

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
UNIVERSITE DES SCIENCES ET DE LA TECHNOLOGIE HOUARI BOUMEDIENE
(U.S.T.H.B)

Faculté d'Electronique et d'Informatique
Département d'Informatique

Thèse

Présentée pour l'obtention du grade de

Magister en Informatique

Spécialité : Intelligence artificielle et base de données avancée

par :

Mr ALLILI Toufik

Sujet

Environnement de développement des
Systèmes d'Information Géographique (SIG) :
Approche basée sur l'intégration des fragments
de modèles

soutenue publiquement le : **29/06/2003** devant le jury composé de :

Mme M.BOUKALA	Maître de conférence (USTHB)	Présidente
Mme Z.ALIMAZIGHI	Maître de conférence (USTHB)	Directrice de thèse
Mr A.K.YELLES CHAOUCHE	Directeur de recherche (CRAAG)	Examinateur
Mme A.BELHADJ-AISSA	Maître de conférence (USTHB)	Examinatrice
Mme F.ADMANE	Attaché de recherche (CERIST)	Invitée

Juin 2003

Remerciements

Je voudrais tout particulièrement remercier :

Mme Z.ALIMAZIGHI, pour avoir accepté de diriger ce travail, pour ses orientations et conseils grâce auxquels ce travail a pu aboutir.

Je tiens à lui exprimer toute ma reconnaissance et ma considération les plus sincères.

Mme M.BOUKALA, pour m'avoir fait l'honneur de présider le jury de cette thèse.

Mme A.BELHADJ, Mme F.ADMANE et Mr AK.YELLES CHAOUICHE, pour avoir accepté de faire partie de ce jury et d'examiner ce travail.

J'adresse mes remerciements au personnel du centre de recherche CRAAG pour l'ambiance agréable, le soutien, l'aide et la bonne humeur. Un vif merci particulièrement à Farida, Leïla, Nadjet, Hamoud, Fethi, Ahmed,

Mes remerciements sont adressés aux membres de l'équipe LSI de l'institut d'Informatique, pour leur soutien et leurs apports dans le cadre des travaux que nous avons menés.

*Merci à baya, pour sa présences, ses encouragements, son aide, ses précieux conseils,
baya, les mots ne suffisent pas pour t'exprimer toute ma reconnaissance et tous mes remerciements.*

Enfin, je remercie tous ceux qui m'accompagnent, famille et amis, pour leur affection, leur patience et leurs encouragements qui m'a permis d'arriver au bout de ce travail.

RESUME

Pour faciliter le développement des systèmes d'information géographique, les concepteurs proposent une approche de modélisation indépendante. Elle consiste à décomposer tout projet SIG en sous systèmes. Elle permettra ainsi, la réutilisation des données déjà existantes que ce soit des fichiers ou des bases de données indépendantes. Le développement de ces nouveaux systèmes repose désormais sur la capacité du système informatique à réaliser l'interopérabilité entre bases de données.

Dans cette étude, nous présentons une approche de développement d'un système d'information géographique basée sur l'intégration des fragments de modèles. La décomposition du système en couches spatiales permettra de retrouver l'ensemble des sous systèmes à développer. Nous aurons ainsi, la possibilité d'utiliser de diverses méthodes de conception et de réintégrer des fragments déjà existants.

Mots clés. Objet géographique, méthodes de conception, couche spatiale, interdépendances spatiales, fragment de modèle, Intégration, SIG.

SOMMAIRE

SOMMAIRE

Introduction générale	1
------------------------------------	---

PARTIE I : Système d'information géographique (SIG) : Etat de l'art

Chapitre I : Généralités sur les Systèmes d'Information Géographique (SIG)	4
I.1. Introduction	4
I.2. Qu'est ce qu'un Système d'Information Géographique (SIG) ?	4
I.3. Domaines d'utilisation des SIG	5
I.4. Base de données géographiques	5
I.4.1. Description géométriques	6
a- Mode raster	6
b- Mode vectoriel	7
I.4.2. Description thématiques (attributaires)	8
I.5. Notion de couche	8
I.5.1. Représentation de la géométrie d'une couche	9
a- Modèle spaghetti	9
b- Modèle réseau	10
c- Modèle topologique	10
I.6. Sources de données	11
I.7. SGBD Géographique et langage de requêtes	12
I.7.1. Modèle entité-relation	12
I.7.2. Modèle relationnel	13
I.7.3. Modèle relationnels étendu aux types abstrait de données	14
I.7.4. Modèle orienté-objet	14
I.8. Les fonctions d'un SIG	16
I.8.1. Sous-système d'acquisition des données	16
I.8.2. Sous-système d'analyse	16
I.8.3. Sous système de gestion et d'interrogation	17
I.8.4. Sous système de présentation	17
I.9. Les outils de développement des SIG	18
I.10. Conclusion	19
 Chapitre II : Concepts Spatial et Temporel	 20
II.1. Introduction	20
II.2. Concept Spatial	20
II.2.1. Structuration continue	20
II.2.2. Structuration discrète	21
a- Choix de la forme géométrique	21
b- Intégration de la géométrie dans la base de données	23

c- Relations spatiales	24
d- Contraintes d'intégrités spatiales	28
II.3. Concept temporel	29
II.3.1. Evolution des objets géographiques	29
II.3.2. Intégration de l'évolution dans la base de données	30
a- Attributs temporels	30
b- Types d'objets temporels	31
c- Types d'associations temporels	31
II.4. Conclusion	33

Chapitre III : Méthodes de conception des SIG : Modèles & outils AGL (Atelier de Génie Logiciels

III.1. Introduction	34
III.2. Spécificités conceptuelles des SIG	34
III.2.1. Concept spatial	35
III.2.2. Concept temporel	36
III.3. Critères d'évaluation des outils AGL	36
III.4. Méthodes de conception des SIG	38
III.4.1. GEO-OOA	39
III.4.2. MADS (Modélisation d'Application à Données Spatio-temporelles)...	43
III.4.3. PERSPECTORY 2000	48
III.5. Conclusion	53

PARTIE II : Environnement de développement des systèmes d'information géographique (SIG)

Chapitre VI : Développement des Systèmes d'Information Géographique (SIG)

IV.1. Introduction	55
IV.2. Cycle de développement d'un projet	55
IV.2.1. La démarche en cascade	55
IV.2.2. La démarche en V	56
IV.2.3. La démarche par prototypage	57
IV.2.4. La démarche en spirale	57
IV.2.5. La démarche tridimensionnelle	58
IV.2.6. La démarche en fontaine	58
IV.3. Les conditions d'un SIG efficace	60
IV.5. Les sept phases de développement d'un projet SIG	61
IV.5.1. Étude d'opportunité	61
IV.5.2. Étude de faisabilité	62
IV.5.3. Modélisation du système	63
IV.5.4. Analyse fonctionnelle	64
IV.5.5. Réalisation technique	64
IV.5.6. Implantation	65
IV.5.7. Exploitation et entretien du SIG	66
IV.6. Conclusion	67

Chapitre V : Intégration de schémas : Concepts et méthodologie	68
V.1. Introduction	68
V.2. Concepts d'intégration de schémas	68
VI.2.1. Diversité de schémas	68
VI.2.2. Propriétés inter-schémas	71
V.3. Méthodologie d'intégration de schémas	71
V.3.1. Les étapes d'intégration	72
V.3.2. Caractéristique du schéma global	73
V.3.3. Les conflits lors de l'intégration	74
V.3.3.1. Conflits de nomination	74
V.3.3.2. Conflits structurels	75
V.3.3.3. Cas d'un SIG	76
V.3.4. Correspondances inter-schémas	77
V.4. Quelques méthodes d'intégration	80
V.5. Conclusion	82
Chapitre VI : Environnement de développement d'un projet SIG :	
<i>Une approche basée sur l'intégration des fragments de modèles..</i>	83
VI.1. Introduction	83
VI.2. Présentation générale de la démarche	84
VI.3. Description de la démarche	87
VI.3.1. Analyse des besoins	87
VI.3.2. Modélisation	91
VI.3.3. Construction du schéma fédéré	93
VI.3.3.1. Intégration des fragments de modèles « Niveau 1 »	93
VI.3.3.2. Fédération des sous-schémas « Niveau 2 »	97
VI.3.4. Réutilisation des fragments de modèles	100
VI.4. Application : un SIG pour la gestion du risque sismique	101
VI.5. Conclusion	107
Conclusion générale	108
Bibliographie	

LISTE DES FIGURES

Figure I.1 : Base de données géographiques	6
Figure I.2 : Mode raster (matricielle)	7
Figure I.3 : Mode vectoriel	8
Figure I.4 : représentation par couche	9
Figure I.5 : Modèles de représentation géométriques	10
Figure I.6 : Modèle entité-relation	12
Figure I.7 : Modèle relationnel	13
Figure I.8 : Les fonctions d'un SIG	16
Figure I.9 : Hiérarchie de logiciels SIG	18
Figure II.1 : Différents types de représentation géométrique	21
Figure II.2 : Échelle et précision de l'information géographique	22
Figure II.3 : Principe de la projection cartographique	23
Figure II.4 : Représentation des primitives géométrique dans la base de données de ArcInfo	24
Figure II.5 : Relation de disjonction	25
Figure II.6 : Relation d'adjacence	25
Figure II.7 : Relation de croisement	25
Figure II.8 : Relation de recouvrement	26
Figure II.9 : Relation d'inclusion	26
Figure II.10 : Relation d'égalité	26
Figure II.11 : Exemple d'agrégation spatiale	27
Figure II.12 : Connectivité dans un réseau	28
Figure II.13 : Cycle de vie d'un objet	29
Figure II.14 : Primitives temporelles	30
Figure II.15 : Type d'attributs et objets temporels	31
Figure II.16 : Type d'association temporel	31
Figure II.17 : Exemple d'association de transition de la méthode MADS	32
Figure III.1 : GEOCLASS	40
Figure III.2 : Structure réseau de GEO-OOA	40
Figure III.3 : Agrégation spatiale de GEO-OOA	41
Figure III.4 : Evolution et versions de l'objet	41
Figure III.5 : Exemple de modélisation selon GEO-OOA	42
Figure III.6 : Exemple de relations topologiques	44
Figure III.7 : Exemple de lien d'agrégation	44
Figure III.8 : Exemple de lien de généralisation	44
Figure III.9 : Exemple de lien de spécialisation	45
Figure III.10 : Exemple de modélisation selon MADS	46
Figure III.11 : Architecture du système MADS	47
Figure III.12 : Editeur de schémas	47
Figure III.13 : Exemple de classe spatiale de PERSPECTORY	48

Figure III.14 : Exemple de lien d'agrégation	49
Figure III.15 : Exemple de l'évolution dans PERSPECTORY	50
Figure III.16 : Exemple de modélisation selon PERSPECTORY	51
Figure III.17 : PERSPECTORY 2000	52
Figure IV.1 : La démarche en cascade	56
Figure IV.2 : La démarche en V	56
Figure IV.4 : La démarche en spirale	57
Figure IV.5 : La démarche tridimensionnelle	58
Figure IV.6 : La démarche en fontaine	59
Figure IV.7 : Les sept phases de développement d'un projet SIG	62
	70
Figure V.1 : Exemple d'équivalence de schémas	70
Figure V.2 : Incompatibilité des spécifications	71
Figure V.3 : Exemple de relation inter-schémas	72
Figure V.4 : Les étapes d'intégration	74
Figure V.5 : Les conflits de nomination	76
Figure V.6 : Exemple de conflit géométrique	79
Figure V.7 : Exemple d'ensemble-attributs équivalents	79
Figure V.8 : Exemple attributs-classe équivalents	79
Figure VI.1 : schéma général du développement d'un projet SIG	84
Figure VI.2 : Document GOMT	88
Figure VI.3 : Processus d'extraction des couches spatiales	89
Figure VI.4 : Structure de la couche spatiale	89
Figure VI.5 : Relations topologiques entre deux couches spatiales	90
Figure VI.6 : Modélisation des couches spatiales	92
Figure VI.7 : Intégration des fragments de modèles	93
Figure VI.8 : Exemple d'homonyme	95
Figure VI.9 : Exemple de classe Pivot dans un réseau routier	96
Figure VI.10 : Fédération des sous-schémas	97
Figure VI.11 : Classe Pivot et lien de généralisation	99
Figure VI.12 : intégration des fragments de modèles existants	100
Figure VI.13 : Processus de pre-intégration des fragments de modèles existants	100
Figure VI.14 : Couches spatiales du système pour la gestion du risque sismique	101
Figure VI.15 : Graphe de topologie	102
Figure VI.16 : Schéma fédéré du SIG pour la gestion du risque sismique	106

LISTE DES TABLES

Table III.1 : Méthodes de conception des SIG	35
Table III.2 : Critères d'évaluation reliés aux fonctionnalités des outils AGL	37
Table III.3 : Critères d'évaluation reliés à l'interface des outils AGL	38
Table III.4 : Type abstrait de données spatiales	43
Table III.5 : Type de relations topologiques	44
Table III.6 : Pictogrammes simples de PVL	48
Table III.7 : Pictogrammes composés de PVL	49
Table III.8 : Pictogrammes temporels de PVL	50
Table VI.1 : Relation topologiques entre deux couches spatiales	90
Table VI.2 : Comparaison des modèles conceptuels selon le modèle géométrique	92
Table VI.3 : classification des conflits	94

INTRODUCTION GENERALE

INTRODUCTON GENERALE

i. Introduction

Dans les années 70, le gouvernement canadien a pris conscience du besoin de « non seulement collecter mais aussi manipuler et analyser l'information géographique, et ceci de façon opportune » (Tomlinson-1990). L'information géographique est d'importance primordiale pour tous ceux qui ont à gérer un espace.

