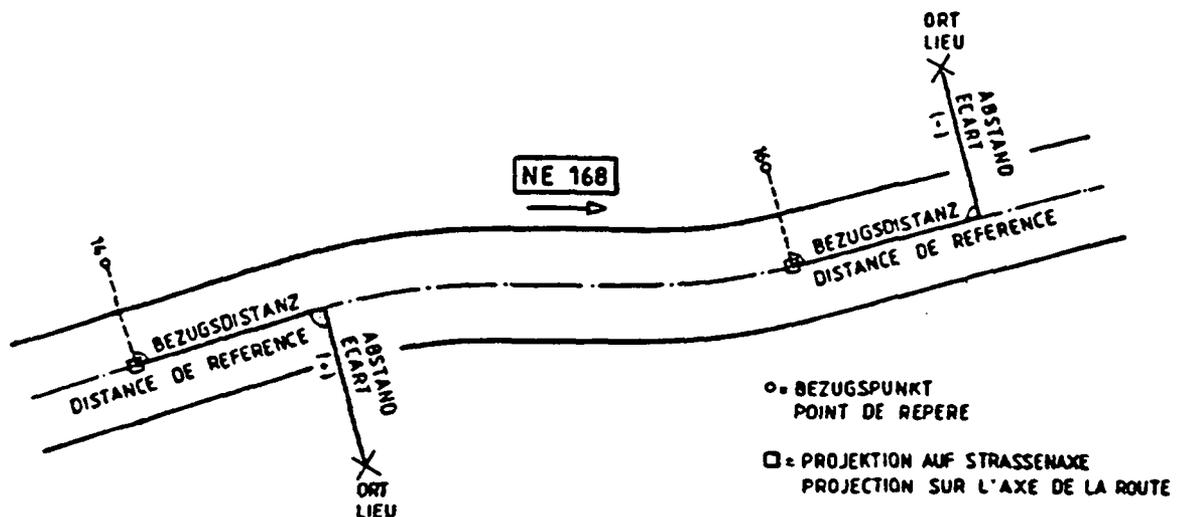


Figure 3.2: Repérage d'un lieu pour la gestion de l'entretien routier, d'après [Petersen]

En Suisse, l'Office fédéral des routes, en collaboration avec plusieurs cantons et avec des partenaires académiques et privés, réalise actuellement la banque de données routières STRADA-DB. Cette banque de données doit faciliter le partage des informations entre les cantons, chargés de l'entretien routier, et la Confédération, qui assume une fonction d'inspection et de coordination. Elle est fondée sur la définition d'un modèle de données commun, extensible par les cantons intéressés, et d'un système de repérage commun, constitué principalement d'axes balisés par des points de repère. Le prototype d'une

première unité de réalisation est en voie d'achèvement, et d'autres projets pilotes sont en cours de réalisation par les cantons.

Deux problèmes spécifiques à une banque de données routières du type de STRADA méritent d'être relevés dans le cadre de cette étude:

- a. Le système de repérage curviligne ne permet pas de visualisation cartographique directe des informations ¹¹.
- b. L'intégration entre les informations de valeur constructive (repérage géographique) et les informations de valeur stratégique (repérage curviligne) est difficile.

Pour résoudre ces deux problèmes, une intégration des deux systèmes de repérage est nécessaire. Des fonctions de transformation d'un système de repérage dans l'autre devraient apporter le niveau d'intégration souhaité.

A titre d'exemple, prenons le cas de la gestion des itinéraires pour convois exceptionnels. Les informations de base sont de nature constructives: gabarits des ouvrages, largeur de la chaussée, etc. Mais l'information n'a de valeur qu'agrégée et exploitée tronçon par tronçon. Une agrégation des informations constructives sur la longueur d'un tronçon de route est donc nécessaire pour décider des restrictions au trafic de convois exceptionnels.

3.4 Analyse statistique géographique

Le géographe doit résoudre quotidiennement des problèmes liés à la démographie, à la péréquation économique, aux déplacements de personnes, etc.

Le géographe s'appuie en particulier sur les nombreuses informations qui lui sont fournies par des recensements, périodiques ou ad hoc, de population, de trafic, etc. L'utilisation directe des données brutes de recensement reste cependant très limitative. C'est par des reclassifications, des comparaisons entre différentes séries temporelles ou différentes localisations que le géographe parvient à mettre en valeur les données brutes dont il dispose [IGUL et EPFL-IM-GM 1992].

Ces traitements informationnels font appel à des fonctions géostatistiques simples ou complexes. Pour juger de la pertinence des résultats obtenus, le géographe les visualise souvent sous forme de cartes thématiques. L'analyse de ces cartes lui permet d'ajuster les paramètres des fonctions utilisées et d'affiner les résultats par étapes successives.

Les limites des méthodes et des outils actuels ont incité les géographes à définir des subdivisions fixes du territoire (secteurs statistiques, "districts" de recensement, grilles statistiques à large maille, etc.). Il serait cependant souhaitable de disposer d'informations brutes directement localisées, qu'il devrait être possible d'agréger ensuite à volonté selon des découpages libres. Une telle flexibilité ne peut cependant être exploitée que dans la mesure où le géographe dispose d'un outil convivial, lui permettant de mesurer interactivement les effets des paramètres qu'il définit et d'en modifier aisément les valeurs.

¹¹ STRADA-DB gère bien la géométrie de la route (types de courbes et paramètres), mais indépendamment du repérage curviligne adopté. Les données nécessaires à sa représentation cartographique sont donc disponibles, mais le développement de fonctions de visualisation et d'interaction cartographiques ne sont pas prévues à l'intérieur de STRADA-DB. C'est pourquoi des projets de gestion et d'exploitation des données de STRADA-DB à l'aide de systèmes de gestion de données localisées sont à l'étude.

Pour permettre des agrégations d'informations à travers le temps, les informations relatives à une époque doivent être conservées à long terme. Il s'agit donc de mettre en oeuvre des outils susceptibles d'assurer cette tâche d'archivage, et de prévoir une documentation adéquate des informations disponibles.

C'est là le cahier des charges assigné au système d'information et d'analyse géographique (SYNAPSE) en cours d'étude dans le cadre d'une recherche commune entre l'Université de Lausanne et l'EPFL [IGUL et EPFL-IM-GM 1992].

3.5. Cartographie

Dans les produits cartographiques traditionnels, la carte a une double fonction:

- a. elle constitue la mémoire assurant la conservation des informations;
- b. elle est un outil de communication des informations mémorisées.

Ces deux fonctions répondent à des cahiers des charges différents, et parfois contradictoires. Ainsi, la fonction "mémoire" impose de conserver des informations sous une forme aussi "brute", aussi proche de l'original que possible. La fonction de communication impose au contraire une présentation bien ciblée des informations [Bertin et Bonin 1977]. Des sélections d'informations doivent être réalisées et des représentations adéquates élaborées. Des déformations géométriques peuvent faciliter la compréhension du message véhiculé (généralisation cartographique).

Dans un système d'information à référence spatiale, ces deux fonctions peuvent être clairement distinguées [Chevallier 1983]. La base de données assume la fonction de mémoire relative à un territoire donné. Sur la base des informations fournies par cette mémoire, il est possible d'élaborer toutes sortes de produits, parmi lesquels les plans et les cartes destinés à la communication d'informations.

Dans cette perspective, l'élaboration d'une carte n'est plus une tâche étroitement liée à la gestion du système d'information, mais constitue une exploitation spécialisée, clairement finalisée, du système d'information. Il est donc raisonnable de considérer la cartographie comme un domaine spécialisé d'utilisation d'informations localisées, dont les objectifs, les activités, les traitements informationnels, le contexte organisationnel sont différents.

Une carte constitue donc un produit à élaborer, et non une mémoire à gérer. Il existe cependant un certain nombre d'informations spécialisées que le cartographe souhaite conserver à plus long terme. Ainsi en va-t-il des textes figurant sur une carte: leur placement est une tâche fastidieuse, difficilement automatisable. Dans le cas de cartes périodiquement mises à jour et rééditées, les mutations relatives aux textes et à leur emplacement restent en général mineures d'une "version" à une autre de la carte. Ces informations spécialisées doivent donc être maintenues à disposition du cartographe.

Dans le contexte de notre étude, l'utilisation d'un système de gestion de données localisées (SGDL) ¹² pour la production cartographique a fait l'objet d'un travail de recherche appliquée particulier [Chevallier, Golay et al. 1989].

¹² Un système de gestion de données localisées (SGDL) est un système informatique dédié à la gestion et à l'exploitation d'informations localisées. Il constitue donc un "SGBD à référence spatiale".

3.6. Conclusion du chapitre

Chacun des domaines présentés a permis de mettre en exergue quelques problèmes particuliers liés à l'utilisation d'informations localisées. Nous pouvons cependant élaborer un inventaire des problèmes communs aux différents domaines. Il s'agit en particulier:

- de l'existence, dans la plupart des domaines spécialisés considérés, de fonctions distinctes de gestion et de production d'informations;
- de l'intégration et de l'interaction étroite de différents niveaux de décision et de différentes échelles de représentation des informations;
- de différentes formes de structuration et de représentation géométrique des informations, correspondant à des objectifs particuliers;
- de l'exploitation de modèles et d'outils mathématiques décrivant le comportement du système;
- de l'importance d'une bonne connaissance des informations disponibles et de leurs caractéristiques (méta-connaissance);
- du besoin de convivialité des outils utilisés par les experts des domaines spécialisés.

La gestion globale et cohérente du territoire implique en outre de nombreuses interactions entre ces domaines spécialisés d'utilisation d'informations localisées. L'aménagement du territoire a une mission importante de coordination et de conciliation entre les différents acteurs de la gestion du territoire.

Les prochains chapitres (4 à 7) mettent en place un cadre théorique permettant de décrire et de mieux comprendre ces problèmes spécifiques aux domaines spécialisés d'utilisation d'informations localisées. Ils suggèrent aussi un cadre conceptuel et organisationnel dans lequel des solutions spécifiques à chaque domaine peuvent être définies.

On trouvera au chapitre 8 l'application du cadre proposé à chacun des domaines décrits dans le présent chapitre.

4. Aspects de la théorie de la modélisation

La réalité peut être considérée comme un continuum informationnel évolutif, que seule une infinité d'informations pourrait décrire. Mais l'homme est capable de distinguer dans la réalité les éléments importants pour chacune de ses activités, et de faire abstraction de tous les détails inutiles. On appelle *modélisation* le processus de transformation des signaux perçus en un modèle pertinent pour une activité donnée [Bédard et Chevallier 1989]. Un *modèle* est le résultat d'un processus de modélisation. Il est une représentation simplifiée, sous forme cognitive ou physique, de certains aspects d'une réalité complexe. Il constitue un système homomorphe de la réalité [Le Moigne 1977].

Quel que soit le domaine spécialisé que nous considérons, l'utilisation pertinente d'informations localisées s'appuie sur une mémoire, susceptible de donner aux intéressés une image de la réalité qui leur évite une observation directe de cette réalité. Une telle observation peut en effet être longue, coûteuse, ou exiger des compétences dont l'intéressé ne dispose pas. Elle peut même être impossible lorsqu'il s'agit d'extrapoler des situations futures, d'évaluer des phénomènes non mesurables dans la réalité tels qu'une crue dix-millénaire, ou d'obtenir des vues de synthèse des informations disponibles. La connaissance emmagasinée dans la mémoire constitue précisément un *modèle* de la réalité.

Ce chapitre rappelle quelques éléments essentiels de la théorie de la modélisation et la complète par de nouveaux aspects nécessaires à la résolution des problèmes qui seront abordés au cours des chapitres suivants.

4.1. Les modèles et leur évolution dans le temps

La réalité évolue dans le temps. Cette évolution peut être lente ou rapide, continue ou discrète. La surrection des montagnes, par exemple, constitue un phénomène ¹³ lent et continu, alors que les mutations foncières constituent un phénomène discret, plus ou moins rapide. Cette évolution est décrite par la *trajectoire* [Le Moigne 1977] du système "réalité" dans le temps (figure 4.1). C'est cette trajectoire qu'il s'agit de connaître. De plus, l'extrapolation de cette trajectoire dans le futur n'est possible que si nous parvenons à la paramétrer, à établir le *programme* du système ¹⁴.

La modélisation de l'ensemble des paramètres pertinents d'une trajectoire est complexe. Elle fait généralement appel à des modèles de différentes natures, dont nous proposons ci-dessous deux classifications.

La première de ces classifications (A) distingue les modèles diachroniques des modèles synchroniques:

1. Les **modèles diachroniques** décrivent explicitement l'évolution de la réalité, que ce soit par des mesures de type "gradient" (débit de vidange d'un réservoir, taux de natalité, etc.) ou par la connaissance des règles d'évolution (équation de Horton décrivant l'infiltration).

¹³ Par phénomène, on entend "tout ce qui se manifeste à la conscience", "tout ce qui est objet d'expérience possible, qui apparaît dans l'espace et dans le temps" [exégèse de Kant dans le Petit Robert].

¹⁴ Cette extrapolation repose généralement sur des hypothèses partielles. Elle n'est pas vérifiable par observation de la réalité, et peut donc donner lieu à diverses variantes de la trajectoire, à divers scénarios d'évolution.

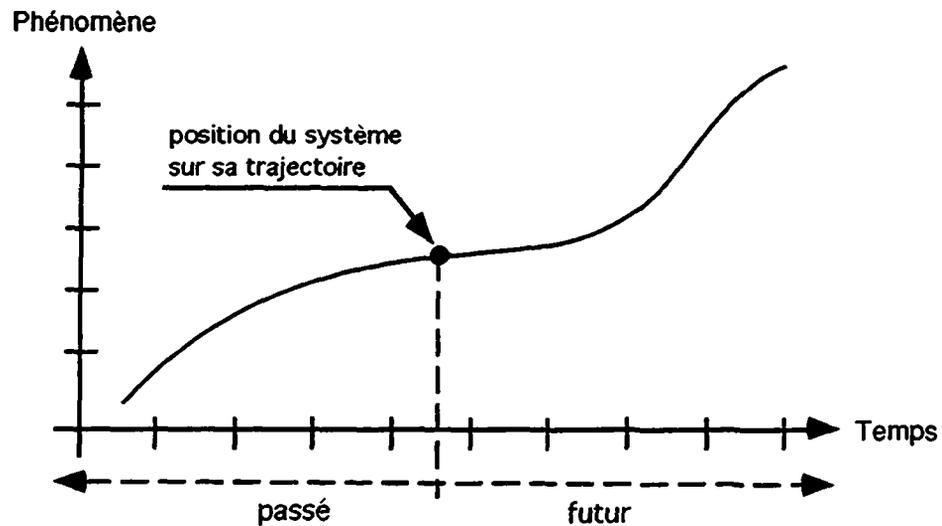


Figure 4.1: Trajectoire d'un système

2. Les **modèles synchroniques** décrivent exclusivement des situations instantanées de la réalité, appelées *états* [Richer et Chevallier 1992] dans la suite de cette étude. Chaque état peut être établi sur la base d'observations directes (propriétaire d'une parcelle par exemple) de la réalité, ou par connaissance des règles de comportement instantané du modèle (courbe hauteur d'eau/débit en un point d'un cours d'eau par exemple). Un *état* constitue en quelque sorte un "arrêt sur image" de la réalité à un instant donné.

Alors que la mise en oeuvre d'un modèle diachronique est tributaire de la connaissance d'un état initial du système, la mise en oeuvre d'un modèle synchronique exige que des procédures de mise à jour soient définies pour assurer un renouvellement des états conforme aux besoins des utilisateurs.

Les exemples décrivant les modèles synchroniques et diachroniques mettent en évidence une seconde classification des modèles (B), qui distingue les modèles observés des modèles de comportement.

1. Les **modèles observés** sont composés de valeurs résultant de la mesure de phénomènes de la réalité. Les processeurs d'espace [Le Moigne 1977] (mises à jour, restitution des informations) et de temps (mémoire) y jouent un rôle prépondérant, à l'encontre des processeurs de forme (transformation des informations).
2. Les **modèles de comportement** sont composés de règles permettant d'inférer les valeurs de certains phénomènes. Les processeurs de forme y jouent un rôle prépondérant.

Les deux classifications proposées sont illustrée par le tableau 4.1 et assorties d'exemples caractéristiques. Elles définissent un cadre cohérent auquel nous pouvons rattacher les notions apparemment contradictoires de *modèles* telles qu'elles sont utilisées en ingénierie et dans le domaine des systèmes d'information:

- Les *cadastres* (juridiques, fiscaux, de conduites, de citernes, ...) sont des systèmes basés sur des modèles observés synchroniques. Leur évolution ne peut pas être paramétrée: des procédures de mise à jour sont donc appliquées en cas de nécessité pour faire coïncider un état du modèle à un état de la réalité.
- Toute mesure directe d'un paramètre dépendant du temps constitue un modèle observé diachronique. Il en va ainsi de mesures de *débit* ou de *vitesse*, dont on peut déduire le niveau d'une rivière ou la position d'un véhicule au cours du temps. En Suisse, le *registre d'état-civil* en est un autre exemple: on y enregistre les changements d'état-civil intervenant au cours du temps.

- Un *modèle de simulation* ¹⁵ du ruissellement tel que URBAN [Leclerc et Schaake 1973] constitue un modèle de comportement. Il est de nature synchronique, bien que le temps en soit un paramètre important: il permet par exemple d'obtenir un débit pour une structure donnée du réseau d'assainissement.
- Un modèle de simulation météorologique permettant de prévoir l'évolution du temps à moyen terme constitue un modèle diachronique, à l'instar de modèles de simulation d'évolution du paysage [Richer et Chevallier 1992].

B ↓	A →	Modèles synchroniques	Modèles diachroniques
Modèles observés		Cadastrés	Mesures de débits, de vitesses Registre d'état-civil
Modèles de comportement		Evaluation de débits de crue	Météorologie Evolution du paysage

Tableau 4.1: Exemples de classification de modèles

Les modèles généralement utilisés dans la plupart des systèmes d'information actuels sont des modèles *observés*. S'ils se révèlent suffisants pour décrire des phénomènes dont l'évolution est lente et discrète (propriété foncière par exemple), ils se révèlent généralement trop réducteurs pour décrire des phénomènes dont l'évolution est continue, rapide et complexe (prévisions météorologiques par exemple). Dans les cas les plus complexes, le modèle requis sera une composition hybride des quatre types de modèles définis ci-dessus: des observations seront complétées par des règles de comportement synchroniques, et des états connus baliseront une trajectoire dont des états complémentaires pourront être interpolés ou extrapolés.

L'évolution des connaissances méthodologiques et de la technologie informatique a considérablement marqué le mode de représentation physique des modèles: les modèles *numériques* ont ainsi succédé aux modèles *analogiques*. Les caractéristiques de chacun de ces deux types de modèles ont des conséquences importantes sur leur mise en oeuvre:

1. les modèles *analogiques*

Les phénomènes que l'on désire modéliser sont représentés par un ensemble indissociable de grandeurs physiques analogiques (tensions électriques, niveaux d'eau, ...). Des valeurs initiales sont assignées à certains paramètres sur la base d'observations. Le comportement du modèle constitue alors un reflet global du comportement réel.

Les *modèles analogiques* de réseaux de liquide sous pression en sont un exemple d'utilisation synchronique. Les *maquettes* utilisées en hydraulique, en particulier pour la modélisation d'écoulements non permanents bi- et tridimensionnels, en constituent un autre exemple d'utilisation (synchronique ou diachronique). Les utilisations diachroniques sont caractérisées par le déclenchement d'un processus de simulation, dans lequel le temps est lui-même un paramètre inséparable des autres, généralement représenté par le temps de la simulation. Un état est alors constitué par les valeurs représentant chaque phénomène à un instant donné de la simulation; aucun état n'est en principe mémorisé.

¹⁵ Nous décrivons plus complètement au chapitre 5 la notion de simulation.

2. les modèles numériques

Chaque phénomène que l'on désire modéliser est représenté par ses propres règles de comportement et valeurs observées sous forme numérique. L'intégration numérique est donc par nature modulaire: chaque module regroupe les règles et valeurs propres à un phénomène. Elle implique en particulier l'extensibilité des modèles, à l'exemple des modèles météorologiques que l'on perfectionne au fur et à mesure de l'avancement des connaissances. En outre, les modèles numériques peuvent intégrer des phénomènes hétérogènes (physiques, statistiques, etc.).

Dans le cas de modèles numériques diachroniques, l'évolution de la réalité est représentée par une succession de transformations discrètes d'un état dans un autre. La figure 4.2 présente l'évolution d'un seul phénomène (une seule échelle de mesure); ce paradigme peut cependant aisément être étendu à un modèle intégrant plusieurs phénomènes. Chaque transformation est instantanée, et le temps est modélisé par la durée, variable, de chaque état. La mémoire du système constitue donc un processeur temporel [Le Moigne 1977].

Comme on le voit sur la figure 4.2, ce découplage entre le temps et les autres phénomènes est conforme à la nature de phénomènes dont l'évolution est discrète (propriété parcellaire, construction d'un bâtiment, accident survenu sur une route, etc.). Elle n'est par contre pas conforme à la nature de phénomènes continus, tels que la croissance des forêts, l'évolution d'un niveau d'eau dans le temps, l'usure d'une route. L'évolution de phénomènes continus devra donc être approximée par des transformations successives, dont l'enchaînement, ainsi que le mode et les paramètres du déclenchement (pas de temps fixe, tolérances sur les écarts entre les valeurs du modèle et la réalité, etc.) devront être définis lors de la mise en oeuvre du système.

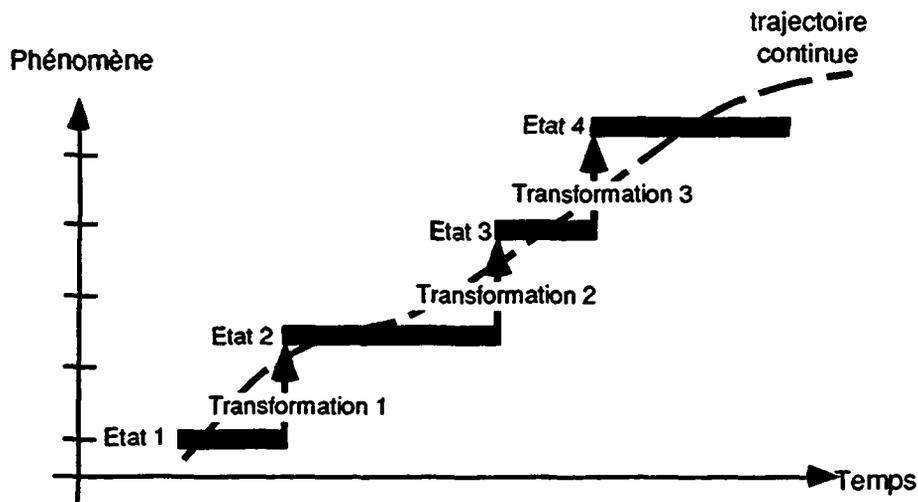


Figure 4.2: Modélisation numérique de l'évolution d'un phénomène continu

Cette dernière décennie, les modèles numériques ont presque totalement supplanté les modèles analogiques. Outre leurs qualités conceptuelles déjà mentionnées, l'adéquation des modèles numériques à la représentation et à la résolution par ordinateur des problèmes constitue sans doute une raison essentielle de cette évolution. Les ordinateurs actuels sont en effet constitués d'un processeur ou d'un réseau de processeurs traitant de manière quasi-instantanée les données stockées dans leur mémoire. Cette architecture et ce fonctionnement sont donc similaires à ceux des modèles numériques décrits ci-dessus. Dans le cadre de notre étude, nous ne prendrons en compte que les modèles numériques.

4.2 Modèles et méta-modèles

A l'exemple de certaines applications de l'intelligence artificielle, on peut imaginer des systèmes intégrant une connaissance relativement figée de la réalité; ainsi en va-t-il d'un système expert pour le diagnostic de pannes dans une machine, où les organes et le fonctionnement de la machine peuvent être modélisés une fois pour toutes. Mais dans la plupart des systèmes traitant de l'information, et en particulier dans les SIRS, la connaissance évolue notablement dans le temps, et l'un des objectifs courants de ces systèmes est précisément de suivre et de gérer une connaissance trop abondante, complexe et changeante pour qu'elle puisse être appréhendée dans sa totalité par l'esprit humain. Ça n'est donc qu'à partir d'un niveau d'abstraction relativement élevé que le modèle de la réalité peut être considéré comme stable.

Ainsi, dans les systèmes de base de données traditionnels, on distingue dans un modèle au moins deux niveaux d'abstraction, afin que ses utilisateurs puissent en comprendre et en maîtriser l'évolution [Tsichritzis et Nierstrasz 1988]:

- un niveau d'abstraction supérieur, pratiquement invariant dans le temps et connu en permanence des utilisateurs, qui constitue un squelette, une *structure* du modèle; la structure constitue en quelque sorte un "modèle du modèle" ou *méta-modèle* de la réalité ¹⁶;
- un niveau d'abstraction inférieur, qui évolue dans le temps à l'image de la réalité, à l'exemple du catalogue d'une bibliothèque, et qui constitue une réalisation ou *instance* du modèle.

Le raisonnement humain est un louvoiement continu entre ces deux niveaux d'abstraction: il part d'une connaissance acquise pour explorer des faits inconnus, qu'il va intégrer à plus ou moins long terme à sa connaissance, et ainsi de suite; on se référera à ce sujet à la théorie des frames [Minsky], citée dans [Barr et Feigenbaum 1981]. La frontière entre structure et instance est donc floue: elle dépend en particulier du niveau d'abstraction pris en considération pour un problème donné. Plus le niveau d'abstraction ¹⁷ est élevé, plus le modèle est stable; ce que l'on appelle *structure* pour un problème donné peut donc parfaitement être considéré comme une simple information à un niveau plus général [Lévine et Pomerol 1989] ! Prenons un exemple dans le domaine de l'aménagement du territoire vaudois: le parlement cantonal détermine (par voie législative) un catalogue de zones d'aménagement (zone à bâtir, zone agricole, etc.), sur lequel les autorités communales se basent à leur tour pour définir le plan des zones communal; dans le domaine du cadastre des conduites, les éléments constructifs inventoriés et localisés avec précision sont soumis à de fréquentes modifications, tout en étant rattachés à des conduites définies dans un cadre moins précis (planification, fonctionnement), mais plus stable. Les mécanismes d'abstraction seront abordés plus en détail au § 4.4.

La conception d'un système traitant des informations passe donc nécessairement par la définition d'un *méta-modèle*, décrivant en particulier la structure des données considérées (méta-modèle des données) ¹⁸, ainsi que les traitements qui leur sont applicables (méta-

¹⁶ Généralement, on appelle simplement *modèle* le méta-modèle ... c'est donc un nouvelle source de confusion autour du concept de *modèle* !

¹⁷ On sera attentif au fait que l'abstraction est un mécanisme applicable sur toute échelle de mesure allant du général au particulier. Ici, nous parlons de l'abstraction liée à la stabilité d'un système d'information (du concept à la réalité). Cette étude abordera aussi l'abstraction liée à la portée des décisions au sein d'une entreprise (du niveau stratégique au niveau opérationnel), ainsi que l'abstraction liée aux phénomènes spatiaux (des petites aux grandes échelles).

¹⁸ En d'autres termes, nous devons définir notre base de données par *intention* et non par *extension*, selon la terminologie suggérée par [Mannino 1990].

modèle des traitements) ¹⁹. Ce méta-modèle permet le dialogue entre concepteurs et utilisateurs du système, et sert de base à la réalisation technique du système. Le mécanisme d'élaboration d'un méta-modèle est illustré par la figure 4.3, qui présente deux étapes successives de modélisation:

1. une étape d'*observation* de la réalité, qui conduit à l'élaboration d'un *modèle-type*, (de nature cognitive), contenant les entités et les traitements de la réalité importants pour l'activité considérée;
2. une étape de *structuration* du modèle-type, qui consiste à en extraire les éléments stables pour constituer le méta-modèle de la réalité.

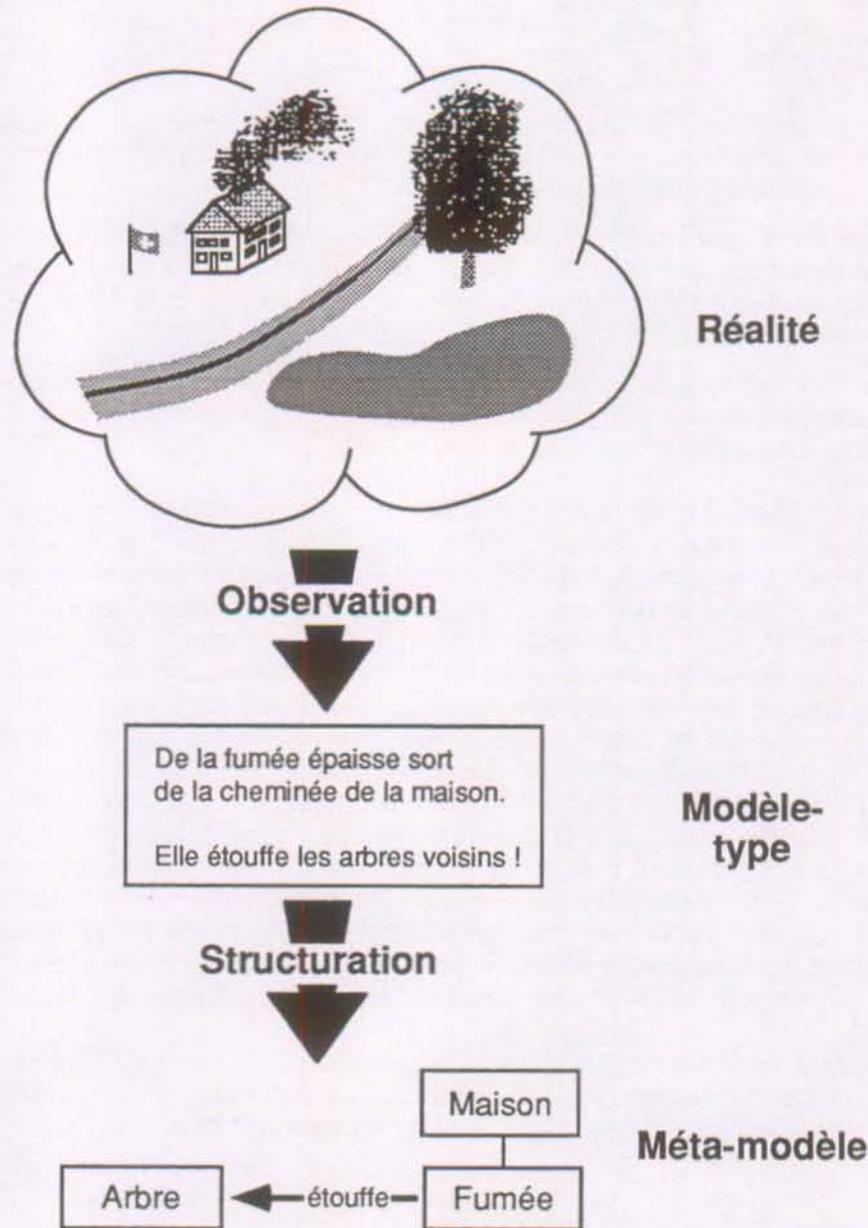


Figure 4.3: Les deux étapes de modélisation de la réalité au méta-modèle

¹⁹ Dans les approches les plus récentes, on tend à éviter la dissociation artificielle entre les données et les traitements en associant à chaque classe d'objets de la réalité aussi bien une structure de données qu'un ensemble de traitements qui lui sont applicables [Caron 1991], [Ward 1990]. Ces approches peuvent être mises en relation avec les recherches liées à la *programmation orientée objets* [Pascoe 1986] et aux *bases de données orientées objets* [Tsichritzis et Nierstrasz 1988].

Lors de l'utilisation opérationnelle du système d'information, l'observateur se basera sur le méta-modèle existant pour élaborer son propre modèle cognitif (figure 4.4); le méta-modèle fixe les échelles de mesure et les règles d'abstraction nécessaires à l'estimation des valeurs, à la sélection et à la simplification des informations, conformément au mécanisme de modélisation décrit par [Bédard et Chevallier 1989].