A cette époque, le monde de l'informatique avait évolué considérablement avec :

- L'accroissement de la mémoire des ordinateurs : cela permet de stocker de plus en plus de données,
- L'amélioration des performances de calculs : on peut travailler avec des volumes de données (des cartes en l'occurrence) de plus en plus conséquents. Il est possible de mettre en place des modèles de la réalité de plus en plus sophistiqués. Ceci démontre peu à peu la supériorité du traitement informatique des données par rapport au traitement manuel des données. L'interactivité graphique s'en trouve améliorée. Pour un opérateur, cela signifie une meilleure compréhension des traitements en cours, et par conséquent plus d'efficacité pour gérer les données en cours de traitement.
- Les coûts des technologies informatiques continuent à diminuer, élargissant progressivement l'éventail des utilisateurs potentiels.

Avec l'essor de la technologie informatique est apparu le besoin de numériser l'information géographique et d'en confier la gestion à l'ordinateur. Ainsi sont nés les systèmes d'information géographique (SIG).

Les SIG ont bénéficié des progrès réalisés indépendamment dans deux branches importantes de l'informatique :

- d'une part la conception assistée par ordinateur (CAO), dont dérivent le dessin assisté par ordinateur (DAO) et la cartographie assistée par ordinateur.
- D'autre part les systèmes de gestion de bases de données (SGBD)

Dans le même temps, une nouvelle technologie fait son entrée dans le domaine civil : la télédétection. Elle permet la production d'une grande quantité d'images de la Terre, ce qui va étoffer l'offre dans le domaine de l'information géographique.

Le domaine des SIG tend à se diffuser dans le monde entier grâce à l'amélioration des technologies et notamment l'accroissement des capacités de collecte de l'information.

ii. Problématique

Historiquement, en terme de modélisation, plusieurs méthodes de conception des SIG ont été développées pour assurer une bonne conception du système. Cette diversité de méthodes conçues dans un contexte applicatif : gestion du territoire (aménagement du territoire, réseaux électriques et téléphoniques, ...), gestion des risques naturels (inondation, sismique, feu de forêts, ...), des applications d'analyse et de planification (mutations urbaines, hydrologie, sciences de la terre, ...), ... , n'a fait qu'accroître l'embarras du concepteur pour choisir la méthode la mieux adaptée à ses besoins. Par ailleurs, l'insuffisance est perçue dans certaines méthodes pour modéliser les concepts spatio-temporels. Cela est dû à la complexité de certains domaines d'applications des SIG.

iii. Solution proposée

La solution au problème posé, consiste à définir une démarche de modélisation qui prendra en considération la variété des concepts introduits dans ces méthodes de conception afin de satisfaire tous les besoins du concepteur des SIG. Cette démarche se base essentiellement sur l'intégration des fragments de modèles conçus par les différentes méthodes de conception des SIG.

Notre approche répond aux objectifs suivants :

- La décomposition du monde réel nous permet bien de contrôler l'analyse spatiale du territoire. On comprend mieux un processus complexe lorsqu'il est réduit en simples blocs d'information
- le concepteur aura à sa disposition une variété de méthodes de conception pour un même projet SIG. Il choisira ainsi pour chaque thème de son projet la méthode adéquate.
- L'architecture proposée dans notre démarche est très ouverte pour prendre en compte des bases de données existantes.

iv. Plan de la thèse

Cette thèse comprend deux grandes parties :

- la première porte sur un état de l'art des systèmes d'information géographique (SIG) divisée en trois chapitres : Des généralités sur les SIG (Définitions, base de données géographiques, SGBD géographiques, outils SIG, ...) sont regroupés dans le premier chapitre. Le deuxième est consacré à définir les concepts spatial et temporel utilisés dans la modélisation des SIG. Et enfin, le troisième chapitre de cette partie synthétise les différentes méthodes développées pour conception des SIG.
- La deuxième partie de cette thèse intitulée « Environnement de développement des SIG » comporte notre approche de développement d'un projet SIG par intégration des modèles. Nous avons précédé ce chapitre par deux autres. Le premier propose les phases de développement d'un projet dans l'ingénierie des systèmes d'information et dans les SIG. Le deuxième comporte les différentes méthodologies d'intégration des schémas. Ces concepts définis dans ces deux chapitres vont nous aider à élaborer une architecture structurée de développement d'un projet SIG par l'intégration des fragments de modèles.

Enfin, nous concluons notre travail par quelques perspectives pour montrer les ouvertures possibles de notre projet.

PARTIE I

Systeme d'information géographique (SIG) :

Etat de l'art

CHAPITRE I

Chapitre I

Généralités sur les Systèmes d'Information Géographique (SIG)

I.1. Introduction

Le concept des Systèmes d'Information Géographique s'est développé à partir des années 70, principalement sur les gros systèmes informatiques, dû au volume de mémoire que prennent les images et les données traitées. Aujourd'hui, les systèmes sont moins coûteux, plus intégrés, ouverts et leur évolution permet de généraliser leur emploi sur des systèmes plus compact et plus accessibles, comme des PC, des Mac ou des stations de travail.

I.2. Qu'est ce qu'un Système d'Information Géographique (SIG) ?

Définition 1 : [PER01]

Le groupe de travail de l'instrumentation géographique sous la tutelle du Conseil National de l'Information Géographique (C.N.I.G, France) donne cette définition des SIG : « Un système d'information géographique est un ensemble coordonné d'opérations généralement informatisées, destiné à transcrire et utiliser un ensemble d'informations géographiques sur un même territoire. Ce dispositif vise particulièrement à combiner au mieux les différentes sources accessibles : base de données, savoir-faire, capacité de traitement selon les applications demandées. ». Cette définition a été éditée dans un travail intitulé « Elément d'un vocabulaire géomatique ».

Définition 2 : [PER01]

Les systèmes d'information géographique (SIG) sont des outils informatiques qui permettent de gérer (stockage, interrogation, manipulation, partage, diffusion, restitution sur écran, cartes ou plans, ...) différents types de données à composante géographique (rues,

fleuves, bâtiments ...). Ils sont utilisés comme outils d'aide à la décision dans des domaines très variés.

Ils assurent :

- l'échange et l'acquisition de données.
- La communication avec l'utilisateur :
Production de cartes, l'affichage à l'écran des plans ou des cartes thématiques munis de légende et de statistiques diverses.
L'interaction avec l'utilisateur qui veut poser des questions, modifier la représentation des cartes, lancer des programmes d'analyse,

I.3. Domaines d'utilisation des SIG

Les SIG sont utilisés aujourd'hui dans un grand nombre de domaines de la vie économique :

- Urbanisme.
- Etudes économiques et socio-économiques : population, emploi, logement, production, géomarketing.
- Planification, aménagement du territoire, protection de l'environnement, gestion des ressources naturelles et des espaces ruraux et forestiers.
- Réseaux de communication et de transport : électricité, eau, routes, fleuves, réseaux souterrains,
- Applications militaires.
- Sciences de la terre : géologie, géodésie, pédologie, géotechnique,

I.4. Base de données géographiques

A l'origine, tout SIG repose sur une base de données géographiques. Des entités appelées objets géographiques sont définis par leur forme (description géométrique) et par un ou plusieurs thèmes (description thématique ou attributive) (figure I.1).

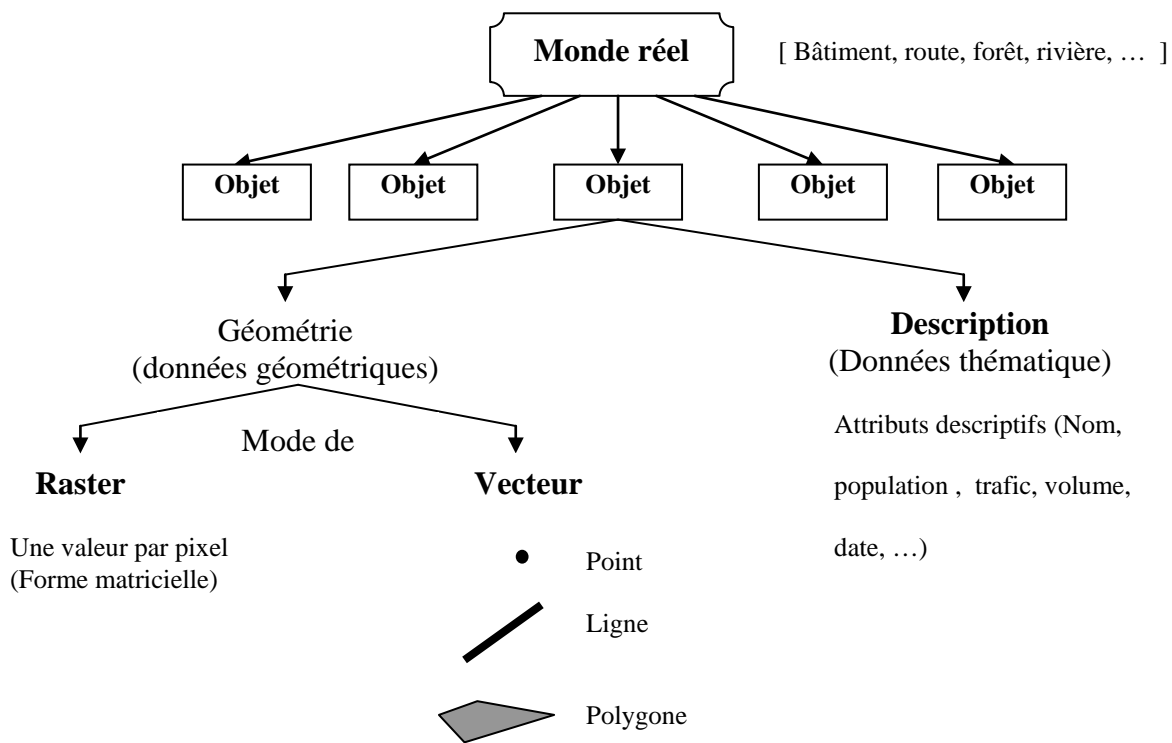


Figure I.1 : Base de données géographiques

1.4.1. Description géométrique

C'est une représentation graphique des objets géographiques qui composent l'espace. Il existe deux (02) modes de représentation :

a- Mode raster (matriciel) :

Ce mode de représentation décompose l'espace avec une grille régulière de cellules (Pixels), ordonnées pour former une matrice régulière. Cette dernière produit un espace continu ou chaque pixel possède une seule valeur alphanumérique qui correspond à l'attribut de l'entité géographique (Figure I.2).

Les données en mode raster se prêtent très bien à certains types de traitements numériques, classiques en traitements d'images : filtres, convolutions, classifications.

L'inconvénient majeur du mode raster réside dans l'importance de la capacité mémoire nécessaire à stocker les données. D'autre part, les erreurs de mesures en distance et en surface des objets géographiques dépendent de la résolution (dimension) du pixel.[1]

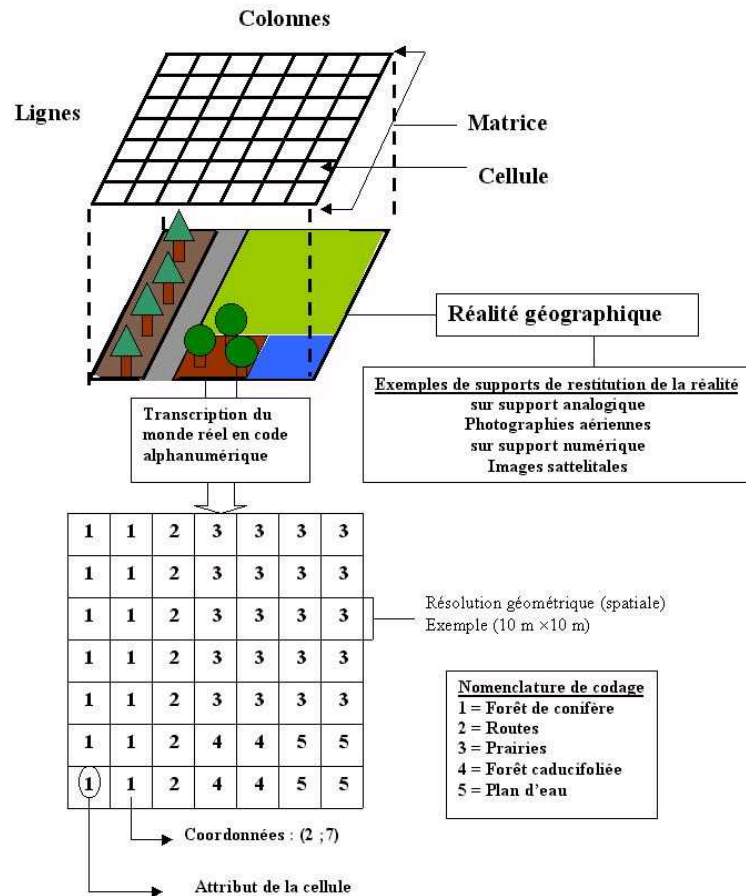


Figure I.2 : Mode raster (matricielle) [PER01]

b- Mode vectoriel :

Le mode vectoriel utilise les concepts géométriques de points, de lignes et de polygones pour représenter les entités distribuées sur l'espace. Chaque entité est représentée par un objet localisé dans un système de coordonnées. Cet objet est relié avec un tableau d'attributs qui contient les valeurs thématiques décrivant l'entité représentée (figure I.3).