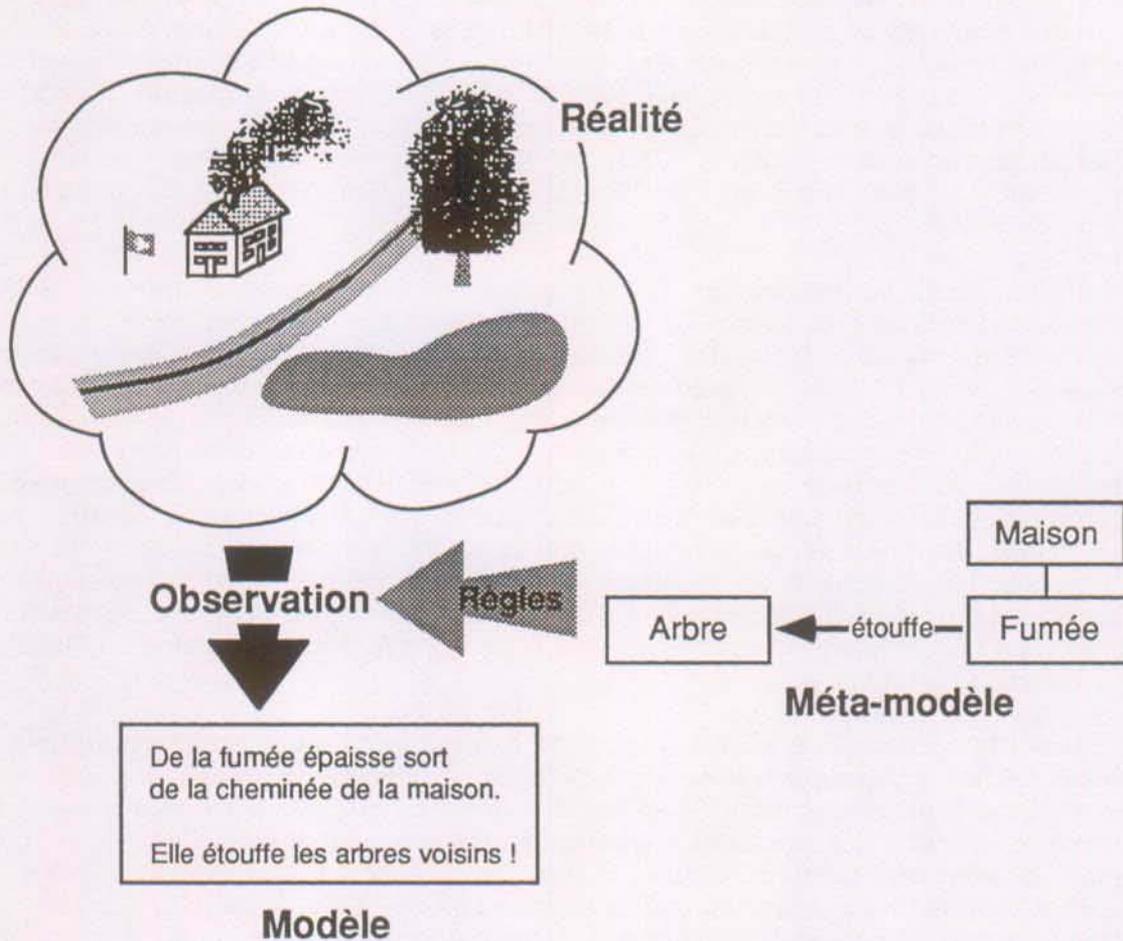


Figure 4.4: Modélisation conforme aux règles du méta-modèle

4.3 Les modèles spatiaux

Un *modèle* a été défini au § 4.1 comme une description des phénomènes de la réalité pertinents pour une activité donnée. Dans un SIRS comme dans tout système traitant des données localisées, le territoire revêt par essence une grande importance, et les autres phénomènes doivent pouvoir lui être rattachés. Nous appellerons *modèle spatial* d'un phénomène de la réalité les éléments du modèle permettant de rattacher le phénomène au territoire.

Imbriqué dans le processus de modélisation décrit par [Bédard et Chevallier 1989], il existe donc un processus spécifique de *modélisation spatiale*, consistant en:

- l'*estimation* de la localisation, de l'étendue et de la forme des phénomènes considérés (niveaux de détail successifs selon [Chevallier et Bédard 1990]);
- la *simplification* du modèle spatial, qui inclut l'*agrégation spatiale* et la *généralisation spatiale* conformément au niveau d'abstraction considéré (cf. § 4.4, mécanismes d'abstraction).

Ce paragraphe présente la nature et les caractéristiques des modèles spatiaux.

4.3.1. Le territoire et la mesure des phénomènes spatiaux: une multiplicité de référentiels

Le territoire nous apparaît intuitivement une réalité très concrète, sur laquelle on vit et on se déplace. Mais cette réalité est trompeuse, car le territoire ne constitue somme toute qu'un concept abstrait, sur lequel s'inscrivent les phénomènes pertinents pour nos activités. Ainsi, nous pouvons nous intéresser à la distance séparant une route d'une maison (accessibilité, bruit, ...), à la superficie d'un bâtiment, à l'adjacence de deux parcelles, etc. Le territoire constitue donc un *référentiel* dans lequel la position, l'orientation, la forme et l'étendue des phénomènes de la réalité peuvent être décrites et au travers duquel les caractéristiques spatiales des objets et leurs relations spatiales peuvent être établies. Cette approche est à la base des nombreux "cadastres polyvalents" qui ont été réalisés ces dernières décennies.

Une telle approche part cependant de l'hypothèse qu'il existerait un modèle unique, *isomorphe*, du territoire réel. Mais nous avons montré au début de ce chapitre qu'un modèle ne peut être qu'une image partielle, *homomorphe*, de la réalité. Chevallier et Bédard [Chevallier et Bédard 1990] ont constaté que la nature des échelles de mesure (nominale, ordinale ou cardinale) des propriétés spatiales dépendent en particulier du niveau de prise de décision considéré (stratégique, tactique ou technique). On ne peut donc pas se contenter de définir sur le territoire un seul et unique référentiel, selon les axes duquel les caractéristiques de tous les phénomènes spatiaux pourraient être mesurées, et sur la base duquel n'importe quelle relation spatiale pourrait être déduite. Chevallier [Chevallier 1983] rappelle par exemple la contradiction qu'il y a à délimiter des zones d'aménagement au stylo-feutre sur un plan à l'échelle du 1:5'000, et d'en déduire par la suite des restrictions d'utilisation d'une propriété dont l'assiette est définie à quelques dixièmes de millimètres près sur un plan au 1:500 !

Nous introduisons donc la notion de *référentiel spatial*, que nous définissons comme un ensemble d'échelles de mesures permettant de situer de manière univoque des phénomènes sur le territoire. Différents référentiels spatiaux peuvent être définis sur une même portion de territoire. Chaque caractéristique spatiale pertinente peut être évaluée de manière univoque au sein d'un même référentiel spatial. Les échelles de mesure et les exigences relatives à la qualité des mesures sont adaptées aux objectifs de la modélisation. Des fonctions compatibles avec la nature des échelles de mesure peuvent en outre être définies²⁰: topologiques dans le cas d'échelles nominales ou ordinales, ou métriques dans le cas d'échelles cardinales.

Le choix d'un ensemble d'échelles de mesure et leur combinaison en un référentiel spatial dépend donc des phénomènes que l'on souhaite modéliser. Les types de référentiels spatiaux suivants sont couramment utilisés:

- Les référentiels *orthogonaux* (x , y et éventuellement z) sont utilisés pour la modélisation spatiale de phénomènes isotropes²¹. La simplicité des changements d'origine en font des référentiels bien adaptés à l'intégration de données relatives à l'ensemble d'un territoire.
- Les référentiels *polaires* combinent deux types de mesures différents (angles et distances). Ils sont avant tout utilisés pour l'observation directe de mesures spatiales. La précision des mesures se dégrade cependant à mesure que l'on s'éloigne du point-origine. On les utilise donc avant tout pour la saisie de mesures dans la réalité. Ces mesures sont généralement transformées dans un référentiel orthogonal pour leur mémorisation et leur exploitation.

²⁰ Au sein d'un même référentiel, une *algèbre* peut être définie [Frank et Kuhn 1986] [Herring 1989], permettant de déduire par le calcul symbolique des faits de la réalité.

²¹ Sans direction préférentielle.

- Les référentiels *curvilignes* sont utilisés pour la modélisation spatiale de phénomènes anisotropes linéaires, tels que les caractéristiques des réseaux routiers et hydrographiques.

Quelques exemples tirés du domaine de la distribution d'eau potable nous permettront d'illustrer la notion de référentiel spatial:

- Un premier référentiel est destiné aux informations nécessaires au repérage des canalisations dans le sous-sol (construction et maintenance). La position des différents objets est mesurée à une dizaine de centimètres près (échelle cardinale non discrétisée), et des distances entre objets peuvent être évaluées.
- Un second référentiel englobe les informations nécessaires au contrôle du fonctionnement du réseau. Les objets sont localisés à quelques mètres près. L'évaluation de distances précises n'est pas possible, mais la connectivité du réseau joue un rôle important. Des fonctions de réseau, telles que l'isolation de biefs, sont possibles.
- Un troisième référentiel est destiné aux calculs de réseau. Seules les connexions des conduites sont évaluées (schémas).
- Les installations peuvent en outre être documentées par des photos obliques non redressées. Ces photos peuvent être numérisées et intégrées à une base de données "multi-médias". Chaque photo constitue alors un référentiel spatial particulier, sur lequel des objets peuvent être localisés et assortis de caractéristiques diverses.

La notion de référentiel spatial est essentielle à la suite de notre étude. Elle permet la définition de modèles spatiaux homomorphes du territoire. Elle implique en particulier que des caractéristiques et relations spatiales ne peuvent être modélisées ou inférées qu'au sein d'un même référentiel spatial.

4.3.2. Typologie et représentation physique des modèles spatiaux

On peut distinguer plusieurs types de modèles spatiaux, susceptibles de décrire des phénomènes de natures différentes. En nous référant en particulier à [Chevallier 1986], et sans prétendre à l'exhaustivité, nous pouvons proposer la typologie suivante:

1. **Les objets spatiaux** sont des modèles spatiaux directement associés à des objets sémantiques. Leur modélisation spatiale est réalisée objet par objet, conformément à des critères d'homogénéité sémantique et éventuellement temporelle, sans que l'on ait à tenir compte de leur voisinage. Ce type de modèle spatial est particulièrement bien adapté à des applications de gestion administrative et technique, dont l'objectif est précisément de restituer une image fidèle des objets de la réalité et de leur évolution.
2. **Les réseaux** sont des modèles spatiaux à une dimension définis dans un espace à deux ou trois dimensions, destinés à la représentation des réseaux de câbles, de conduites ou de rivières. Leur configuration topologique peut être soumise à des contraintes particulières de sens d'écoulement, de connectivité, etc.
3. **Les mosaïques** sont des classifications de l'espace conformément à des critères sémantiques, spatiaux ou temporels. Elles constituent donc des partitions de l'espace (généralement bidimensionnel) en unités spatiales élémentaires.

Il est impératif que le critère de classification retenu soit mesuré sur une échelle discrétisée²² homogène, définie sur l'ensemble du territoire considéré, faute de quoi le critère retenu ne définira pas une partition de l'espace ! Il ne serait par

22 Il est en effet impossible de définir une mosaïque sur un domaine de valeurs continu. Le passage d'une échelle continue (généralement cardinale) à une échelle discrétisée (généralement ordinale) constitue ainsi un problème courant de l'analyse spatiale ou thématique (élaboration de cartes thématiques en géographie socio-économique par exemple) [IGUL et EPFL-IM-GM 1992].

exemple pas adéquat de classifier le territoire en "forêt" et en "surface agricole", avec la forêt définie en termes de densité et de hauteur de peuplement, avec la surface agricole définie en termes de qualité des sols, et avec une surface bâtie qui n'appartiendrait ni à l'une, ni à l'autre des deux classes considérées ²³.

Les mosaïques sont aussi appelées fréquemment *couches* ou *cartes*, termes imagés qui prêtent cependant dangereusement à confusion. Ce type de modèle d'information est particulièrement adapté à l'analyse spatiale.

4. Les *carroyages* sont des découpages du territoire en *mailles* régulières (en principe carrées, mais des découpages triangulaires ou hexagonaux sont aussi possibles). La valeur de différents phénomènes peut être estimée à l'intérieur de chaque maille, mais le processus de modélisation reste incomplet, puisqu'aucune sélection ni simplification sémantique ou spatiale ne sont réalisées. Un carroyage est donc adapté à la mesure brute d'une valeur sur l'ensemble de l'espace, telle qu'elle est réalisée en télédétection par exemple. La grandeur des mailles doit être faible par rapport à l'étendue des phénomènes à saisir, pour en permettre la classification ultérieure, nécessaire à l'achèvement du processus de modélisation.
5. Les *modèles d'observations ponctuelles* sont constitués d'un semis de points d'échantillonnage, qui peuvent être répartis régulièrement ou irrégulièrement (conformément à différents critères dépendants du domaine considéré) dans l'espace; un exemple de modèle d'observations ponctuelles régulier est donné par la grille d'échantillonnage à l'hectare de la statistique suisse des superficies, et un exemple de modèle d'observations ponctuelles irrégulier par le jeu des sondages réalisés lors d'une campagne de mesures pédologiques. Il est fondamental de distinguer les modèles d'observations ponctuelles, dans lesquels les phénomènes sont mesurés en divers points, des modèles en mosaïque ou en carroyage, dans lesquels des valeurs moyennes sont mesurées sur des surfaces.
6. Les *modèles algorithmiques* déterminent la valeur d'un phénomène en chaque point du territoire sur la base de quelques points dont la valeur est mesurée, ainsi que d'un algorithme et de paramètres d'interpolation. Le modèle d'interpolation peut être de nature physique ou de nature stochastique (krigeage), suivant la nature du phénomène modélisé. Les modèles numériques de terrain en sont des exemples caractéristiques.

Pour la modélisation spatiale d'un phénomène, on s'efforcera de choisir le type de modèle spatial le mieux adapté. Dans ce domaine, les solutions de facilité peuvent se révéler dangereuses. A titre d'exemple, le rattachement d'informations socio-économiques telles que la densité d'habitation à des entités politiques telles que les communes est fréquente en géographie humaine, mais peut conduire à dangereusement biaiser les informations [Buogo et Stucky 1992].

Des difficultés importantes surviennent encore lorsque le phénomène à modéliser fait appel aux caractéristiques de plusieurs types de modèles spatiaux. Ainsi en va-t-il des *bassins versants* dans le domaine de l'assainissement, qui constituent une partition du territoire (mosaïque), mais auxquels on souhaite associer des informations sémantiques complexes propres aux objets spatiaux. La gestion explicite des *relations topologiques* (cf. § 4.3.4) permet de résoudre un tel problème.

En outre, l'adéquation d'un type de modèle à un phénomène dépend de l'objectif de la modélisation, et donc du domaine d'application considéré. L'exploitation de données spatiales dans un but différent de celui pour lequel elles ont été saisies peut donc impliquer un changement du type de modèle spatial. Une telle transformation exige généralement l'apport d'expertise ou d'informations supplémentaires. A titre d'exemple, une conduite d'assainissement peut être modélisée comme un *objet spatial* à des fins de gestion des

23 La présence d'une classe "autre" ou "reste" est le symptôme fréquent d'une classification inadéquate !

constructions, alors qu'elle sera modélisée comme un réseau pour en gérer le fonctionnement. Là encore, la gestion explicite des relations topologiques peut permettre, sous certaines conditions, de concilier ces contraintes.

L'analyse spatiale recourt aussi fréquemment à la transformation d'un modèle d'information dans un autre. En voici quelques exemples:

- La *spatialisation*²⁴ d'un modèle d'observations ponctuelles en un modèle par mosaïque exige que des informations soient données sur le type de répartition spatiale du phénomène considéré.
- L'agrégation d'objets spatiaux sur la base d'une mosaïque exige le choix préalable des critères d'agrégation (plusieurs zones forestières peuvent par exemple constituer un seul et même objet "forêt").
- La classification d'un carroyage permet d'obtenir une mosaïque ou d'isoler des objets spatiaux, conformément aux critères et à la méthode de classification choisis.

Malheureusement, la problématique fondamentale du choix d'un modèle spatial adapté au phénomène à décrire est souvent masquée par des préoccupations d'ordre essentiellement technologique. Ainsi, le cloisonnement des technologies "vecteur" et "raster" a laissé une empreinte durable sur le traitement informatique des données localisées, malgré la récente apparition sur le marché d'un nombre croissant de systèmes informatiques *intégrés* ou *hybrides*, dans lesquels les représentations "vecteurs" et "raster" coexistent.

Il existe d'ailleurs encore bien d'autres représentations des modèles spatiaux que les "vecteurs" et "raster" mentionnées ci-dessus, à l'exemple des quadtree (2D) ou des octree (3D) [Laurini 1989], qui constituent une agrégation partielle d'une représentation "raster" de dimension correspondante. Dans une autre direction, l'intégration de l'imagerie vidéo dans des systèmes "multi-média" se révèle prometteuse pour la gestion et l'exploitation d'informations non structurées [Chevallier 1992].

Ces différentes technologies étant toujours mieux intégrées, il devient possible de choisir la mieux adaptée pour représenter les modèles spatiaux d'un système d'information. Le tableau 4.2 esquisse quelques solutions possibles, ainsi que leurs avantages et leurs inconvénients. Nous rappelons que ce tableau ne prétend pas à l'exhaustivité, et que d'autres représentations, aussi bien que d'autres types de modèles spatiaux pourraient être ajoutés.

Ce tableau permet de situer les différentes solutions possibles pour représenter les modèles spatiaux, et montre combien les choix technologiques sont distincts des choix conceptuels. Quelques remarques sont cependant encore nécessaires pour préciser le cadre défini:

1. La représentation par "raster" est caractérisée par une grandeur de maille qui implique une échelle de mesure cardinale discrétisée des caractéristiques spatiales des phénomènes, alors que la représentation par vecteurs implique une échelle de mesure cardinale non discrétisée [Chevallier et Bédard 1990].
2. La représentation par vecteurs exige dans tous les cas une agrégation de l'information préalable à la saisie: l'utilisateur doit en effet lui-même *délimiter* dans l'espace géographique le phénomène qu'il veut décrire. La modélisation "raster" se base au contraire sur des surfaces élémentaires arbitraires (les pixels), sur lesquelles les phénomènes à modéliser sont évalués; ce n'est que lors de l'utilisation du système que des "zones" pour lesquelles un phénomène prend des valeurs caractéristiques peuvent être identifiées. La représentation par vecteurs est donc adaptée à l'agrégation d'informations par objets, alors que la représentation "raster" est adaptée à l'agrégation d'informations par valeurs.

24 Généralisation à l'ensemble de l'espace.

Modèle d'information	Solutions technologiques envisagées
Objets spatiaux	<p>La description ou la délimitation d'objets spatiaux par des vecteurs est particulièrement bien adaptée à la modélisation d'entités artificielles (bâtiments, canalisations, ...) ou juridiques (parcelles, zones, ...).</p> <p>La technologie vectorielle est par contre mal adaptée à la modélisation d'objets naturels, aux limites mal définies et difficiles à tenir à jour ²⁵.</p>
Réseau	La technologie vectorielle se prête particulièrement bien à la modélisation des connexions caractéristiques d'un réseau.
Mosaïque	<p>La représentation d'une mosaïque peut faire appel aussi bien à la technologie vectorielles qu'à la technologie raster:</p> <ul style="list-style-type: none"> • Des blocs de pixels adjacents constituent le résultat privilégié de la classification d'images satellitaires. Ces blocs peuvent être agrégés par valeur pour représenter les zones d'une mosaïque. • Des mailles irrégulières peuvent être délimitées par des vecteurs (réseaux de polygones par exemple).
Carroyage	<p>Chaque maille peut être aisément représentée par un pixel d'une matrice de points (raster).</p> <p>Lorsque la forme de la maille a une importance déterminante (réseau triangulé par exemple), la délimitation de chaque maille par des vecteurs est possible.</p>
Modèle d'observations ponctuelles (échantillonnage régulier)	<p>La technologie raster se prête bien à la représentation d'un tel modèle: chaque pixel de la trame représente alors un point d'échantillonnage.</p> <p>En technologie vectorielle, il est possible de définir un semis de points régulier.</p>
Modèle d'observations ponctuelles (échantillonnage irrégulier)	Un tel modèle peut être représenté par un semis de points quelconque en technologie vectorielle, ou par un maillage fin dont seuls quelques pixels ont une valeur significative en technologie raster.
Modèle algorithmique	<p>Le choix d'une technologie dépend aussi bien de la nature des phénomènes représentés par le modèle que de la forme des résultats attendus. Il existe par exemple des modèles numériques d'altitude (MNA) adaptés à la technologie raster, et des MNA adaptés à la technologie vectorielle.</p> <p>Dans tous les cas, des données de calage doivent être définies, sous forme de points et de lignes de structure ou de rupture (vecteurs) ou d'un ensemble de pixels (raster).</p>

Tableau 4.2: Adéquation des modèles de données aux modèles d'information

3. La représentation par vecteurs permet par essence de décrire des structures linéaires, avec lesquels il est possible de délimiter des surfaces. La représentation "raster" décrit par essence des surfaces, dont les limites communes peuvent ensuite être mises en évidence [Aalders 1990]. Dans un espace bidimensionnel, la représentation par vecteurs est donc a priori mieux adaptée à la représentation de phénomènes linéaires, alors que la représentation "raster" est mieux adapté à représenter les surfaces.

²⁵ La modélisation d'objets naturels peut être appréhendée par les fractales [Mandelbrot 1989]. L'intégration des concepts et des solutions numériques relatives aux fractales dans des systèmes de traitement des informations localisées est une préoccupation de recherches parallèles.

4. La mise à jour pixel par pixel d'un modèle représenté par un "raster" est généralement fastidieuse et difficile, contrairement à la mise à jour d'un modèle représenté vectoriellement. La représentation par vecteurs se prête donc bien à une mise à jour continue, comme on la rencontre dans les applications de gestion (orientées objets), alors que la représentation "raster" est mieux adaptée à des mises à jour globales, pratiquées dans le cas des carroyages.

4.3.3 Utilité des modèles spatiaux

Les "cartes" et les "plans" constituent simultanément le support d'information et le document de consultation des systèmes d'information du territoire traditionnels [Chevallier 1983]. Cette double fonction tend à masquer la nature des informations que l'utilisateur peut en tirer:

- D'une part, ces documents permettent à l'utilisateur d'accéder visuellement aux informations qui y sont conservées: il est ainsi facile de repérer sur un plan des conduites la rue des Acacias, et d'examiner le diamètre de l'égout qui le parcourt. Une carte ou un plan a donc un rôle d'*index spatial*.
- D'autre part, on peut aisément lire sur un plan les caractéristiques spatiales des ses composantes (localisation, dimension, forme) et leurs relations spatiales: quels sont les bâtiments voisins du 15 de la rue des Acacias ? à quelle conduite d'adduction d'eau potable ces bâtiments sont-ils connectés ? quelle est la longueur de chaque branchement ? Une représentation graphique adéquate permet de lever des ambiguïtés gênantes.

De telles informations doivent bien sûr aussi être fournies par les systèmes d'information numériques, dans lesquels le support d'information (base de données) est distinct des documents graphiques qui peuvent en être extraits. Si on retrouve bien la fonction d'indexation spatiale dans le processus d'interaction visuelle sur un écran graphique, les nombreuses représentations visuelles possibles, dont le contenu, l'échelle et la qualité graphiques peuvent varier considérablement, ne permettent plus une modélisation univoque de la géométrie par le biais d'artifices graphiques.

On attend cependant d'un système informatique qu'il soit capable d'agrèger automatiquement de telles informations. Pour ce faire, deux solutions sont possibles:

- a) disposer des fonctionnalités permettant de déduire en tout temps ces caractéristiques et relations de la métrique, par calcul: l'agrégation des informations est effectuée lors de chaque utilisation.
- b) stocker explicitement dans la base de données chaque caractéristique ou relation géométrique pertinente pour l'application considérée: l'agrégation des informations est alors effectuée lors de leur saisie.

La solution a) est par essence adéquate pour l'agrégation d'informations métriques (paramètres de courbes et de surfaces, distances, etc.), qui sont mesurées sur des échelles intervalle ou ratio [Chevallier et Bédard 1990] de précision limitée. Cette solution est par contre insuffisante pour l'agrégation univoque d'informations *topologiques* (dimension topologique d'une composante, adjacence, connexion, inclusion, disjonction, etc.²⁶), qui sont mesurées sur une échelle nominale ou ordinale et qui ne peuvent par conséquent être déduites de manière univoque d'une métrique²⁷ de précision limitée. Conformément aux notions présentées au § 4.3.1, les relations métriques et topologiques ne sont pas évaluées

26 Une étude exhaustive des relations topologiques est proposée dans [Champoux 1991].

27 L'imprécision de la métrique dépend bien de la nature même de son échelle de mesure, et les imprécisions supplémentaires induites par la représentation numérique ne constituent donc dans ce cadre pas un problème essentiel, contrairement à l'approche technologique proposée dans [Frank et Kuhn 1986].

dans le même référentiel spatial. La solution b) est donc indispensable à l'agrégation univoque des informations topologiques.

La solution b) exige la gestion d'un plus grand nombre d'informations que la solution a); les structures de données y sont ainsi notablement plus lourdes et complexes. En échange, les ambiguïtés inhérentes aux modèles purement métriques sont résolues une fois pour toutes par le responsable de la saisie des données, et non a posteriori par chaque utilisateur.

4.3.4. Topologie et modèles spatiaux

La plupart des systèmes informatiques modernes permettent de "gérer la topologie". Cette expression imprécise et fourre-tout cache en fait deux approches fondamentalement différentes de la topologie:

- a. La première approche est d'essence *technologique* [Frank et Kuhn 1986][Herring 1989]. Elle a pour objectif de lever les ambiguïtés liées à la représentation numérique des données, et de diminuer les temps de réponse des systèmes lorsque des agrégations d'informations topologiques sont requises. La solution la plus couramment utilisée de nos jours consiste en l'intégration de toutes les primitives géométriques en un graphe planaire total et rigoureux, par agrégation des éléments duquel tous les objets sont décrits (figure 4.5); le système informatique TIGRIS constitue actuellement l'exemple le mieux achevé d'une telle approche.

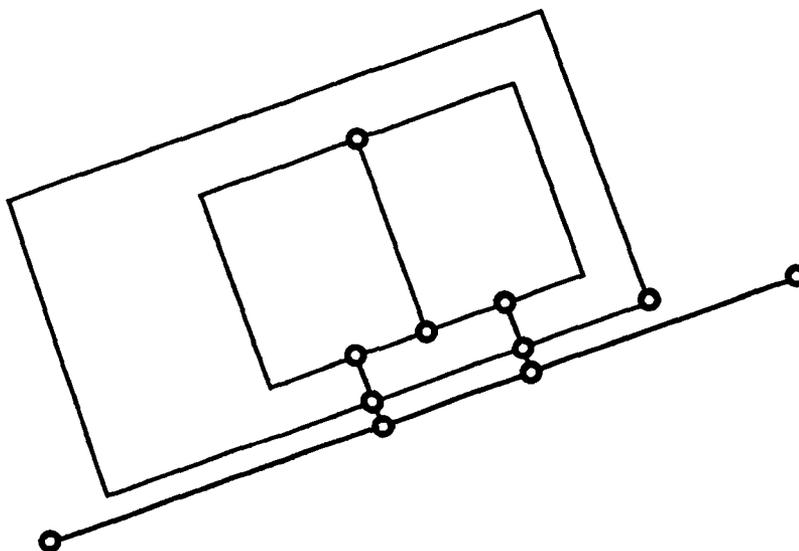


Figure 4.5: Description d'objets spatiaux à l'aide d'un graphe planaire

Aucun indice - hormis bien sûr les temps de réponse !- ne devrait trahir aux utilisateurs de tels systèmes informatiques la présence de structures topologiques. Ces systèmes utilisent un mécanisme d'"oracles" (décisions arbitraires, basées sur la métrique) [Frank et Kuhn 1986] pour construire la topologie en "temps réel" lors de la saisie.

Un défaut important d'un tel système est son incapacité à refléter la connaissance réelle que l'on a de la réalité, et en particulier à prendre la *méconnaissance* en considération: une conduite est-elle connectée à une autre ? peut-être ne le sait-on pas lors de la saisie; un système informatique "à oracles" vous obligera cependant à décider ! Une voie aérienne passe-t-elle au-dessus de la ligne du trolleybus no 5 ? Que cela vous intéresse ou non, un système informatique "à oracles" intégrera ce fait à son modèle, alors qu'il ne reflète qu'indirectement la réalité.

- b. La seconde approche de la topologie répond directement aux *besoins en information des utilisateurs*, et reflète exactement la connaissance que l'on a de la réalité. Dans une telle approche, on ne recueille lors de la saisie que l'information que l'on connaît effectivement: deux conduites peuvent par exemple

être connectées, déconnectées, ou on peut ne pas le savoir... ou encore n'y pas être intéressé ! La gestion de la topologie n'est donc pas ici un procédé informatique, mais avant tout un processus cognitif.

Dans cet esprit, il est donc nécessaire d'intégrer au méta-modèle du système les structures topologiques pertinentes pour les besoins des utilisateurs (que veut-on connaître ?), et subsidiairement les contraintes qui s'appliquent à ces relations.

On notera que le modèle topologique peut parfaitement dépendre du niveau d'abstraction considéré: la connaissance de la connexion de conduites est importante pour comprendre le fonctionnement d'un réseau d'égouts, alors qu'elle est parfaitement superflue pour repérer une telle conduite dans la rue !

Ce processus cognitif est néanmoins inscrit dans un contexte spatial, et il est donc étroitement lié à la métrique, avec laquelle des contraintes de *vraisemblance* doivent être respectées: il serait de toute évidence difficile d'admettre la connexion de deux conduites dont les extrémités seraient éloignées de plusieurs mètres ! Pour gérer efficacement la topologie dans un système numérique, une interface graphique interactive permettant de visualiser la géométrie et de lui associer des contrôles de vraisemblance adéquats est donc nécessaire. Un exemple d'une telle approche est donné par le système informatique APIC de la société française Polilog.

Bien que les deux approches de la topologie présentées ci-dessus ne soient généralement pas clairement distinguées dans la littérature actuelle, elles poursuivent des objectifs complémentaires et distincts:

- l'approche technologique tend à résoudre les ambiguïtés liées à la gestion numérique des informations, et à améliorer les performances de traitement informatique. Elle ne permet pas de gérer des informations qui ne seraient pas intrinsèquement liées à la métrique.
- l'approche cognitive permet de modéliser des caractéristiques spatiales qui ne peuvent pas être représentées de manière appropriée par la métrique.

La modélisation de différents référentiels spatiaux au sein d'un même système (cf. § 4.3.2) est susceptible d'apporter une solution satisfaisant ces deux objectifs²⁸. Les échelles de mesure de chaque référentiel peuvent en effet être adaptées à la nature des phénomènes et des relations à gérer. A notre connaissance, on ne trouve cependant pas sur le marché d'aujourd'hui de système informatique dont les structures internes permettraient de modéliser parallèlement plusieurs référentiels spatiaux. Mais les systèmes informatiques disponibles aujourd'hui proposent différents modes de représentation interne de la topologie: graphe planaire du System 9 de Computervision, "relations automatiques" utilisées dans APIC de Polilog, ... Ces caractéristiques doivent être soigneusement évaluées lors de la l'acquisition d'un système informatique.

Parmi les SGDL actuels, on distingue en particulier [Pornon 1989]:

- a. les systèmes capables de gérer la *nodalité*, c'est à dire de modéliser le rattachement de deux lignes à un même noeud; de tels systèmes permettent de représenter en particulier la connectivité dans les réseaux industriels;
- b. les systèmes capables de modéliser l'adjacence de deux surfaces le long d'une même ligne, adaptés à la gestion de réseaux polygonaux (couverture du sol, zones d'aménagement, etc.), et en particulier à la gestion cadastrale (parcelles);

28 La coexistence de plusieurs *sets topologiques* au sein d'un même système informatique (système TIGRIS) ou les *représentations topologiques multiples* proposées par [Bruegger et Kuhn 1990] constituent un pas dans la direction souhaitée.

- c. les systèmes dans lesquels différents objets peuvent être décrits par une même surface (par exemple, assiette d'un droit de bâtir défini sur l'ensemble d'une parcelle), adaptés aux tâches les plus complexes de l'analyse spatiale.

En outre, les composantes géométriques sont généralement associées à un ensemble de contraintes classées en *types topologiques*, dont le système assure la cohérence permanente. La classification en types topologiques est essentiellement basée sur la *dimension* du modèle spatial: on distingue les *points* (0 dimension), les *lignes* (1 dimension), les *surfaces* (2 dimensions), voire encore les *volumes* (3 dimensions). L'existence possible de "trous" dans les surfaces (enclaves), le non-recouvrement de surfaces de même type, ainsi que le non-recoupement de lignes de même type constituent d'autres exemples de contraintes qui peuvent être associées à des types topologiques particuliers.

4.4. Mécanismes d'abstraction

Au début de ce chapitre (§ 4.1), nous avons évoqué la démarche d'abstraction qui nous permet de distinguer dans la réalité les phénomènes pertinents pour un problème donné. Nous avons aussi constaté que le raisonnement humain est basé sur un va-et-vient continu entre différents niveaux d'abstraction.

Dans le domaine administratif cependant, les systèmes d'information considérés se situent généralement dans une plage d'abstraction relativement étroite ²⁹ (figure 4.6), dans laquelle la *classification* d'une instance (l'EPFL, par exemple) en une composante du méta-modèle (université, par exemple) et son mécanisme inverse, l'*instanciation*, suffisent à parcourir la plage d'abstraction considérée. Le registre foncier est un bon exemple d'un tel système, qui consiste à décrire la réalité en termes de *personnes*, de *parcelles*, et de *droits des personnes sur les parcelles*.

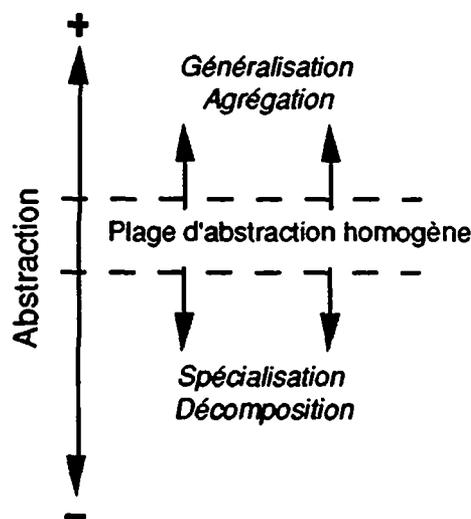


Figure 4.6: Niveaux et mécanismes d'abstraction

²⁹ Les niveaux d'abstraction interviennent non seulement dans le *fonctionnement* des systèmes d'information localisés, comme envisagé ici, mais aussi dans le cycle d'abstraction lié à la *conception* d'un système d'information (niveaux d'abstraction successifs [Caron 1991]). Il s'agit alors de niveaux d'abstraction propres au système de représentation [Le Moigne 1977], et non au système d'information lui-même.

Mais les systèmes traitant des informations localisées, et particulièrement les SIRS, sont des systèmes généralement plus complexes: ils peuvent en effet intégrer l'ensemble des informations pertinentes sur un territoire donné, qui peuvent aller de la planification d'installations ou d'aménagements à leur positionnement précis et détaillé à des fins de construction ou de maintenance [Chevallier et Bédard 1990]. Les SIRS couvrent donc une plage d'abstraction relativement large; ils doivent assurer la coexistence cohérente des différents niveaux d'abstraction, et offrir les mécanismes nécessaires pour passer d'un niveau à l'autre.

Ces mécanismes d'abstraction et la structure des informations sur lesquels ils s'appuient sont l'objet du présent paragraphe.