Le mode vectoriel répond au souci de représenter un objet de manière aussi exacte que possible. L'espace des coordonnées est calculé avec précision. Le volume de la base de données est en fonction de la complexité des lignes, des polygones et de la densité des points. En contre partie, les croisements des objets géographiques sont délicats et nécessitent une topologie parfaite : des erreurs (comme un polygone mal fermé) peuvent poser problème.[1]

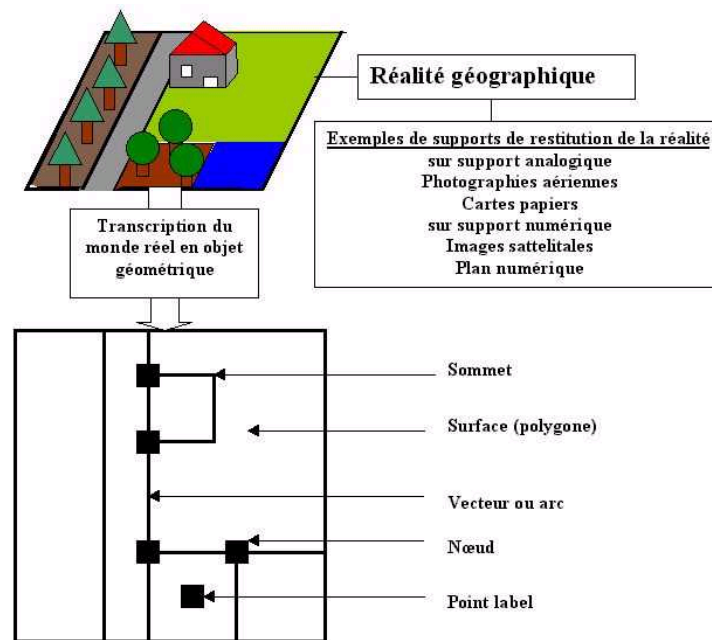


Figure I.3 : Mode vectoriel [PER01]

I.4.2. Description thématique (attributaire)

Les données thématiques des objets géographiques complètent la description du monde réel. Par exemple, une route est représentée par une ligne, il peut être intéressant de renseigner la forme par le nom de la route, la densité du trafic, revêtement,

I.5. Notion de couche [SCH96][UNI00]

On appelle couche (couverture spatiale), un ensemble d'un ou plusieurs objets géographiques de même type qui correspondent à un thème donné (Figure I.4). On met sur une même couche un seul type d'objet géographique. Ainsi on voit assez de couches qui possèdent à la fois des points, des lignes et des polygones qui pourront être recombinaées pour être analysées ou simplement visualisées.

La séparation en couches permet d'améliorer la gestion de la base de données et l'efficacité des requêtes.

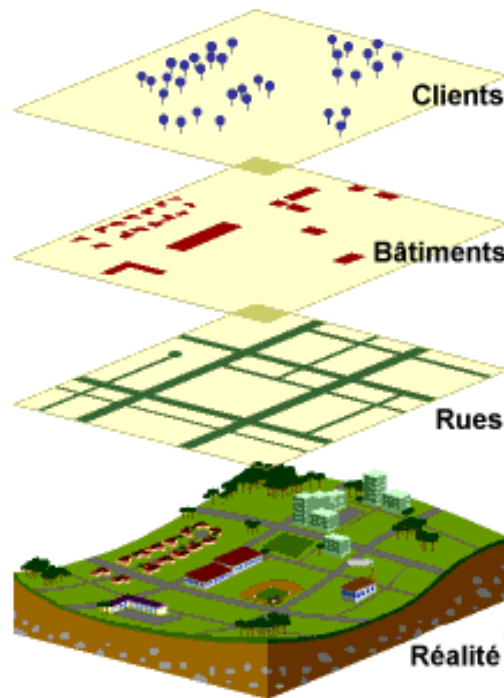


Figure I.4 : représentation par couche (image ESRI⁽¹⁾)

1.5.1. Représentation de la géométrie d'une couche

La représentation d'une couche varie suivant la richesse des dépendances topologiques entre les objets géographiques de la couche. Il existe trois modèles très courants :

a- Modèle spaghetti

Dans ce modèle, aucune relation topologique n'est représentée. La géométrie d'un objet géographique est stockée indépendamment de celle des autres objets. Il n'existe aucun moyen de savoir si deux objets partagent un point (deux lignes connexes), une ligne (deux polygones adjacents). Un point est décrit par ses coordonnées. Une ligne est décrite par une liste de coordonnées, celles de ses sommets. Un polygone est également décrit par la liste des coordonnées de ses sommets (Figure I.5-a).

Le modèle spaghetti est une structure efficace lorsqu'on recherche un objet particulier. Il est bien adapté à la saisie de thèmes séparés. Par contre, il est inefficace pour toutes les opérations dites topologiques (inclusion, adjacence, ...). Il faut à chaque fois calculer la topologie des objets géographiques.

(1) ESRI : un des principaux éditeurs sur le marché des logiciels SIG

b- Modèle réseau

On représente toutes les relations entre objets de type ligne et entre lignes et point. Une nouvelle primitive appelée nœud remplace le point. Il permet la description de la topologie d'un réseau (deux lignes sont connexes, si elles partagent un nœud) et rend ainsi plus efficace les opérations topologiques telles que le test de connexité de deux lignes et les opérations de calcul des réseaux (calcul du chemin le plus court) (Figure I.5-b).

c- Modèle topologique

On représente toutes les relations entre lignes et entre lignes et points, mais également toute les relations entre polygones, entre polygones et lignes et entre polygones et points. Le graphe associé à toute la géométrie doit être planaire (à toute extrémité d'une ligne correspond un nœud du graphe ; à toute ligne correspond une arête du graphe) (Figure I.5-c).

L'inconvénient d'une telle structure est sa lourdeur. Chaque fois qu'on insère un nouvel objet géographique, il faut recalculer toutes les intersections de la géométrie de cet objet avec la géométrie existante.

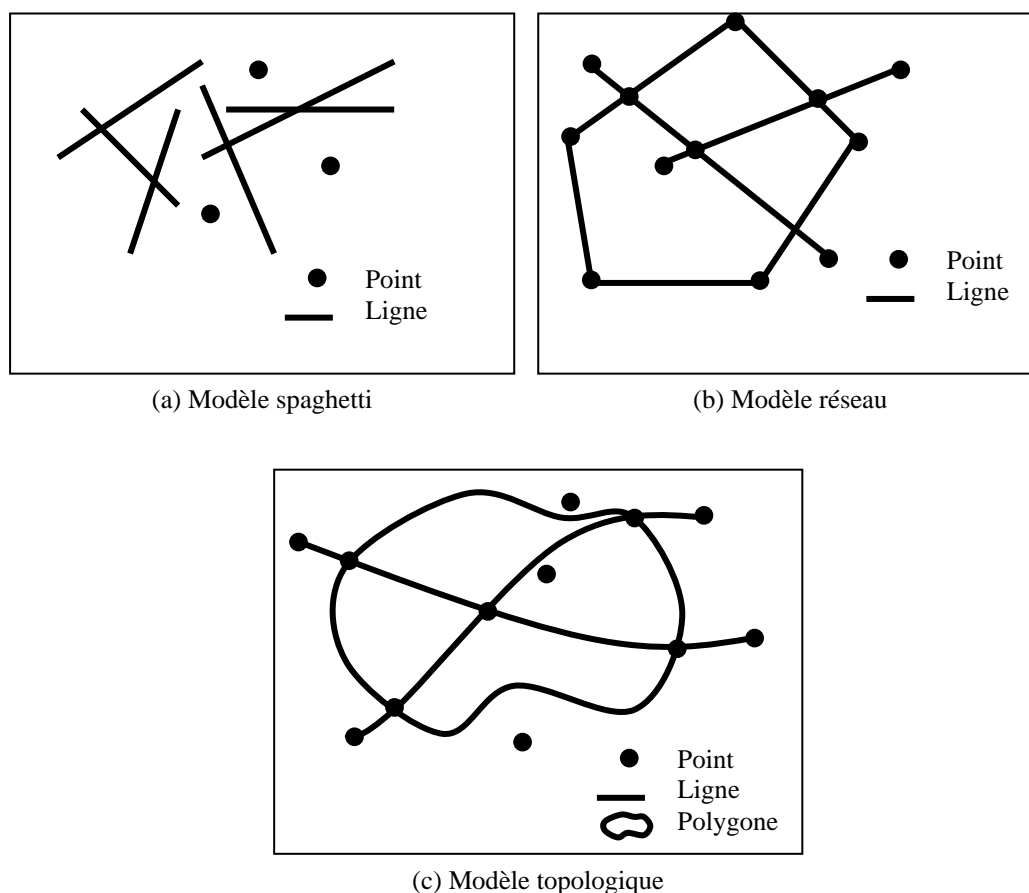


Figure I.5 : Modèles de représentation géométrique

I.6. Sources de données

Il existe quatre (04) types de données considérées comme sources fondamentales d'informations géographiques :

I.6.1. Données localisées

Les données localisées sont de larges données issues d'enquête, de recensements, de l'observation de terrain, des statistiques, Une fois chargées, elles doivent être réduites et ajustées pour être au mieux exploitées. L'utilisateur pourra ensuite choisir entre différents traitements tels que la modélisation du terrain, la localisation et l'analyse de déformations,

I.6.2. La carte

La carte est la plus ancienne source d'informations géographiques. Il existe deux types de carte :

- Les cartes topographiques : ce sont des cartes « générales » dans le sens où elles ne représentent pas un but spécifique
- Les cartes thématiques : ce sont des cartes contenant des informations sur un sujet ou un thème précis.

I.6.3. La photographie

La photographie (généralement aérienne) est un moyen très utilisé pour obtenir de l'information spatiale. Une photographie prise au cours d'une reconnaissance du terrain, est une source de connaissances permettant d'établir les relations spatiales. Se sont cependant, des données analogiques comportant des imprécisions dues à l'altitude, à la couverture nuageuses, Le travail du photo-interprète consiste à étudier des photographies pour y connaître certains objets et en déduire des éléments d'information. Pour cela, il a besoin d'avoir à sa disposition toutes les informations géographiques qui sont rattachées soit au lieu de la prise de vue aérienne, soit aux objets à reconnaître.

I.6.4. Image satellitaire

Comme son nom l'indique, c'est une image reçue depuis l'espace. L'image satellitaire brute n'est pas utilisable. Elle comporte des déformations dues à la géométrie de prise de vue, à la topographie et à l'interface de propagation qui est l'atmosphère. La télédétection consiste alors, à corriger, gérer et exploiter ces images.

I.7. SGBD Géographique et langage de requêtes

Un SGBD est un outil de gestion de base de données qui facilite le stockage, l'organisation et la gestion des données.

Comment représenter dans le SGBD les objets géographiques ? Cela dépendra du modèle conceptuel de données auquel obéit le SGBD.

La modélisation de la base de données s'inscrit dans le processus global de planification du SIG. Il s'agit d'une action qui concerne à la fois les données et les traitements.

La phase de modélisation doit être réalisée avec rigueur, car elle détermine l'architecture de la base de données et par conséquent les opérations d'analyse.

I.7.1. Modèle entité-relation :

La modélisation entité-relation proposée par CHEN 1977, est très utilisée en géomatique et en informatique de base de données pour produire les schémas de données.

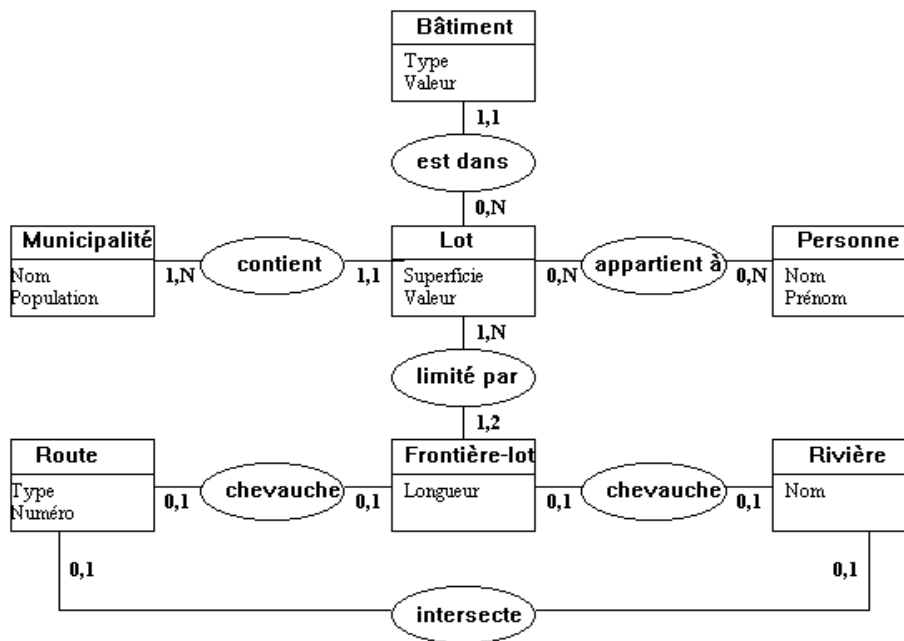


Figure I.6 : Modèle entité-relation

Le schéma conceptuel permet d'exprimer les relations entre entités situées sur des couches différentes. Ainsi le schéma précédent (Figure I.6) énonce le fait que les routes et les lots peuvent parfois avoir une frontière commune.

I.7.2. Modèle relationnel :

Proposé par E.F Codd au début des années 70, en usage depuis le début des années 80 (Oracle, EMPRESS, Informix, Access,...). C'est le modèle le plus utilisé dans les SIG, les données y sont représentées sous forme de tables utilisant certains champs comme lien. Cette approche peut paraître simple, offre une grande souplesse et une fiabilité permettant aux SIG de s'adapter à tous les cas de figure.

Le schéma conceptuel présenté en (Figure I.6) se traduit par le schéma relationnel suivant (Figure I.7) :

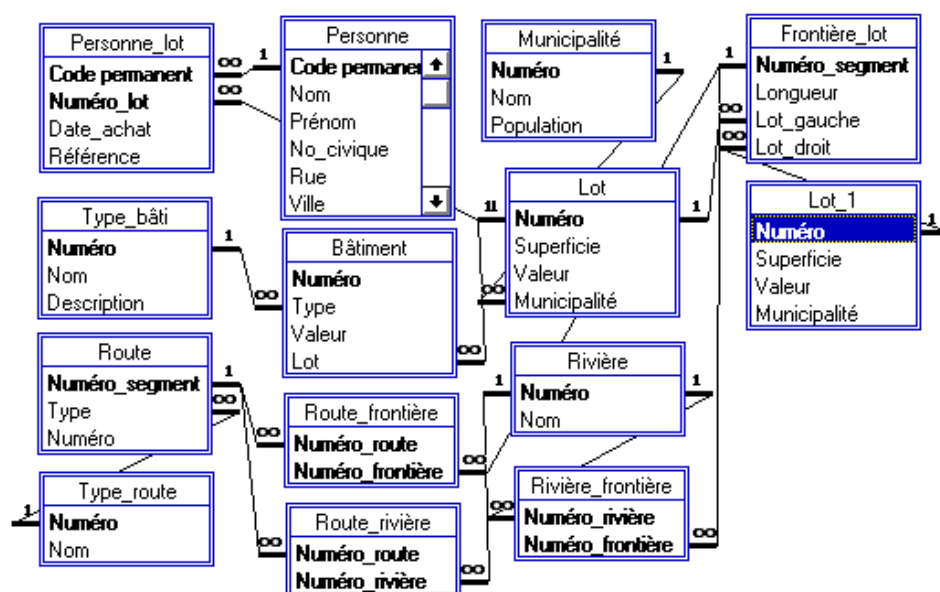


Figure I.7 : Modèle relationnel

Dans l'approche relationnelle, les requêtes d'un SGBD géographique vont s'appuyer sur le standard SQL qui est alors étendu en fonction des besoins pour les requêtes et le traitement de l'information géographique.

Exemple :

```
SELECT numéro, valeur
FROM Bâtiment
WHERE valeur > 100000;
```

"Afficher le numéro et la valeur des bâtiments dont la valeur est supérieure à 100000\$"

I.7.3. Modèle relationnel étendu aux types abstraits de données : [SCH96]

Il est possible de coupler un SGBD relationnel avec un système spécifique qui traite l'information spatiale. On ajoute alors aux attributs d'une relation, des attributs de nouveaux types de données comme des points, des lignes et des polygones ainsi que de nouvelles opérations sur ces types et notamment sur le langage de requêtes.

Exemple :

Réseau routier :

Route

Num-route : integer,

Nom-route : string

Tronçon

Num-tronçon : integer,

Nb-voies : integer,

Num-route : integer,

Géométrie : ligne

Ville

Nom-ville : string,

Géométrie : polygone

I.7.4. Modèle orienté-objet : [SCH96]

Les fournisseurs de SGBD ont compris l'immense importance de ce type particulier d'information et commencent à proposer des solutions intéressantes pour stocker, analyser et gérer cette information. En fait, la nouvelle orientation des SGBD est de fournir des outils pour gérer tout type de données, ce qui nécessite d'étendre le modèle relationnel le plus souvent utilisé. On voit donc apparaître des SGBD orientés objets qui doivent permettre de gérer des types de données complexes. Deux des principaux fournisseurs de SGBD actuels que sont Oracle et Informix proposent actuellement leurs serveurs universels (Universal Server), qui font partie de cette nouvelle génération de SGBD, et permettant l'intégration de données complexes.

Le modèle orienté objet est basé sur les concepts d'objet et de classes d'objets. Les objets sont les instances de classes. Les classes ont des propriétés structurelles (attributs) et des propriétés comportementales (méthodes). Les interrelations entre classes d'objet permettent de définir des notions d'héritage, d'agrégation, de généralisation, de spécialisation,

Le modèle orienté-objet représente les objets géographiques de façon homogène par un ensemble de classes (chaque couche est représentée par une classe) décrites par des attributs et des relations, instanciées par des objets.

Exemple de classes :

Réseau routier :

Class Route

Tuple(Num-route : integer,
 Nom-route : string,
 Tronçons : set(Tronçon),
 Intersect (l : ligne) : bool)

Class Tronçon

Tuple(Num-tronçon : integer,
 Nb-voies : integer,
 Route : Route,
 Géométrie : ligne,
 Length (l : ligne) : real)

Class Ville

Tuple(Nom-ville : string,
 Routes : set(Route),
 Géométrie : polygone,
 Adjacent (z : polygone) : bool)

La plupart des SGBD orienté-objet récents fournissent un langage de requêtes dont la syntaxe n'est pas sans rappeler celle de SQL mais qui est plus puissant. Le groupe d'éditeurs de SGBD orienté-objet réunis au sein de l'ODMG (Object Database Management Group) a choisi comme norme de langage **OQL** (langage fonctionnel avec une syntaxe SQL).

Exemple :

- Dessiner la route national 66

```
SELECT t.géométrie
FROM r in Routes, t in r.tronçons
WHERE r.nom-route = 'N66'
```

I.8. Les fonctions d'un SIG

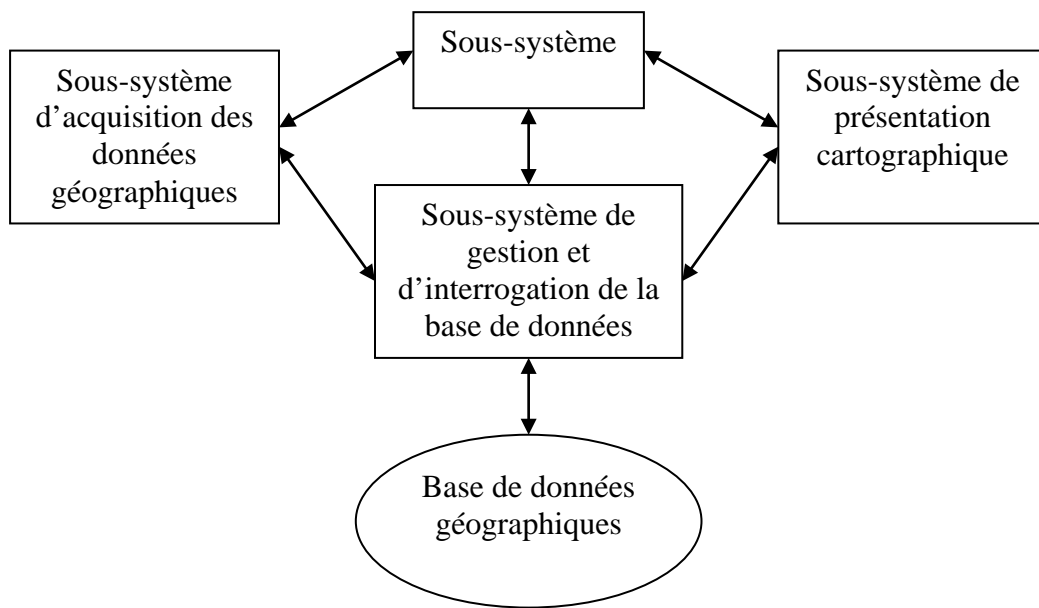


Figure I.8 : Les fonctions d'un SIG [2]

I.8.1. Sous-système d'acquisition des données

Il existe différents procédés pour l'intégration des données dans un SIG : la digitalisation, la scannérisation, l'intégration directe via des passerelles depuis d'autres logiciels, l'acquisition par GPS,

La création d'une base de données géographiques nécessite les quatre (04) phases suivantes :

- Entrée des données spatialement référencées.
- Pré-traitement.
- Entrée des données descriptives.
- Mise en relation des données spatiales et descriptives.

I.8.2. Sous-système d'analyse

Le sous-système propose des opérations permettant à tout utilisateur de créer de l'information à partir de la base de données géographiques. Il s'agit de :

- La construction de la base de données : concevoir le schéma conceptuel.
- L'entrée des données dans le système selon les contraintes : topologiques, géométriques et attributaires.
- La mise en place de ces données en différentes couches.

1.8.3. Sous système de gestion et d'interrogation

Le sous système de gestion et d'interrogation organise les données de manière à ce qu'elles soient facilement accessibles à l'utilisateur pour des analyses afin de permettre des mises à jour sur la base de données spatiales.

L'interrogation se fait par un langage de requête muni de fonctions qui permettent le traitement de la géométrie (ex : GSQL)

Quelques requêtes spatiales :

- ***Localisation*** : Qu'y a-t-il à cet endroit ?
- ***Condition*** : Ou est ce ?
- ***Distribution*** : Quelle est la répartition spatiale de ce ?
- ***Evolution*** : Qu'est ce qui à changé depuis ?
- ***Modélisation*** : Que se passe t-il si ?
- ***Optimisation*** : Quelle est la meilleur façon de ?

1.8.4. Sous système de présentation

Les produits qu'un SIG est capable de fournir sont :

- Des tableaux,
- Des graphiques, figures, histogrammes,
- Des cartes,
- Des images, en particulier des visualisations en trois dimensions, voir des images animées.

Pour élaborer des documents cartographiques, le SIG doit permettre à l'utilisateur de :

- Choisir la zone à représenter,
- Ajuster l'échelle,
- Choisir les couleurs et les symboles de représentation,
- Placer du texte, des annotations,
- Créer une légende,
- Compléter l'habillage de la carte par des cadres, logos, orientations,

I.9. Les outils de développement des SIG [ALL01]

Depuis une dizaine d'années, l'utilisation des Systèmes d'Information Géographique (SIG) croît dans de nombreux domaines d'activités. Les logiciels de cartographie et d'analyse spatiale sont adaptés à tous ceux qui ont à gérer des informations localisées quelle que soit leur échelle de travail. Il existe dans la demande de solutions géoinformatiques toute une hiérarchie d'utilisateurs dont les besoins et les compétences sont très différentes. Chaque utilisateur doit donc pouvoir trouver sur le marché l'outil correspondant à ces besoins.

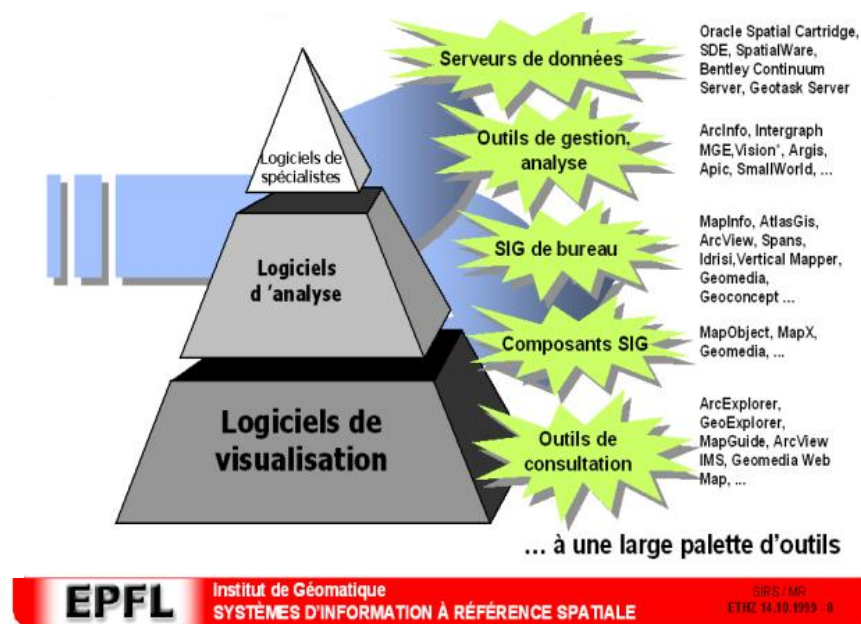


Figure I.9 : Hiérarchie de logiciels SIG

Parmi les logiciels SIG proposés sur le marché, on retrouve (Figure I.9) :

- **Outils de gestion et d'analyse** : ArcInfo, Intergraph, MGE, Vision, SmallWorld, ...
- **SIG de bureau** : MapInfo, ArcView, AtlasGis, Spans, Idrisi, Géomédia, Geoconcept,
- **Composants SIG** : MapObject, MapX, Geomedia,
- **Outils de consultation**: ArcExplorer, GeoExplorer, MapGuide, ArcView, IMS, Geomedia Web,
- **Serveurs de données** : Oracle Spatial, SDE, SpatialWare, Bentley Continuum Server, Geotask Server,

I.10. Conclusion

Dans ce chapitre, nous avons mis en évidence les concepts de base définissant le domaine des Systèmes d'information géographique. Il a été consacré à définir :

- Le domaine SIG : sa définition, domaine d'utilisation, ses fonctions et ses outils.
- La base de données géographique.
- Le système de gestion des bases de données géographique et son langage de requêtes.

Le chapitre suivant complète l'introduction au domaine SIG en introduisant les concepts : spatial et temporel.

CHAPITRE II

Chapitre II

Concepts Spatial et Temporel

II.1. Introduction

L'information géographique comme elle est définie dans le chapitre I, n'est pas suffisamment structurée pour qu'elle soit utilisée efficacement dans un SIG. Il est nécessaire de connaître certaines propriétés géométriques et temporelles des objets géographiques (leur structure, topologie, évolution, ...) pour modéliser complètement l'information géographique. Ces propriétés sont présentées dans ce chapitre selon deux (02) axes : Concept Spatial et Concept Temporel.

II.2. Concept Spatial

Les données géographiques possèdent des propriétés spécifiques qui les distinguent des informations généralement utilisées par les bases de données traditionnelles. Leur saisie impose de suivre une procédure préalable de structuration cartographique (représentation discrète et continue), de choix d'un système de projection et d'identification d'un système de coordonnées. [CLA98]

Les éléments de géométrie utilisés dans la modélisation des ces données sont définis et organisés selon un modèle géométrique. Pour cela, deux catégories de structure sont utilisées :

II.2.1. *Structuration continue* [CLA98]

Ce mode de structuration décompose l'espace avec une grille régulière de cellules (Pixels) ordonnées pour former une matrice régulière. Cette dernière produit un espace continu où chaque pixel possède une seule valeur alphanumérique qui correspond à l'attribut de l'entité géographique (Figure II.1).