Les mécanismes d'abstraction ont préoccupé de nombreux auteurs, venant aussi bien des domaines de la psychologie [Perner 1991], de l'informatique, de l'intelligence artificielle, des systèmes d'information [Brodie, Mylopoulos et al. 1984], et plus directement des SIRS [Frank et Hudson 1988]. Ce dernier auteur distingue les mécanismes d'*agrégation*, consistant à combiner les composantes d'un modèle, de *généralisation*, consistant à combiner les composantes d'un méta-modèle, et de *classification*, consistant à rattacher une composante du modèle (instance) à une composante du méta-modèle. Ces définitions restent à nos yeux trop réductrices, dans la mesure où elles se fondent sur l'idée propre aux bases de données que la frontière entre la structure du modèle et ses instances peut être définie sans ambiguïté. Or, la coexistence au sein d'un SIRS de niveaux d'abstraction très différents implique que des informations "stables" pour un terme donné (et par conséquent éléments de la structure) sont "évolutives" pour un terme plus éloigné (et donc des instances).

Nous suggérons donc de prendre en considération les mécanismes d'abstraction suivants:

- a. La *généralisation* est un mécanisme d'abstraction qui tend à diminuer la *diversité* de l'information.

Si il est par exemple pertinent pour l'évaluation des émissions acoustiques du trafic routier de distinguer les poids lourds des automobiles et des motocycles, il est par contre suffisant de dénombrer des véhicules pour déterminer des itinéraires de délestage. C'est une généralisation qui nous permet d'assimiler poids lourds, automobiles et motocycles à des véhicules.

On constate que la généralisation ne diminue en rien la *quantité* d'information traitée: que l'on considère les véhicules X et Y ou le camion X et la voiture Y, nous distinguons toujours deux objets de la réalité !

Le mécanisme inverse de la généralisation est la *spécialisation*, qui vise la diversification du modèle considéré.

- b. L'*agrégation* est un mécanisme d'abstraction qui tend à diminuer la *quantité* d'informations considérées ³⁰.

Ainsi, le visiteur de l'EPFL souhaite en principe pouvoir distinguer le bâtiment du génie rural des autres bâtiments de l'EPFL, alors que le facteur des PTT pourra se contenter d'apporter son courrier à l'EPFL. Tous les bâtiments de l'EPFL sont dans ce cas regroupés en une seule entité du même type.

30 Le lecteur sera attentif à ce que le terme d'*agrégation* revêt un sens plus vaste et plus général dans le domaine de l'aide à la décision: on entend par *agrégation* toutes combinaisons finalisées d'informations pertinentes pour la prise de décisions.

Les relations définies entre les instances d'un modèle définissent aussi des agrégations d'informations: ainsi, les droits de propriété gérés par le registre foncier sont des agrégations de parcelles et de personnes.

Le mécanisme inverse de l'agrégation est la *décomposition*, qui tend à plus détailler le modèle considéré.

Dans ce contexte, la *classification* constitue un mécanisme d'abstraction hybride, qui intègre:

- le rattachement d'un objet instancié (de durée de vie limitée) à son concept atemporel;
- une généralisation du concept atemporel.

A titre d'exemple, l'interprétation du prédicat "Les bâtiments de l'EPFL ont été inaugurés en 1976" implique de rattacher l'EPFL à un concept atemporel, puis à rattacher ce concept à celui plus général de "bâtiment".

Les mécanismes d'abstraction énumérés ci-dessus ont été définis dans un contexte exclusivement sémantique. Mais des mécanismes similaires peuvent être appliqués à des modèles spatiaux:

- a. La *généralisation spatiale* tend à diminuer la *complexité* des modèles spatiaux, en simplifiant la forme (figure 4.7), voire même en ne conservant que leur position [Chevallier et Bédard 1990]. Cette définition est conforme à la notion de *généralisation cartographique*.

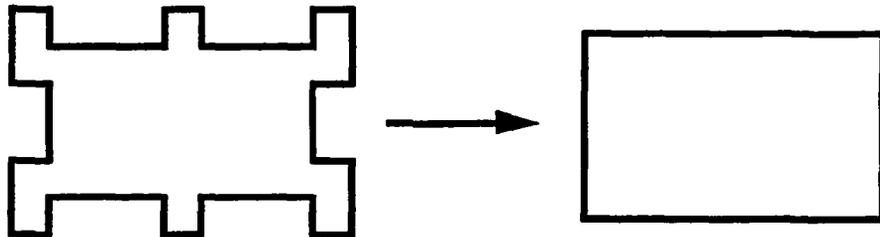


Figure 4.7: Généralisation spatiale

- b. L'agrégation spatiale tend à diminuer la *quantité* des modèles spatiaux, en amalgamant des modèles représentant différentes parties d'une même entité de la réalité. L'agrégation peut être réalisée par *réunion* ou par *intersection* des modèles agrégés.

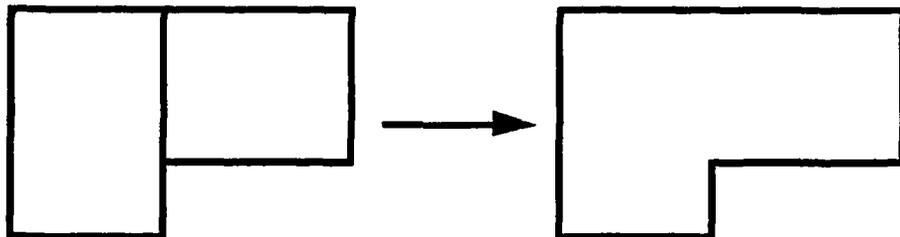


Figure 4.8: Agrégation spatiale (réunion)

De même, on peut définir une généralisation et une agrégation temporelles:

- a. La *généralisation temporelle* tend à diminuer la *granularité* de l'échelle de mesure du temps: on exprimera par exemple qu'un évènement est attendu pour 1991 plutôt que pour le 1er août 1991.
- b. L'agrégation temporelle consiste à regrouper des intervalles de temps.

A l'extrémité supérieure de l'échelle d'abstraction représentée sur la figure 4.6, on trouve des modèles à la fois très généraux et agrégés, essentiellement adaptés aux préoccupations de niveau stratégique. A l'extrémité inférieure de l'échelle, on trouve au contraire des modèles très diversifiés et décomposés, adaptés à des tâches de gestion technique ou administrative. On constate donc que pour une tâche donnée, un modèle doit être situé à un niveau de généralisation et à un niveau d'agrégation cohérents, situés dans une plage d'abstraction homogène.

Cette cohérence entre le niveau de généralisation et d'agrégation doit en outre s'étendre aux modèles spatiaux et temporels: à un niveau d'approche plus abstrait (planification) correspondent en règle générale une plus faible résolution spatiale (petite échelle) [Chevallier et Bédard 1990] et temporelle.

Les modèles utilisés pour le dessin assisté par ordinateurs (DAO) constituent généralement des contre-exemples typiques: à la fois très peu diversifiés (toute l'information est réduite à des symboles, des traits et des trames) et très détaillés (pas d'agrégation des éléments du dessin), leurs éléments doivent être classifiés et agrégés pour être intégrés à un système d'information. Ces opérations impliquent un apport informationnel considérable... et sont par conséquent coûteuses ! Les images satellites brutes (non classifiées) constituent un exemple similaire de modèle très peu diversifié (composé de pixels) et très décomposé (chaque pixel constitue un élément indépendant); la classification de l'image consiste à regrouper les pixels en classes et à les agréger en zones homogènes, à un niveau d'abstraction conforme au but recherché. En appliquant ces concepts aux origines de notre monde, on constate que le chaos est fait de toutes petites particules faiblement diversifiées; il ne fallut que six jours au Seigneur, le premier grand Concepteur de systèmes, pour le diversifier et le modeler en un monde fonctionnel et cohérent, de l'échelle des atomes à celle de l'univers [Genèse 1, 1-27].

Si l'on en revient aux préoccupations beaucoup plus concrètes de l'informatique, on constate que les systèmes de bases de données actuels ont été développés avant tout pour des applications dans le domaine administratif, couvrant une plage d'abstraction homogène dans lesquelles un seul niveau de généralisation est utilisé: la classification des instances en éléments du schéma. Le modèle relationnel ne distingue ainsi pas la généralisation de l'agrégation ³¹. Les modèles et formalismes sémantiques développés ultérieurement pour permettre une description moins biaisée de la réalité ne sont généralement guère mieux adaptés à la modélisation de la généralisation; c'est en particulier le cas du modèle entité-relation [Chen 1976] et de ses représentations formelles les plus courantes, telles que le formalisme individuel [Tardieu, Rochfeld et al. 1984]. Des recherches récentes tendent néanmoins à corriger cette lacune et permettent une modélisation distincte des agrégations et des généralisations [Frank 1982] [Spencer, Teorey et al. 1990]. La description des généralisations proposée par Frank a ainsi été reprise dans le modèle et le formalisme en cours de développement à l'Université Laval [Caron 1991].

Un avenir particulier semble actuellement promis aux systèmes de bases de données orientés objets. Entre autres qualités, ces systèmes intègrent les concepts des bases de données et ceux de la programmation orientée objets [Pascoe 1986], permettant ainsi de modéliser le comportement aussi bien que la structure des entités de la réalité, et introduisant la notion d'héritage entre classes d'entités. Ces systèmes constituent ainsi de toute évidence une contribution importante de la technologie informatique à la réduction du biais existant encore entre les modèles cognitifs et les modèles informatiques ³². Tous les problèmes ne sont cependant pas encore résolus, et le mécanisme informatique de

31 La modélisation des généralisations n'est possible qu'avec l'aide de fonctions ad hoc, permettant d'effectuer des jointures assorties de contraintes sémantiques conformes à la nature de la généralisation.

32 Brodie et al. ont réalisé une synthèse fondamentale des différentes approches de la modélisation proposées dans les domaines des banques de données, de l'informatique et de l'intelligence artificielle dans [Brodie, Mylopoulos et al. 1984]. Cette synthèse met bien en évidence les hiatus existant entre les différentes approches.

l'héritage ne constitue pas un reflet sans biais du mécanisme cognitif de *généralisation*. Une étude plus détaillée devrait permettre d'appréhender les fonctions de représentation à mettre en oeuvre.

Dans la mesure où agrégation et généralisation constituent des mécanismes cognitifs distincts, il est judicieux que les formalismes sémantiques utilisés permettent aussi de les représenter de manière distincte. Ce n'est que lors de l'implémentation de ces modèles dans des systèmes de bases de données opérationnels qu'on peut admettre de les biaiser, à l'exemple de relations assorties de mécanismes de contrôle des cardinalités des jointures dans le modèle relationnel.

4.5. Conclusion du chapitre

Ce chapitre a permis d'introduire plusieurs notions nouvelles relatives à la modélisation des informations localisées:

- Un canevas de classification des modèles a été proposé, qui intègre aussi bien la notion courante de *modèle de données* que celle de *modèle de simulation*.
- La notion nouvelle de *référentiel spatial* a été définie. Un référentiel spatial regroupe des modèles spatiaux mesurés sur des échelles de nature et de qualité homogènes. Des opérations conformes à la nature des échelles de mesure utilisées peuvent être appliquées au sein d'un même référentiel.
- Une *typologie des modèles spatiaux* a été élaborée, qui reflète la nature des informations considérées, plutôt que la représentation informatique adoptée.
- Une nouvelle approche de la *topologie*, mieux adaptée que l'approche classique aux besoins courants des utilisateurs, a été proposée.
- Les mécanismes d'abstraction classiques d'*agrégation* et de *généralisation* ont été définis et étendus à la référence spatiale.

Ces notions seront utilisées dans les prochains chapitres pour résoudre les problèmes de modélisation des données et des traitements dans un système traitant des informations localisées.

5. Problèmes et solutions relatifs aux traitements

Le chapitre 4 a présenté les modèles numériques comme des enchaînements de transformations permettant de déduire d'un état initial d'un système d'autres états de ce système. Nous avons montré que de tels enchaînements de transformations décrivent le *comportement* du système dans le temps.

On soupçonne cependant intuitivement qu'il existe encore d'autres formes de traitements pertinentes pour l'utilisation d'informations localisées dans des domaines spécialisés: mises à jour de la mémoire, représentation et visualisation des informations, calage de modèles numériques pour n'en citer que quelques-uns. Pour permettre une mise en oeuvre efficace de systèmes de traitement de l'information localisée dans des domaines spécialisés, nous devons être à même d'identifier et de décrire ces traitements.

Ce chapitre a donc pour objectif de présenter une synthèse des problèmes liés aux traitements de données localisées et de définir les voies de solutions nécessaires.

5.1. Systèmes de prise de décision: de l'agrégation des informations à l'intégration des décisions

Nous partons du paradigme du système général, pertinent dès lors que nous définissons un cadre non seulement structuré, mais clairement finalisé [Le Moigne 1977]. Nous en montrerons cependant les limites pour notre approche, et adopterons un modèle systémique mieux adapté aux objectifs de cette étude.

A son niveau le plus achevé, Le Moigne décrit le système général comme "une intervention finalisante dans un environnement". Les processeurs du système, agents de cette intervention, peuvent être rattachés à l'un des trois sous-systèmes suivants:

- Le *sous-système opérant* a pour mission de transformer des flux intrants de matière, d'énergie ou d'information en flux extrants, conformément aux décisions du système de pilotage.
- Le *sous-système de pilotage* dirige et contrôle le sous-système opérant. Il est lui-même constitué d'un sous-système de décision-sélection chargé d'agrèger les décisions, d'un sous-système d'imagination-conception porteur de l'intelligence du système, et d'un sous-système de finalisation chargé de déterminer les objectifs du système sur la base des problèmes perçus de la réalité.
- Le *sous-système d'information* met à disposition du système de pilotage les informations qui lui permettent de fonder ses décisions.

Qu'il s'agisse de construire un nouvel égout, d'entretenir une route ou de procéder à des analyses géostatistiques, on peut considérer chaque domaine spécialisé comme un système isomorphe du système général. Ces domaines spécialisés traduisent bien la diversité des préoccupations des nombreux intervenants sur le territoire. A ces différents objectifs correspondent en particulier des extrants différents du système, et une nature et un fonctionnement différents du système opérant.

Tous ces intervenants partagent cependant un objectif commun: prendre les "meilleures" décisions possibles. La gestion d'informations décrivant la réalité et son comportement ne constitue ainsi pas un but en soi: elle permet d'étayer les décisions prises, de raccourcir les délais de décision. Nous allons donc circonscrire notre réflexion à un *système de prise de*

décisions³³. La figure 5.1 en est une représentation conforme à la symbolique adoptée par [Le Moigne]. On y distingue formellement les *informations-représentation* des *informations-décision*. Selon Le Moigne en effet, une *décision* est "un objet ayant la même apparence qu'une information, le même support (des signaux), doté d'une propriété exclusive complémentaire: il est présumé devoir provoquer une action prédéfinie, une modification connue du comportement du ou des processeurs qui le reçoivent en le différenciant parmi leurs intrants".

Il est alors nécessaire de distinguer les *décisions* produites par le sous-système opérant, destinées à l'environnement, des *méta-décisions* (décisions sur les décisions) produites par le sous-système de pilotage, permettant la régulation du système. Par exemple, l'octroi d'une autorisation de construire constitue une décision administrative à l'intention d'un requérant externe au système de décision, alors que la convocation d'une séance destinée à entériner cet octroi constitue une méta-décision.

En outre, les objectifs de chaque prise de décision sont déterminés par les objectifs propres à chaque domaine d'activité. Ces objectifs sont donc modélisés comme des intrants du sous-système de pilotage.

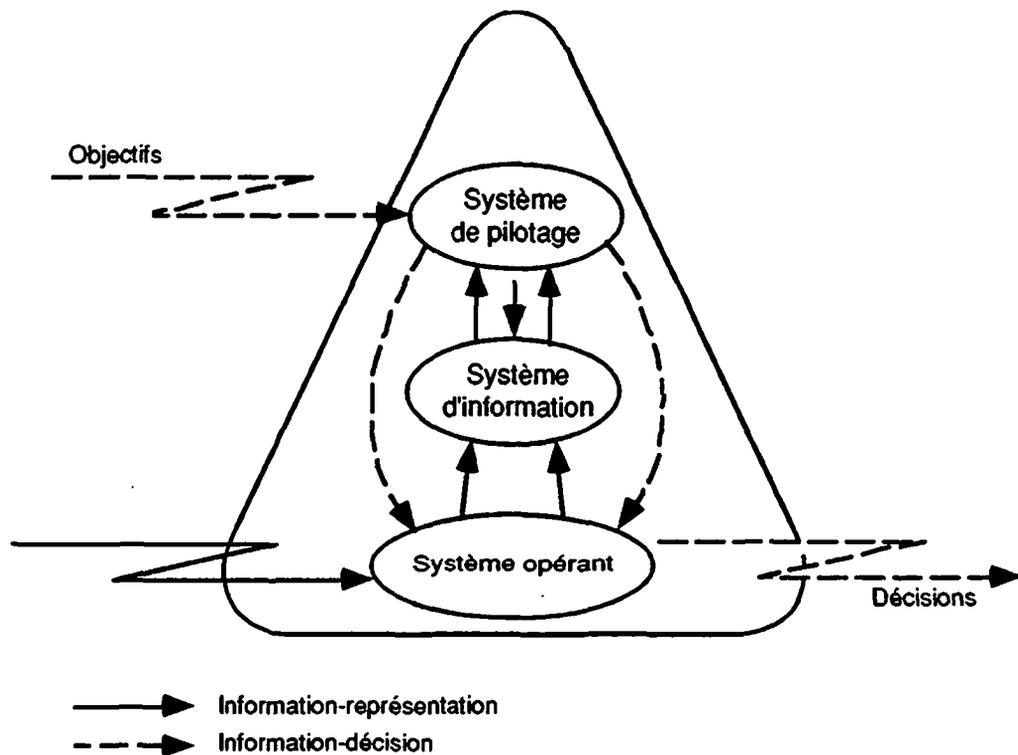


Figure 5.1: **Système de prise de décision (dérivé de [Le Moigne])**

L'approche par le système général nous a permis de définir le système de prise de décisions, d'en situer les objectifs et d'en décrire les intrants et les extrants. L'ensemble de la problématique de la prise de décisions dépasse cependant le cadre de cette étude, et aucun des trois sous-systèmes de [Le Moigne] n'en constitue une délimitation satisfaisante. Peut-on alors définir un modèle du système de prise de décision plus pertinent pour délimiter le champ de notre étude ? Un paradigme auquel nous nous référerons encore fréquemment est celui de la pyramide de décision ([Davis et Olson 1985], appliqué en particulier dans [Chevallier et Bédard 1990] aux systèmes d'information à référence spatiale). La figure 5.2 propose un modèle du système de prise de décision constitué d'un ensemble de processeurs

33 Nous ne tiendrons pas encore compte dans ce chapitre de la répartition des rôles entre l'homme et l'ordinateur. C'est pourquoi nous considérons bien un système de prise de décisions, et non pas seulement un système (informatique) d'aide à la décision.

répartis entre les trois niveaux (stratégique, tactique et opérationnel) de la pyramide de décision ³⁴. Les flux entre le système et son environnement, ainsi qu'entre les différents niveaux de décision sont représentés en gras.

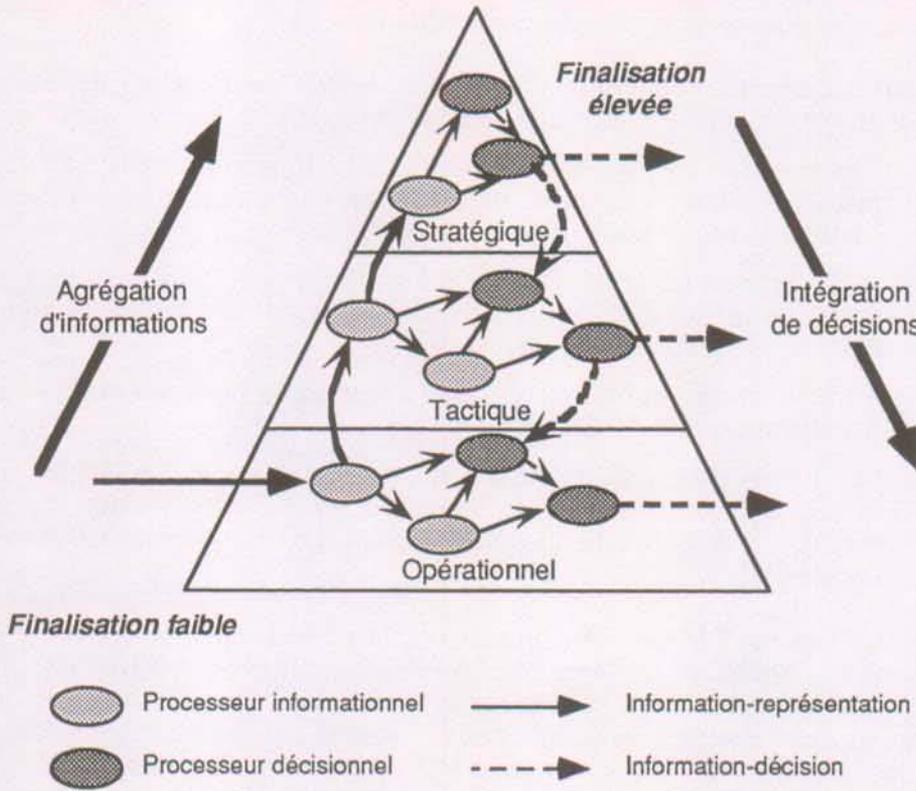


Figure 5.2: **Modèle du système de prise de décision adopté pour notre étude**

Nous distinguons en premier lieu les *processeurs décisionnels*, produisant au moins un flux d'informations-décision, des *processeurs informationnels*, ne produisant que des flux d'informations-représentation. Nous serons amenés dans cette étude à nous intéresser tout particulièrement aux processeurs informationnels et aux flux d'information-représentation qui leur sont liés. Les processeurs décisionnels sont déjà étudiés dans le cadre de recherches parallèles, consacrées à l'utilisation des SIRS dans des domaines spécialisés (voir en particulier [Leclerc et Chevallier 1986]) et plus généralement aux systèmes d'aide à la décision à référence spatiale (voir en particulier [Chevallier 1992]).

Les niveaux supérieurs de la pyramide sont aussi souvent les plus finalisés. Pour étayer les décisions qui y sont prises, des informations de synthèse sont agrégées à partir des niveaux inférieurs. On a donc représenté, à gauche de la pyramide, des processeurs informationnels chargés d'agrégier l'information pour la rendre accessible au niveau de décision supérieur. Ces processeurs constituent le *sous-système d'agrégation des informations*.

À droite de la pyramide, on a représenté des processeurs décisionnels. Les décisions de niveau supérieur peuvent s'appliquer à l'environnement, et déterminer le cadre dans lequel les décisions de niveau inférieur seront prises. On constate donc un enchaînement des processeurs décisionnels du haut vers le bas de la pyramide, permettant l'intégration cohérente des décisions du niveau stratégique au niveau tactique. Ces processeurs constituent le *sous-système d'intégration des décisions*.

À l'intérieur de chaque niveau, on constate une direction générale des flux d'informations et de décisions de la gauche vers la droite de la pyramide. Elle traduit des changements de

³⁴ La représentation par [Le Moigne] du système général par un triangle au coins arrondis est sans rapport avec la pyramide de [Davis et Olson].

représentation des informations et des décisions pour les rendre conformes à la nature des décisions à prendre. Les échanges d'informations-représentation avec l'environnement se font de manière préférentielle au niveau opérationnel: l'information y est la plus "neutre", la moins finalisée; l'utilisation des données cadastrales à grande échelle comme bases des premiers "cadastres polyvalents" en est un exemple typique.

Les quatre phases du processus de décision de H. A. Simon, reprises par [Lévine et Pomerol 1989], peuvent être aisément décrites sur la base de ce modèle:

1. l'information ou le renseignement (*intelligence*) consiste à alimenter la mémoire du sous-système d'agrégation des informations en informations pertinentes pour le domaine considéré;
2. la conception (*design*) consiste à élaborer des scénarios en agrégeant et en enrichissant les informations mémorisées, en étroite interaction entre les deux sous-systèmes considérés;
3. le choix (*choice*) entre les différents scénarios conçus relève du sous-système d'intégration des décisions;
4. l'évaluation des choix précédents (*review*) exige à nouveau une étroite interaction entre les deux sous-systèmes; elle permet d'améliorer la connaissance générale du système (informations disponibles et expertise du décideur).

Une illustration de ce modèle nous est fournie par la procédure d'octroi d'autorisations de construire dans le canton de Genève. Une demande d'autorisation déposée en zone de développement est tributaire de l'élaboration préalable d'un plan localisé de quartier (niveau tactique). A cet effet, les informations nécessaires sur la situation existante, sur les contraintes du voisinage, etc. sont agrégées, et les décisions du niveau stratégique (plan directeur, plan des zones, ...) intégrées. Le plan localisé de quartier est alors adopté (décision de niveau tactique), puis l'autorité exécutive statue sur l'octroi du permis de construire (décision de niveau opérationnel).

Ce modèle est représenté de manière simplifiée à la figure 5.3. Il nous permet finalement de décrire simplement l'objet de notre étude. Il s'agit du *sous-système d'agrégation des informations du système de prise de décision*. Ce sous-système est piloté par le sous-système d'intégration des décisions. De la qualité de leur interaction dépend la capacité du système à traiter des questions du type "what if".

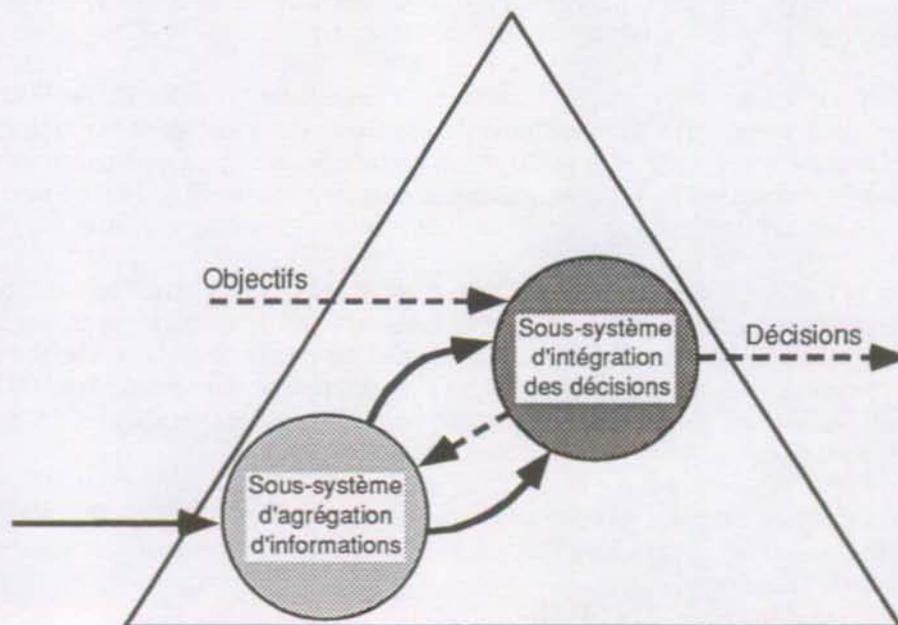


Figure 5.3: Modèle simplifié du système de prise de décision

Nous convenons cependant d'une restriction au cas le plus général du sous-système d'agrégation d'informations. Notre environnement se révèle changeant, ou tout au moins la connaissance que nous en avons. Cette évolution de la connaissance se traduit par la création de nouvelles méthodes d'agrégation des informations. L'organisation habituelle des systèmes de prise de décision nous autorise cependant à ne pas prendre en compte cette dynamique créatrice dans notre étude. En effet, les décideurs ne créent en principe pas eux-mêmes de nouvelles méthodes, mais choisissent une méthode parmi celles qui ont été développées dans un autre temps par d'autres intervenants (instituts ou départements de recherche, universités, etc.). Nous pouvons donc accepter de recourir à nouveau au système de représentation pour intégrer de nouvelles méthodes au système de prise de décision.

En outre, le système de prise de décision est ouvert à de nombreux domaines d'application (aménagement du territoire, construction d'infrastructures, ...). Mais c'est la finalité unique de la prise de décisions qui est pertinente pour notre étude. Nous considérerons donc que le système d'agrégation d'informations, à finalité permanente dans un environnement permanent, est soumis à un équilibre homéostatique [Le Moigne 1977].

5.2. Typologies des traitements d'un système d'agrégation d'informations

Le sous-système d'agrégation d'informations a pour objectif de fournir au sous-système d'intégration des décisions des images pertinentes de la réalité actuelle, ainsi que de son évolution vraisemblable sous différentes hypothèses.

Le rôle et la nature des modèles à mettre en oeuvre pour apporter aux décideurs la connaissance nécessaire à étayer leurs décisions, ainsi que les mécanismes d'abstraction qui leur sont liés, ont été étudiés au chapitre 4. Il s'agit dans la suite de ce chapitre d'inventorier et de classer les traitements d'un système d'agrégation d'informations localisées et d'en suggérer des solutions de mise en oeuvre. Nous définissons un *traitement* comme l'action d'un processeur.

Le cadre de cette classification nous est donné par la figure 5.4, déjà proposée au chapitre 4. Cette figure montre la trajectoire d'un système dans le temps. Ce système peut être aussi bien la réalité elle-même qu'un modèle pertinent pour un décideur. Nous proposons de répartir l'ensemble des traitements survenant dans un système d'agrégation d'informations localisées dans les trois classes suivantes:

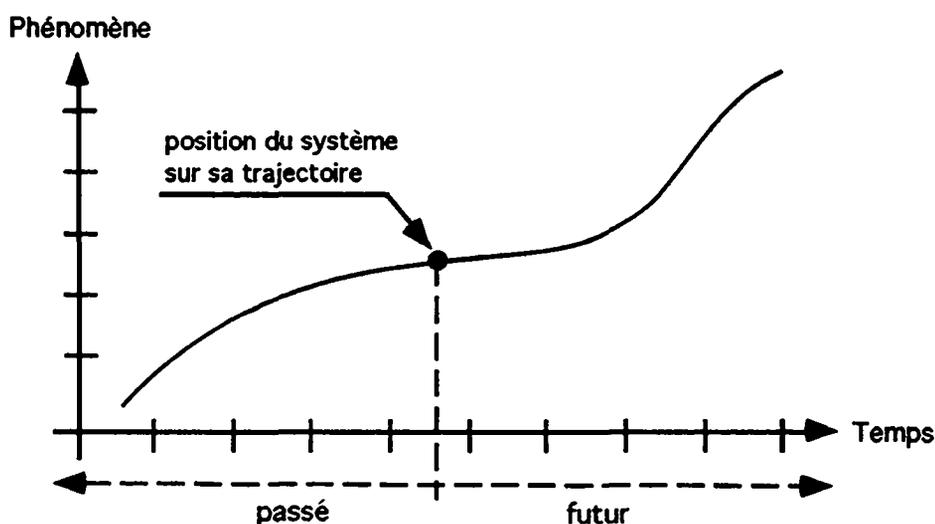


Figure 5.4: Trajectoire d'un système

1. Déplacement du modèle sur sa trajectoire

Les traitements coucourant au déplacement d'un modèle sur sa trajectoire reflètent l'évolution de la réalité. Ces traitements réalisent les transformations propres aux modèles de comportement diachroniques (voir chapitre 4).

2. Changement d'échelle de mesure des phénomènes

Un phénomène de la réalité peut être mesuré selon différentes échelles. Ces échelles sont étroitement dépendantes de la *structure* générale des données et des modèles d'information utilisés. Ainsi, la position et la forme géométriques d'un bâtiment devront être mesurées pour l'intégrer au cadastre suisse, alors que la mesure de sa superficie et la définition de la parcelle dont il dépend suffiront à son inscription au registre foncier.

De nombreux traitements ont pour objet un ou des changements d'échelle de mesure, appelés aussi *changements de représentation* dans le domaine des systèmes experts [Lévine et Pomerol 1989]. Ils sont le reflet des modèles de comportement synchroniques définis au chapitre 4. On peut distinguer différents cas particuliers de changement de représentation:

a. les changements de *niveau d'abstraction*

Ce sont des cas particuliers de changement de représentation, pour lesquels les mécanismes d'abstraction présentés au chapitre 4 sont utilisés pour passer à un niveau d'abstraction de l'information plus ou moins élevé de la pyramide de décision (stratégique, tactique ou opérationnel). Le niveau d'abstraction est caractérisé en particulier par la résolution de l'échelle de mesure des phénomènes et par l'échelle de modélisation des objets à référence spatiale [Chevallier et Bédard 1990].

Des informations de niveau opérationnel seront généralement mesurées sur des échelles détaillées; la forme et l'étendue précise des objets à référence spatiale seront généralement connues (grande échelle). A l'opposé, des informations de niveau stratégique seront généralement mesurées sur des échelles nominales ou ordinales discrètes; seule leur éventuelle position dans l'espace sera modélisée (petite échelle).

b. les changements du *type de modèle d'information*

Les informations peuvent être représentées par différents types de modèles d'information. Les phénomènes spatiaux peuvent en particulier être représentés par des objets spatiaux, des carroyages, etc. (voir chapitre 4.2). La substance informationnelle véhiculée peut varier considérablement d'un type de modèle à un autre.

c. les changements de *référentiel spatial*

Un même modèle spatial peut être mesuré dans différents référentiels spatiaux (voir § 4.3.1). Le passage d'un référentiel à l'autre peut être nécessaire pour la mémorisation à long terme d'un modèle, pour sa combinaison avec d'autres modèles ou pour les besoins de son exploitation. La représentation cartographique d'une base de données routières exige par exemple le changement d'un référentiel curviligne en un référentiel orthogonal.

Ces traitements peuvent être simples et algorithmiques, comme le calcul de la superficie d'une parcelle si l'on en connaît la forme et l'étendue. D'autres traitements peuvent se révéler beaucoup plus complexes, et exiger un apport d'expertise, voire de nouvelles observations de la réalité; un exemple en est la détermination de sous-bassins versants en hydrologie, qui fait appel à l'expertise d'un hydrologue d'une part, et à une connaissance précise de la topographie.