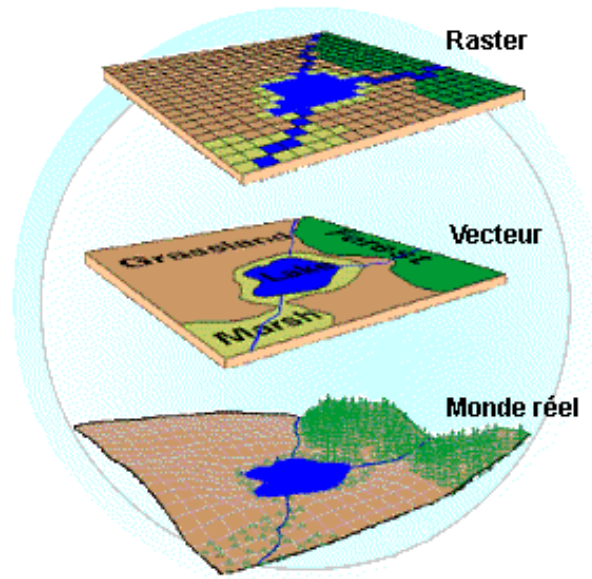


Figure II.1 : Différents types de représentation géométrique (Image ESRI⁽¹⁾)

II.2.2. Structuration discrète [CLA98]

Le monde réel peut être vu comme un ensemble d'objets statiques ou dynamiques. On peut aussi le voir comme un ensemble de couches superposées dont chacune d'elles regroupe des objets de même type.

Dans le cas discret, les objets qui composent le monde réel sont représentés par un modèle géométrique. Ceci se base essentiellement sur des primitives géométriques de type : Point, Ligne, polygone (surface). Par exemple, un bâtiment peut être vu comme objet de type point, une route comme une ligne et une parcelle comme surface. (Figure II.1).

e- Choix de la forme géométrique

La représentation de l'information dans la base de données est déterminée par des choix de perception et d'abstraction de la réalité, des choix dictés par les besoins des applications à réaliser.

Les formes géométrique décrivant l'objet dépendent des niveaux de représentation. Ces derniers, se caractérisent essentiellement par l'échelle de représentation et la résolution d'acquisition.

a.1- Echelle

Par définition, l'échelle d'une carte est définie par le rapport constant existant entre les longueurs mesurées sur la carte et les longueurs correspondantes à la mesure réelle de

(1) ESRI : un des principaux éditeurs sur le marché des logiciels SIG

l'objet [3]. Elle est exprimé par une fraction telle que 1/50 000. Ce qui signifie que 1mm sur la carte représente 50 000mm, soit 50m sur le terrain (FigureII.2).

La détermination de la géométrie correspondante à un objet du monde réel dépend du niveau de représentation choisie. Une représentation à petite échelle décrit la zone à étudier de façon moins détaillée. Alors que la représentation à grande échelle décrit la zone de façon plus détaillée. Par exemple, un bâtiment sera représenté par la surface correspondante à son emprise à une représentation plus détaillée (grande échelle) et par un point donnant juste sa position à une représentation moins détaillée (petite échelle) (FigureII.2).

a.2- Précision

La précision exprime la qualité géométrique des données : si l'on considère la carte comme une matrice de points (pixels), chaque point est porteur d'une information visuelle (couleur, appartenance à un objet géographique, proximité d'un point décrit (point géodésique par exemple) (Figure II.2). [3]

L'échelle et la précision de l'information géographique représentée sont aussi très souvent confondues par les utilisateurs alors que ce sont deux notions fondamentalement différentes (Figure II.2).

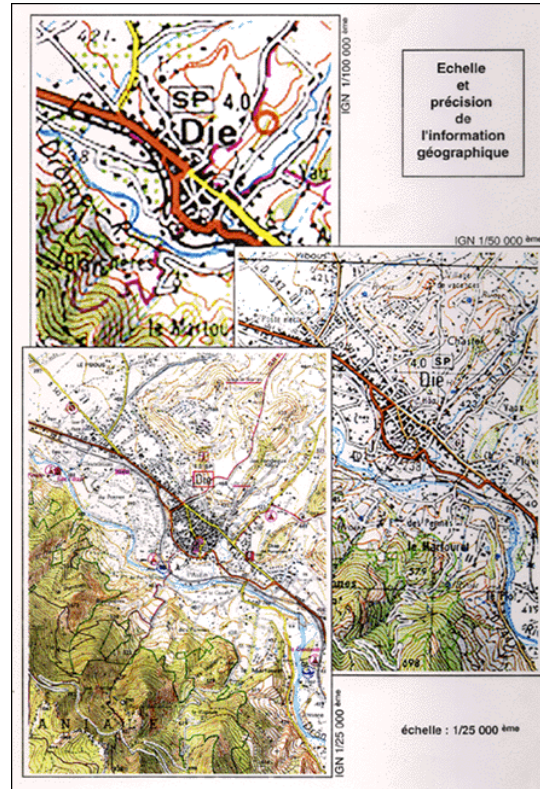


Figure II.2 : Échelle et précision de l'information géographique (Carte IGN⁽¹⁾)

f- Intégration de la géométrie dans la base de données

Tout objet peut être localisé dans l'espace selon trois systèmes différents : les coordonnées cartésiennes, les coordonnées géographiques, les coordonnées dans un système de projection. [LAU93]

b.1. Les coordonnées cartésiennes

Les coordonnées cartésiennes correspondent aux distances X, Y, Z selon trois axes orthogonaux passant par le centre de la terre. Les axes Ox et Oy sont dans le plan de l'équateur, Oz est confondu avec l'axe de rotation de la terre, le plan des axes Ox, Oy contient le méridien origine de Greenwich.

b.2. Les coordonnées géographiques

Les coordonnées géographiques (latitude, longitude, altitude) sont très répandues. La latitude est l'angle que forme la droite passant par le centre de la terre avec le plan de l'équateur. La longitude est l'angle de cette droite avec le plan contenant le méridien de référence (Greenwich) et l'axe de rotation (plan perpendiculaire au plan de l'équateur). L'altitude n'est pas exprimée depuis le centre de la terre (difficile à situer), mais depuis la surface de référence dit géoïde).

b.3. Les coordonnées dans un système de projection

Les coordonnées cartésiennes et les coordonnées géographiques ne peuvent être représentées que sur un globe (trois dimensions). Pour pouvoir réaliser des représentations sur cartes à deux dimensions, il est nécessaire de projeter les coordonnées sur un plan.

Passer des coordonnées géographiques en trois dimensions aux coordonnées en projection à deux dimensions nécessite d'abord de ramener tous les points de la surface de la terre sur l'ellipsoïde de référence en vue de simplifier les calculs de projections (Figure II.3).

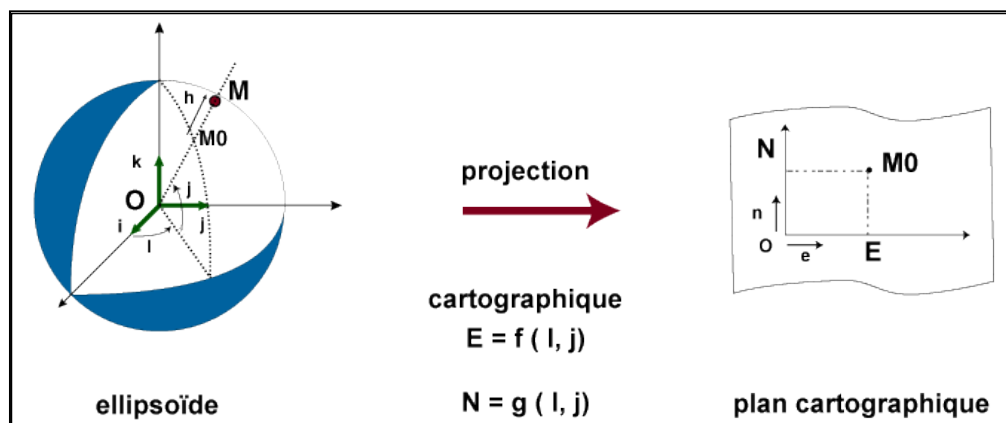


Figure II.3 : Principe de la projection cartographique [3]

L'intégration de la géométrie des objets dans la base de données se fait à partir d'un système de coordonnées. Généralement, un **point** est représenté par ses coordonnées (X, Y) tandis qu'une **ligne** est représentée par un ensemble de points qui la composent. Un **polygone** quant à lui est composé d'un ensemble de points et de lignes (Figure II.4).

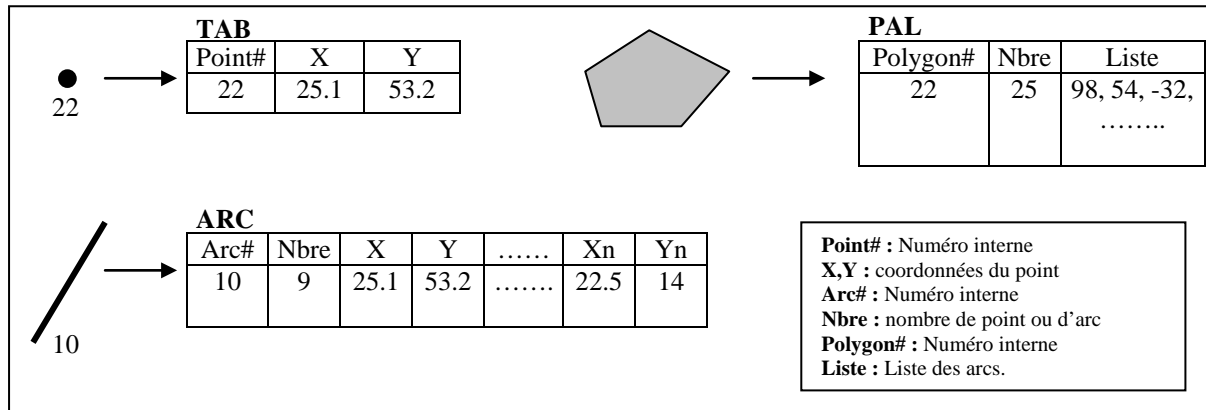


Figure II.4 : Représentation des primitives géométriques dans la base de données de ArcInfo [ALL01]

c- Relations spatiales [PAR97][KAR00]

Une relation spatiale représente un lien entre deux ou plusieurs objets spatiaux et éventuellement d'autres objets non spatiaux, dont la sémantique est une contrainte d'intégrité spatiale définie [PAR97]. Par exemple, la relation « Adjacence » relie des objets de type bâtiment dont la contrainte d'intégrité est « se trouvant à une distance réciproque de moins de 300m ».

Les relations spatiales peuvent être de différents types : topologique, réseau et agrégation.

c.1. Relations topologiques

Il existe un grand nombre de relations topologiques. Avant de présenter les différents types de relations, nous allons définir deux notions mathématiques.

- a- A^0 : intérieur de l'objet A.
- b- $dim(A)$: dimension de l'objet A ou $dim(A) \in \{0,1,2\}$.
 - $dim(A) = 0$: A est de type Point.
 - $dim(A) = 1$: A est de type Line.
 - $dim(A) = 2$: A est de type Polygone

- **Disjonction :**

La disjonction est définie par le fait qu’aucun partage n’existe entre les objets. Elle est valable pour tout type d’objets spatiaux (Figure II.5).

Relation mathématique :

A, B deux objets spatiaux

A disjoint de B = VRAI $\Leftrightarrow A \cap B = \phi$

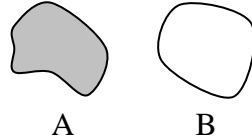


Figure II.5 : Relation de disjonction

- **Adjacence :**

Deux objets spatiaux sont adjacents si est seulement si il existe un partage entre eux sans intérieur commun (Figure II.6).

Relation mathématique :

A, B deux objets spatiaux, A et B \notin classe d’objets de type Point.

A adjacent à B = VRAI $\Leftrightarrow (A^\circ \cap B^\circ = \phi) \wedge (A \cap B \neq \phi)$

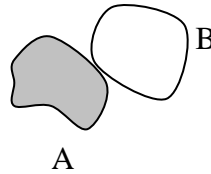


Figure II.6 : Relation d’adjacence

- **Croisement :**

C’est un partage d’une partie de l’intérieur des objets spatiaux telle que la dimension de la partie partagée est strictement inférieure à la plus grande dimension des deux objets en relation (Figure II.7).

Relation mathématique :

A \in classe Ligne, B \in classe Line ou Polygone

(A croise B) = VRAI $\Leftrightarrow (\dim(A^\circ \cap B^\circ) = \max(\dim(A^\circ), \dim(B^\circ)) - 1) \wedge (A \cap B \neq A) \wedge (A \cap B \neq B)$

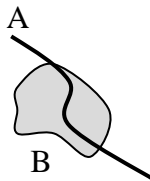


Figure II.7 : Relation de croisement

- **Recouvrement (chevauchement) :**

Le recouvrement est le partage de deux objets spatiaux de même type telle que la dimension de la partie partagée est égale à la plus grande dimension des deux objets en relation (Figure II.8).

Relation mathématique :

$A, B \in \text{classe Ligne ou Polygone}$

$A \text{ chevauche } B = \text{VRAI} \Leftrightarrow (\dim(A^\circ) = \dim(B^\circ) = \dim(A^\circ \cap B^\circ)) \wedge (A \cap B \neq A) \wedge (A \cap B \neq B)$

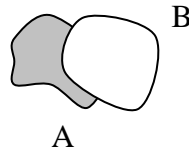


Figure II.8 : Relation de recouvrement

- **Inclusion :**

Un objet spatial est inclus dans l'autre si et seulement si la totalité de l'intérieur de l'un correspond à une partie de l'autre (figure II.9).