3. Changement d'échelle de mesure du temps

Notre calendrier constitue une échelle de mesure du temps que l'on peut considérer comme absolue dans le contexte d'un système de prise de décision. Il existe néanmoins différentes définitions du temps: sommes-nous intéressés au moment où un phénomène naît ou meurt dans la réalité ? au moment où ce phénomène est saisi ou supprimé d'un modèle ? Dans le cadre de cette étude, nous distinguerons deux concepts distincts du temps, inspirés des propositions de [Bilodeau, Caron et al. 1990]:

- le temps effectif, dans lequel nous mesurons la naissance et la mort des phénomènes dans la réalité;
- le temps du modèle, dans lequel nous mesurons l'entrée et la sortie du phénomène dans le modèle considéré.

A chacun de ces concepts, nous attachons sa propre échelle de mesure du temps. Ainsi:

- la mise à jour d'un modèle implique l'entrée ou la sortie de phénomènes déjà nés ou morts dans la réalité;
- la réalisation ou la mise en vigueur d'objets projetés correspond à la naissance dans la réalité d'objets déjà entrés dans un modèle.

Chacun de ces traitements implique donc un changement de l'échelle de mesure du temps: il conduit à mesurer sur l'une des échelles du temps un phénomène déjà mesuré dans une autre échelle.

5.3. Classification de quelques fonctions essentielles des systèmes d'agrégation d'informations

Les systèmes d'agrégation d'information localisées font appel à des fonctions telles que *levé*, *restitution*, *simulation*, *paramétrisation*, *prévision*. Dans ce paragraphe, ces fonctions sont décrites conformément à la typologie des traitements définie au paragraphe précédent.

1. Levé

Un *levé* consiste en la saisie "sur le terrain" des informations nécessaires à la modélisation des objets spatiaux. Le terme lui-même dévoile des racines dans les sciences géomatiques. Selon notre classification, le levé est une méthode d'*observation* de la réalité, destinée à élaborer ou actualiser un modèle. Il s'agit donc d'un changement d'échelle du temps, assorti d'un changement de niveau d'abstraction.

Si l'élaboration initiale d'un modèle est un cas simple d'observation et de modélisation, son actualisation ou *mise à jour* est une démarche plus complexe. Elle réalise le déplacement d'un modèle sur sa trajectoire, sur la base d'un programme d'observations de la réalité. Ce programme est généralement aisé à établir pour des phénomènes discrets (il consiste en des réactions aux mutations de la réalité); il exige par contre la définition de déclencheurs adéquats dans le cas des phénomènes continus: période de temps (mise à jour périodique), constats d'insuffisance, etc.

2. Restitution

Une *restitution* consiste en l'élaboration d'un produit (écran d'ordinateur, dessin automatique, fichier d'interface, croquis autographe, etc.) destiné à *communiquer* tout ou partie d'un modèle vers l'environnement du système. Le produit restitué sera interprété par son destinataire sous la forme d'un nouveau modèle, physique ou cognitif. Une restitution vise donc un changement d'échelle de mesure du temps, éventuellement assorti d'un changement du niveau d'abstraction.

Une restitution peut véhiculer aussi bien des informations-décision que des informations-représentation. Dans le cas des systèmes d'agrégation d'informations, une restitution fournit au système d'intégration des décisions les informations-représentation nécessaires à la prise de décisions.

3. Simulation

Selon [Le Moigne 1977], une *simulation* consiste à faire varier les intrants d'un système dont on connaît l'équation d'état pour en évaluer les extrants. Une simulation met donc en oeuvre un modèle de comportement, représentant soit le déplacement du modèle sur sa trajectoire, soit un changement d'échelle de mesure des phénomènes. La notion de variation des intrants est très importante dans cette définition, car elle implique une double incertitude:

- On distingue une première incertitude sur les *intrants*, qui peuvent n'être que des hypothèses, ou une fourchette de valeurs vraisemblables. Souvent, on recherche de manière heuristique les intrants correspondant au "plus mauvais cas possible": on peut citer de nombreux exemples, comme le choix de la pluie critique pour une simulation de débits de crue ou les cas de charge dans un calcul de déformation de structures en génie civil.
- Une seconde incertitude est liée à la sensibilité du comportement simulé: quels sont les conséquences d'une erreur d'estimation des intrants sur le résultat de la simulation ? L'application répétée du processus simulé sur une fourchette de valeurs des intrants est souvent nécessaire pour évaluer ces conséquences.

Une simulation constitue donc un ensemble complexe de traitements élémentaires, dont les résultats doivent être comparés et éventuellement agrégés; les intrants et extrants de chaque variante considérée doivent donc être conservés en mémoire jusqu'à la fin de la simulation.

4. Paramétrisation de modèles de comportement

Une simulation est fondée sur l'hypothèse que l'équation d'état du modèle est connue, en d'autres termes que d'un état quelconque du système à un temps t_0 , on peut compléter cet état (modèle synchronique), ou déduire l'état du système à un temps t quelconque (modèle diachronique).

De nombreux modèles de comportement complexes ne sont cependant constitués que d'un réseau de processus, qui doivent être paramétrés en fonction des cas d'espèce rencontrés. Cette paramétrisation nécessite la connaissance préalable des intrants et des extrants du système pour un nombre de cas dépendant de la sensibilité du modèle et de la résolution attendue. Les paramètres ainsi évalués seront à leur tour utilisés pour évaluer des extrants inconnus.

La paramétrisation est donc une fonction de régulation du modèle. Dans le cas de modèles diachroniques, elle consiste à établir le programme du système sur la base d'une trajectoire connue.

Un exemple en est offert par la paramétrisation d'un modèle de simulation des débits de crue "temps réel" à l'aide d'un modèle classique complexe, telle qu'elle est réalisée entre autres à la Communauté urbaine de Québec.

5. Prévision

La *prévision* consiste à évaluer un état futur d'un système en connaissant l'état présent et des règles d'évolution. Il s'agit d'un cas particulier de simulation, strictement diachronique, dont les intrants sont connus.

Une prévision est donc un déplacement du modèle sur sa trajectoire, sur la base d'un modèle diachronique du comportement.

5.4. Modélisation des traitements dans les systèmes d'agrégation d'informations localisées

La définition d'un système repose en principe sur une double spécification: celle de sa structure organique (processeurs et flux) et celle de son programme (enchaînement d'états au cours du temps). J. Mélése, cité par [Le Moigne 1977], propose ainsi d'appeler O.graphe les représentations graphiques de la structure organique d'un système, et P.graphe les représentations graphiques de son programme.

Les formalismes aptes à représenter le programme d'un système sont rares; sans prétendre à l'exhaustivité, on peut citer les traditionnels organigrammes utilisés depuis longtemps en programmation, les réseaux de Petri utilisés pour la description des traitements dans la méthode MERISE, ainsi que les diagrammes PERT utilisés en gestion de projets³⁵. Ces formalismes n'ont toutefois connu de succès que dans des domaines bien délimités. Cette lacune dans la modélisation des programmes n'est toutefois pas l'apanage des systèmes à référence spatiale, et nous ne chercherons donc pas à la combler dans le cadre de cette étude. Nous nous contenterons d'en constater les conséquences fâcheuses sur la conception de nombreux systèmes actuels: prise en compte insuffisante de la composante temporelle, systèmes d'information développés sans concept de mise à jour, etc.

Par contre, de nombreuses recherches ont conduit à la définition de formalismes permettant de représenter graphiquement la structure organique d'un système. Certains de ces formalismes sont plus spécifiquement destinés à la représentation des processeurs et de leur action (traitements), alors que d'autres sont plus spécifiquement destinés à la représentation des flux et de leurs interdépendances (structures de données). Les modèles décrits à l'aide de ces formalismes n'en sont pas moins des représentations complémentaires, des "vues externes" d'un seul et même modèle: la structure organique du système considéré.

Nous constatons cependant que les recherches menées jusqu'à ce jour dans le domaine des systèmes d'information à référence spatiale se sont plus préoccupées de modélisation des structures de données que de modélisation des traitements [Bédard et Paquette 1990; Bédard et Prince 1990; Paquette 1990; Caron 1991]. Nous voyons deux raisons à une telle anamorphose:

- a. Les systèmes d'information à référence spatiale jouent, au sein d'une organisation donnée, un rôle institutionnel de mémoire permanente des informations relatives au territoire³⁶. Dans de tels systèmes, la définition des structures de données joue donc un rôle prioritaire, alors que les traitements sont relativement simples et typés (saisie, mise à jour et restitution des informations).
- b. Au niveau conceptuel le plus général, les traitements mis en oeuvre dans les systèmes à référence spatiale ne se distinguent a priori guère de ceux des autres systèmes (gérer les données, restituer, simuler, etc.). Ces traitements sont bien plutôt le fait de tout système de prise de décisions. C'est dans la nature des informations traitées qu'intervient en premier lieu la référence spatiale (voir chapitre 6).

Nous discuterons dans la suite de ce paragraphe la modélisation des traitements dans un système d'agrégation d'informations localisées, alors que nous consacrerons le chapitre 6 à la modélisation des structures de données.

³⁵ Dans le domaine des SIRS, [Caron 1991] propose de définir le déclencheur de chaque traitement dans le dictionnaire du système. C'est une ébauche de modélisation du programme, qui peut se révéler suffisante pour les SIRS dans lesquels les traitements sont relativement indépendants les uns des autres.

³⁶ Ces aspects organisationnels seront discutés en détail au chapitre 7.

Contrairement aux systèmes d'information à référence spatiale, ambivalents [Chevallier 1983], les systèmes d'agrégation d'information sont clairement finalisés. Des données y sont gérées en perspective des traitements qui les exploitent. Les traitements jouent par conséquent un rôle prépondérant pour la modélisation de la structure organique; les données ne seront modélisées qu'en termes d'intrants et d'extrants de ces traitements.

Au niveau conceptuel, nous avons déjà constaté que les traitements à référence spatiale ne se distinguent a priori guère des autres. Ils pourront donc être modélisés sur la base d'un formalisme classique, tel que les diagrammes de flux de données (DFD) [Gane et Sarson 1979]. Ceux-ci permettent une approche top-down de la modélisation des traitements, du concept général à la spécification informatique³⁷. Ils sont utilisés dans de nombreuses recherches récentes ou en cours [Caron 1991]. Mais tout autre formalisme dont les composantes sémantiques sont des traitements associés à des flux intrants et extrants peut se révéler pertinent; [Champoux 1991] a par exemple étudié les fonctions d'*analyse spatiale*, importantes pour les systèmes d'agrégation d'informations localisées, et a proposé le formalisme DENOT pour les modéliser.

Nous ne suggérons pas de formalisme spécifique pour la modélisation des traitements dans les systèmes d'agrégation d'informations localisées. Toutefois, nous proposons quelques règles de modélisation applicables quel que soit le formalisme employé:

1. Dès les premières étapes de la modélisation, une distinction claire entre les trois types de traitements définis au § 5.3 doit être obtenue:
 - déplacement sur la trajectoire
 - changement d'échelle de mesure des phénomènes
 - changement d'échelle de mesure du temps

Cette règle est la conséquence de la classification fondamentale des traitements proposée dans cette étude.

2. Lorsqu'un état peut être obtenu par des méthodes différentes (observation de la réalité, différents modèles du comportement, ...), les données de cet état seront assorties de méta-informations spécifiant les traitements qui ont conduit à leur création.

Les données d'un état peuvent être obtenues par différentes méthodes. Chaque donnée dont la source pourrait se révéler ambiguë doit être qualifiée par des méta-informations, de telle sorte que le sous-système d'agrégation des décisions puisse fonder ses choix.

Ces méta-informations peuvent être limitées à la désignation de la méthode d'obtention de la donnée (photogrammétrie, calcul à l'aide du modèle URBAN, etc.), ou être accompagnées des paramètres nécessaires à renouveler la donnée (historisation de tous les paramètres d'un calcul par exemple).

³⁷ On restera cependant attentif à ce qu'il s'agit à l'origine d'un outil de génie logiciel destiné à concevoir des "programmes", et non d'un outil de conception de systèmes, destiné à analyser et à planifier les activités d'une organisation: leur utilisation sans un discernement et une expertise suffisante peuvent se révéler un exercice de style sans grande utilité pour l'objectif poursuivi !

3. Les distinctions suivantes sont réservées au niveau organisationnel:

- distinction entre les traitements réalisés par l'homme et les traitements réalisés par la machine;
- distinction entre traitements spatiaux ³⁸ et traitements non spatiaux, nécessaire au choix d'un SGDL susceptible d'exécuter le traitement désiré.

La modélisation d'un système de prise de décision est généralement entreprise dans la perspective de l'informatisation de certaines tâches. La tentation est alors grande de réduire le champ de l'étude au domaine a priori destiné à être automatisé. Le respect des deux principes proposés permet d'éviter des choix organisationnels prématurés.

On trouvera au chapitre 8 un exemple de modèle de traitements sous forme de diagramme de flux de données, réalisé conformément aux principes proposés ci-dessus.

**5.5. Approche du niveau organisationnel:
les systèmes d'aide à la décision et les systèmes experts**

Dans ce chapitre, nous avons introduit la notion de *système de prise de décision*, et non pas seulement celle de *système d'aide à la décision*. En d'autres termes, nous n'avons pas encore distingué le rôle dévolu à l'homme, garant de l'expertise du système, et le rôle dévolu à la machine, susceptible de réaliser rapidement les agrégations d'information requises par l'expert. Cette distinction est la pierre angulaire de la phase organisationnelle de la conception d'un système.

Alors qu'il n'y a guère plus qu'une décennie, l'ordinateur n'était qu'un habile manipulateur de données numériques, l'avènement des *systèmes experts* laisse entrevoir un champ d'action beaucoup plus vaste à la machine. Mais c'est sans doute la formalisation de la connaissance des experts qui constitue aujourd'hui le principal obstacle à la mise en oeuvre de systèmes experts. Dans le cadre des systèmes de prise de décision, une bonne expertise du domaine spécialisé considéré est nécessaire pour réaliser des agrégations complexes d'informations, et pour décider le choix d'un scénario parmi d'autres.

On peut estimer que le rôle des systèmes d'aide à la décision, et par conséquent celui des systèmes experts, va continuer à s'accroître dans l'avenir, grâce au développement de l'informatique et des méthodes de formalisation de la connaissance; dans le domaine de l'ingénierie, on peut citer les travaux conduits par Claude Marche dans le cadre du projet CASTOR [Robert et Marche 1989]. Dans cette attente, l'informatique peut d'ores et déjà apporter un soutien considérable à la prise de décisions. Elle peut en effet assumer les tâches répétitives et astreignantes aisément automatisables, laissant aux spécialistes du domaine considéré le soin d'appliquer leur expertise.

C'est à nos yeux l'objectif qui doit être actuellement dévolu aux systèmes d'aide à la décision. Il ne s'agit pas tant de vouloir faire résoudre à la machine des problèmes relevant de la seule compétence d'un expert, que de présenter à l'expert sous la forme la plus aisément interprétable la connaissance nécessaire à la résolution d'un problème.

³⁸ On entend par *traitement spatial* un traitement faisant intervenir la position, la forme ou l'étendue d'un ou de plusieurs objets spatiaux.

Plusieurs travaux de recherche corroborent ce postulat. Lévine [Lévine et Pomerol 1989] introduit le concept de *système interactif d'aide à la décision (SIAD)*, dans lequel l'interactivité entre le système informatique et l'expert joue un rôle déterminant. [Chevallier 1992] préconise l'utilisation de *systèmes d'aide à la décision à référence spatiale (SADRS)* pour présenter la connaissance sous une forme aisément interprétable, en s'appuyant par exemple sur les concepts et la technologie *multi-médias*. Et de manière percutante, Lord Chorley rappelait, dans son allocution d'ouverture au congrès URISA 1989, que *la sagesse est le fruit d'une bonne intégration de la connaissance*.

Dans le cadre de la présente étude, [Buogo 1990] a réalisé une maquette informatique d'un *système de simulation pour le drainage urbain*. L'ordinateur y prend en charge la plupart des tâches d'agrégation d'informations nécessaires, sous le pilotage d'un expert auquel revient la responsabilité des principales décisions. L'importance d'une interface flexible et performante entre le système informatique et l'expert a été clairement mise en évidence.

5.6. Conclusion du chapitre

Dans ce chapitre, nous avons proposé de découper un système de prise de décision en deux sous-systèmes: le sous-système d'agrégation des informations et le sous-système d'intégration des décisions. Ce découpage se révèle en effet particulièrement fécond et structurant pour décrire des systèmes de prise de décision ³⁹. Il se révèle en outre étroitement cohérent avec le paradigme de la pyramide de prise de décision.

Dans ce cadre, les traitements d'informations consistent essentiellement

1. à représenter le déplacement du système sur sa trajectoire;
2. à changer l'échelle de mesure des phénomènes considérés, en particulier pour changer le niveau d'abstraction du modèle;
3. à transférer des informations d'un modèle à un autre, c'est à dire à changer l'échelle de mesure du temps, étroitement associée à chaque modèle.

Sur cette base, nous avons suggéré un ensemble de règles pour la modélisation conceptuelle et organisationnelle des traitements.

Nous avons enfin constaté que les outils informatiques doivent tendre à décharger les experts de travaux de routine pour leur permettre de se concentrer sur les décisions à prendre. Mais un système informatique n'est pas apte à remplacer un expert.

On trouvera au chapitre 8 plusieurs exemples d'utilisation du cadre systémique proposé.

³⁹ Nous avons en particulier constaté qu'un système d'agrégation d'informations ne constitue pas un système d'information au sens de [Le Moigne], mais un système opérant produisant des informations.

6. Problèmes et solutions relatifs aux données

Au chapitre précédent, on a rappelé que la conception d'un système implique en particulier la modélisation de sa structure organique. Cette structure inclut aussi bien les processeurs que les flux d'informations reliant ces processeurs.

Dans ce même chapitre, les problèmes et solutions relatifs à la modélisation des traitements, actions des processeurs, ont été discutés. Il nous reste donc à discuter les problèmes et solutions relatifs à la modélisation des flux d'informations circulant entre les processeurs du système. Ces flux d'informations sont représentés physiquement par les données traitées par le système [Chevallier 1983].

Nous avons cependant constaté au chapitre 4 que les données d'un système ne peuvent généralement être considérées comme stables qu'à des niveaux d'abstraction supérieurs, qui constituent le *méta-modèle des données* du système. Ce sont de tels méta-modèles que nous devons être à même de décrire.

Ce chapitre a pour objectifs de combler certaines lacunes importantes des formalismes actuellement utilisés pour la modélisation des systèmes traitant des données localisées, et d'en définir les règles essentielles de mise en oeuvre.

6.1. Les formalismes géosémantiques

Les formalismes utilisés jusqu'à ce jour pour la modélisation des données dans les systèmes d'information à référence spatiale ont généralement reçu des dénominations faisant référence à leur origine (en général, le formalisme entité-relation ou l'un de ses dérivés); Paquette [Paquette 1990] parle par exemple d'un "formalisme entité-relation étendu", à l'instar de plusieurs chercheurs dans le domaine des bases de données classiques. Caron [Caron 1991] intègre au formalisme entité-relation plusieurs nouveaux "modules" (généralisations, agrégations, gestion du temps, etc.), ce qui l'a conduit à dénommer son formalisme "MODUL-R".

Mais tous ces formalismes poursuivent un objectif commun: faciliter la modélisation conceptuelle des systèmes d'information à référence spatiale. Ils obéissent à des cahiers des charges similaires. Il serait donc à nos yeux souhaitable que les efforts conjugués de différentes universités pour définir un formalisme adapté à la modélisation conceptuelle des systèmes à référence spatiale se reflètent dans une dénomination commune unique. Nous proposons de donner le nom de *formalismes géosémantiques* à la famille de formalismes destinés à la modélisation conceptuelle de systèmes d'information localisées. Les propositions formulées dans ce chapitre ne constituent pas un formalisme à elles seules: elles se veulent plutôt des contributions susceptibles de compléter les formalismes géosémantiques actuels.

6.2. Cahier des charges des formalismes géosémantiques

Les problèmes de la modélisation de systèmes traitant des informations localisées ont été abordés au chapitre 4. Sur cette base, nous proposons un *cahier des charges* des formalismes géosémantiques, sous la forme des postulats suivants:

- a. Les formalismes géosémantiques sont des formalismes à vocation conceptuelle. Ils doivent garantir une modélisation de la réalité aussi peu biaisée que possible, sans intégrer de contraintes propres à l'implémentation des modèles dans quelque système informatique que ce soit.

- b. Un méta-modèle physique (tel que décrit à l'aide d'un formalisme géosémantique) doit être un reflet univoque du méta-modèle cognitif de son concepteur, et il doit réciproquement être interprétable en un modèle cognitif univoque [Kuhn 1989].
- c. "On pense le monde en termes de faits sur des objets" [Barr et Feigenbaum 1981]. Un formalisme géosémantique doit permettre l'application aisée de cet axiome, et doit en particulier favoriser la discrétisation du territoire en *objets spatiaux*.
- d. Les objets spatiaux peuvent appartenir à des niveaux d'abstraction très hétérogènes. Les mécanismes d'abstraction (généralisation et spécialisation, agrégation et décomposition) jouent donc un rôle déterminant, qu'il est important de prendre en compte dans les méta-modèles.
- e. La référence spatiale permet d'une part d'accéder aux informations par l'identification d'objets spatiaux sur une carte ou un écran, et d'autre part de déduire d'innombrables relations spatiales entre ces objets. Un formalisme géosémantique doit permettre la modélisation explicite de ces besoins en informations spatiales. Il doit en particulier intégrer sous une forme adéquate la notion de *référentiels spatiaux* multiples présentés au chapitre 4.
- f. Le temps joue un rôle considérable dans les systèmes traitant des informations localisées (voir § 4.1). La mémorisation par le système de plusieurs états échelonnés dans le temps est indispensable pour extrapoler ou interpoler de nouveaux états [Leclerc, Golay et al. 1989] [Richer et Chevallier 1992]. Un formalisme géosémantique doit donc être à même de gérer plusieurs états successifs des mêmes objets.

6.3. Apports et lacunes des formalismes actuels

Une analyse critique exhaustive des qualités et des défauts des principaux formalismes et de leurs dérivés actuellement utilisés pour la conception et l'implémentation des systèmes traitant des informations localisées dépasse les objectifs de cette étude. Ces formalismes ont été succinctement présentés au chapitre 2. Qu'il nous soit simplement permis d'en inventorier ci-dessous quelques lacunes en regard du cahier des charges établi au § 6.1.

Les formalismes géosémantiques actuels sont basés sur le formalisme *individuel* [Tardieu, Rochfeld et al. 1984], de la famille des formalismes *entité/relation* [Chen 1976]. Conçus initialement pour faciliter l'implémentation de bases de données dans le domaine administratif, à un niveau d'abstraction relativement bien défini et homogène, les versions originales souffraient de lacunes que de nombreuses recherches contribuent à combler, comme en témoigne l'activité de l'E/R Institute. C'est dans le domaine des structures de données supportant les mécanismes d'abstraction (généralisation et agrégation) que les améliorations les plus pertinentes pour notre étude ont été apportées (voir entre autres [Teorey 1990]). Une synthèse des concepts les plus récents a débouché sur une proposition de "modèle E/R standard" [Spencer, Teorey et al. 1990], qui servira de référence à notre étude.

Ces concepts ont été intégrés au fur et à mesure de leur publication dans les formalismes géosémantiques en cours d'élaboration. Ainsi, le formalisme MODUL-R [Caron 1991] intègre-t-il différents modules, comprenant en particulier les concepts de généralisation, d'agrégation et de contraintes interrelationnelles.

Pour combler les lacunes subsistant encore, Paquette [Paquette 1990] a proposé la technique de substitution de sous-modèles (SSM), qui permet de remplacer les structures de données complexes décrivant la référence spatiale d'un type d'objet par un simple pictogramme. Le catalogue des sous-modèles proposés par Paquette ne répond cependant pas aux exigences de la description conceptuelle de modèles spatiaux; il rattache en particulier aux types d'objets des caractéristiques propres à des *relations* entre les types d'objets, voire à des *contraintes* sur ces relations.

Caron [Caron 1991] lève ces lacunes en postulant la distinction entre un *modèle conceptuel des données* (MCD), décrivant les entités de la réalité, et un *modèle conceptuel cartographique* (MCC), dans lequel est défini la représentation graphique des entités. La dimension spatiale des entités constitue l'articulation entre le MCD et le MCC; c'est cette dimension que décrivent les pictogrammes, dont il ne subsiste que les quatre types correspondant respectivement aux objets ponctuels (0 dimension), linéaires (1 dimension), surfaciques (2 dimensions) et volumiques (3 dimensions).

Cette idée ouvre des perspectives intéressantes pour la spécification de représentations cartographiques, mais néglige à nos yeux la dimension fondamentalement sémantique de certaines contraintes spatiales ⁴⁰, que les *contraintes interrelationnelles* proposées ne suffisent pas à modéliser. Ainsi, à titre d'exemple, l'affirmation que deux rivières ne peuvent se croiser est d'abord une spécification sémantique, et a pour simple conséquence que les traits qui les représentent sur une carte ne doivent pas se recouper ! La distinction entre des entités cartographiques et des entités "réelles" est à nos yeux arbitraire: toute entité pertinente pour un utilisateur et un objectif donnés revêt en effet par définition une valeur sémantique. N'y aurait-il pas là l'indice d'une relation d'une autre nature entre la sémantique et la référence spatiale, où celle-ci serait une solution, parmi d'autres possibles, de gérer des caractéristiques et des relations métriques et topologiques ? C'est cette intuition que nous tenterons d'étayer dans la suite de ce chapitre, à l'appui du concept de *référentiel spatial* défini au chapitre 4.

L'étude de Caron constitue néanmoins la définition d'un formalisme géosémantique la plus avancée dont nous disposons aujourd'hui. Ce formalisme satisfait d'ores et déjà les points a, b, c d et f du cahier des charges proposé au § 6.2. Il représente en outre le fruit de recherches menées durant plusieurs années à l'université Laval, dans le contexte desquelles nous souhaitons intégrer la présente étude. Sauf indication contraire, nous nous baserons donc sur le formalisme MODUL-R.

6.4. Approche du niveau conceptuel

Conformément au cycle d'abstraction lié à la conception de systèmes d'information [Tardieu, Rochfeld et al. 1984] [Tabourier 1986], nous devons spécifier au niveau conceptuel les structures d'information les plus stables d'une organisation. Ces structures reflètent donc les missions de l'organisation, indépendamment des solutions organisationnelles et matérielles existantes ou préconisées.

Dans le cas de systèmes traitant des informations localisées, ce niveau conceptuel peut par conséquent inclure des caractéristiques spatiales telles que l'orientation, la forme et l'étendue des objets, ainsi que des relations spatiales entre objets telles que la distance, l'adjacence, la connexion. Ces caractéristiques et relations relèvent donc naturellement d'un modèle conceptuel des données. Le recours à des méthodes cartographiques (calcul géométrique) pour gérer ces caractéristiques et relations est par contre un choix qui relève à nos yeux du niveau organisationnel.

⁴⁰ Caron propose cependant de décrire ces contraintes au niveau du dictionnaire du système, dont le modèle conceptuel est une "vue", une représentation partielle.

Néanmoins, un enrichissement inconsidéré du modèle conceptuel des données en rendrait l'interprétation impossible. C'est cette contrainte qui a précisément incité Caron à reléguer certaines informations dans le modèle conceptuel cartographique ou encore dans le dictionnaire des données. Une telle solution relègue malheureusement à l'arrière-plan des informations dont la valeur sémantique peut être très importante. Nous préconisons donc:

1. de représenter un modèle conceptuel des données complexe par plusieurs diagrammes de contenu différent ("vues externes")⁴¹;
2. de reléguer les spécifications cartographiques à l'étape organisationnelle de la définition du système.

Les composantes sémantiques proposées par le formalisme MODUL-R peuvent être utilisées sans restriction pour l'établissement du modèle conceptuel des données suggéré. Les caractéristiques spatiales sont modélisées principalement par la dimension topologique de l'entité (représentée par l'un des quatre pictogrammes définis dans MODUL-R⁴²). Des relations spatiales, telles que distance ou connexion, peuvent en principe être modélisées comme n'importe quelle autre relation, et être assorties de cardinalités et de contraintes interrelationnelles.

Nous souhaiterions cependant encore pouvoir exprimer au niveau conceptuel certaines contraintes intervenant dans la composition d'agrégats d'entités spatiales. Nous pouvons illustrer cette exigence sur la base des exemples suivants:

- a. On utilise en hydrologie des bassins versants que l'on divise en sous-bassins; plusieurs jeux de sous-bassins peuvent coexister, mais les sous-bassins d'un même jeu ne peuvent se superposer.
- b. Les canaux et rivières d'un même réseau hydrographique doivent être interconnectés; ils forment donc un graphe connexe, dans lequel les circuits sont prohibés.
- c. Un jeu opérationnel de conduites, de vannes, de bouches d'incendie, etc., constitue un réseau de distribution d'eau. Chaque vanne et chaque conduite constituant le réseau doivent être connectées à ce réseau, qui doit lui-même constituer un graphe connexe (si le graphe n'est pas connexe, on a alors deux réseaux distincts). Il peut exister des conduites isolées déconnectées du réseau, inutilisées mais toujours enfouies dans le sous-sol.

Conformément au formalisme MODUL-R, les bassins versants, les réseaux hydrographiques et les réseaux de distribution d'eau potable décrits ci-dessus constituent des *agrégats*. Ces agrégations sont soumises à des ensembles complexes de *contraintes* :

1. Certaines contraintes s'appliquent globalement à chaque occurrence d'agrégat. L'absence de circuit dans un réseau hydrographique en est un exemple.
2. D'autres contraintes portent sur l'intégration dans un agrégat d'une occurrence d'une entité agrégée. Au sein d'un réseau donné par exemple, une vanne doit être connectée au réseau.
3. D'autres contraintes enfin portent sur la création de sous-ensembles d'entités agrégées, que [Spencer, Teorey et al. 1990] appelle une *catégorie* de l'entité initiale. Ainsi, des sous-bassins versants d'un même jeu ne peuvent se recouper.

41 Le recours à une hiérarchie de vues partielles, de niveau toujours plus détaillé, préconisée par Caron, gagne donc encore en pertinence.

42 Les pictogrammes ne font que représenter le concept de dimension topologique. Ils ne sont définis ni par la substitution de sous-modèles (SSM) de [Paquette 1990], ni comme charnières vers le MCC.

L'expression formelle de ces contraintes à l'aide des composantes sémantiques de MODUL-R (agrégats, cardinalités, contraintes interrelationnelles, ...) conduirait à des structures complexes. Caron suggère de les décrire textuellement dans le dictionnaire. L'absence de formalisation des contraintes d'agrégation dans le modèle conceptuel et leur relégation dans le dictionnaire ne permettent malheureusement pas d'en garantir une bonne appréhension cognitive, qui se révélera néanmoins nécessaire pour le passage aux modèles organisationnel et logique.

Nous proposons donc de redéfinir la composante sémantique *agrégat* de MODUL-R pour prendre en compte les *catégories* d'entités composantes et définir les *contraintes d'agrégation spatiales*.

6.5. Nouvelle composante sémantique d'un modèle de données: l'agrégat

6.5.1. Définition

Un *agrégat* est une composante sémantique d'un formalisme géosémantique. A l'instar d'une relation, un agrégat est associé à une ou plusieurs entités ou autres agrégats, qui forment sa *collection*. Chaque association exprime qu'une occurrence de l'agrégat est composée d'occurrences de la composante associée. Chaque association peut être de nature sémantique ou spatiale.

Un *type d'agrégat* est défini par:

1. son *nom* (à l'instar des *entités* et des *relations*);
2. la liste de ses *attributs* (à l'instar des *entités* et des *relations*);
3. la liste des *contraintes globales* d'agrégation s'appliquant à chaque agrégat, à l'exemple de l'exclusion des circuits dans un réseau hydrographique; une définition de ces contraintes globales est présentée au § 6.4.2;
4. les *contraintes partielles* d'agrégation de chaque occurrence d'une entité dans un agrégat, à l'exemple d'un puisard qui doit impérativement être rattaché à une conduite dans un réseau d'assainissement; la nature de ces contraintes partielles est présentée au § 6.4.3;
5. les *cardinalités de l'agrégat* dans une entité, qui expriment combien d'occurrences de l'entité peuvent être associées à une occurrence de l'agrégat;
6. un éventuel *opérateur spatial d'ensemble* (union, intersection), qui doit être appliqué à tout agrégat spatialement associé à l'agrégat considéré, définissant le modèle spatial objet de l'agrégation (voir chapitre 4.4).

La figure 6.1 propose une nouvelle notation de l'agrégat, permettant de transcrire les contraintes; les numéros renvoient à la définition ci-dessus. La notation se distingue très clairement de celle de la généralisation, contrairement à la notation des *agrégations* proposée par Caron. Elle permet en outre une meilleure flexibilité de la notation pour la transcription de modèles complexes. La notation hybride proposée rappelle que l'agrégat est une combinaison d'une entité et de ses relations de composition.

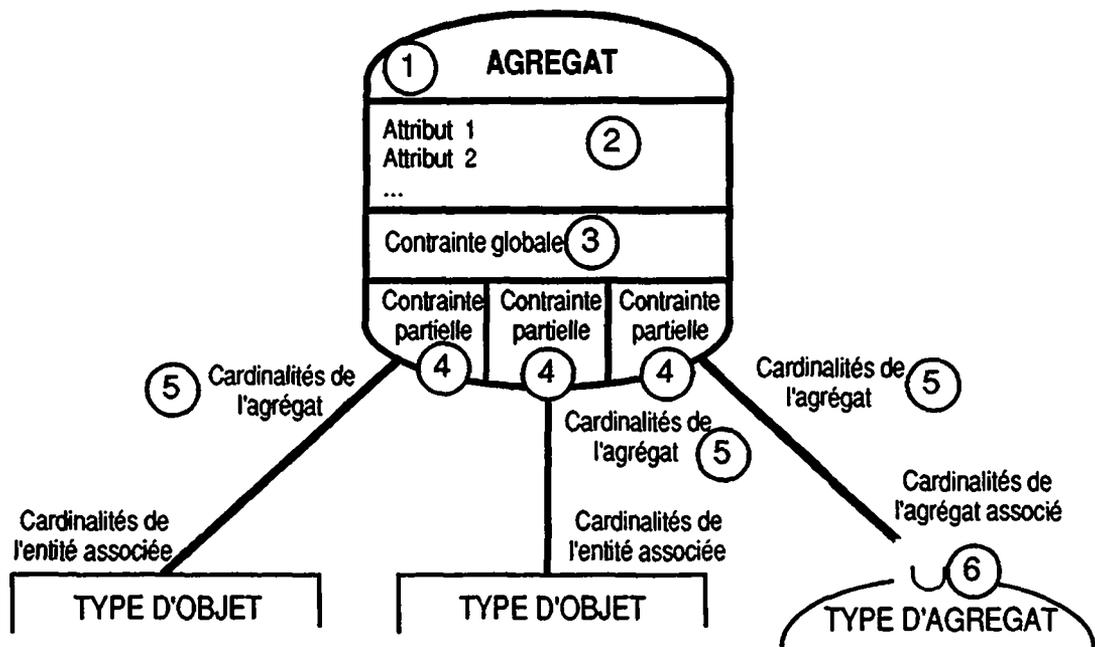


Figure 6.1: Notation suggérée de l'agrégat

Les règles structurelles suivantes sont applicables:

1. La récursivité de l'agrégation est interdite: un agrégat ne peut être associé à lui-même, directement ou indirectement.
2. Un agrégat peut être associé par des types de relations classiques à d'autres entités ou agrégats.
3. Un agrégat peut lui-même être une composante d'un autre agrégat, à l'instar d'une entité. Si l'agrégation est de nature spatiale, le modèle spatial considéré doit être défini par un opérateur ensembliste (réunion, agrégation) sur le set des points du modèle agrégé [Egenhofer et Franzosa 1991]. Par exemple, l'agrégation de deux bâtiments adjacents en un seul bâtiment est une réunion spatiale, alors que le mur mitoyen commun à ces deux bâtiments correspond à l'intersection des deux modèles (voir chapitre 4.4 et les exemples d'application ci-dessous).

6.5.2. Contraintes d'agrégation globales

Les contraintes d'agrégation globales peuvent être de nature purement sémantique ou de nature spatiale.

Il n'entre pas dans le cadre de cette étude de proposer un catalogue exhaustif des contraintes d'agrégation globales qui peuvent être définies. Dans le cadre de notre étude, des exemples ont essentiellement été identifiés dans le domaine de la théorie des graphes: pas de boucle ou circuit, pas de bief, etc...

Nous ne suggérons pas de notation particulière de ces contraintes. Une description textuelle de la contrainte réduite à quelques mots-clés paraît adéquate, dans la mesure où une description plus détaillée peut être intégrée au dictionnaire. Dans les cas complexes ou répétitifs, la définition d'icônes définies dans une légende associée au modèle est aussi envisageable.

On trouvera plus loin quelques exemples d'application du concept d'agrégat, dans lesquels des contraintes d'agrégation globales sont définies.

6.5.3. Contraintes d'agrégation partielles

Les contraintes d'agrégation partielles spécifient les conditions que doit remplir chaque occurrence de la composante considérée (entité ou agrégat) pour pouvoir participer à l'agrégat modélisé.

Dans le domaine spatial, les contraintes les plus pertinentes semblent être d'ordre topologique, même si des contraintes métriques sont aussi envisageables (distance minimale ou maximale entre les points d'un réseau par exemple).

Selon [Champoux 1991], toutes les interdépendances topologiques entre deux objets spatiaux peuvent être décrites par la composition des deux associations "adjacence" (incidence de l'enveloppe d'un objet sur un autre) et "inclusion" (incidence de l'intérieur d'un objet sur un autre). Ces deux associations peuvent prendre les valeurs *nulle*, *partielle* ou *totale*.

Une contrainte d'agrégation partielle peut donc en particulier spécifier des valeurs obligatoires ou interdites de chacune de ces deux associations. On sera attentif à ce qu'on considère l'association de chaque occurrence de la composante choisie avec une occurrence complète de l'agrégat ⁴³; on considère à cet effet l'agrégat comme un set de points [Egenhofer et Franzosa 1991], quels que soient les modèles spatiaux qui le composent ⁴⁴.

Dans le cadre de notre étude, nous avons temporairement adopté une représentation par icônes des contraintes d'agrégation partielle de nature topologique. Ces représentations peuvent être complétées par une description formelle dans le dictionnaire. Nous ne proposons cependant pas de catalogue exhaustif d'icônes, au vu du trop grand nombre de possibilités existantes ⁴⁵. La notation par icônes n'est donc envisageable que pour des modèles dans lesquels existe un catalogue restreint et typé de contraintes; une *légende des icônes* devra être impérativement associée à chaque modèle de données.

Des exemples d'icônes de contraintes d'agrégation partielles sont proposés dans les exemples ci-dessous.

6.5.4. Exemples d'application des agrégats

Le concept d'agrégat défini ci-dessus est illustré dans ce paragraphe par quelques "cas d'école". Un exemple plus complet, issu du domaine du drainage urbain, est proposé au chapitre 9.

43 L'association inverse (de l'agrégat avec une occurrence de la composante considérée) n'est pas prise en considération, du fait que chaque occurrence de chaque composante est testée par rapport à l'agrégat auquel elle participe.

44 Ces modèles spatiaux peuvent en effet être de dimensions différentes, à l'exemple de vanes (points) et de conduites (lignes) composant un réseau.

45 Une rapide évaluation, tenant compte des contraintes possibles, des compositions de contraintes entre inclusion et adjacence, et finalement des différentes dimensions des modèles spatiaux pouvant intervenir nous conduit à estimer le nombre d'icônes possibles à plus de 300.

6.5.4.1. Lot de parcelles

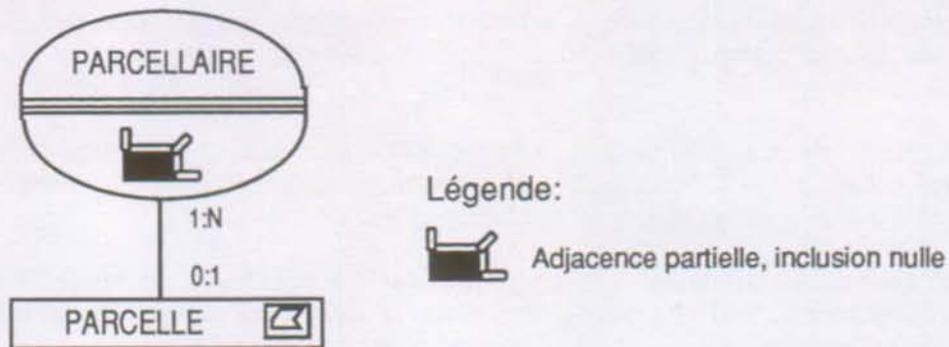


Figure 6.2: Exemple d'un modèle *lot de parcelles*

Un lot de parcelles est un ensemble de parcelles qui ne se chevauchent pas (figure 6.2). On constate que des parcelles peuvent ne pas appartenir à un lot (parcelles radiées par exemple).

6.5.4.2. Subdivision cadastrale d'une commune

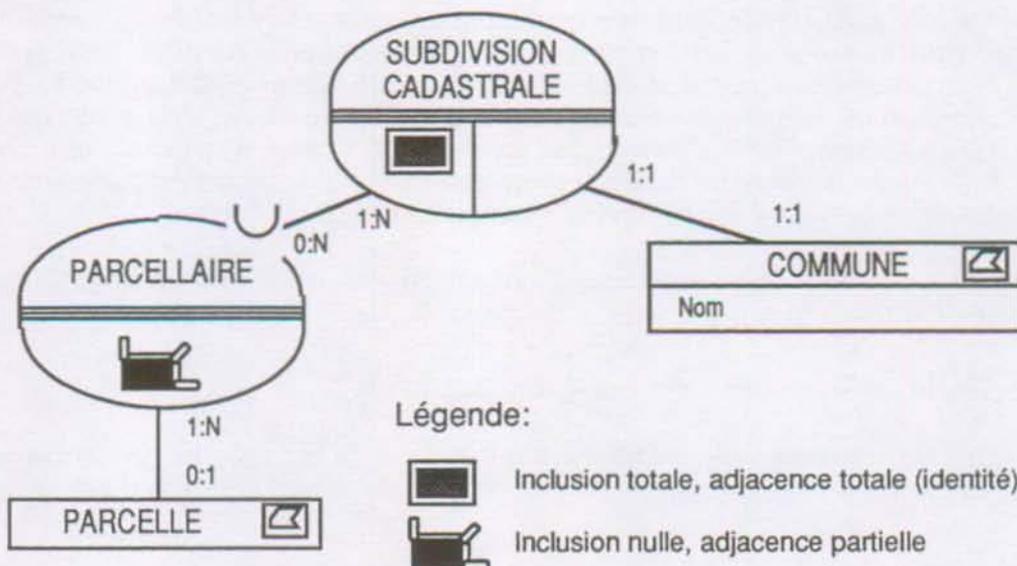


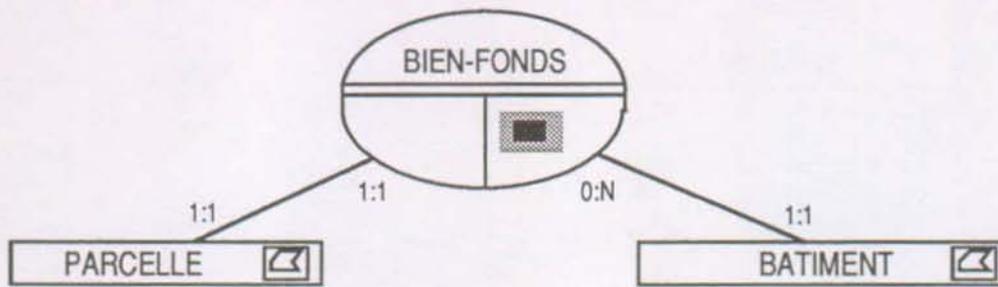
Figure 6.3: Exemple d'un modèle *subdivision cadastrale*

Cet exemple (figure 6.3) exprime que le parcellaire (ensemble des parcelles d'une commune) constitue une subdivision d'une commune, sans recouvrement ni lacune (mosaïque).

On modélise deux entités spatiales a priori indépendantes, les *parcelles* et les *communes*. Cette indépendance traduit les origines pas toujours cohérentes des limites politiques et des limites des droits réels.

Néanmoins, les parcelles peuvent être assemblées dans une structure parcellaire sans chevauchement. A une commune est associée exactement une subdivision cadastrale, constituée d'une réunion des parcelles d'une structure parcellaire recouvrant exactement le territoire communal.

6.5.4.3. Bien-fonds



Légende:

 Inclusion totale, adjacence totale

Figure 6.4: Exemple d'un modèle bien-fonds

Un bien-fonds (figure 6.4) est modélisé comme une parcelle sur laquelle un ensemble de bâtiments est totalement inclus.

6.5.4.4. Réseau d'adduction d'eau potable



Légende:

 Adjacence partielle, inclusion nulle

 Inclusion totale

Figure 6.5: Exemple d'un modèle réseau d'adduction d'eau potable

Un réseau d'adduction d'eau potable est un agrégat de vannes et de conduites. Les vannes sont totalement incluses, et les conduites partiellement adjacentes.

6.5.4.5. Schéma d'un réseau d'évacuation des eaux pluviales

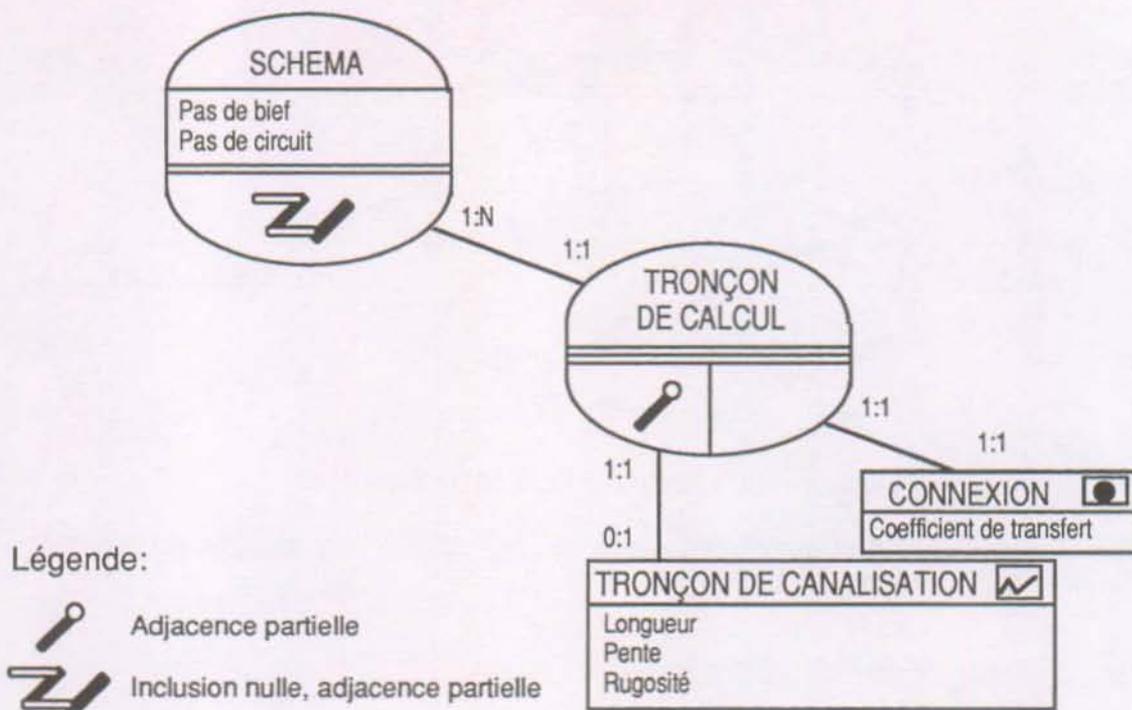


Figure 6.6: Exemple d'un modèle schéma d'un réseau d'évacuation des eaux pluviales

Le schéma de réseau est modélisé comme un agrégat de *tronçons de calcul* partiellement adjacents. Chaque tronçon de calcul est composé d'exactly un tronçon de canalisation adjacent à un point de connexion.

Le schéma est en outre soumis à des contraintes d'agrégation globales: les biefs et les circuits sont exclus.

6.5.4.6. Structure d'un bassin versant

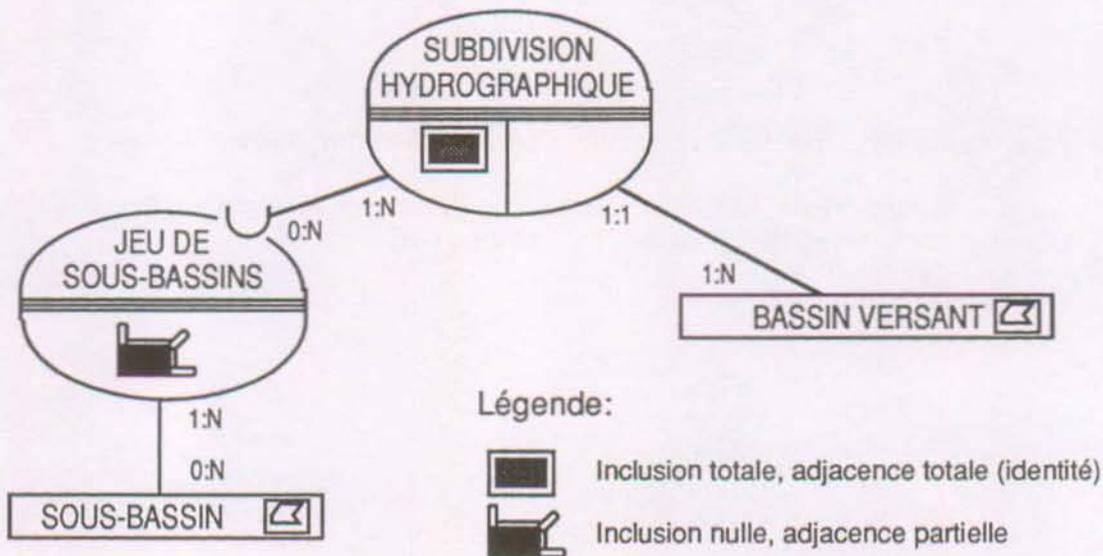


Figure 6.7: Exemple d'un modèle bassin versant

La division d'un versant en sous-bassins présente des analogies avec la structure parcellaire d'une commune. Néanmoins, on constate que plusieurs jeux de sous-bassins peuvent coexister, correspondant à différents modèles hydrologiques du bassin.

6.5.4.7. Mur mitoyen

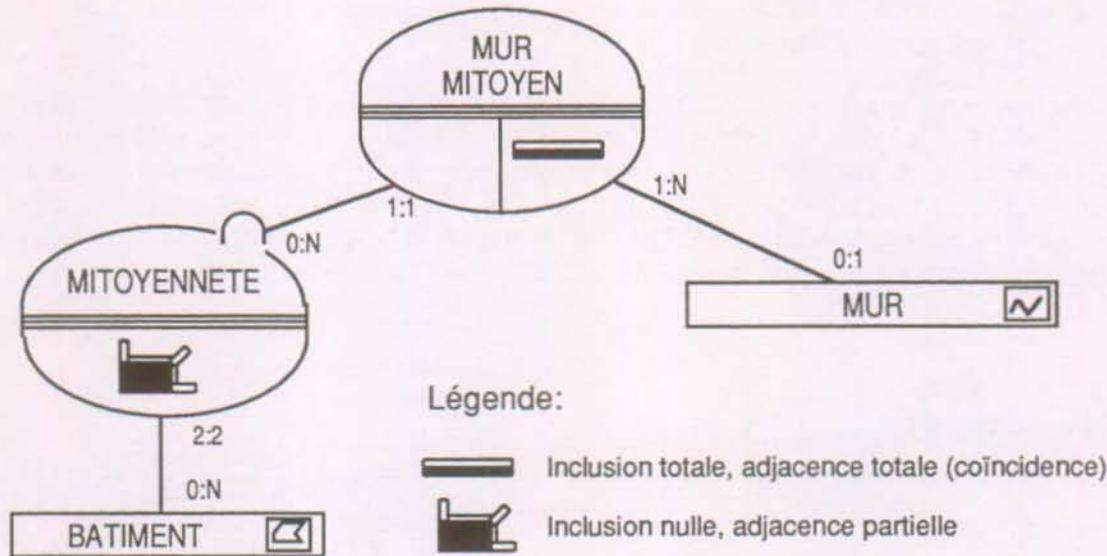


Figure 6.8: Exemple d'un modèle mur mitoyen

Un mur mitoyen constitue un cas particulier d'agrégation par intersection des composantes d'un agrégat. L'intersection d'exactly deux bâtiments adjacents qui coïncident avec un mur peut constituer un mur mitoyen.

6.5.4.8. Utilisation d'agrégats dans un contexte multi-médias

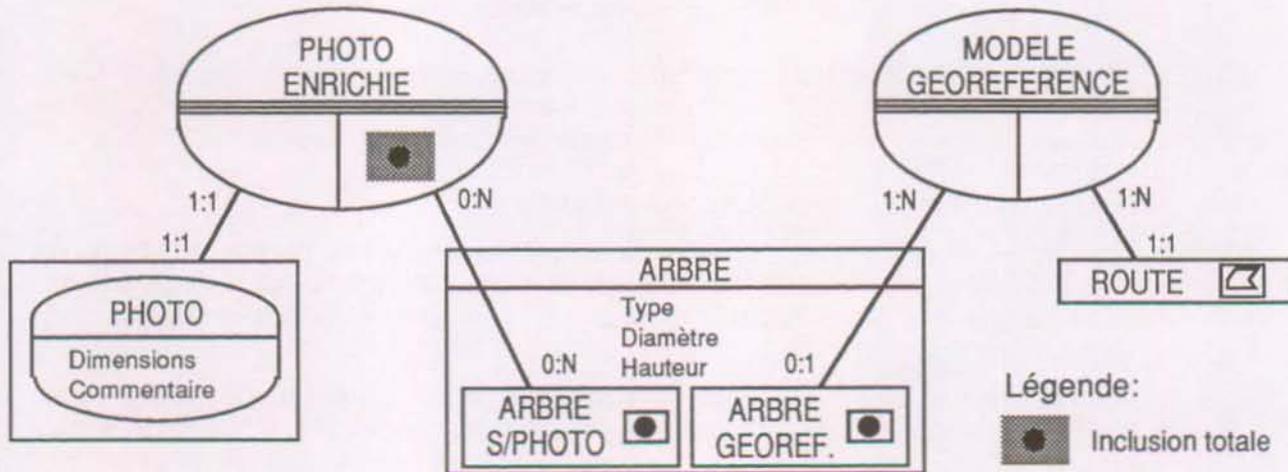


Figure 6.9: Exemple de l'utilisation d'agrégats dans un contexte multi-médias

Des supports d'information "non structurés", tels que photos et vidéos, peuvent aujourd'hui être intégrés à des bases de données multi-médias. Un formalisme géosémantique doit donc être à même de décrire les modèles sous-jacents.

La figure 6.9 en présente un exemple, dans lequel des arbres peuvent être localisés aussi bien sur des photos (qui constituent chacune un référentiel spatial) que dans un espace géoréférencé. Une notation est proposée pour la composante non structurée "photo". Un point localisant un arbre doit obligatoirement être sur la photo considérée.

6.6. Approche du niveau organisationnel

Le passage du niveau conceptuel au niveau organisationnel implique la prise en compte de choix organisationnels: répartition des tâches entre l'homme et la machine, niveau des services rendus par la machine, etc. Les modèles conceptuels élaborés conformément aux principes généraux énoncés au § 6.3 vont nous permettre de définir plusieurs types de modèles organisationnels, dont nous allons esquisser dans ce paragraphe la structure générale et la démarche d'élaboration.

En principe, on déduira du modèle conceptuel des données (MCD) un *modèle organisationnel des données* (MOD), attribuant les responsabilités liées à la gestion et à l'exploitation des données. Nous ne nous attarderons pas sur cette démarche documentée dans la littérature scientifique et technique. On déduira aussi du MCD un *modèle organisationnel cartographique* (MOC) par *référentiel spatial* que nous identifierons sur la base du MCD. Cette architecture de modèles est représentée sur la figure 6.10.

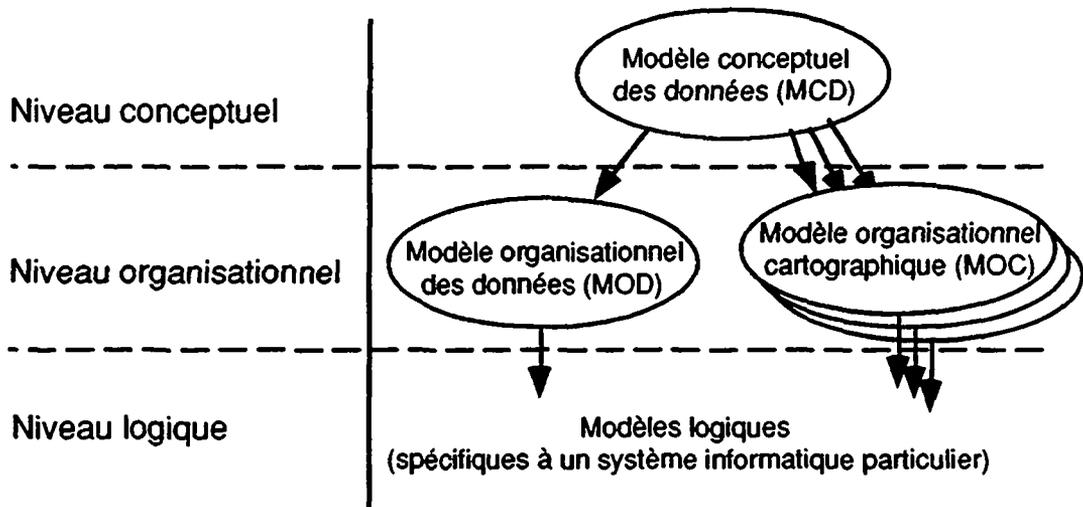


Figure 6.10: Architecture de modèles proposée

La première étape de modélisation organisationnelle cartographique consiste à identifier sur le MCD quelles entités, quelles relations et quels agrégats nous allons assortir d'une référence spatiale et gérer à l'aide d'un système de gestion de données localisées.

Nous proposons les principes de modélisation suivants:

- Les entités assorties d'un pictogramme de référence spatiale dans le MCD y sont déjà identifiées comme entités à référence spatiale. Les entités ayant des attributs spatiaux tels que pente, superficie, longueur, ... devraient dans tous les cas faire partie de cette catégorie ⁴⁶.
- Les relations spatiales (distance, angles, connexions, ...) ainsi que les agrégats assortis d'attributs ou de contraintes spatiales sont susceptibles d'être gérés par le biais de la référence spatiale ⁴⁷. Nous proposons de représenter les associations correspondantes par un double trait, à l'exemple de la figure 6.11. Les entités

⁴⁶ A l'exclusion des attributs assortis de valeurs officielles gérées indépendamment des mutations spatiales, tels que la superficie des immeubles utilisée par le Registre foncier comme base de la taxation immobilière.

⁴⁷ Si des valeurs ne sont mesurées qu'occasionnellement, on pourra envisager de gérer la relation ou l'agrégat sans recourir à la référence spatiale. Des relations ou agrégats ne faisant intervenir que des mesures ordinales ou nominales (topologie), tels que des schémas de réseau, peuvent souvent être gérés relativement aisément à l'aide d'un système de base de données classique. La complexité des contraintes à vérifier, les exigences des traitements et le contexte général du système et de l'entreprise sont des critères déterminants pour le choix d'un outil de gestion.

mis en relation doivent alors elles aussi être assorties d'un pictogramme de référence spatiale.



Figure 6.11: Notation suggérée des relations spatiales lors de la modélisation organisationnelle

Sur cette base, on verra apparaître sur le MOC des "réseaux de composants" (entités, relations et agrégats) interconnectés par des doubles traits. **Chaque réseau définit un référentiel spatial potentiel.**

L'homogénéité de la référence spatiale doit encore être validée pour chaque réseau identifié. Conformément au chapitre 4.3.1, les échelles de mesure de toutes les caractéristiques spatiales doivent être de même nature et de même qualité au sein d'un référentiel donné. Toutes les composantes doivent appartenir à un même niveau d'abstraction, et avoir la même résolution spatiale. Si ces principes d'homogénéité ne sont pas vérifiés, le réseau concerné devra être scindé en deux ou plusieurs réseaux, par dédoublement des modèles spatiaux de certaines entités-charnières. Un tel mécanisme de scission est illustré par la figure 6.12.

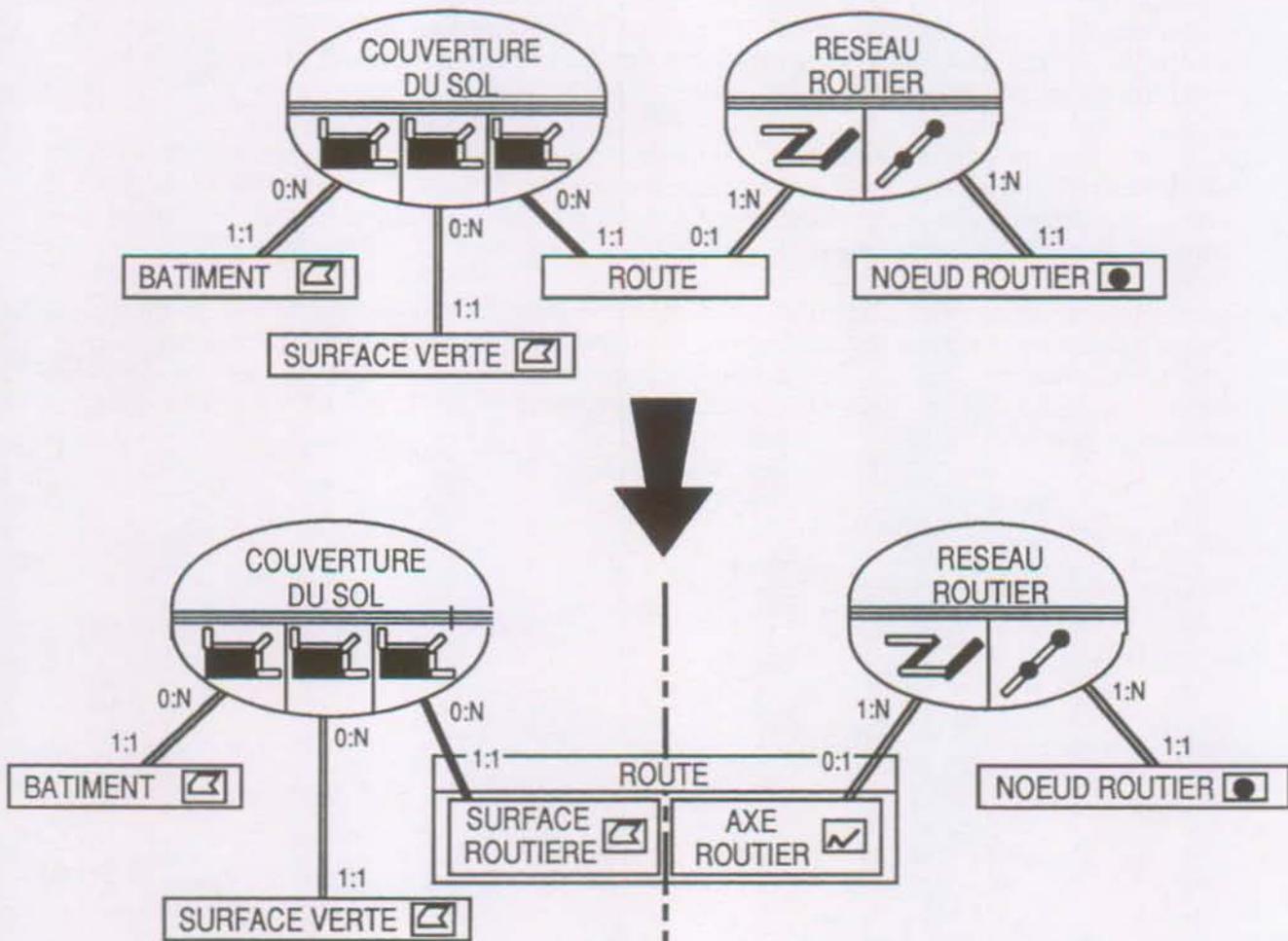


Figure 6.12: Exemple de scission d'un réseau hétérogène de relations spatiales

A l'issue de l'étape de modélisation organisationnelle, les MOD et MOC obtenus pourront être transformés en modèles logiques, conformément aux caractéristiques du système informatique cible.

Les règles de transformation entre un formalisme E/R et des modèles de données en réseau ou relationnels sont bien connus [Tardieu, Rochfeld et al. 1984] [Tabourier 1986] [Zehnder 1985]. Caron [Caron 1991] a défini le formalisme géosémantique MODUL-R dans la perspective de modèles logiques *orientés objets*. Un *modèle logique cartographique* (MLC) doit quant à lui spécifier de quelle manière les modèles spatiaux seront implémentés dans un SGDL donné.

Malheureusement, les plupart des systèmes de gestion de données localisées actuels ne sont capables de gérer qu'un seul référentiel spatial. Seuls quelques systèmes font exception, comme TIGRIS, capable de gérer plusieurs sets topologiques dans un même référentiel métrique, et SICAD, qui offre un "espace schématique" en plus de l'"espace géographique" standard. Dans l'état actuel de la technologie, les différents référentiels spatiaux identifiés au niveau *logique* devront donc être souvent fondus en un seul référentiel spatial lors de la réalisation du système. Il en résulte un danger élevé d'extraire des informations non pertinentes du système.

6.7. Conclusion du chapitre

Ce chapitre nous a permis de choisir un formalisme géosémantique pour la modélisation des structures de données à référence spatiale. Ce formalisme remplit aussi bien les exigences de la modélisation des SIRS que celle de systèmes d'agrégation d'information localisées. Il reprend pour l'essentiel les concepts du formalisme MODUL-R, dont il redéfinit la composante sémantique *agrégat*.

Assorti d'une nouvelle démarche de modélisation, ce formalisme permet de définir plusieurs référentiels spatiaux sur la base d'un modèle conceptuel de données unique. Il prend ainsi en compte le rôle fondamental de la référence spatiale: la gestion efficace des caractéristiques et relations spatiales pertinentes pour un système donné.

Dans l'attente de systèmes informatiques capables de gérer plusieurs référentiels spatiaux, des solutions transitoires devront encore être trouvées. La plus simple, mais aussi la plus coûteuse, consiste à gérer l'ensemble des informations spatiales conformément aux exigences de qualité les plus élevées identifiées dans le système, à l'instar des anciens cadastres polyvalents.

7. Problèmes et solutions relatifs aux organisations

Dans les années soixante, lors de l'avènement de l'informatique dans les entreprises, les problèmes liés au traitement de l'information étaient identifiés et résolus au coup par coup par des "applications" ad hoc [Tardieu, Rochfeld et al. 1984]. Les insuffisances d'une telle approche, qui négligeait les interdépendances entre les applications informatiques, sont rapidement apparues. Elles ont conduit les entreprises à adopter des démarches plus systémiques, décrivant dans une première étape le *système d'information d'entreprise* dans lequel les applications informatiques doivent ultérieurement s'inscrire. Une telle démarche a été formalisée par quelques méthodes qui ont rencontré un vif succès commercial, en particulier MERISE [Tardieu, Rochfeld et al. 1984] et Information Engineering [Martin et McClure 1985]. Ces méthodes continuent aujourd'hui à être développées et adaptées à l'évolution des connaissances méthodologiques et technologiques.