Relation mathématique :

$A, B \text{ deux objets spatiaux, } B \in \text{classe Polygone}$

$A \text{ est inclus dans } B = \text{VRAI} \Leftrightarrow A \cap B = A$

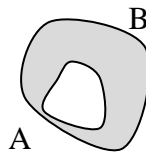


Figure II.9 : Relation d'inclusion

- **Egalité :**

Cette relation est valable pour des objets de même type tel que les deux objets en relation partagent la totalité de leur intérieur et de leur contour (Figure II.10).

Relation mathématique :

$A, B \text{ deux objets spatiaux de même type}$

$A \text{ Egale à } B = \text{VRAI} \Leftrightarrow (A \cap B = A) \wedge (A \cap B = B)$

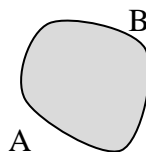


Figure II.10 : Relation d'égalité

c.2- Agrégation spatiale

Par définition, une agrégation est une relation binaire spécifiant que les objets de la classe cible sont des composants de la classe source. L'agrégation permet de représenter des associations de type Maître/Esclave, Tout/Partie, Composé/Composants.

Une agrégation spatiale exprime le fait que l'entité géographique composite est l'union géométrique de ses composants. Par exemple, l'entité wilaya est reliée par une agrégation spatiale à l'entité géographique commune (Figure II.11). [PAR97][KPS97]

Ce type d'agrégation définit certaines relations topologiques dans lesquelles on retrouve :

- **Couverture** : la région de l'espace occupée par l'objet composite est égale ou incluse dans l'union de ses parties.
- **Contenance** : l'objet composite regroupe des objets de type quelconque sans contraintes topologiques.
- **Partition** : l'objet composite regroupe le même type d'objets et les parties forment une partition complète de l'objet.



Figure II.11 : Exemple d'agrégation spatiale

c.3- Connectivité [1]

La relation de connectivité permet de connecter des liens (arcs) et des sommets (nœuds) pour former des réseaux linéaires (Figure II.12).

Généralement, la connectivité est utilisée pour modéliser des applications de tous domaines où les réseaux sont applicables. Par exemple, un réseau routier contient des routes représentées par des arcs et des carrefours représentés par nœuds qui sont reliés pour former un réseau linéaire planaire.

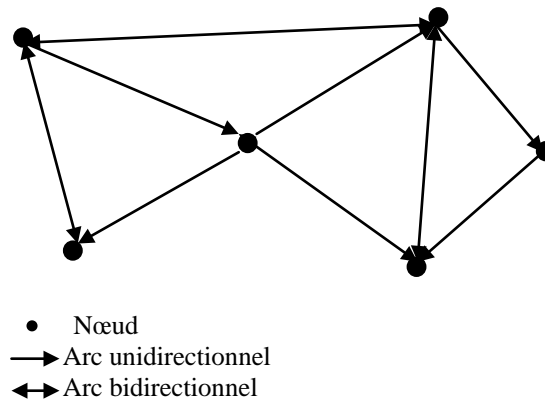


Figure II.12 : Connectivité dans un réseau

d. Contraintes d'intégrités spatiales [KAR00]

La vérification de la cohérence est un des soucis majeurs dans les bases de données ; elle s'effectue entre autres par la mise en oeuvre de contraintes d'intégrité. Pour les bases de données géographiques, les contraintes d'intégrité touchent la géométrie et la topologie, et sont liées au type de représentation géométrique.

Les contraintes d'intégrités spatiales peuvent être de type :

- **Sémantiques**

Les contraintes d'intégrités sémantiques s'appliquent aux bases de données validées par la sémantique des objets géographiques qui doivent être stockés. Un exemple de cette contrainte est la règle qui ne permet pas à un bâtiment d'être arrêté par un segment de rue.

- **Topologiques**

Le principe des contraintes d'intégrités topologiques est de définir formellement des règles spatiales qui peuvent être appliquées aux objets géographiques et aux relations topologiques. Par exemple, l'adjacence est définie comme « se trouve à une distance réciproque de moins de 300m »

- **Définies par l'utilisateur**

Ce sont des contraintes imposées par l'utilisateur pour certaines limitations dans son analyse. Par exemple, la distance entre une station de gaz et une école doit être supérieure à 300m.

II.3. Concept temporel

Depuis quelques années, le besoin et la nécessité d'une gestion temporelle de l'Information Géographique se fait ressentir dans la communauté des utilisateurs de SIG. De nombreuses applications relevant de différents domaines exigent en effet l'utilisation de données référencées dans le temps.

II.3.1. Evolution des objets géographiques

La connaissance de l'évolution temporelle des données, en particulier spatiales est souvent indispensable pour comprendre la dynamique des phénomènes du monde réel.

Dans un SIG, la temporalité exprime l'existence d'un objet géographique dans le temps. Elle permet de garder trace de son cycle de vie défini par des évènements de création, suspension, réactivation et de destruction (Figure II.13). [PAR97]

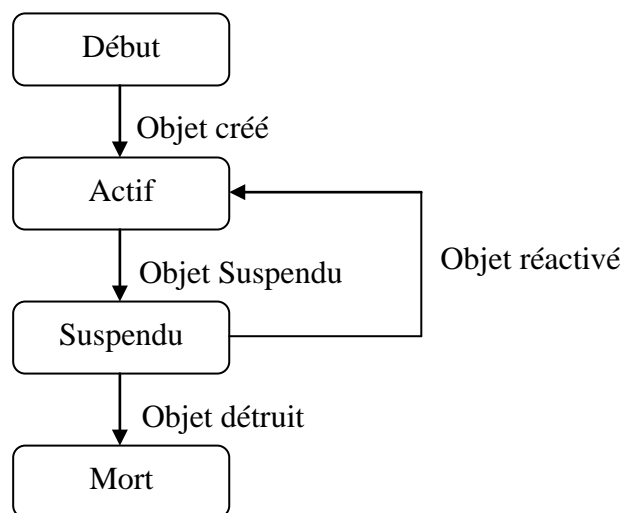


Figure II.13 : Cycle de vie d'un objet

Dans le temps, un objet évolue. Cela est dû au changement des propriétés de l'objet. Dans ce cas, on peut avoir deux types de changements [BED00] :

- a- **Descriptif** : il survient suite au changement de valeurs des attributs de l'objet. La temporalité est définie au niveau des attributs. on parlera ainsi, d'attributs temporels. Par exemple, le volume d'un barrage peut augmenter ou diminuer durant le temps. l'attribut volume sera défini alors comme étant un attribut temporel.

- b- **Spatial** : il survient suite au changement de la position ou de la forme de l'objet. On parlera dans ce cas d'une évolution de l'objet dans son type. Par exemple un bâtiment peut subir des changements dans sa forme par des nouvelles constructions.

II.3.2. Intégration de l'évolution dans la base de données

La composante temps dans une base de données est généralement gérée par une structure de données de type « Date ». Mais bien souvent, l'utilisateur sera plus intéressé de connaître l'historique et l'état des objets qui évoluent. Cela nécessite une représentation spécifique de la dimension temporelle.

Les SIG actuels expriment l'évolution des objets dans leur base de données par des primitives temporelles de type instant et intervalle. Un objet géographique peut alors avoir deux types d'existence dans l'espace : instantanée ou durable. Par exemple, un séisme possède une existence instantanée qui nécessite d'enregistrer qu'une seule date (Figure II.14-a). A l'inverse, une intervention sur les lieux du séisme peut aller d'un à plusieurs jours. Ce qui nécessite alors d'enregistrer deux dates (début et fin) (Figure II.14-b).

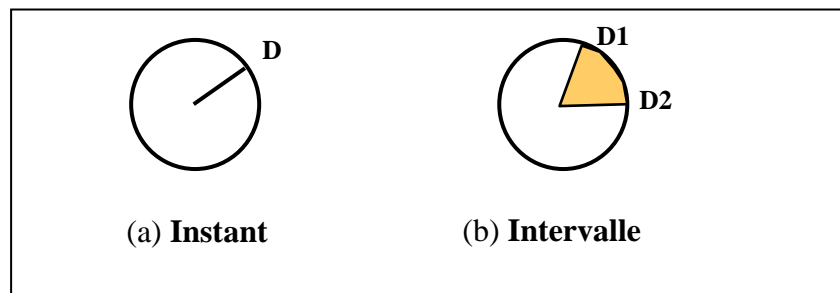


Figure II.14 : Primitives temporelles [PAR97][BED00]

La modélisation de l'existence est basée sur la définition de nouveaux concepts :

d- Attributs temporels

Un attribut temporel est un attribut dont la valeur est susceptible d'être modifiée au cours du temps. Il représente l'évolution temporelle d'un objet dans sa valeur. Par exemple, la valeur marchande d'un bâtiment peut varier dans le temps (Figure II.15-a).

Un attribut temporel est généralement identifié dans la plupart des diagrammes par la présence d'un pictogramme dont la valeur appartient au domaine d'intervalle {instant, intervalle} (Figure II.15-a).

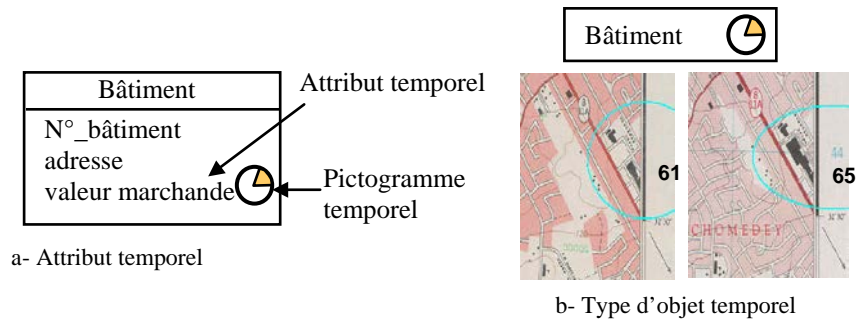


Figure II.15 : Type d'attributs et objets temporels

e- Types d'objets temporels [PAR97]

La temporalité concerne l'existence des objets dans leur type plutôt que de leur valeur. Elle représente l'évolution spatiale d'un objet. La temporalité sera alors définie sur la structure de l'objet (Figure II.15-b).

Un objet temporel peut avoir des attributs temporel ou non-temporels. Il n'y a pas de contraintes par défaut entre la durée de validité d'un attribut temporel et le cycle de vie de l'objet auquel il appartient.

La représentation des types d'objets temporels est visuelle sur les diagrammes par les mêmes pictogrammes que ceux définis pour les attributs temporels et qui appartiennent au même intervalle de valeurs (Figure II.15-b).

f- Types d'associations temporelles [PAR97]

Déclarer un type d'association temporelle permet de garder trace du cycle de vie de ses instances. Celles-ci peuvent être créées, suspendues, réactivées et détruites. Un type d'association temporelle peut relier des types d'objet temporels ou non-temporels, et vice versa. La sémantique serait de garder toutes les instances passées, présentes et futures. Par exemple, dans la figure II.16, le type d'association Entretien est temporel: dès lors, les couples (barrage, compagnie) passés, présents, et futurs sont gardés dans la base de données.

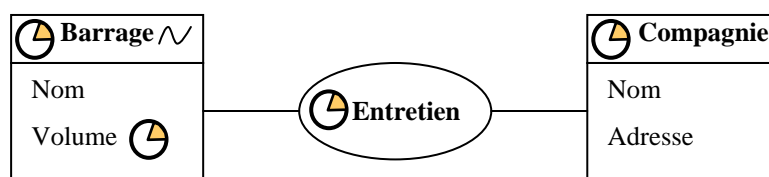


Figure II.16 : Type d'association temporel [PAR97]

c.1. Agrégation temporelle

L'agrégation temporelle est caractérisée par le fait que le type d'objet composé est temporel par sa sémantique : les objets composants représentent soit des instantanés, soit des périodes de la vie de l'objet composé. Cette association est très utile lors de l'intégration d'une base de données temporelle avec une autre non-temporelle décrivant les mêmes objets.

c.2. Association de transition

Ce type d'association modélise le changement de classe des objets : on dit qu'un objet subit une transition lorsque l'objet migre de la population d'un type d'objet source vers la population d'un type d'objet cible. Comme l'objet qui subit la transition garde son identité (puisque'il représente la même entité du monde réel), les types d'objet source et cible doivent appartenir à la même hiérarchie de généralisation. S'il est utile pour l'application de mémoriser cette transformation, une relation de transition peut être décrite. Par exemple, dans la figure II.17, l'association de transition « Devient » permet de garder trace des transformations des parcelles cultivées en parcelles bâties.

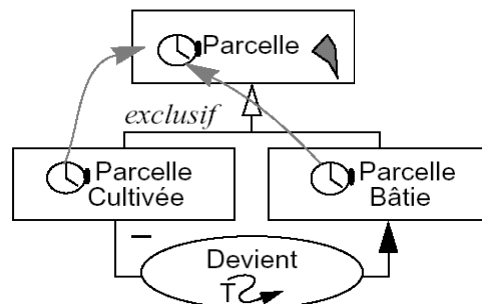


Figure II.17 : Exemple d'association de transition de la méthode MADS [PAR97]

c.3. Association de génération

Ce type d'association modélise les processus qui donnent lieu à l'émergence de nouveaux objets : une instance d'un type d'objet source génère une instance ou plusieurs instances d'un type d'objet cible. Cette association permet de modéliser des relations de causalité et de temporalité liées à l'apparition et à la disparition d'entités dans le monde réel.

Exemple d'association de génération : dans une application cadastrale, une parcelle peut être divisée pour donner naissance à plusieurs parcelles plus petites ou au contraire, plusieurs parcelles peuvent être fusionnées pour en créer une plus grande.

II.4. Conclusion

L'étude effectuée dans ce chapitre, a montré la nécessité de connaître certains aspects liés à la représentation spatiale des objets géographiques et de leur évolution dans le temps pour développer un SIG.