Les entreprises qui gèrent, exploitent ou produisent des informations localisées se heurtent cependant à un obstacle important: le *facteur intégrant* du territoire. Les préoccupations écologistes actuelles nous rappellent sans cesse les conséquences multiples de toute décision sectorielle relative à la gestion et à l'aménagement du territoire. A une nécessaire intégration des décisions doit donc répondre une intégration des informations, qui acquièrent une importante valeur ajoutée par les recoupements spatiaux qu'elles rendent possible. Chevallier [Chevallier 1983] a déjà souligné la nature hétérogène des données localisées ainsi que l'ambivalence des fonctions qu'on leur applique (saisie finalisée et exploitation pour des objectifs variés).

Les données localisées sont donc soumises à un double impératif d'intégration: avec les données des entreprises concernées d'une part, et avec les autres données relatives à une même portion de territoire d'autre part. La conception générale d'un système d'information répondant à toutes ces contraintes d'intégration serait une gageure difficilement surmontable, de par le nombre d'intervenants impliqués et la complexité de leurs relations.

Dans un tel contexte, ce chapitre a pour objectif de dégager une structure organisationnelle générale des systèmes traitant des informations localisées, un canevas permettant de les découper, puis d'en identifier et d'en classifier les éléments.

7.1. Découpage d'une entreprise en domaines de gestion

Des méthodes telles que MERISE préconisent de découper le système d'information de l'entreprise en *domaines* pour lesquels des projets informatiques puissent être conduits de manière autonome. Une telle approche peut être rendue nécessaire par la complexité des activités de l'entreprise et des informations qu'elles mettent en jeu.

Un tel découpage ne devrait cependant pas dégrader la cohérence du système d'information de l'entreprise. Conformément à une approche systémique du système d'information, MERISE propose donc de retenir un découpage qui minimise les flux d'information entre les domaines identifiés [Tardieu, Rochfeld et al. 1984]. Malheureusement, un tel inventaire des flux se révèle très difficile à mettre en oeuvre au stade initial de l'analyse [Tabourier 1986]. Dans le cas des informations à référence spatiale, ce problème est encore rendu plus aigu par l'abondance et la complexité de leurs flux [Chevallier 1983].

D'autres critères de découpage doivent donc être mis en oeuvre. Ainsi, Tabourier propose un découpage initial du système opérant, et non du système d'information. Les choix de gestion propres au niveau conceptuel doivent en effet prendre en compte les objectifs opératoires de l'entreprise, et non pas être calqués sur les procédures informationnelles existantes. Ce découpage du système opérant peut être fondé sur un inventaire des activités de l'entreprise, que l'on regroupe en *domaines de gestion* répondant à des objectifs

homogènes, à des préoccupations communes: gestion d'une ressource, conduite d'activités opérationnelles, etc. . Ainsi, il peut se révéler pertinent de regrouper des activités d'une entreprise de services industriels telles que *distribuer de l'eau, distribuer du gaz et distribuer de l'électricité* dans un même domaine de gestion *distribuer un fluide*. Une exigence complémentaire est que ces activités répondent à une même *logique de pilotage*: il s'agit d'éviter de regrouper des activités dont les horizons décisionnels seraient hétérogènes (court et long terme) ou dont les logiques décisionnelles seraient conflictuelles.

Nous proposons d'étayer et d'étendre ces critères en les rattachant aux composantes du système général [Le Moigne 1977] décrit au chapitre 5 (figure 5.1): le sous-système opérant, le sous-système d'information et le sous-système de pilotage. Chaque domaine constitue en effet un sous-système du système entreprise. Il s'agit de garantir au sein de chaque domaine l'homogénéité des sous-systèmes opérant, de pilotage et d'information:

1. L'homogénéité du sous-système opérant est liée à l'identité des finalités opératoires des activités qu'il regroupe, postulée par Tabourier. On admettra en effet que pour parvenir à un même objectif, on fait en règle générale appel aux mêmes processeurs. Ainsi, la gestion du cadastre se révèle similaire à la gestion d'un réseau de conduites, alors que cette même gestion d'un réseau de conduites est fonctionnellement très différente d'un calcul hydraulique.
2. L'homogénéité du sous-système de pilotage est liée à la cohérence de la logique de pilotage des activités qu'il regroupe, toujours postulée par Tabourier. L'horizon décisionnel des activités doit en particulier se révéler homogène: ainsi, l'octroi d'une autorisation de construire obéit-il à une logique de pilotage à court terme, alors que la gestion d'un plan directeur d'aménagement obéit à une logique de pilotage à long terme.
3. L'homogénéité du sous-système d'information implique l'intégration d'informations de même nature, qui répondent à de mêmes règles de gestion de l'information. Le regroupement en un même domaine d'activités se basant sur des informations à référence spatiale répond par exemple à ce critère d'homogénéité informationnelle. Ce troisième critère n'a pas été pris en compte par Tabourier. Il n'en est pas moins particulièrement important pour les systèmes intégrant des informations localisées, qui exigent le recours à des *compétences* particulières en matière de gestion d'informations.

Nous allons maintenant appliquer ces critères au cas particulier d'entreprises gérant ou utilisant des informations localisées. Nous pourrions ainsi suggérer un canevas de découpage des systèmes traitant des informations localisées sur un territoire donné.

7.2. Systèmes d'information à référence spatiale: des sous-systèmes du système d'information d'entreprise

La gestion et le traitement d'informations localisées font appel à des compétences et à une technologie spécifiques. Conformément au critère d'homogénéité du sous-système d'information postulé au paragraphe 7.1 (critère no 3), il est donc pertinent d'envisager le regroupement des activités d'une entreprise produisant ou utilisant des informations localisées au sein d'un même domaine de gestion.

Prenons l'exemple d'une société de services industriels chargée de distribuer l'eau, le gaz et l'électricité à une communauté urbaine. Elle a besoin d'informations sur la position des conduites souterraines des différents réseaux pour assurer cette mission. Le seul critère de finalités opératoires homogènes (critère no 1) pourrait nous conduire à inclure la gestion des informations localisées correspondantes dans un domaine de gestion propre à chaque fluide. Le critère d'homogénéité des informations nous conduit cependant à distinguer un domaine de gestion relatif aux *plans de réseau*.

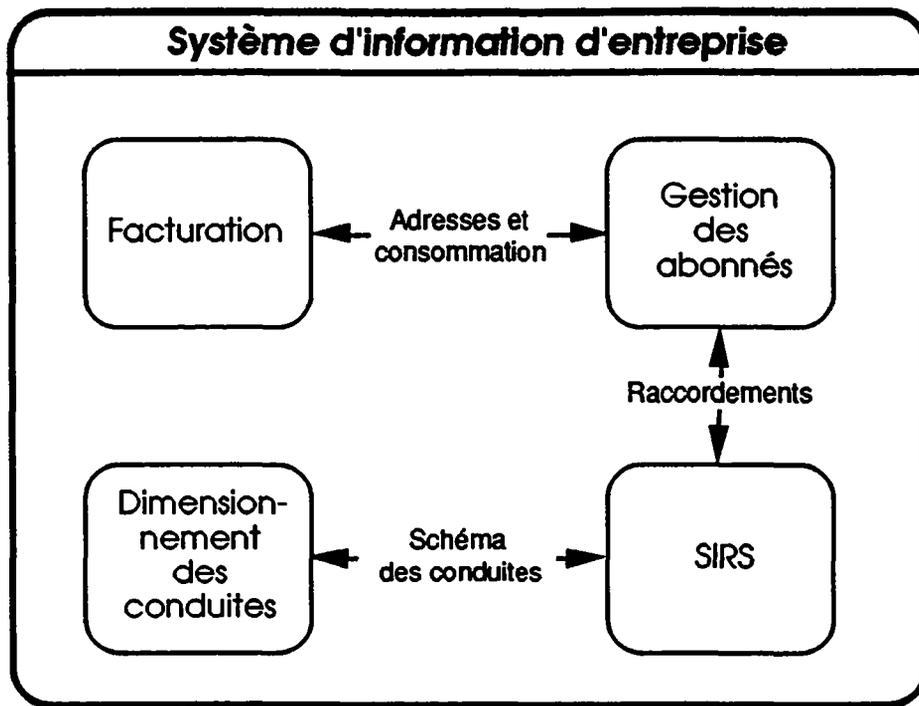


Figure 7.1: Structure partielle du SI d'une entreprise de services industriels

Plus généralement, le domaine de gestion regroupant les activités "fortement spécialisées" de l'entreprise se reflétera dans un sous-système du système d'information d'entreprise: le *système d'information à référence spatiale (SIRS)* ⁴⁸. Une telle architecture est illustrée par la figure 7.1.

Ce découpage justifie *a posteriori* les SIRS existant déjà dans de nombreuses entreprises et administrations. Dans de nombreux cas, ces SIRS ne reposent en effet sur aucune approche globale du système d'information d'entreprise, sur aucun *schéma directeur*, mais sur l'intuition et l'expérience que les problèmes spécifiques aux informations localisées doivent être résolus dans un système d'information ad hoc ⁴⁹.

Nous avons évoqué au paragraphe 7.1 le critère "merisien" de découpage du système d'information d'entreprise par minimisation des flux d'information entre sous-systèmes. Nous avons remarqué que ce critère n'était pas applicable en pratique, du fait qu'il n'était guère possible d'inventorier ces flux et encore moins de les pondérer au stade de la modélisation conceptuelle. Il serait néanmoins pertinent de tenter de vérifier ce critère *a posteriori*. Force nous est de constater que les flux entre un SIRS et d'autres sous-systèmes d'information de l'entreprise restent en règle générale relativement importants ⁵⁰, du fait du *facteur intégrant* que constitue le territoire. La mise sur pied d'un SIRS a donc des conséquences significatives sur les autres sous-systèmes d'information de l'entreprise. Ces conséquences doivent être résolues par des solutions organisationnelles adéquates (redondance contrôlée des données en particulier).

48 Dans certaines approches, on considère au contraire un système d'information comme un sous-système d'un système d'information à référence spatiale. Mais il s'agit là d'une définition *ensembliste* et non *systémique* du système d'information: en effet, les fonctions d'un SI constituent généralement un sous-ensemble des fonctions d'un SIRS.

49 Peut-être existe-t-il aussi une raison plus triviale à un tel découpage: l'existence de systèmes informatiques dédiés à la gestion d'informations localisées...

50 Il s'agit d'une estimation qualitative, et non quantitative.

7.3. Le SIT: un concept fédérateur des SIRS sur un territoire donné

La gestion du territoire nécessite l'intégration d'informations provenant de sources très différentes. Celles-ci peuvent concerner la propriété foncière, l'affectation du sol, les infrastructures existantes, etc. Le territoire joue donc un rôle de facteur intégrant non seulement à l'intérieur d'une entreprise, mais aussi entre tous les organismes impliqués dans la gestion du territoire.

D'un point de vue systémique, la définition d'un SIRS comme un sous-système d'un système d'information d'entreprise est donc trop réductrice. Un SIRS doit prendre en considération l'ensemble des données localisées pertinentes sur un territoire donné. Une telle solution se heurte cependant à de multiples problèmes:

- La complexité d'un tel SIRS global en rend la planification et la maîtrise très difficiles.
- L'indépendance des structures décisionnelles impliquées (nombreux partenaires publics et privés) peut en paralyser le développement et l'exploitation.

La gestion des informations sur le territoire doit donc être abordée sous-système par sous-système. Nous distinguons essentiellement deux critères permettant de circonscrire de tels sous-systèmes:

1. l'unité de *décision*, constituée par une entreprise, une administration ou un de leurs services au bénéfice d'une large autonomie de décision;
2. l'unité *spatiale*, constituée par l'étendue du territoire considéré; il s'agit généralement d'une entité politique (en Suisse: commune, arrondissement, district, canton; au Québec: municipalité, MRC, communauté urbaine) constituant la juridiction de nombreux services publics.

La combinaison sans biais de ces deux critères n'est cependant guère possible, du fait de l'hétérogénéité des juridictions ⁵¹ considérées. Nous distinguons donc deux types d'organisation des SIRS sur un territoire donné:

1. Les SIRS fédérés dans le cadre d'un partenariat.

Dans une telle organisation, chaque gestionnaire du territoire est un partenaire qui met ses informations à disposition des autres intéressés. Cette formule s'accommode en principe très bien d'une décentralisation des systèmes d'information et de leurs outils informatiques. De tels SIRS sont donc avant tout des concepts organisationnels. La figure 7.2 illustre une telle organisation et un ensemble (non exhaustif) de flux d'informations qui peuvent en résulter.

Le système d'information du territoire à Genève (SITG) ⁵² constitue un modèle d'une telle approche. Il regroupe les principaux fournisseurs de données localisées du Canton. Le SITG ne dispose d'aucun budget ni d'infrastructure informatique propres; ce sont les initiatives concertées des différents partenaires qui en déterminent le développement.

⁵¹ Portion de territoire sur lequel s'étend l'autorité d'un organisme.

⁵² Institué par arrêté du Conseil d'Etat de la République et Canton du Genève du 25 mars 1991.

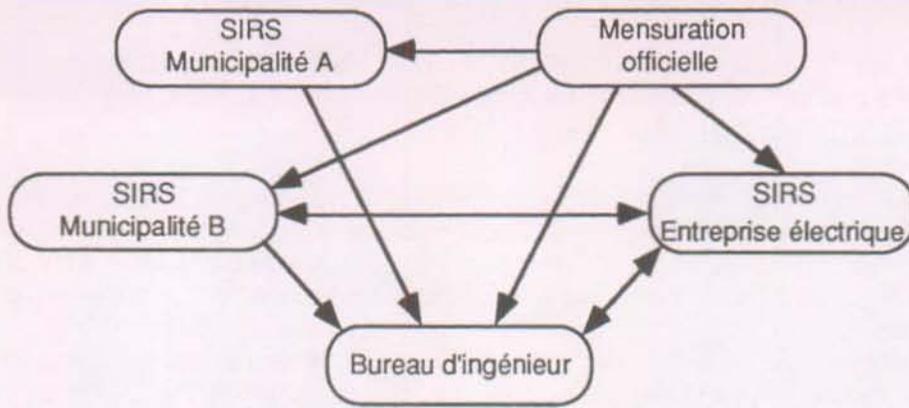


Figure 7.2: **SIRS fédérés dans le cadre d'un partenariat**

2. *Les SIRS fédérés autour d'un SIRS institutionnel.*

Une telle organisation s'articule autour d'un SIRS conçu et développé dans le but de constituer une mémoire collective relative à un territoire donné. Une telle approche par unité spatiale implique une organisation, une infrastructure et un budget ad hoc. Ce SIRS gère les données spatiales d'intérêt commun au profit des entreprises et services qui désirent s'y rattacher. Il s'agit donc d'un système d'information à vocation purement institutionnelle. Les traitements spécialisés restent décentralisés auprès des entreprises et administrations intéressées. La figure 7.3 illustre une telle organisation des SIRS.

A titre d'exemple, nous pouvons citer le système cantonal d'information du territoire (ScIT) vaudois, en voie de conception et de réalisation sous les auspices de l'Etat de Vaud. Un autre exemple en est la base géographique de la Communauté urbaine de Lyon (COURLY).

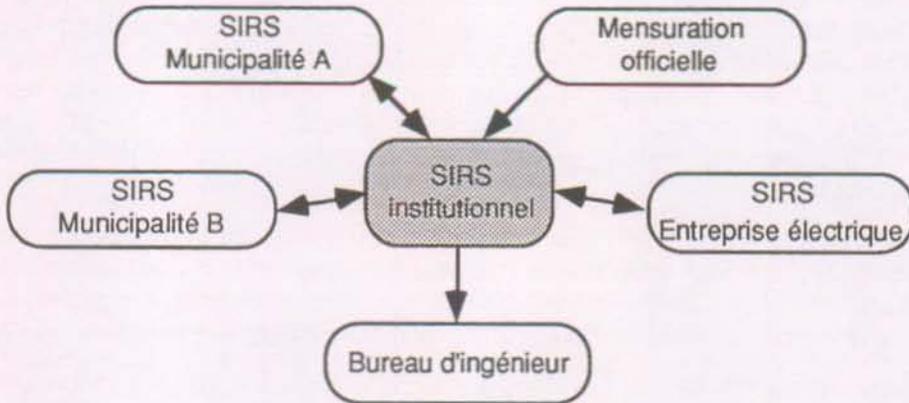


Figure 7.3: **Organisation des SIRS autour d'un SIRS institutionnel**

Ces deux modèles d'organisation – partenariat ou SIRS institutionnel – constituent tous deux une approche fédérée des SIRS d'un territoire donné. L'existence d'une organisation de coordination institutionnalisée distingue cependant le second modèle du premier. C'est à une telle organisation fédérée des SIRS que nous suggérons de réserver le nom de *système d'information du territoire (SIT)*, accréditant ainsi l'unicité généralement reconnue au SIT d'un territoire donné.

7.4. Systèmes de prise de décision à référence spatiale: des systèmes spécialisés exploitant et produisant des informations localisées

Parmi les critères de découpage d'une entreprise en domaines de gestion proposés au paragraphe 7.1, seuls les critères d'homogénéité du système opérant et du système d'information ont retenu notre attention jusqu'ici. Nous allons examiner dans ce paragraphe les conséquences liées au second critère: l'homogénéité du système de pilotage.

Ce critère implique que les activités regroupées dans un même domaine de gestion obéissent à une même logique de pilotage, à un même terme et à un même niveau stratégique des décisions prises. L'examen des activités d'une organisation liées à la référence spatiale nous permet de distinguer deux types d'activités très différentes quant à leur mode de pilotage:

1. Activités de gestion

Les activités de gestion des aménagements et des infrastructures s'inscrivent dans une démarche à long terme. Elles font appel à des tâches souvent répétitives, dont l'exécution rationnelle et rapide détermine en grande partie la valeur et le rendement de ces aménagements et infrastructures.

Les processeurs et les flux d'informations liés à ces activités sont essentiellement des consultations et des mises à jour de la mémoire d'entreprise (historique des interventions, état des canalisations, etc.). Ils définissent donc un *sous-système du système d'information d'entreprise*.

2. Activités de planification et de construction

Les activités de planification et de construction des aménagements et des infrastructures sont généralement réalisées dans le cadre de projets clairement finalisés et planifiés. L'élaboration d'une étude directrice d'aménagement, ainsi que le dimensionnement et la construction d'un réseau de canalisations en sont des exemples caractéristiques. Le délai de réalisation d'un projet est relativement court comparativement à l'impact des décisions prises. Le pilotage d'un projet n'obéit par conséquent pas aux mêmes règles que le pilotage d'une activité de gestion.

Les processeurs et les flux d'information liés à ces activités ont essentiellement pour rôle la prise de décisions complexes, étayées par l'agrégation de nombreuses informations. Ils définissent donc un *sous-système du système opérant de l'entreprise*, isomorphe du *système de prise de décision* présenté au chapitre 5.1.

Les informations nécessaires à un projet sont agrégées à partir du système d'information d'entreprise, dont la mémoire est une *base de données permanente (BDP)* au sens de [Chocat 1981]. La mémoire du système de prise de décision constitue la *base de données de travail (BDT)* définie par Chocat. Les informations agrégées et les décisions prises en cours de projet enrichissent progressivement la BDT. En fin de projet, les informations nécessaires pour des activités ultérieures sont consignées dans le système d'information d'entreprise.

Si nous considérons une répartition possible des tâches entre l'homme et la machine, nous constatons que des *banques de données* permettent aux spécialistes concernés (aménagistes, ingénieurs, etc.) de consulter et de mettre à jour la mémoire d'entreprise pour accomplir les activités de gestion décrites ci-dessus. Les activités de planification et de construction recourent quant à elles à des *systèmes d'aide à la décision* pour aider les experts du domaine concerné à agréger les informations et à inférer les décisions nécessaires à l'accomplissement de leurs missions. Les *modules spécialisés* décrits dans [Leclerc et Chevallier 1986] poursuivent des buts similaires. Les systèmes de simulation utilisés dans différents domaines de l'ingénierie en sont des exemples caractéristiques: calcul de réseaux d'eau potable, calcul de réseaux d'égouts, modèles de propagation du bruit, etc. On parlera de *systèmes d'aide à la décision à référence spatiale (SADRS)* [Chevallier

1992] dans les cas où la référence spatiale des informations joue un rôle déterminant dans le processus de prise de décision.

La distinction entre "systèmes d'information à référence spatiale" et "systèmes de prise de décision" se reflète aussi dans les compétences attendues de leurs concepteurs et de leurs exploitants, ainsi que dans le cahier des charges des outils informatiques utilisés:

- La conception et l'exploitation de *systèmes d'information à référence spatiale* implique l'engagement de personnel disposant de compétences particulières dans les techniques de modélisation et de gestion de données à référence spatiale.

Ces systèmes nécessitent en outre la mise en oeuvre de systèmes de gestion de données localisées (SGDL), assurant la mémorisation des données localisées et offrant une gamme de fonctions permettant en particulier une gestion fiable et efficace de ces données.

- La conception et l'exploitation des *systèmes de prise de décision à référence spatiale* exigent avant tout des compétences dans le domaine spécialisé considéré (aménagistes, ingénieurs, etc.).

Ces systèmes nécessitent la mise en oeuvre de systèmes informatiques permettant aux experts d'un domaine spécialisé d'appliquer et d'étayer leurs compétences. Dans le domaine de l'ingénierie, il peut s'agir de systèmes de conception assistée par ordinateur (CAO), tels que CASTOR [Lessard, Leahy et al. 1989]; des *systèmes experts à référence spatiale* sont aussi en cours de définition [Robert 1989]. Dans le domaine de la planification et de l'aménagement du territoire, des outils d'analyse spatiale tels que les *systèmes d'information géographiques (SIG)* sont fréquemment utilisés.

7.5. Synthèse et conclusion du chapitre: un modèle organisationnel global

Sur la base des critères de découpage d'une organisation en domaines de gestion, nous avons suggéré une organisation fédérée des SIRS sur un territoire donné. Cette organisation fédérée s'appuie soit sur un partenariat entre les organismes impliqués dans la gestion du territoire, soit sur la consignation de toutes les informations intéressant plusieurs partenaires dans un SIRS institutionnel. En outre, les *systèmes d'information à référence spatiale (SIRS)* doivent être clairement distingués des *systèmes de prise de décisions à référence spatiale (SPDRS)*.

La figure 7.4 propose un modèle global d'une organisation fédérée des SI, des SIRS et des SPDRS pour un territoire donné. On y recense différents organismes disposant d'un SIRS, parmi d'autres sous-systèmes d'information (une organisation sous forme de partenariat est aussi possible, comme l'a montré le § 7.3). Les informations d'intérêt commun sont gérées dans le cadre d'un SIRS institutionnel. Plusieurs partenaires mettent en oeuvre des systèmes de prise de décision destinés à la résolution de problèmes spécialisés.

Nous avons relevé au paragraphe 7.2 que le territoire représente un important *facteur intégrant* des informations qui s'y rapportent. Par conséquent, tout découpage du système en domaines de gestion implique des flux d'information importants entre les sous-systèmes considérés.

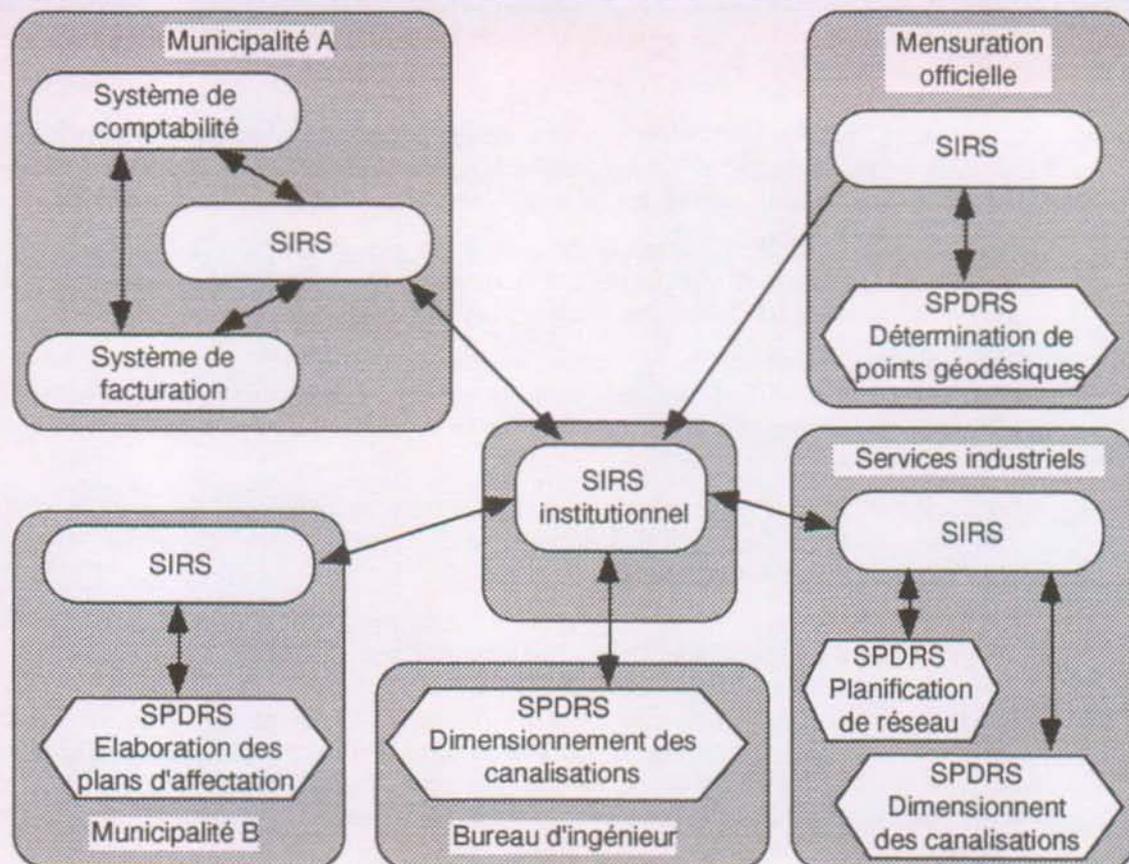


Figure 7.4: Modèle global d'organisation fédérée

Afin d'éviter que ces nombreux flux ne constituent des goulots d'étranglement rédhibitoires, chaque sous-système considéré doit respecter un ensemble de contraintes minimales. Ces contraintes sont (d'après [Chevallier 1983] et [Leclerc et Chevallier 1986]):

1. un modèle d'information commun, comprenant:
 - le catalogue des données gérées dans le système;
 - la définition univoque de l'ensemble de ces données;
 - des indications sur la qualité des données (méta-données) disponibles (précision, fiabilité, disponibilité, état de mise à jour, etc...), déterminant l'usage que l'on peut faire de ces données.

La méthode FRAMBOISE [Chevallier, Golay et al. 1989] destinée à l'intégration de nouveaux modules d'information dans un SIRS, ainsi que [Nyerges 1989], ont mis en évidence les difficultés pour fusionner différents modèles d'information conçus indépendamment les uns des autres.

2. des référentiels spatiaux communs s'appuyant sur des bases géodésiques uniques.
En Suisse par exemple, la mensuration officielle met à disposition des intéressés un système de points fixes matérialisés, dont les coordonnées sont connues dans le système de projection suisse.
3. des interfaces communes, susceptibles d'offrir des échanges d'information performants entre les systèmes considérés.

Différentes normes d'échanges d'informations localisées sont en voie de test ou de réalisation dans différentes régions du monde (SDTS aux USA, NTF au Royaume-Uni, EDIGEO en France, et INTERLIS en Suisse).

Le transfert d'informations d'un système à un autre, et en particulier la constitution d'une base de données de travail à partir d'une base de données permanente, peut nécessiter des changements des échelles de mesure des phénomènes considérés (voir § 5.2). Il peut alors s'agir de transformations *informationnelles* (dans le cas où l'expert doit apporter des compléments d'informations) ou de transformations *formelles* (dans le cas de transformations automatisables), conformément au cadre défini dans [Leclerc, Golay et al. 1989]. Ces transformations, conformes aux objectifs du système de prise de décision, sont réalisées par le sous-système d'agrégation des informations.



8. Application aux domaines d'utilisation considérés

Les chapitres 4 à 7 ont contribué à la définition d'un cadre conceptuel et organisationnel de la gestion et du traitement d'informations localisées. Il s'agit dans le présent chapitre de montrer en quoi ce cadre s'applique aux domaines d'utilisation spécialisés présentés au chapitre 3. En d'autres termes, nous voulons démontrer la *pertinence* du cadre proposé.

Nous examinerons en particulier la pertinence:

1. du paradigme du système de prise de décision;
2. du modèle d'organisation fédérée proposé;
3. des principes de modélisation des traitements proposés;
4. des principes de modélisation des données proposés.

Le domaine d'utilisation de l'*assainissement urbain*, à la base de notre étude, constitue l'objet essentiel de cette analyse (§ 8.1). Seuls quelques aspects significatifs des autres domaines d'utilisation seront mentionnés dans les paragraphes suivants.

8.1. Assainissement urbain

8.1.1. Application du paradigme du système de prise de décision à l'assainissement urbain

Nous avons constaté au chapitre 3 que les études de projets conduisent généralement à des décisions de haut niveau – *stratégique et tactique* –, alors que la mise en oeuvre de projets est tributaire de décisions à plus court terme, de niveau *opérationnel*.

Ce constat est illustré par la figure 8.1, qui représente le cheminement des agrégations d'informations et des prises de décision conformément au paradigme du système de prise de décision présenté au chapitre 5.

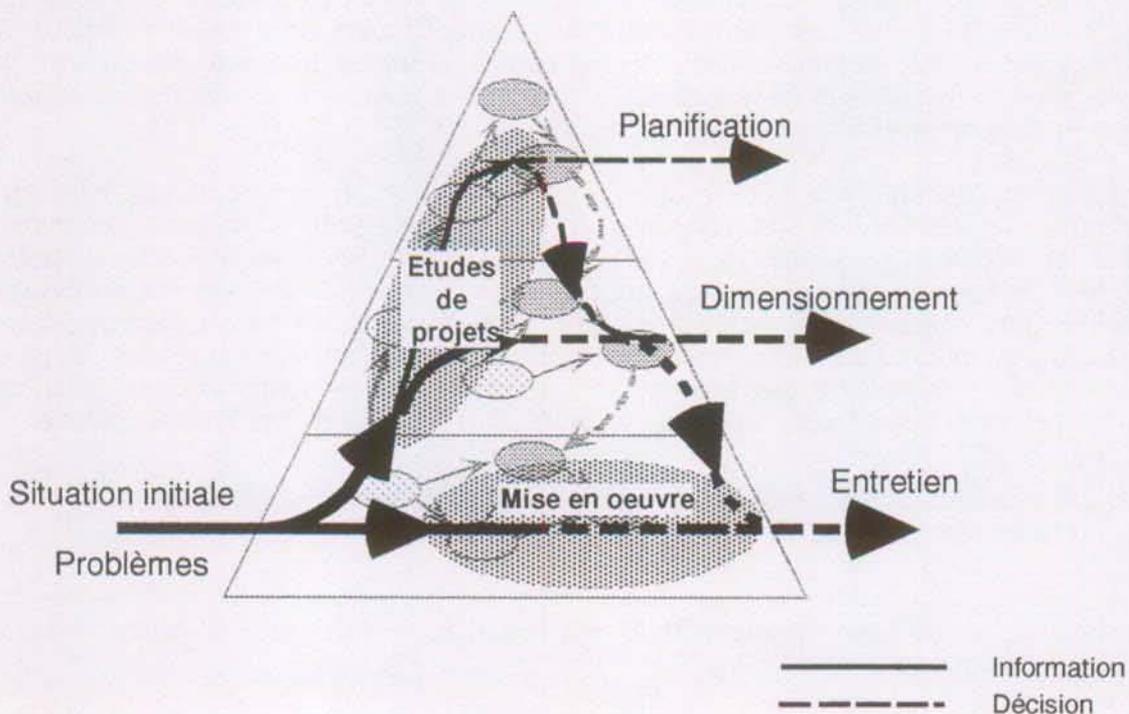


Figure 8.1: Prises de décisions dans le domaine de l'assainissement urbain

Les décisions relatives à la planification et au dimensionnement de canalisations en particulier se basent sur des agrégations complexes d'informations, qui impliquent des changements du niveau d'abstraction et d'importants apports informationnels. Le choix parmi les variantes et les scénarios élaborés peut aussi se révéler complexe, dans la mesure où la définition des critères de choix et leur pondération dépend du contexte propre à chaque projet. Les décisions des niveaux inférieurs doivent respecter les décisions de niveau supérieur. Dans le canton de Vaud par exemple, chaque commune doit tenir un *plan à long terme des canalisations*, auquel les réalisations ultérieures doivent se conformer.

L'épaisseur des flux représente le volume d'informations impliqués; celui-ci est d'autant plus faible que l'on s'approche du niveau stratégique.

8.1.2. Application du modèle d'organisation fédérée proposé à l'assainissement urbain

Nous avons évoqué au chapitre 3 le nombre important d'informations localisées utilisées par l'ingénieur urbain dans la plupart de ses activités.

En raison de la périodicité de l'activité de prévision [Chocat 1981], l'ingénieur doit disposer en permanence des informations nécessaires. Il doit donc disposer d'un *système d'information*, dans lequel soient mémorisées les informations qui lui sont utiles à moyen ou long terme et capable de les lui mettre à disposition sous une forme aisément interprétable. Ce système d'information intègre des informations localisées tout comme des informations non localisées, telles que des informations de planification de projets ou des listes de prix unitaires. Conformément au critère d'homogénéité du système d'information proposé au chapitre 7, on peut donc s'attendre à ce que l'ingénieur mette en oeuvre plusieurs sous-systèmes d'information. Parmi ces différents sous-systèmes figure en particulier un *système d'information à référence spatiale (SIRS)*. La gestion d'un tel SIRS nécessite non seulement des compétences dans le domaine de l'assainissement urbain, mais encore la maîtrise de la gestion d'informations localisées.

Pour chaque projet qu'il étudie, l'ingénieur réalise des agrégations d'informations conformes au paradigme du *système de prise de décision*. Il se base sur un sous-ensemble des informations disponibles dans le système d'information. Ces informations doivent généralement être complétées par des informations acquises spécialement pour le projet, auprès de fournisseurs particuliers ou par relevé "in situ". Ces informations sont utiles pour la durée du projet. Les résultats obtenus viennent progressivement les enrichir. A la fin du projet, les informations utiles à l'entretien des infrastructures réalisées ou à d'autres projets sont mémorisées dans le système d'information, alors que les autres informations du projet sont archivées.

L'organisation fédérée des SIRS proposée au chapitre 7 répond donc aux besoins de l'ingénieur urbain. Celui-ci doit disposer d'un réseau organisé de SIRS dans lesquels il va puiser les informations localisées qui lui sont nécessaires pour ses missions d'étude de projets et de mise en oeuvre. Alors qu'une exploitation directe du SIRS et des autres sous-systèmes d'information peut suffire aux activités de mise en oeuvre, qui font appel à des agrégations d'informations généralement élémentaires, les activités d'études de projets font généralement appel à des transformations et agrégations d'informations complexes, qu'il est pertinent de rattacher à des *systèmes de prise de décision à référence spatiale*.

8.1.3. Application des principes de modélisation des traitements proposés à l'assainissement urbain

Nous restreindrons notre analyse aux traitements d'informations nécessaires à la planification et au dimensionnement de réseaux d'assainissements à l'aide du modèle URBAN présenté au paragraphe 3.1.2.

Nous voulons premièrement replacer URBAN dans son contexte de modèle de simulation.

Nous avons constaté au paragraphe 5.3.3 qu'une simulation consiste à faire varier des intrants pour en évaluer les extrants. Nous avons aussi constaté que la nécessité de faire varier les intrants s'explique par une double incertitude sur les intrants d'une part et sur la sensibilité du comportement d'autre part. Cette double incertitude est caractéristique des modèles hydrologiques: les pluies, leur répartition spatiale, le taux d'infiltration, etc. sont des paramètres difficiles à estimer, et les conséquences d'une erreur d'estimation sur l'un de ces paramètres dépend en particulier de la structure du bassin versant et du réseau d'assainissement considérés.

En outre, dans son fonctionnement détaillé, URBAN, à l'instar d'autres modèles hydrologiques, paraît essentiellement *diachronique*: une simulation démarre à un instant t_0 pour se terminer généralement après un ou plusieurs épisodes pluvieux. Mais le domaine d'application de URBAN n'est pas le temps réel, mais un temps fictif. Le but de URBAN n'est pas de prévoir quelle quantité d'eau sera acheminée vers un exutoire à un instant donné, mais bien d'établir un diagnostic du dimensionnement existant ou prévu d'un réseau d'assainissement. L'ingénieur cherchera par exemple à évaluer le temps de retour d'une insuffisance du réseau pour une configuration donnée. URBAN est donc utilisé avant tout comme un modèle du comportement *synchronique* d'un réseau ⁵³.

Nous avons énuméré au paragraphe 3.1.2 (figure 3.1) les transformations informationnelles nécessaires à la préparation des intrants du modèle URBAN. Trois classes de transformations ont été définies: *schématisation*, *paramétrisation* et *discrétisation*. Nous constatons qu'il s'agit pour l'essentiel de *changements d'échelle de mesure des phénomènes*, conformément à la typologie des traitements proposée au paragraphe 5.2.2. Ces changements d'échelle de mesure peuvent impliquer simultanément des changements de *niveau d'abstraction*, de *modèle d'information* et de *référentiel spatial*. Conformément aux concepts développés dans cette étude, nous pouvons esquisser les définitions suivantes:

1. Nous définissons la schématisation comme un passage à un niveau d'abstraction supérieur de la représentation spatiale de phénomènes.

La schématisation répond néanmoins à des objectifs et à des critères sémantiques; ainsi, la schématisation de bassins versants en segments de ruissellement obéit-elle au critère de l'homogénéité globale des paramètres de pente, de rugosité, de taux d'infiltration, etc.

Afin qu'un modèle soit situé dans une plage d'abstraction homogène, conformément au postulat émis au paragraphe 4.4, la schématisation implique aussi souvent un *changement de référentiel spatial* correspondant à une précision plus faible, voire à une mesure des seules caractéristiques et relations topologiques. Ainsi en va-t-il des segments de transport et de ruissellement, dont seuls la connectivité et les paramètres sémantiques sont pertinents pour le calcul par le modèle URBAN.

2. Nous définissons la paramétrisation comme un passage à un niveau d'abstraction supérieur des caractéristiques sémantiques de phénomènes localisés.

La paramétrisation est une opération consécutive à la schématisation; elle consiste à attribuer une valeur globale à une caractéristique sémantique sur l'ensemble d'une composante spatiale résultant de la schématisation. A titre d'exemple, la paramétrisation de l'équation d'infiltration de Horton implique l'évaluation de deux coefficients caractérisant l'ensemble d'un segment de ruissellement.

⁵³ Il existe des modèles à vocation diachronique: on peut citer par exemple les modèles destinés au pilotage en temps réel d'un réseau d'assainissement, sur la base des pluies détectées par radar.

3. Nous définissons la discrétisation comme un passage à un niveau d'abstraction supérieur des caractéristiques temporelles de phénomènes localisés.

La discrétisation d'une pluie consiste par exemple à définir un nombre limité d'intervalles de temps pour chacun desquels la quantité de pluie sera évaluée et supposée d'intensité constante.

Nous constatons encore une lacune dans la classification des transformations informationnelles proposées dans [Leclerc, Golay et al. 1989]. La *schématisation des pluies* qui y est décrite ne correspond en effet pas à la définition que nous venons de donner du traitement de schématisation. Plus que de changer de niveau d'abstraction, il s'agit en effet de passer d'un modèle d'observations ponctuelles (défini au paragraphe 4.3.2), défini par la localisation des pluviographes et de dimension topologique nulle, à un modèle d'information en mosaïque, constitué de polygones de Thiessen et de dimension topologique égale à 2. Nous proposons par conséquent de définir une quatrième classe de transformations informationnelles:

4. Nous définissons la spatialisation comme un passage à un modèle d'information de dimension topologique supérieure.

La spatialisation d'un phénomène nécessite souvent un apport informationnel important, à l'exemple du pédologue qui délimite des zones pédologiques sur la base de quelques profils, auxquels s'ajoutent la connaissance de la topographie, de l'hydrographie, de la couverture du sol, etc.

La modélisation conceptuelle des traitements doit mettre en évidence ces transformations informationnelles, conformément au postulat 1 énoncé au paragraphe 5.1. La figure 8.2 présente un modèle conceptuel global des traitements ⁵⁴ d'un système d'agrégation d'informations faisant appel à URBAN.

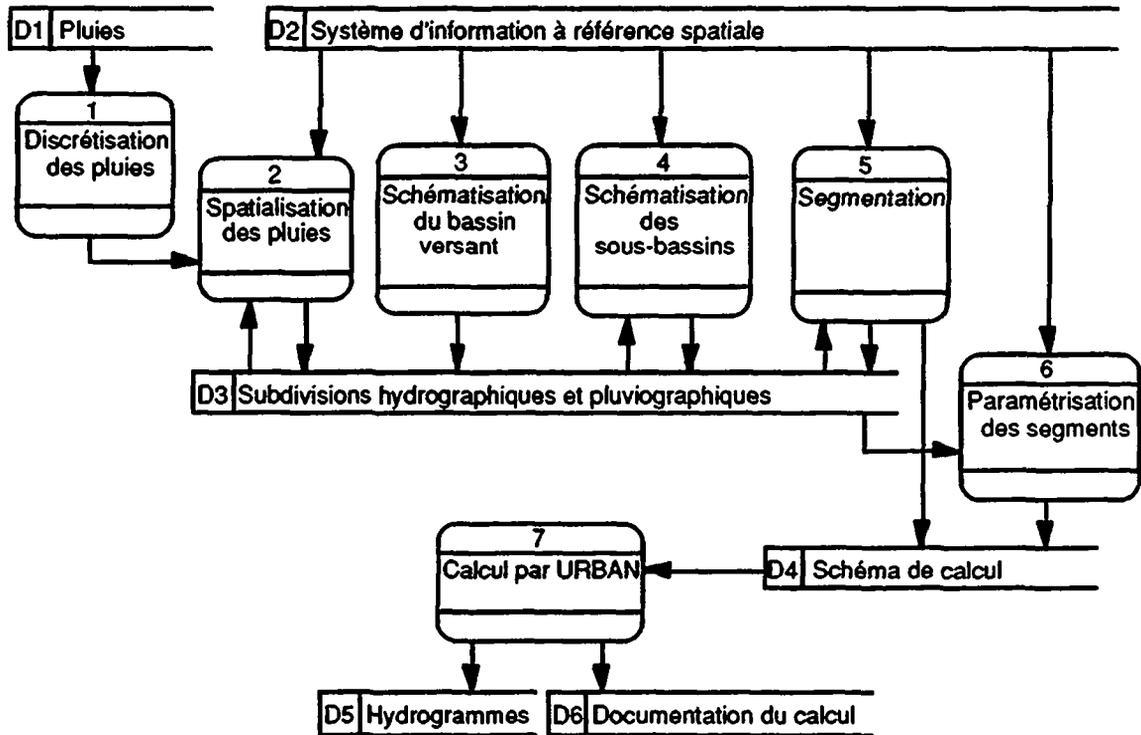


Figure 8.2: **Modèle de traitements de niveau conceptuel pour URBAN**
(DFD - entités externes négligées)

⁵⁴ Le formalisme utilisé est un diagramme de flux de données (DFD) conforme à [Gane et Sarson 1979].

8.1.4. Application des principes de modélisation des données proposés à l'assainissement urbain

Nous trouvons à la figure 8.3 un modèle organisationnel d'un système d'agrégation d'informations faisant appel au modèle hydrologique URBAN. Il ne se distingue d'un modèle de niveau conceptuel que par la représentation des associations gérées spatialement par un double trait.

Ce modèle met en évidence le rôle de la nouvelle composante sémantique *agrégat*. Les entités spatiales, telles que BASSIN VERSANT, SOUS-BASSIN, ELEMENT DE RUISSELLEMENT, etc., peuvent être mémorisées sans contrainte autre que le respect de leur dimension topologique. C'est au sein des agrégats que la plupart des contraintes spatiales sont définies.

Nous pouvons relever les points suivants:

- a. Chaque occurrence de l'agrégat SIMULATION regroupe toutes les données utilisées pour un calcul, ainsi que les résultats obtenus. Il est ainsi possible d'accéder en tout temps à l'ensemble des variantes de calcul réalisées, sans accumulation de données redondantes. Il s'agit d'un élément capital de méta-information.
- b. Les référentiels spatiaux suivants apparaissent dans ce modèle organisationnel:

- **subdivisions "hydrographie et assainissement" et "pluviographie"**

Dans la mesure où ces deux subdivisions sont caractérisées par des échelles de mesure et une qualité homogènes, elles peuvent être associées en un même référentiel spatial.

- **réseau d'assainissement**

Le réseau d'assainissement, résultant de la schématisation des canalisations existantes, est susceptible de constituer un référentiel indépendant. Néanmoins, ses caractéristiques sont similaires à celles du référentiel subdivisions "hydrographie et assainissement" et "pluviographie"; on peut donc envisager de les réunir en un seul référentiel spatial.

- **schéma**

Le schéma de connexion des conduites constitue un référentiel exclusivement topologique. Il serait néanmoins envisageable de gérer ces connexions sans interface graphique. Le rapport entre la convivialité apportée par l'interface graphique et le coût de l'investissement technique nécessaire constitue un critère décisif à appliquer lors de l'étape de modélisation organisationnelle.

- **topographie, hydrographie, couverture du sol et infrastructures**

Ce référentiel spatial regroupe les informations de base relatives à la nature et à l'équipement du territoire. Ces informations sont généralement gérées et mises à disposition par le biais de SIRS institutionnels.

Ce référentiel spatial est structurellement indépendant du reste du modèle; il lui est cependant étroitement lié par les traitements informationnels présentés au paragraphe 8.1.3.

- c. Ce modèle met aussi en évidence les lacunes liées au manque d'intégration des modèles de données et de traitements. Cette lacune est particulièrement sensible pour des systèmes d'agrégation d'informations, dans lesquels la structure des données est étroitement dépendante des traitements qui la sous-tendent.

Une meilleure intégration des formalismes utilisés doit permettre de réaliser des modèles représentant aussi bien les types d'objet à référence spatiale que la nature des mécanismes d'abstraction (généralisation, agrégation, ...) mis en oeuvre pour passer d'un type d'objet à un autre.

8.2 Aménagement du territoire

Le domaine de l'aménagement du territoire nous paraît particulièrement pertinent pour illustrer le paradigme du système de prise de décision. L'aménagement du territoire constitue en effet un univers très complexe, dans lequel les niveaux stratégique, tactique et opérationnel sont très interdépendants.

La figure 8.4 place les principales activités de l'aménagement du territoire sur la pyramide de décision, et met en évidence les principaux flux d'informations. On constate que l'aménagement du territoire constitue par essence un système de prise de décision.

- Au niveau stratégique sont définis les principes de l'aménagement; ils doivent répondre aux objectifs politiques de l'aménagement (densification ou extension des villes, industrie ou logement, etc.), et conduisent à l'élaboration de *plans directeurs* d'aménagement.
- Au niveau tactique, les principes d'aménagement définis dans les plans directeurs sont traduits en lois et normes, qui prennent la forme de *plans d'aménagement* aux spécificités et dénominations diverses: plans d'occupation du sol, plans de zones, plans d'affectation, etc.
- Au niveau opérationnel, les normes d'aménagement sont appliquées et interprétées pour délivrer – ou refuser – les permis de construire requis par les intéressés.

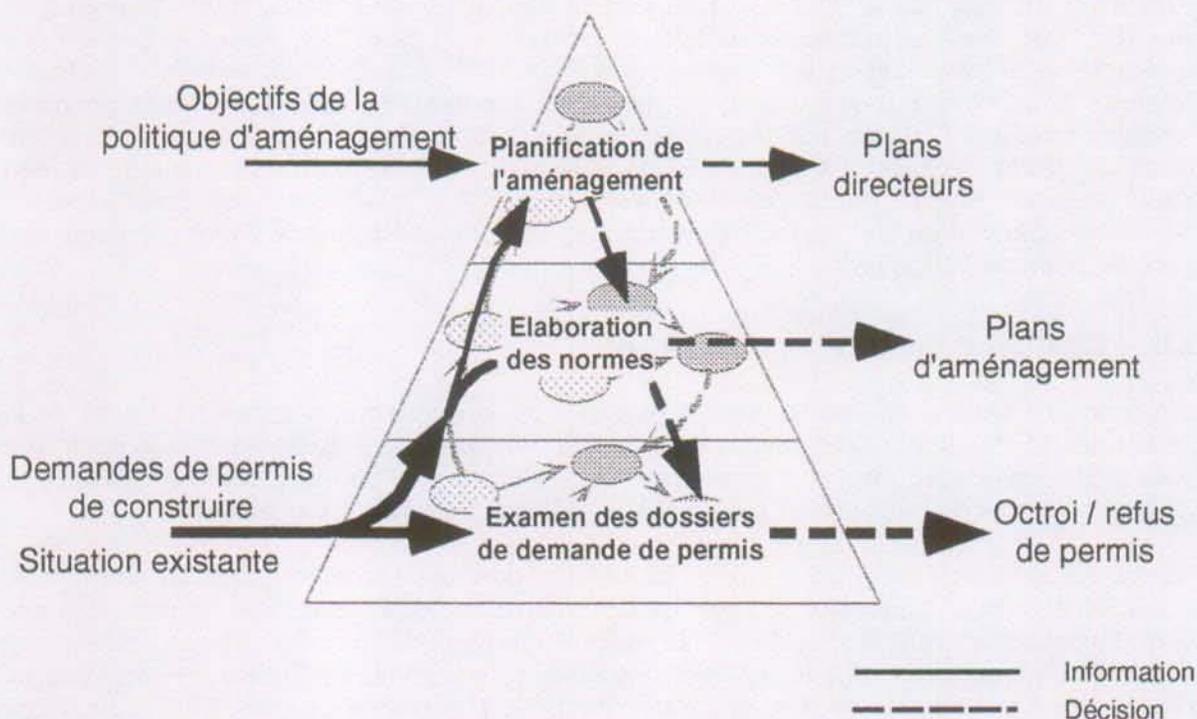


Figure 8.4: Prises de décisions dans le domaine de l'aménagement du territoire

Ce modèle de l'aménagement du territoire a quelques implications qui méritent d'être relevées dans le cadre de cette étude:

- Les plans d'aménagement ont une vocation tactique, et non opérationnelle. Ils sont donc généralement conçus à moyenne échelle, entre le 1:1'000 et le 1:5'000 dans la plupart des cas. La prise de décisions opérationnelles (octroi ou refus de permis de construire) exige donc une *interprétation* du plan d'aménagement pour chaque cas d'espèce. Il s'agit en particulier de délimiter les assiettes des restrictions de droit public à la construction avec une précision égale à celle des assiettes des droits réels fonciers [Chevallier 1983]. Afin de garantir une sécurité élevée du droit, il serait souhaitable que cette interprétation du droit soit effectuée

de manière systématique lors de l'élaboration ou de la révision d'un plan d'aménagement. Il semble malheureusement que le coût de telles opérations serait dans de nombreux cas sans commune mesure avec le bénéfice attendu.

- L'élaboration de plans directeurs et de plans d'aménagement constituent des études de projets dont le déroulement est comparable à la planification et au dimensionnement des réseaux d'assainissement discutés au paragraphe 8.1.2. Cette activité nécessite des agrégations d'informations et des choix complexes, relevant des systèmes de prise de décision. L'examen des demandes de permis de construire relève quant à lui de la mise en oeuvre des projets d'aménagement. La consultation d'informations simples, qui peuvent être fournies directement par le système d'information, se révèle suffisante pour la plupart des cas.
- La figure 8.4 montre que l'élaboration des plans d'aménagement et des plans directeurs prend en compte l'état existant. Il s'agit en particulier de tenir compte des normes déjà en vigueur, ainsi que des projets à l'étude ou en cours de procédure. L'aménagiste doit donc disposer d'informations très complètes sur les informations existantes. La *méta-information* joue ainsi un rôle capital en aménagement du territoire.

En conclusion à ce paragraphe relatif à l'aménagement du territoire, risquons une remarque relative aux délais d'octroi des permis de construire. Ceux-ci font dans de nombreuses régions et localités de Suisse le désespoir des entrepreneurs et promoteurs immobiliers. Le dépôt de demande de permis de construire dans des zones "sensibles" peut ainsi avoir un effet "gigogne": l'octroi du permis peut nécessiter l'élaboration d'un plan de quartier, qui peut lui-même nécessiter la révision du plan des zones... Est-il donc envisageable de simplifier de telles procédures ? Il n'entre pas dans le cadre de notre étude de nous prononcer sur les exigences de l'aménagement du territoire. Nous pouvons cependant déduire de la figure 8.3 que pour mettre en oeuvre un aménagement du territoire cohérent, il est nécessaire que les décisions s'enchaînent harmonieusement du niveau stratégique au niveau opérationnel. Il s'agit donc de choisir entre un aménagement rationnel, mais coûteux, et un aménagement bon marché, mais peu cohérent. Ce choix est, bien sûr, d'ordre politique !

8.3. Gestion de l'entretien routier

La gestion de l'entretien routier peut être située au niveau intermédiaire (tactique) de la pyramide de décision. Elle consiste à mettre sur pied des programmes d'entretien (goudronnages, inspections et entretien d'ouvrages d'art, etc.) conformément à des options stratégiques déjà définies (priorité accordée à l'entretien préventif par exemple).

Comme nous l'avons mentionné lors de la présentation de la problématique au chapitre 3, la gestion de l'entretien routier se base sur des informations localisées par rapport à un axe de maintenance suivant le parcours de la route. La plupart de ces informations doivent être conservées à long terme, et sont utilisées périodiquement pour la planification des travaux d'entretien. *Les banques de données routières (BDR) présentées au chapitre 3 constituent donc le noyau d'un SIRS destiné à la gestion de l'entretien routier.*

Cas particulier de BDR, le logiciel STRADA est assorti des fonctions nécessaires à la gestion des informations routières [Petersen, Marschal et al. 1991]. Des fonctions d'exploitation sous forme de questions élémentaires ou complexes sont prévues pour de prochaines étapes de réalisation. Le concept de STRADA prévoit en outre l'intégration des BDR de chaque canton au sein d'une organisation fédérée sous la responsabilité de l'Office fédéral des routes, afin de faciliter l'échange des informations nécessaires à la planification générale, à la coordination et à la gestion comptable de l'entretien des routes nationales.

Mais l'interaction de ces BDR avec les SIRS d'autres domaines dans le cadre d'une organisation fédérée est encore à l'étude ([INSER SA, CIGC-EPFL et al. 1992] en particulier). Nous évoquerons ici trois cas dans lesquels une telle interaction serait nécessaire:

1. la constitution et la mise à jour de la BDR sur la base des informations disponibles dans d'autres SIRS, à l'exemple du système de repérage, de la géométrie de la route (succession de segments de droites, d'arcs de cercles, de clothoïdes par exemple), des lieux de noeuds (position des intersections d'axes) et des caractéristiques géométriques de la route (largeur par exemple).
2. la prise en compte d'informations de niveau opérationnel (localisation et état des canalisations de drainage par exemple) dans les programmes d'entretien et la mise en oeuvre au niveau opérationnel des programmes d'entretien élaborés (programme d'intervention et localisation des installations à inspecter par exemple).
3. la visualisation et la gestion des informations sur la base d'un référentiel géographique, ainsi que la production cartographique.

L'agrégation d'informations provenant de ces divers SIRS est complexe. Elle fait appel à des référentiels de nature fondamentalement différente (géographique et curviligne), ainsi qu'à des modèles d'information différents. Il nous paraît donc opportun de considérer un tel système de gestion de l'entretien routier comme un *système de prise de décision à référence spatiale*, distinct des SIRS (BDR ou autre) dans lesquels il puise ses informations de base et retourne les résultats pertinents à long terme.

Les traitements nécessaires à l'intégration des informations des BDR et des SIRS classiques sont essentiellement des *changements de référentiel spatial* (d'un système de repérage curviligne à un système de repérage géographique et réciproquement). Ces changements de référentiel spatial sont généralement accompagnés de *changements du niveau d'abstraction* (du niveau opérationnel au niveau tactique et réciproquement). Le changement de référentiel spatial devrait pouvoir être aisément automatisé. Le changement de niveau d'abstraction implique par contre un apport informationnel important, pour lequel l'intervention d'un expert paraît indispensable.

8.4. Analyse statistique géographique

L'analyse statistique géographique a pour objectif l'agrégation et la présentation d'informations aux décideurs (politiciens, chefs d'entreprise, etc.), qui auront la responsabilité de résoudre les problèmes mis en évidence [IGUL et EPFL-IM-GM 1992]. Elle s'inscrit donc, selon le paradigme du système de prise de décision présenté au chapitre 5, à l'intérieur d'un *système d'agrégation d'informations à référence spatiale*.

Les informations utilisées par le géographe pour entreprendre une analyse sont essentiellement des données de recensements et d'inventaires de diverses époques, localisées sur le territoire. Ces informations ont donc un triple volet sémantique, spatial et temporel. Elles sont mémorisées à long terme dans des systèmes d'information statistiques ou à référence spatiale, d'où le géographe peut les extraire. L'analyse spatiale est donc particulièrement tributaire d'une organisation fédérée efficace des systèmes d'information – et des SIRS en particulier.

Les limites des méthodes et outils anciens ont donné de l'analyse statistique géographique une image de productrice de cartes thématiques en tous genres. En outre, la difficulté de réaliser des agrégations spatiales flexibles a conduit les géographes à rapporter leurs informations à un nombre restreint d'entités spatiales clairement définies et délimitées, à l'exemple des communes. Les recherches actuelles tendent à éliminer les conséquences de ces restrictions méthodologiques et technologiques:

- la carte n'est plus considérée comme un produit, mais comme un support d'interaction.

- grâce à la localisation détaillée de nombreuses informations de recensement, des agrégations variées peuvent être réalisées en fonction des besoins de l'analyse.

Cet enrichissement fonctionnel de l'analyse statistique géographique induit une complexité croissante des traitements appliqués. Il est donc indispensable de conserver la trace de ces traitements, en associant à chaque donnée mémorisée la référence au traitement qui l'a engendrée.

Les traitements de l'analyse statistique géographique couvrent l'ensemble des types recensés au chapitre 5. Mentionnons en particulier:

- l'analyse de séries temporelles, qui conduit à la définition d'un modèle du comportement diachronique du phénomène considéré et permet d'en déduire la trajectoire, puis d'en interpoler ou extrapoler la valeur pour différentes époques.
- les changements de niveaux d'abstraction, qui peuvent impliquer l'agrégation de deux ou plusieurs variables statistiques, la reclassification d'une variable ou l'agrégation d'entités spatiales.
- les changements de type de modèle d'information, tel que le passage d'une grille statistique (carroyage) à des secteurs statistiques (mosaïque).

8.5. Cartographie

A l'instar du domaine de l'analyse statistique géographique, la cartographie s'inscrit dans des *systèmes d'agrégation d'informations à référence spatiale*. Les cartes élaborées sont destinées à des décideurs, des concepteurs, etc., dont les besoins en informations déterminent les caractéristiques de chaque carte (contenu, échelle, graphisme).

On distinguera sur la pyramide de décision les cartes à vocation opérationnelle, telles que les cartes topographiques nationales, peu finalisées et pour lesquelles la précision géométrique constitue une exigence importante, des cartes thématiques à vocation stratégique, très finalisées et dont la géométrie ne représente qu'un support de communication, à l'exemple de l'Atlas de la Suisse.

Comme dans les autres domaines présentés dans ce chapitre, la cartographie se base sur les informations mémorisées dans des systèmes d'information à référence spatiale. Ces informations sont structurées conformément aux exigences propres du système d'information. La structure des données de la carte élaborée doit au contraire être axée sur la représentation cartographique: symboles, traits de différentes textures, trames et hachurages divers, et écritures.

Les traitements d'un système spécialisé de cartographie ont donc pour objectif la sélection et la mise en valeur d'informations pertinentes pour la prise de décisions. Cette mise en valeur nécessite une transformation formelle de la structure de données du SIRS vers la structure de données cartographiques [Chevallier, Golay et al. 1989]. Mais elle fait aussi appel à l'expertise du cartographe [Bertin 1967; Bertin et Bonin 1977] pour assurer la sélection des informations à représenter et pour réaliser des transformations informationnelles telles que des généralisations et agrégations spatiales, des changements de référentiel spatial, etc.

Certaines informations intégrées à un produit cartographique peuvent être conservées pour faciliter l'élaboration ultérieure de produits similaires. On peut mentionner en particulier la localisation des écritures sur une carte, dont la détermination est fastidieuse et difficilement automatisable.

A titre d'exemple, notons la pratique fréquente en Suisse d'utiliser des logiciels de dessin tels que Auto CAD pour élaborer des produits cartographiques extraits des bases de données de la mensuration officielle.

8.6. Conclusion du chapitre

Ce chapitre a permis de montrer que les concepts théoriques développés dans les chapitres 4 à 7 trouvent leur application dans plusieurs domaines spécialisés d'utilisation d'informations localisées.

En particulier, le paradigme du système de prise de décision, composé des sous-systèmes d'agrégation des informations et d'intégration des décisions, semble particulièrement pertinent pour décrire ces domaines spécialisés et leurs contraintes de mise en oeuvre.

Nous avons déjà noté le rôle de *facteur intégrant* joué par le territoire entre les systèmes d'information et leurs domaines spécialisés d'utilisation. L'organisation fédérée des SIRS est une réponse efficace à cette contrainte.

La typologie des traitements proposée constitue un canevas dans lequel les traitements propres à chaque domaine peuvent être situés. Le formalisme de description des structures de données n'a été testé que pour la mise en oeuvre d'un modèle hydrologique particulier. Sa validation et son affinage nécessiteraient encore d'autres tests, qui pourront faire l'objet d'études ultérieures.

Certaines conséquences plus générales de notre étude n'entraient pas dans le canevas du présent chapitre. Elles sont l'objet du prochain chapitre 9.

9. Implications de l'étude sur les méthodes de conception et les outils informatiques

Nous avons identifié quelques conséquences générales de notre étude, qui s'appliquent aussi bien aux systèmes d'information qu'aux systèmes de prise de décision à référence spatiale.

Ce chapitre recense ces implications et les regroupe en trois sous-chapitres:

1. implications sur les méthodes de conception des systèmes d'information et de prise de décision à référence spatiale;
2. implications sur l'exploitation des systèmes d'information et de prise de décision à référence spatiale;
3. implications sur les outils technologiques destinés à la gestion et au traitement d'informations localisées.

9.1. Implications de l'étude sur les méthodes de conception des systèmes d'information et de prise de décision à référence spatiale

Les chapitres 5 et 6 constituent une forme de *cahier des charges* pour la modélisation des traitements et des données dans des systèmes d'agrégation d'informations à référence spatiale. Des suggestions pour la description formelle des structures de données sont encore présentées au chapitre 6.

Un "mode d'emploi" des formalismes proposés est en outre associé à ces différents chapitres (chapitres 5.4, 5.5, 6.4, 6.5 et 6.6). A l'Université Laval, Bédard conçoit et développe un outil intégré d'analyse et de conception de SIRS (Integrated Relational Dictionary System- IRDS), auquel une démarche d'utilisation est associée. Mais il manque encore – dans cette étude comme dans la littérature actuelle – une véritable *méthode d'analyse et de conception* des systèmes gérant et traitant des informations localisées, comparable à des méthodes telles que MERISE ou IEM dans le domaine de l'informatique de gestion.

Faute d'une méthode adaptée à leurs besoins, les concepteurs de SIRS et de leurs applications spécialisées adaptent tant bien que mal les méthodes traditionnelles aux particularités de la référence spatiale. Mais la mise au point d'une méthode complète et cohérente d'analyse et de conception de "systèmes à référence spatiale" est attendue impatiemment par les sociétés de services actives dans ce domaine.

Notre étude nous permet cependant de formuler quelques contraintes qu'une telle méthode et ses outils associés devra respecter:

a. Approche par les traitements des systèmes de prise de décision

Les démarches préconisées à ce jour pour l'analyse et la conception de systèmes à référence spatiale ont généralement privilégié les approches par les données.

Une telle approche est sans doute pertinente pour les SIRS, mémoires du territoire faisant principalement appel à des traitements simples et typés (mise à jour, recherche sur critères, etc.). L'approche par les données est par contre mal adaptée à la conception de systèmes d'agrégation d'informations et de prise de décision, clairement finalisés, et dont les traitements peuvent se révéler complexes et variés.

Une méthode de conception générale des systèmes à référence spatiale devra donc proposer les deux types d'approche, que le concepteur choisira conformément à la nature du système auquel il sera confronté.

b. Intégration des modèles de traitements et de données

La modélisation d'un système à référence spatiale par des modèles de données et de traitements totalement distincts néglige les interdépendances étroites qui peuvent exister entre les données et leur "comportement". A titre d'exemple, la modélisation d'un système de calcul de réseau d'assainissement, tel qu'il a été esquissé au chapitre 8.1, gagnerait sans doute en clarté si les traitements transformant les données pouvaient être superposés au modèle des données (figure 8.3). Les traitements contribuant à des changements de niveau d'abstraction pourraient en particulier être mis en évidence.

Mais l'intégration des données et des traitements reste encore très lacunaire dans les outils de conception de systèmes actuels (ateliers de génie logiciel). A ce jour, les orientations suivantes peuvent être dégagées:

1. la plupart des méthodes de conception qui proposent des outils de modélisation distincts pour les données et les traitements tentent de garantir la cohérence de ces deux modèles par *validation* réciproque [Tardieu, Rochfeld et al. 1984];
2. certains auteurs proposent d'ajouter la notion d'activité au modèle de données [Caron 1991] ou d'isoler des comportements propres à certaines entités dans les DFD [Ward 1990], dans la perspective d'implémenter le modèle conceptuel défini dans un système informatique orienté objets.

Le développement de formalismes permettant d'intégrer dans un même schéma les structures de données et les traitements serait donc particulièrement opportun pour la conception de systèmes à référence spatiale.

c. Maîtrise de la complexité par les *hyperschémas*

Le formalisme MODUL-R [Caron 1991], destiné à la description de structures de données à référence spatiale, permet de modéliser un grand nombre de concepts. Pour maîtriser la complexité des modèles réalisés, il définit cinq niveaux d'abstraction successifs, allant d'une description globale et générale du système à une description très détaillée mais limitée à des sous-systèmes.

Les extensions que nous avons définies à ce formalisme, ainsi que l'intégration des structures de données et des traitements que nous avons suggérée au point b ci-dessus, pourraient aboutir à la création de modèles trop complexes pour être aisément interprétables.

Nous proposons donc de ne pas limiter les produits de la conception à quelques modèles au contenu défini de manière rigide (modèles conceptuels de données, modèles conceptuels cartographiques, modèles conceptuels de traitements, modèles organisationnels de tous types, etc.), mais de ne normaliser que le catalogue des composantes sémantiques à définir lors de chaque étape de la conception. Il n'existerait par conséquent plus qu'un modèle, "le" modèle du système, dont l'ensemble des composantes serait défini dans le dictionnaire du système.

Pour présenter ce modèle au concepteur et lui permettre de le développer et de le tenir à jour, les composantes définies dans le dictionnaire pourraient être présentées sous formes de schémas de nature et de contenu flexibles, adaptés aux besoins spécifiques de chaque système et de chaque étape de conception.

Des liens croisés entre les différents schémas définis devraient en outre permettre au concepteur une navigation intuitive et aisée d'un schéma à l'autre. Cette analogie avec les hypertextes et les logiciels navigationnels du type de HyperCard nous conduit à nommer *hyperschémas* ces vues flexibles du modèle d'un système à référence spatiale.