De plus, la diversité de concepts marqués montre aussi qu'un SIG n'est plus le système connu depuis 30ans comme étant un outil de gestion des données à composante géographique, mais plutôt un système d'information spécifique pour la gestion de l'information géographique. Ce qui a conduit les chercheurs à concevoir des méthodes de conception des SIG pour la gestion de l'information géographique.

CHAPITRE III

Chapitre III

Méthodes de conception des SIG : Modèles & outils AGL (Atelier de Génie Logiciels)

III.1. Introduction

En terme de modélisation, plusieurs méthodes de conception ainsi que des formalismes de représentation de la référence spatiale et temporelle ont fait l'objet de plusieurs recherches, parmi lesquelles on retrouve plus récemment Modul-R [CAR93][BED96], MADS [PAR 97], GEO-OOA [KPS97], CONGOO[PAN96], POLLEN [GAY97] Ces méthodes se basent essentiellement sur des formalismes conçus pour gérer les données classiques.

A l'apparition de l'approche objet et du génie logiciel, certaines méthodes de conception des SIG ont fait l'objet d'études approfondies pour intégrer des outils AGL (Ateliers de Génie Logiciel) dans leurs concepts de modélisation des données spatio-temporelles. On retrouve des outils AGL développés par les méthodes GEO-OOA [KPS97], MADS[PAR97], PERSPECTORY (extension de la méthode Modul-R) [BED00],

La modélisation des données et des traitements est une phase importante dans le processus d'élaboration des SIG, et l'utilisation d'un outil d'aide à la conception se révèle généralement nécessaire.

III.2. Spécificités conceptuelles des SIG

Plusieurs méthodes d'analyse et de conception ont été développées ces dernières années. Elles offrent de nombreuses solutions pour la structuration des systèmes d'information géographique, et intègrent toute la complexité liée à la gestion des données spatiales et temporelles.

Ces méthodes dérivent essentiellement des différents formalismes de conception des systèmes d'information. Elles consistent à enrichir leur niveau conceptuel afin de prendre en compte la représentation spatiale et temporelle des objets géographiques. On citera :

Méthode	Formalisme
GEO-ER	ER
Modul-R	ER
GEO-OOA	Objet/OOA
GEO-OM	Objet/OMT
POLLEN	Objet/OMT
MADS	ER et OO
PERSPECTORY	Objet/UML

Table III.1 : Méthodes de conception des SIG [CLA97]

La solution proposée par ces méthodes se base essentiellement sur les concepts spatial et temporel que nous allons définir dans les paragraphes suivants.

III.2.1. Concept spatial

Les méthodes de conception des SIG intègrent la spatialité des objets géographiques dans tous les concepts du modèle (classes, associations et attributs), mais elles diffèrent par leur façon d'intégrer cette spatialité dans leur schéma.

- Certaines méthodes associent la spatialité aux attributs. Les objets spatiaux sont décrits par des classes possédant un ou plusieurs attributs de type "GEOMETRY". Ceci est un type abstrait de données ayant pour structure un ensemble de primitives géométriques (Point, Ligne, Polygone, ...). ex : GEO2.
- D'autres méthodes représentent explicitement la spatialité dans le modèle par un ensemble de classes géométriques : Point, Ligne, Polygone. La spatialité des objets est définie alors par un lien d'association avec l'un des types de classes géométriques. ex : POLLEN et GEO-OM.
- D'autres encore représentent la spatialité par des pictogrammes associés aux classes. Elles utilisent une palette d'icônes de types spatiaux très complète : Simple (Point, Ligne, Polygone), complexe (Point et Ligne, ...), alternatif (Ligne ou Polygone, ...), multiple (un ou plusieurs géométries). ex : Modul-R, MADS, GEO-OOA.

Pour ce qui est de la topologie, trois besoins importants doivent être satisfaits par les méthodes de conception :

- Pouvoir contraindre des objets géographiques à respecter un lien topologique donné (affirmation de l'existence du lien).

- Pouvoir définir une sémantique propre pour ces liens (Par exemple, l'adjacence pourrait être définie comme « se trouvant à une distance réciproque de moins de 300m »).
- Pouvoir décrire explicitement ces liens dans le schéma conceptuel.

La plupart des méthodes expriment les liens (relations) topologiques par des associations classiques (ex : MADS), ou bien par des contraintes d'intégrités spatiales (ex : CONGOO), ou alors par un concept d'agrégation spatiale (ex : GEO2, GEO-OOA, MADS).

III.2.2. Concept temporel

La plupart des méthodes de conception des SIG utilisent les primitives de base **instant** et **intervalle** pour décrire la dimension temporelle dans le schéma conceptuel. Certaines méthodes représentent la composante temps dans le diagramme de classes par des pictogrammes temporels liés aux attributs qui changent de valeurs suite à un événement temporel (évolution descriptive des objets), ou bien à la structure de l'objet qui évolue spatialement dans le temps (changement de position ou de forme). On retrouve dans ce cas les méthodes GEO-OOA, MADS, PERSPECTORY, Par contre, d'autres méthodes préfèrent représenter les primitives temporelles associées aux attributs temporels sous forme de classes indépendantes. Par exemple : POLLEN.

III.3. Critères d'évaluation des outils AGL [POU97]

Les outils connus sous le nom AGL (Atelier de Génie Logiciel) en anglais CASE (Computer Aided Software Engineering) ont pour objectif d'aider les concepteurs dans leur développement des systèmes. Ces outils interviennent dans une ou plusieurs phases du cycle de vie d'un logiciel : conception, programmation, test,

Il existe essentiellement deux types d'AGL selon la nature des outils intégrés :

- les AGLs de conception (upper-case) : ces ateliers s'intéressent plus particulièrement aux phases d'analyse et de conception du processus logiciel. Ils intègrent généralement des outils pour l'édition de diagrammes, des dictionnaires de données, des outils pour l'édition des rapports, des générateurs de code, Exemple : PowerDesigner, Designer2000.

- Les AGLs de développement (lower-case) : ces ateliers s'intéressent aux phases d'implémentation et de test du processus logiciel. Ils intègrent généralement des éditeurs, des générateurs d'interfaces homme/machine, des SGBD, des compilateurs, optimisateurs, Exemple : Windev.

L'architecture standard des outils AGL utilisée lors de la mise en œuvre des systèmes d'information, doit néanmoins vérifier certains critères qui permettent d'indiquer comment et avec quels efforts, en pratique, la modélisation sera réalisée. Les tables 2 et 3 ci-dessous résument les différents critères qui sont liés aux fonctionnalités particulières des outils AGLs (table III.2) et à leurs interfaces (table III.3).

Critères	Description
Intégralité du processus	Réalisation de toutes les étapes de modélisation des données soit le conceptuel (MCD), logique (MLD) et physique (MPD) et également celle des traitements.
Conversion	Transformation simple, directe et automatique d'une étape de conception à une autre.
Interopérabilité	Conversion des schémas vers une large diversité de SGBD
Dictionnaire	Création et mise à jour automatique du dictionnaire de données. Permettre la génération automatique de fichiers rapports.
Vérification de l'intégrité	Vérification du schéma selon des règles proposées à l'utilisateur et activées ou non par celui-ci. Avertissements si certaines règles ne sont pas respectées et génération d'un rapport.
Reverse Engineering	Création automatique du modèle logique à partir du modèle physique de divers SGBD.
Génération de codes variés	Création automatique de codes dans des langages évolués (VBA, Pbuilder, Java, ...).
Lien Internet	Génération de codes HTML (texte et graphique)

Table III.2 : Critères d'évaluation reliés aux fonctionnalités des outils AGL

Critères	Description
Conception	Introduction directe et intuitive des entités, attributs et relations. Interaction entre le graphique et les formulaires de saisie. Accès direct au dictionnaire de données.
Capacité d'adaptation	Développement ou adaptation de son, graphisme et de certaines de ses fonctionnalités.
Convivialité et interface	Convivialité de l'interface et aide en directe (simple, compréhensible et efficace).
Flexibilité du dessin	Souplesse pour le dessin de conception des modèles et accessibilité à des outils de dessins complémentaires.
Robustesse	Stabilité du système et vitesse d'exécution raisonnable
Compatibilité	Echange avec d'autres outils AGL au niveau du dessin et du dictionnaire de données.
Plate-forme	Fonctionnement sur plusieurs plates-formes informatiques (PC, MAC, Unix).
Langue	Disponibilité des logiciels en différentes langues de travail.

Table III.3 : Critères d'évaluation reliés à l'interface des outils AGL

III.4. Méthodes de conception des SIG

Parmi les méthodes de conception des SIG développées, nous avons choisi celles qui possèdent un outil AGL (GEO-OOA, MADS, PERSPECTORY). Avant d'aborder ces méthodes, voici un scénario qui servira d'exemple illustratif :

« *La gestion d'une catastrophe naturelle tel qu'un séisme, nécessite une consultation de tout un patrimoine de données. Lors d'un séisme, nos stations de surveillance et avec l'intervention du personnel qualifié enregistrent les caractéristiques du séisme : temps origine (date, heure, minute, second), localisation (épicentre), magnitude, intensité (en relation avec les dégâts engendrés par le séisme). Ces informations permettent d'élaborer deux types de cartes : carte de sismicité, carte des intensités, qui vont nous aider à déterminer les zones à risques.*

Les différents séismes enregistrés sont engendrés par des failles. Ces dernières sont connues par leurs mécanismes de rupture, trace (longueur), ainsi que leur localisation. L'objectif d'intégrer les caractéristiques des failles dans le système, est de définir les zones les plus fracturées qui peuvent générer des séismes.

Pour réduire les dommages lors d'un séisme, il est nécessaire d'avoir des informations sur les organismes d'intervention et de secours : nom, adresse, téléphone, fax, emplacement géographique sur la carte, pour une intervention rapide sur les lieux de la catastrophe.

Enfin, les différents résultats obtenus et leur interprétation sont représentés sur la carte du découpage administratif des wilayas (daïras et communes). On associe à cette carte, des données alphanumériques telles que : surface, densité de population, ainsi que la distribution du type de construction (bâti dominant) par wilayas. Ces informations vont nous permettre d'avoir une idée sur la répartition de la population dans le but d'estimer la perte humaine et les dégâts en cas d'un séisme ». [ALL02]

III.4.1. GEO-OOA

GEO-OOA est une méthode orientée objet mise au point par Kôster, Pagl et Six [KPS97]. Elle propose d'enrichir le diagramme de classes de la méthode OOA (Object Oriented Analysis) de Coad et Yourdon [COA91] par un ensemble de classes spécifiques (Géoclasses et réseau) et des liens d'associations pour représenter les relations topologiques. Cette extension permet au concepteur des SIG de modéliser la représentation spatiale et temporelle des objets géographiques dans un modèle conceptuel.

a- Concept spatial

Pour modéliser les objets géographiques, la méthode utilise dans son diagramme de classes deux (02) types de classes :

- Des classes appelées « GEOCLASS » qui sont au nombre de quatre (04) : Point, Ligne, Région (Polygone) et raster (Figure III.1). Ces dernières admettent des opérations élémentaires tels que test d'inclusion, d'intersection,

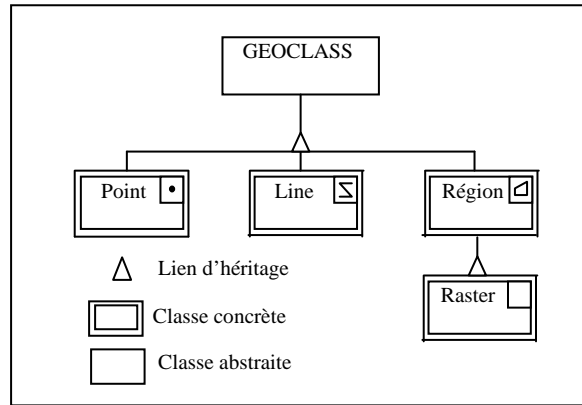


Figure III.1 : GEOCLASS

- Des classes réseaux NODE et LINK qui spécialisent la classe abstraite Network (Figure III.2). Cette dernière admet des opérations élémentaires telles que : test de connectivité, d'analyse de flux, recherche du chemin, Ce type de classe est souvent utilisé pour modéliser des domaines ayant une représentation graphique de type réseau. Par exemple, un réseau électrique est représenté par des classes réseaux dont les nœuds correspondent à des transformateurs ou des bâtiments reliés par des liens correspondants à divers types de lignes électriques.

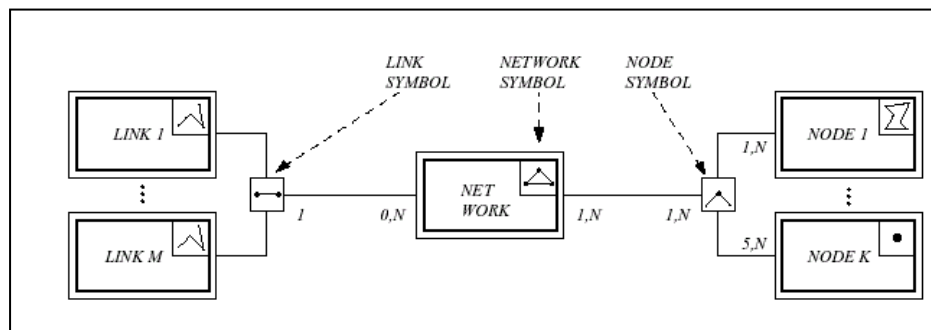


Figure III.2 : Structure réseau de GEO-OOA

Associations spatiales

La méthode GEO-OOA propose d'enrichir l'ensemble des associations classiques de la méthode OOA par d'autres types d'associations qui représentent les relations topologiques existantes entre les objets géographiques. Elle définit un autre type d'agrégation appelée « agrégation spatiale » qui permet de décrire les relations de (Figure III.3) :

- **Couverture :** la région de l'espace occupée par l'objet composite est égale ou incluse dans l'union de ses parties.