9.2 Implications de l'étude sur l'exploitation des systèmes d'information et de prise de décision à référence spatiale

a. Rôle et importance de la méta-connaissance

La plupart des domaines spécialisés considérés dans notre étude couvrent plusieurs niveaux de la pyramide de décision (stratégique, tactique et opérationnel). Si à un niveau donné, le détail des informations du niveau inférieur n'est évidemment pas pertinent, il est par contre généralement nécessaire d'avoir une vue d'ensemble de ces informations de niveau inférieur. Un exemple caractéristique en est le besoin de connaître la localisation des demandes de permis de construire déjà traitées ou en cours de procédure lors de l'élaboration de plans d'aménagement.

Ainsi, la prise de décisions se base sur une connaissance agrégée des informations de niveau de décision inférieur. Cette méta-connaissance est particulièrement importante dans les systèmes de prise de décision, et le concepteur lui prêtera une attention particulière lors de la modélisation.

b. Echanges d'informations et compatibilité des modèles

La structure fédérée des SIRS préconisée au chapitre 7 implique de nombreux échanges d'informations entre des systèmes dont les modèles de données peuvent être très différents.

Il ne suffit donc pas de régler les problèmes organisationnels et techniques des échanges, comme le permettent les interfaces normalisées en cours de développement dans de nombreux pays (SDTS aux Etats-Unis, EDIGEO en France, NIF en Grande-Bretagne, INTERLIS en Suisse, etc.). Il faut surtout assurer la cohérence des modèles d'informations entre les systèmes impliqués dans un échange.

[Bezos 1990] propose l'élaboration d'un modèle global, dont chaque système constitue une vue externe. [Chevallier, Golay et al. 1989] préconise une autre solution, qui consiste à injecter dans le modèle-but les structures du modèle-source impliquées dans l'échange. Cette dernière solution, sans doute plus "minimaliste", est néanmoins plus facile à mettre en oeuvre dans un réseau complexe de systèmes fédérés.

c. La gestion et le traitement des informations à référence spatiale: une formation fondamentale pour tous les gestionnaires du territoire

Notre étude a montré que la gestion et le traitement des informations à référence spatiale concerne tous les professionnels impliqués dans la gestion du territoire, que ce soit à titre de fournisseurs d'informations (SIRS) ou d'utilisateurs spécialisés (systèmes de prise de décision à référence spatiale).

Une formation solide en matière de systèmes d'information et de systèmes de prise de décision à référence spatiale devrait par conséquent faire partie de tout programme de formation dans un domaine relevant de la gestion du territoire.

9.3. Implications de l'étude sur les outils technologiques destinés à la gestion et au traitement d'informations localisées

a. Rôles respectifs des SGDL et de la CAO

Les systèmes d'information à référence spatiale et les systèmes de prise de décision à référence spatiale (ou plus spécifiquement les systèmes d'agrégation d'informations localisées) ont des exigences très différentes sur les systèmes informatiques à mettre en oeuvre:

- a. Les systèmes d'information à référence spatiale doivent assurer une gestion à long terme des données mémorisées.

Ils doivent donc offrir une sécurité d'exploitation maximale, tributaire de mesures de sécurité et de protection des données sévères. Les mutations peuvent en particulier s'étendre sur un laps de temps important, durant lequel elles doivent pouvoir être purement et simplement annulées.

Les systèmes de gestion de données localisées (SGDL) répondent à un tel cahier des charges. Le coeur de ces systèmes est constitué par un système de base de données à référence spatiale, intégrant notamment des structures géométriques prédéfinies et un langage de manipulation des données spatiales assorti d'un set minimal de fonctions permettant la manipulation graphique interactive des données.

- b. Les systèmes d'agrégation d'informations localisées ne sont pas soumis aux mêmes exigences de sécurité des données que les SIRS. Ils doivent avant tout offrir à l'ingénieur ou à l'expert une interface flexible, lui permettant d'exprimer son expertise sans contrainte inutile ni opération superflue [Buogo 1990].

Ce cahier des charges correspond à celui d'un système de conception assistée par ordinateur (CAO). De tels systèmes ont déjà été créés dans divers domaines spécialisés utilisant des informations localisées, à l'exemple du système CASTOR développé à l'Ecole polytechnique de Montréal [Lessard, Leahy et al. 1989].

L'utilisation d'un SGDL pour réaliser un système d'agrégation d'informations a été testée par Buogo. La flexibilité des structures de données permet généralement de réaliser rapidement la base de données et d'importer les informations de base à partir de SIRS. Mais les développements fonctionnels restent en général difficiles, et les performances souvent limitées des éditeurs graphiques auront tôt fait de renvoyer l'ingénieur intéressé à ses règles et ses plumes !

La mise en place d'un système adéquat de CAO peut par contre avoir des conséquences favorables sur les méthodes de travail de l'ingénieur. Déchargé des tâches fastidieuses de dessin et de calcul, il pourra enrichir aisément le jeu de variantes étudiées.

b. Référentiels spatiaux multiples

Néanmoins, aucun système informatique actuellement disponible sur le marché ne semble à même de répondre à une exigence clairement posée par cette étude: la gestion de plusieurs référentiels spatiaux au sein d'une même base de données. La complexité croissante des systèmes à référence spatiale et la généralisation des systèmes à plusieurs niveaux de décision vont pourtant inévitablement renforcer cette exigence.

En attendant l'apparition sur le marché de solutions technologiques adéquates, les concepteurs de systèmes devront continuer à chercher des solutions organisationnelles satisfaisantes. Un exemple en est l'interdiction de l'affichage simultané à l'écran de "couches" appartenant à des référentiels spatiaux différents.

10. Conclusion

Conformément à l'objectif général posé dans notre introduction, nous avons tenté de définir un cadre général permettant aux ingénieurs et aux gestionnaires du territoire d'intégrer de manière efficace et cohérente les SIRS à leurs propres réalisations.

Cette conclusion fait la synthèse des principaux résultats obtenus. Elle met en particulier en évidence les apports de notre étude à la problématique abordée. Elle souligne aussi les amendements qu'elle suggère ou implique dans la manière de concevoir et d'exploiter des systèmes gérant et traitant des informations localisées dans des domaines spécialisés de la gestion du territoire.

La démarche fondamentalement globale et systémique qui a été adoptée dans cette étude implique que de nombreux problèmes soient soulevés et placés dans une perspective générale cohérente. Les problèmes peuvent ensuite être mieux décrits et formalisés, puis résolus dans le respect de leurs interdépendances. Les problèmes soulevés par une telle démarche sont cependant souvent trop abondants et complexes pour qu'on puisse envisager de les résoudre tous et immédiatement !

Les problèmes identifiés mais non résolus dans le cadre de notre étude sont donc finalement passés en revue. Des perspectives et des suggestions de recherches ultérieures sont évoquées, et la nécessité d'une bonne coordination académique entre ces recherches soulignée.

10.1. Résultats de l'étude

Les résultats les plus significatifs de notre étude sont passés en revue ci-dessous. Ils sont classés selon les trois facettes du "triangle des SIRS" présentés dans notre introduction: les méthodes, l'homme et la technologie.

a. Résultats rattachés à la facette *méthodes*

- Un modèle des systèmes de prise de décision a été développé (chapitre 5).

Ce modèle prend en compte les différents niveaux de décision (stratégique, tactique, opérationnel). Il évite par contre soigneusement la distinction d'ordre purement organisationnel entre les rôles respectifs de l'homme et de l'ordinateur. Il s'avère pertinent pour décrire et analyser les activités des différents domaines spécialisés considérés dans notre étude, et modéliser les flux d'informations et de décisions intrantes et extrantes.

Ce modèle montre en particulier que les ingénieurs et gestionnaires du territoire sont des utilisateurs importants des informations des SIRS. Mais c'est conformément au paradigme des systèmes de prise de décision que les ingénieurs et les gestionnaires du territoire conduisent leurs études et appliquent leur expertise. Les SIRS constituent néanmoins des mémoires institutionnelles indispensables à la mise en oeuvre de systèmes de prise de décision.

- Sur cette base, une classification des transformations et des traitements d'informations mis en oeuvre dans les systèmes d'agrégation d'informations localisées est proposée (chapitre 5). Cette classification met en évidence l'importance des mécanismes d'abstraction sémantiques et spatiaux, qui sont étudiés au chapitre 4.

- Les traitements étudiés peuvent impliquer des changements de niveau d'abstraction, de modèle spatial ou de référentiel spatial. Ils peuvent encore représenter l'évolution d'un modèle dans le temps. Ces notions sont développées au chapitre 4:
 - La notion de référentiel spatial, entièrement nouvelle, constitue une description formelle de contraintes intuitives bien connues, telles que l'incohérence de mesurer des distances entre des objets localisés avec une précision très hétérogène.
 - Une classification des types de *modèles spatiaux* est proposée, qui se distingue des classifications traditionnelles basées sur la distinction des technologies "vecteur" et "raster".
 - Le rôle de la *topologie*, souvent considérée dans la littérature comme un simple auxiliaire technologique, est redéfini comme un lien informationnel entre les objets spatiaux et les autres types de modèles spatiaux (mosaïques et réseaux en particulier).
 - le concept de *modèle* est généralisé de manière à englober aussi bien la notion de *modèle observé*, provenant du domaine des systèmes d'information, que celle de *modèle de comportement*, provenant des calculs de simulation en ingénierie.
- Quelques enrichissements sémantiques des formalismes de description des structures de données sont proposées, à l'exemple de la notion d'*agrégat*, porteur en particulier de contraintes d'agrégation spatiales.
- La démarche de modélisation des systèmes à référence spatiale est redéfinie à la lumière des nouvelles notions introduites. Une distinction fondamentale est en particulier faite entre la notion d'information localisée, de niveau conceptuel, et la notion de donnée géométrique, de niveau organisationnel.
- La complexité de la modélisation conceptuelle des systèmes à référence spatiale, ainsi que les besoins d'intégration des modèles de données et de traitements traditionnels, nous ont conduit à suggérer l'élaboration d'outils CASE permettant d'extraire d'un modèle évolutif différents *hyperschémas* correspondant à diverses étapes et préoccupations de la conception, et de naviguer aisément d'un hyperschéma à un autre.

b. Résultats rattachés à la facette *homme*

Nous rattachons en particulier à cette facette les résultats relatifs aux *organisations*.

Un ensemble cohérent de critères de découpage d'une entreprise en domaines d'activité a été suggéré. Ces critères nous ont permis de formuler et de démontrer la pertinence des propositions suivantes:

- Une entreprise qui gère ou utilise des informations localisées doit définir un SIRS comme sous-système de son système d'information d'entreprise.
- Une organisation fédérée des SIRS sur un territoire donné constitue une réponse efficace aux besoins d'intégration des informations localisées.
- Les systèmes de prise de décision constituent des sous-systèmes du système opérant de l'entreprise. Ils peuvent mettre en oeuvre des systèmes informatiques d'aide à la décision à référence spatiale (SADRS).

c. Résultat rattaché à la facette *technologie*

Les systèmes de gestion de données localisées (SGDL) sont susceptibles de répondre aux exigences informatiques des SIRS. Les systèmes d'agrégation d'informations nécessitent par contre la mise en oeuvre de logiciels axés sur le traitement interactif graphique assisté par l'ordinateur; ce sont les systèmes de conception assisté par ordinateur (CAO) qui répondent le mieux à un tel cahier des charges.

10.2. Perspectives de recherches et d'applications

a. Développement d'une méthode de conception de systèmes à référence spatiale

Des efforts de recherche intensifs sont investis pour élaborer des outils d'aide à la conception de systèmes d'information à référence spatiale. Une démarche de mise en oeuvre est bien sûr associée à de tels outils.

Mais il manque aux concepteurs de SIRS comme aux concepteurs des systèmes d'aide à la décision à référence spatiale une méthode générale d'analyse et de conception qui tienne compte des spécificités de la référence spatiale.

Cette méthode devrait en outre tenir compte des différences fondamentales entre les SIRS et les systèmes d'aides à la décision à référence spatiale; ces derniers exigent en effet une approche étroitement interdisciplinaire entre les concepteurs du système et les spécialistes (ingénieurs, géographes, aménagistes, ...) compétents dans le domaine d'application considéré. La méthode devrait encore prendre en compte les interactions entre les deux types de systèmes: comment en particulier élaborer la base de données de projet d'un système d'aide à la décision sur la base d'un SIRS ?

La mise au point d'une telle méthode est très attendue des concepteurs de systèmes, en particulier du secteur privé.

b. Validation et intégration des nouvelles composantes géosémantiques proposées

Les nouvelles composantes géosémantiques proposées pour la description des structures de données (agrégat, relation spatiale) n'ont été testées que pour la mise en oeuvre d'un modèle hydrologique particulier. Les notations n'ont en particulier pas été validées. Leur affinage nécessiterait encore d'autres tests, allant de la conception à la réalisation, intégrant d'autres domaines spécialisés, et offrant une comparaison avec d'autres formalismes.

Une fois cet affinage réalisé, il conviendrait d'intégrer ces nouvelles composantes à un formalisme tel que MODUL-R, qui constitue aujourd'hui un "standard de fait".

c. Coordination des recherches et développements en matière de SIRS

La conception de systèmes à référence spatiale nécessite l'emploi des formalismes adéquats, la disponibilité des outils informatiques intégrés correspondants, et l'intégration de l'ensemble dans le cadre d'une méthode rigoureuse. Une telle intégration nécessite une coordination des efforts de développement, permettant:

- la validation des révisions des formalismes;
- l'intégration de ces révisions dans les outils informatiques correspondants;
- la mise à jour de la méthode globale.

La coordination des travaux de recherche et de développement, la normalisation de la méthode et des formalismes développés ainsi que leur promotion devraient être assurées par un organisme reconnu sur le plan international. L'exemple de "I'E/R Institute", qui assume un tel rôle en faveur du modèle entité/relation, mériterait d'être étudié.

Bibliographie

- Aalders (1990). IVth international Conference on Spatial Data Handling, Zürich.
- Barr, A. et E. A. Feigenbaum, éd. (1981). The Handbook of artificial intelligence. Stanford, California, HeurisTech Press.
- Bédard, Y. et J.-J. Chevallier (1989). Processus de modélisation dans les systèmes d'information à référence spatiale. Québec, Université Laval - Laboratoire de SIRS.
- Bédard, Y. et J.-J. Chevallier (1989). Processus de modélisation dans les systèmes d'information à référence spatiale. Université Laval - Laboratoire de SIRS.
- Bédard, Y. et F. Paquette (1988). Extending Entity/Relationship Formalism for Spatial Information Systems. Ste-Foy, Québec, Laboratoire de Systèmes d'information à référence spatiale, Université Laval.
- Bédard, Y. et F. Paquette (1990). Extending entity/relationship formalism for spatial information systems. Auto Carto 9, Baltimore, Maryland, USA.
- Bédard, Y. et J. Prince (1990). Outil d'inventaire et d'analyse des données spatiales d'une organisation. Québec, Université Laval - Laboratoire de SIRS.
- Bertin, J. (1967). Sémiologie graphique. Paris, Editions Gauthier-Villars.
- Bertin, J. et S. Bonin (1977). La graphique et le traitement graphique de l'information. Paris, Flammarion.
- Bezos, A. (1990). STEP technology for product data definition. Workshop of the Vth European Conference, Basel, Suisse, AM/FM International-European Division.
- Bilodeau, J.-M., C. Caron, et al. (1990). Gestion des événements temporels dans les SIRS. Université Laval - Laboratoire de SIRS.
- Brassel, K., éd. (1990). Proceedings of the fourth symposium on spatial data handling. Zürich.
- Brodie, M. L., J. Mylopoulos, et al., éd. (1984). On conceptual modelling : perspectives from artificial intelligence, databases, and programming languages. Topics in information systems. New York, Springer-Verlag.
- Bruegger, B. P. et W. Kuhn (1990). Multiple Topological Representations. Document de travail remis par les auteurs.
- Buogo, A. (1990). Conception et maquettage d'un système de simulation pour le drainage urbain. Lausanne, Ecole Polytechnique Fédérale.
- Buogo, A. et J.-C. Stucky (1992). SYNAPSE - projet d'un système d'information et d'analyse géographique. Géodésie et Mensuration, EPFL et Institut de géographie, UNIL.
- Caron, C. (1991). Nouveau formalisme de modélisation conceptuelle adapté aux SIRS. Université Laval - Laboratoire de SIRS.
- Champoux, P. (1991). Etude sur les fonctions d'analyse spatiale à utiliser dans un SIRS appliqué à l'exploitation minière. Université Laval - Laboratoire de SIRS.

Bibliographie

- Chen, P. P.-S. (1976). "The Entity-Relationship Model: Toward a Unified View of Data." ACM TODS (March).
- Chevallier, J.-J. (1983). Une approche systémique des systèmes d'information du territoire et de leur intégrité. Ecole Polytechnique Fédérale.
- Chevallier, J.-J. (1986). Systèmes d'information du territoire. CREM, Martigny.
- Chevallier, J.-J. (1992). S'informer ou décider ? des SIRS aux SADRS. GIS National Conference, Ottawa.
- Chevallier, J.-J. et Y. Bédard (1990). "Classification des types de référence spatiale utilisés dans les systèmes d'information à référence spatiale (SIRS)." MPG 90(11), Sigwerb, Villmergen, Suisse.
- Chevallier, J.-J., F. Golay, et al. (1989). Mise en oeuvre des systèmes interactifs-graphiques pour la gestion des informations sur le territoire: Partie 1: la méthode FRAMBOISE. Lausanne, Géodésie et Mensuration, EPFL. 29 + annexes.
- Chevallier, J.-J., F. Golay, et al. (1989). Mise en oeuvre des systèmes interactifs-graphiques pour la gestion des informations sur le territoire: Partie 2: la représentation des données à référence spatiale. Lausanne, Géodésie et Mensuration, EPFL. 29 + annexes.
- Chocat, B. (1981). Un système d'aide à la prévision et la conception des équipements d'assainissement. INSA-Lyon et université Claude Bernard Lyon 1.
- Date, C. J. (1987). An Introduction to Database Systems. Reading, Mass., Addison Wesley.
- Davis, G. B. et M. H. Olson (1985). Management Information Systems: Conceptual Foundations, Structure and Development. USA, MacGraw Hill.
- De Rosnay, J. (1975). Le Macroscope. Paris, Editions du Seuil.
- Egenhofer, M. J. (1989). A formal definition of binary topological relationships. Orono, Maine, Université du Maine à Orono.
- Egenhofer, M. J. et R. D. Franzosa (1991). "Point-set topological spatial relations." GIS Journal 5(2), Taylor & Francis, Londres.
- Frank, A. (1982). Datenstrukturen für Landinformationssysteme - Semantische, topologische und räumliche Beziehungen in Daten der Geo-Wissenschaften. Eidgenössische Technische Hochschule Zürich.
- Frank, A. et D. Hudson (1988). Theories for Spatial Information Systems. Orono, Maine, Université du Maine à Orono.
- Frank, A. et W. Kuhn (1986). Cell graphs: a provable correct Method for the Storage of Geometry. Second international Symposium on spatial Data Handling.
- Gane, C. et T. Sarson (1979). Structured Systems Analysis: Tools and Techniques. Englewood Cliffs, NJ, USA, Prentice Hall.
- Herring, J. (1989). The Definition and Development of a topological spatial Data System. Seminar on Photogrammetry and Land Informations Systems, Lausanne, Presses Polytechniques Romandes, Lausanne.
- IGUL et EPFL-IM-GM (1992). SYNAPSE: projet d'un système d'information et d'analyse géographique. Institut de géographie de l'Université de Lausanne et unité Géodésie et mensuration de l'Ecole polytechnique fédérale de Lausanne.

- INSER SA, CIGC-EPFL, et al. (1992). Projet pilote "STRADA-DB / ARGIS 4GE" - Etude préliminaire. Département des travaux publics du Canton de Genève.
- Kuhn, W. (1989). Interaktion mit raumbezogenen Informationssystemen: vom Konstruieren zum Editieren geometrischer Modelle. Mitteilungen. Zürich, Institut für Geodäsie und Photogrammetrie, ETHZ.
- Laurini, R. (1989). L'ingénierie des connaissances spatiales. Paris, Hermès.
- Le Moigne, J.-L. (1977). La Théorie du système général. Paris, Presses universitaires de France.
- Leclerc, G. et J.-J. Chevallier (1986). SIURS et génie urbain, vers un concept d'organisation. Centre de développement technologique, Ecole Polytechnique de Montréal.
- Leclerc, G., F. Golay, et al. (1989). "Les systèmes d'information du territoire et leur exploitation en génie urbain." MPG 89(5), Sigwerb, Villmergen, Suisse.
- Leclerc, G. et J. C. Schaake jr (1973). Methodology for Assessing the Potential Impact of Urban Development on Urban Runoff and the Relative Efficiency of Runoff Control Alternatives. M.I.T. Technical Report. Cambridge, Mass., Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics, M.I.T.
- Lessard, G., D. Leahy, et al. (1989). CASTOR, Version 2.0, Manuel de l'utilisateur. Montréal, Ecole Polytechnique.
- Lévine, P. et J.-C. Pomerol (1989). Systèmes interactifs d'aide à la décision et systèmes experts. Paris, Hermès.
- Mandelbrot, B. (1989). Les objets fractals. Survol du langage fractal. Paris, Flammarion.
- Mannino, M. V. (1990). Panel on Extensional/Intensional Database Systems. Proceedings of the 9th International Conference on Entity-Relationship Approach, Lausanne, Suisse, COSMOPRESS, Genève.
- Martin, J. et C. L. McClure (1985). Structured techniques: The basis for CASE. Englewood Cliffs, New Jersey, USA, Prentice Hall.
- Nyerges, T. L. (1989). "Schema integration analysis for the development of GIS databases." GIS Journal 3(2), Taylor & Francis, Londres.
- Olle, T. W., J. Hagelstein, et al. (1988). Information Systems Methodologies: a Framework for understanding. Addison Wesley.
- Paquette, F. (1990). Conception d'un système d'information à référence spatiale pour la gestion et l'aménagement de la forêt Montmorency. Université Laval - Laboratoire de SIRS.
- Pascoe, G. A. (1986). "Elements of Object-Oriented Programming." BYTE (August 1986).
- Perner, J. (1991). Understanding the Representational Mind. London, England, The MIT Press.
- Petersen, G., C. Marschal, et al. (1991). STRADA-DB. Aperçu général. Office fédéral des routes; Rosenthaler + Partner AG.
- Pornon, H. (1989). Systèmes et logiciels de cartographie assistée par ordinateur. Paris, Editions Hermès.
- Pullar, D. et M. J. Egenhofer (1989). Toward formal definitions of topological relations among spatial objects. Orono, Maine, Université du Maine à Orono.

Bibliographie

- Richer, O. et J.-J. Chevallier (1992). SIRS pour la récolte forestière: outil d'inventaire ou moyen de planification? GIS National Conference, Ottawa.
- Robert, B. (1989). Un système expert d'aide à la conception hydraulique: élaboration d'une stratégie et d'une méthode de résolution de problèmes hydrauliques. Montréal, Ecole Polytechnique.
- Robert, B. et C. Marche (1989). "An expert system to assist designers of hydraulic systems." Eng. Appl. of AI 2(march).
- Spencer, R., T. Teorey, et al. (1990). ER Standards Proposal. Proceedings of the 9th International Conference on Entity-Relationship Approach, Lausanne, Suisse, COSMOPRESS, Genève.
- Tabourier, Y. (1986). De l'autre côté de Merise. Systèmes d'information et modèles d'entreprise. Paris, Les Editions d'Organisation.
- Tardieu, H., A. Rochfeld, et al. (1984). La méthode MERISE, principes et outils. Paris, Les éditions d'organisation.
- Teorey, T. J. (1990). Database modeling and design: the entity-relationship approach. San Mateo, California, USA, Morgan Kaufmann.
- Tsichritzis, C. et O. M. Nierstrasz (1988). Fitting Round Objects Into Square Databases. Source exacte inconnue.
- Walesh, S. G. (1989). Urban surface water management. New York, Toronto, Wiley.
- Ward, P. T. (1990). The use of current-generation CASE tools in object-oriented analysis and design. CASE on trial. Chichester, England, Jon Wiley & Sons.
- Zehnder, C. A. (1985). Informationssysteme und Datenbanken. Stuttgart, B.G.Teubner.

Index

- abscisse curviligne 16
- abstraction 25
 - (mécanisme d') 25, 37
- activités 10, 11, 68
- adjacence 9, 33
- agrégat 57, 58
- agrégation 13, 18, 37, 39, 40, 54
 - spatiale 27
- aménagement du territoire 1, 15, 83
- analyse spatiale 9, 30, 50, 85
- assainissement urbain 11, 77
- banque de données 7
- banque de données routières 46, 84
- base de données 1, 7, 10, 18, 29, 33, 92, 95
 - de travail 14, 72, 75
 - permanente 14, 72, 75
- bassin 13
 - versant 13, 30, 56, 63, 79
- BDP 14, 72
- BDR 84
- BDT 14, 72
- cadastre 22
- cardinalité 57
- carroyage 30, 31
- carte 18
- cartes thématiques 17
- cartographie 18
- CASE 10
- classification 38
- communication 18
- comportement 22, 24, 41
- composantes sémantiques 8
- conception 8, 10
- connaissance 25, 34
- connectivité 35
- connexion 33
- contraintes 55, 56
 - globales 57, 58
 - partielles 57, 59
- contraintes inter-relationnelles 9
- décision 42
- décomposition 38, 54
- découpage 67
- DENOT 9
- DFD 8, 50
- diagrammes de flux de données 8, 50
- dictionnaire 10, 56
- dimension topologique 9, 33
- discrétisation 13, 80
- disjonction 33
- domaines 67
- estimation 27
- état 22, 23, 54
- états 23
- expertise 13, 15
- formalisme 7, 9, 10, 49, 50
- formalismes 2, 8, 53
 - entité-relation 8
 - géosémantique 53, 54, 55, 57
 - individuel 8, 39
- frames 25
- généralisation 8, 18, 37, 39, 40, 54
 - spatiale 27
- géographe 17
- gestion de l'entretien routier 16, 84
- gestion du temps 9
- héritage 40
- hydrologie 13
- hyperschémas 90
- inclusion 9, 33
- informations localisées 1, 2
- ingénierie des connaissances 7
- instance 25
- intelligence artificielle 25
- interface 74
- levé 47
- localisation 27
- maille 30
- maquette 15
- maquettes 23
- MCC 55
- MCD 55
- mémoire 18, 21, 24
- méta-connaissance 15, 16, 19, 91
- méta-information 50, 84
- méta-modèle 25, 26, 27, 35, 36, 37, 53, 54
 - des données 25
 - des traitements 25
- méthodes 3
- mise à jour 33, 47
- MLC 66
- MOC 64
- MOD 64
- modèle 21, 25, 27, 28, 45
 - algorithmique 30
 - analogique 23
 - conceptuel cartographique 9, 55, 56
 - conceptuel des données 9, 10, 55, 56
 - de comportement 22
 - de données 16
 - de simulation 23
 - diachronique 21, 22, 46, 48
 - d'observations ponctuelles 30, 31
 - logique cartographique 66
 - mathématique 13
 - numérique 24
 - observé 22
 - organisationnel cartographique 64
 - organisationnel des données 64
 - spatial 27, 30, 38, 40
 - synchronique 22, 46
- modèle-type 26
- modélisation 5, 7, 9, 21, 26, 30, 49, 53, 64
 - spatiale 27
- modélisation (échelle de) 46
- MODUL-R 10

- mosaïque 29, 30
- multi-média 7, 31, 63
- mutation 18
- niveau organisationnel 64
- nodalité 35
- notation 8
- objectif 2
- objectifs 2
- objet spatial 29, 30
- opérateur spatial 57
- oracle 34
- organisation 3
- orienté objets 7, 10
- outil méthodologique 2
- paramétrisation 13, 79
- partenariat 70, 71
- partition du territoire 30
- permis de construire 15
- phénomène 21, 22, 23, 24, 27, 28, 31, 46
- pictogramme 54, 64
- plans d'aménagement 15, 16
- plans directeurs 15, 16
- polygones de Thiessen 13
- prévision 48
- processeur 24, 41, 43
- programme 21, 49
- pyramide de décision 12, 15
- raster 31
- recherche 3, 10
- référentiel 28
- référentiel spatial 28, 29, 40, 46, 55
- règles structurelles 8
- représentation (changement de) 46
- réseau 29, 31
- restitution 47, 48
- SADRS 72
- schématisation 13, 79
- segments
 - de ruissellement 13
 - de transport 13
- sets de points 9
- SGBD 7
- SGDL 18, 35, 51, 66, 73, 92, 94
- simplification 27
- simulation 13, 23, 48
- SIRS 1, 2, 3, 7, 8, 25
 - fédérés 70, 71
 - institutionnel 71
- sous-bassin 13, 46, 56
- sous-modèles 9
- spatialisation 31, 80
- spécialisation 37, 54
- STRADA-DB 16
- structuration 26
- structure 25
- structure organique 49, 50
- substitution de sous-modèles 8
- système
 - d'agrégation d'informations 43, 45
 - d'aide à la décision à référence spatiale 52, 72
 - d'information à référence spatiale 1
 - d'intégration des décisions 43, 45
 - de base de données 7, 25
 - de gestion de données localisées 18, 92, 94
 - de prise de décision 75
 - de repérage 16, 17
 - de représentation 45
 - expert 25, 51
 - général 41
 - homomorphe 21
 - interactif d'aide à la décision 52
 - systémique 7
- temps
 - du modèle 47
 - effectif 47
- topologie 9, 33, 34, 40, 55
- topologiques (relations) 9
- traitement 2, 5, 9, 13, 17, 45, 46, 47, 49, 50
- trajectoire 21, 23
- transformation 17, 24
 - formelle 14, 75
 - informationnelle 14, 75
- URBAN 13
- vecteurs 31

Remerciements

La rédaction d'une thèse est une course de fond dont le concurrent-chercheur émerge parfois avec le souffle court, regrettant en vrac l'arrêt trop important au ravitaillement, les erreurs d'alimentation, le thé trop sucré, l'entraînement insuffisant et les chaussures un peu trop justes.

Mais qu'importe ! La ligne d'arrivée est franchie. Oh ! pas en première position sans doute ! Mais est-ce bien là l'essentiel ? Une fois la douleur des derniers kilomètres dissipée, force est de constater qu'on ressort de l'épreuve un peu différent – ou plutôt avec un regard différent sur la course et sa propre participation.

On constate que c'est avant tout contre soi-même qu'il faut se battre. Contre les distractions de la course. Contre ses paresse. Contre le bon sens et la raison qui nous disent "à quoi bon continuer ?!". Ce combat solitaire serait sans issue si l'on ne rencontrait sur sa route les encouragements, les renseignements, les conseils, de quelque spectateur bienveillant. Parfois même, un coureur expérimenté fait un bout de course à vos côtés.

Je suis profondément reconnaissant à Jean-Jacques d'avoir rempli ce rôle d'instigateur et d'entraîneur, toujours présent dans les moments critiques, malgré l'itinéraire tourmenté de la course. Merci aussi à Guy de m'avoir guidé dans les tronçons ardu de la science de l'ingénieur, à laquelle ma formation de géomètre ne me prédestinait guère; il m'a en outre permis de découvrir outre Atlantique d'autres horizons, à la fois scientifiques, géographiques et humains.

Le Professeur Howald a bien voulu être le "chef de la course" durant les dernières étapes; il a su m'inciter à puiser les forces nécessaires pour mener la course à son terme. J'ai de plus trouvé dans l'unité *Géodésie et mensuration* de l'EPFL, et particulièrement auprès de Madame Gautier, un cadre, une atmosphère et des encouragements chaleureux.

De nombreux autres appuis ont encore joué un rôle déterminant pour l'aboutissement de mon travail: le Professeur Yvan Bédard du Centre de Géomatique de l'Université Laval et Claude Caron à l'EPFL ont enrichi cette étude de leurs idées. Alain Buogo a apporté de nombreuses pièces à mon puzzle. Mes collègues de SIT-Conseil m'ont aidé et encouragé à persévérer malgré les migraines données à notre directrice et planificatrice Béatrice. Nos clients et partenaires n'ont pas seulement su faire preuve de patience à l'égard de mes trop nombreux retards; ils m'ont surtout apporté de nombreux exemples et enrichissements dans des domaines tels que les réseaux industriels, l'aménagement du territoire ou l'entretien routier.

Je ne saurais mentionner sans lacune ni oublier les nombreux relecteurs, correcteurs, traducteurs de ce rapport. Leurs critiques et leurs suggestions ont constitué un apport important à cette étude.

Mais toute cette infrastructure serait vaine sans un "service de ravitaillement" moral à la hauteur des ambitions poursuivies. Carmen a assumé cette tâche durant six longues années, seule d'abord, puis "assistée" de Jean, de Denis et dernièrement de Diane. En retour, c'est surtout d'un mari fatigué et d'un père absent qu'ils ont pâti...

Mais que l'exaltation de l'aboutissement ne dissimule pas les limites du modeste coureur amateur. Les nombreuses perspectives ouvertes n'en masquent que bien mal les imperfections et les lacunes. Les voies esquissées méritent d'être critiquées, approfondies peut-être, étayées sans doute. Je remercie donc ceux qui, par leurs propres apports et réflexions, sauront mettre les résultats de ce travail en valeur.

Aperçu biographique

Né en 1958, François Golay a suivi toute sa scolarité dans le canton de Vaud, où il a obtenu une maturité fédérale classique en 1976. Il a ensuite obtenu un diplôme d'ingénieur du génie rural et géomètre de l'Ecole polytechnique fédérale de Lausanne (Suisse) en 1982.

Resté dans cette même école comme adjoint scientifique et chargé de cours, il a participé à plusieurs études et travaux de recherche dans le domaine des systèmes d'information à référence spatiale. Durant cette période, il a pu effectuer un séjour de recherche de 4 mois à l'Ecole polytechnique de l'Université de Montréal, et a eu des contacts réguliers avec le Centre de géomatique de l'Université Laval à Québec.

Il a maintenant rejoint à Genève un bureau de consultants spécialisé dans la conception et la mise en oeuvre de systèmes d'information à référence spatiale.

Marié et père de trois enfants, il habite à Cossonay, près de Lausanne.