- **Contenance** : l'objet composite regroupe des objets de type quelconque sans contraintes topologiques.
- **Partition** : l'objet composite regroupe le même type de géoclasses et les parties forment une partition complète de l'objet.

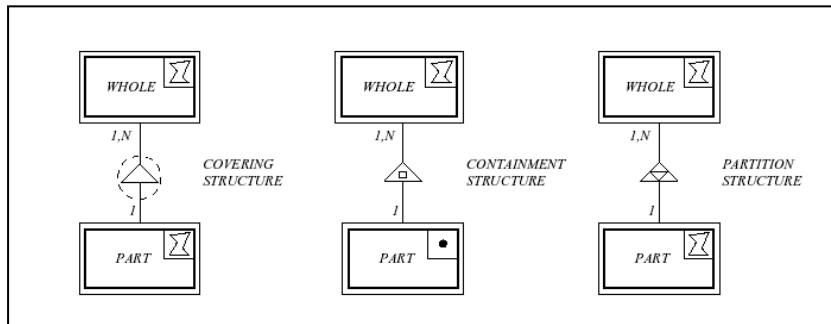


Figure III.3 : Agrégation spatiale de GEO-OOA

b- Concept temporel

La méthode propose d'introduire la dimension temporelle au niveau des classes et de leurs caractéristiques. Un objet temporel, instance d'une classe temporelle, est représenté par un ensemble de versions qui permet de suivre son évolution dans le temps. La création de ces versions est provoquée par un événement temporel associé à une opération modifiant tout attribut temporel de l'objet à une estampille t. De plus, cette opération fournit un lien de succession assurant la gestion des objets en gardant trace des ancêtres et des descendants de l'objet. Par contre, les opérations atemporelles associées aux classes temporelles agissent sur l'ensemble des versions de l'objet (Figure III.4).

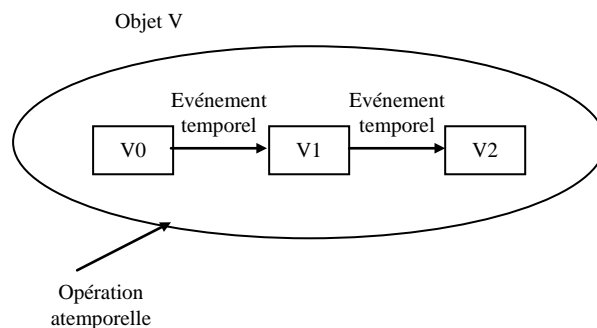


Figure III.4 : Evolution et versions de l'objet

c- Exemple illustratif

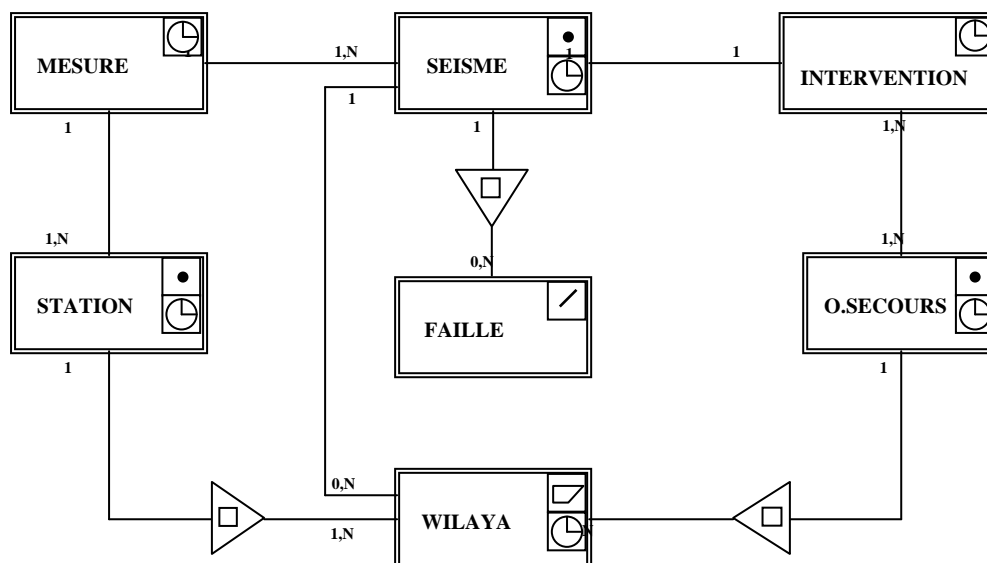


Figure III.5 : Exemple de modélisation selon GEO-OOA

d- Développement par l'outil AGL

La méthode GEO-OOA bénéficie d'un environnement de type AGL constitué :

- d'une boîte à outils : le méta-modèle GEO-OOA qui décrit les définitions de toutes les primitives GEO-OOA.
- d'un éditeur graphique utilisant des pictogrammes pour représenter le spatial et le temporel afin de permettre de spécifier le modèle d'une application.
- d'un système de gestion de modèles : chaque modèle peut être stocké sur support persistant.
- d'un convertisseur graphique vers un langage cible.

Discussion

L'approche GEO-OOA est bien une méthode de conception étendue aux problèmes de la gestion des données à références spatiale et temporelle. Tous les concepts des SIG sont bien définis. La méthode dispose d'une grande variété de classes spatiales (Géoclasse et Réseau) qui expriment la spatialité des objets géographiques. Elle dispose aussi d'un ensemble de relations topologiques (agrégation spatiale, réseau, ...) qui permet de définir les dépendances existantes entre les objets. Mais ces relations ne permettent pas de modéliser les différentes topologies possibles comme par exemple : l'adjacence, et le croisement.

La vision temporelle est définie par l'état d'un ensemble de versions de l'objet à une estampille t, et par la succession de ces versions (historique), mais il reste que cette notion n'est pas définie sur une base d'instant ou d'intervalle.

III.4.2. MADS (Modélisation d'Application à Données Spatio-temporelles)

MADS est une méthode orientée objet développée par C.Parent, S.Spaccapietra et E.Zimányi [PAR97] dans un contexte applicatif de gestion et aménagement du territoire : gestion pétrolière en Colombie, gestion du réseau urbain des claires et usées de la ville de Genève, Les auteurs de cette méthode proposent un modèle conceptuel de données sous forme d'entité-association conçu selon trois (03) axes : structurel, spatial et temporel.

a- Concept structurel

Le modèle de MADS proposé est de type Objet-relation étendu, qui offre les concepts usuels : objet, association, attribut, lien de généralisation, lien d'agrégation ainsi qu'un ensemble de contraintes d'intégrités associées.

b- Concept spatial

La méthode MADS propose une hiérarchie de type de représentation spatiale des objets. Cette représentation est décrite dans la classe par un attribut prédéfini « GEOMETRY » dont le domaine de valeur est un type abstrait spatial (table III.4).












Type spatial	Picto	Définition
géo		tout type spatial défini ci-dessous dans ce tableau
géosimple		tout type spatial simple (un point, une ligne, une ligne orientée ou une surface simple)
point		un point
ligne		toute ligne qu'elle soit droite, courbe, brisée, fermée ou non, orientée ou non.
ligne orientée		toute ligne orientée qu'elle soit droite, courbe, brisée, fermée ou non
surface simple		toute surface connexe (avec éventuellement des trous)
géocomposé		toute composition de types spatiaux simples
semis		une collection de points
graphe		une collection de lignes
digraphe		une collection de lignes orientées
surface complexe		une collection de surfaces simples

Table III.4 : Type abstrait de données spatiales

Associations spatiales

MADS ne se limite pas seulement à la représentation spatiale des objets, elle offre aussi des relations spatiales reliant ces objets et qui peuvent être de différents types : topologiques (table III.5, Figure III.6), agrégation (Figure III.7), généralisation (Figure III.8) et spécialisation (Figure III.9).

Type spatial	Picto	Définition
Disjonction		Aucun partage
Adjacence		Partage sans intérieur commun
Croisement		Partage d'une partie de l'intérieur
recouvrement		partage d'une partie de l'intérieur (valable pour des types spatiaux de même dimension)
Inclusion		la totalité de l'intérieur de l'un correspond à une partie de l'intérieur de l'autre
Egalité		Partage de la totalité (valable pour des types spatiaux de même dimension)

Table III .5 : Type de relations topologiques

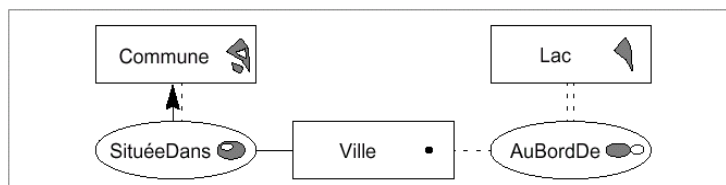


Figure III.6 : Exemple de relations topologiques

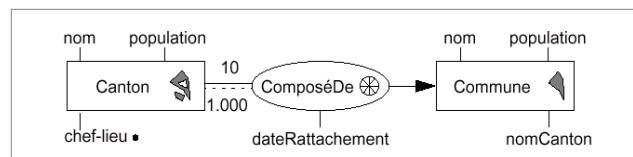
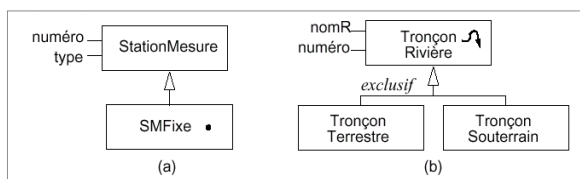
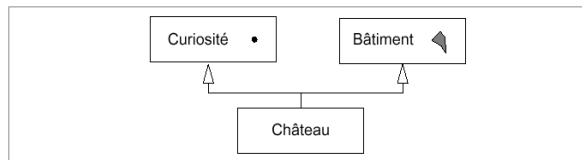


Figure III.7 : Exemple de lien d'agrégation



(a) le lien de généralisation peut relier des types d'objet spatiaux et non-spatiaux.
 (b) chaque objet de la sous classe a une spatialité héritée de la super classe.

Figure III.8 : Exemple de lien de généralisation



La spatialité des instances de la classe château peut avoir une des spatialités héritées (point ou surface) en spécifiant le point de vue considéré (Curiosité, Bâtiment)

Figure III.9 : Exemple de lien de spécialisation

D'autres types de relations peuvent être définies explicitement par l'utilisateur à l'aide des méthodes associées aux types abstraits spatiaux. Par exemple, une relation spatiale de proximité « ProcheDe » peut être définie entre deux types d'objet spatiaux ville et lac avec la contrainte d'intégrité spatiale distance (ville,lac) < 5km.

c- Concept temporel

La temporalité dans MADS est associée à tout concept du modèle : attribut, objet et association.

- Elle conserve les différentes valeurs d'un attribut prises au cours du temps.
- Elle conserve l'existence des objets dans leur type plutôt que leurs valeurs. Ce qui permet de garder trace de cycle de vie des objets définis par des événements de création, de réactivation et de destruction.
- De même, elle garde trace du cycle de vie des instances d'une association. Celle-ci peut être créée, suspendue, réactivée ou détruite.

De plus, la méthode MADS offre divers types d'associations temporelles. On peut citer :

- **Agrégation temporelle** : l'objet composant représente soit des instantanés, soit des périodes de vie de l'objet composé.
- **Agrégation instantanée** : l'objet composé est temporel tandis que le type d'objet composant ne l'est pas forcément.
- **Généralisation** : l'objet de la sous-classe possède deux cycles de vie. L'un hérite de la super classe qui décrit la vie totale de l'objet, et l'autre propre à l'objet.
- **Association de transition** : modélise le changement de classe des objets.
- **Association de génération** : modélise les processus qui donnent lieu à l'émergence de nouveaux objets.

d- Exemple illustratif

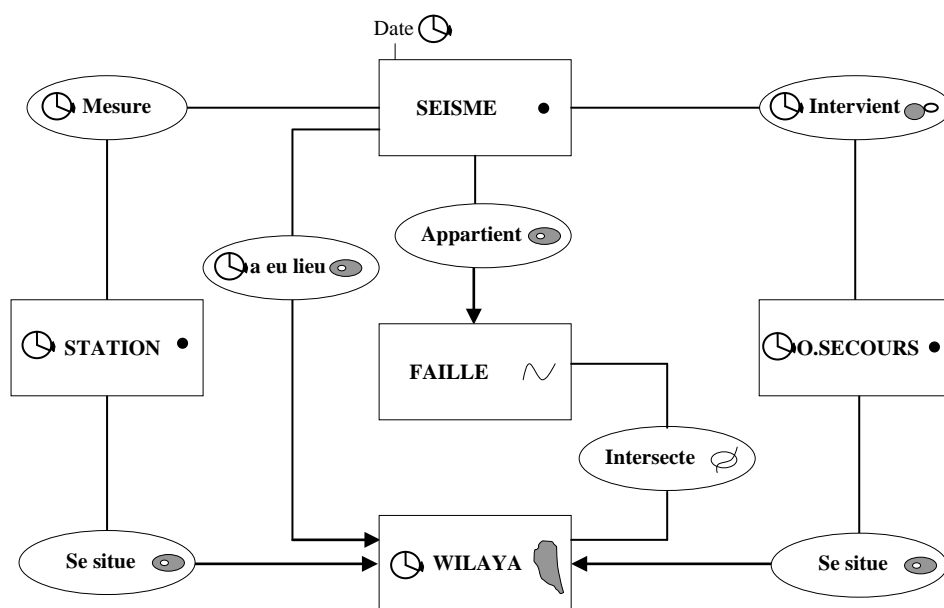


Figure III.10 : Exemple de modélisation selon MADS

e- Développement de l'outil AGL

Au delà des qualités du modèle conceptuel, MADS a réalisé un outil AGL très complet qui permet aux utilisateurs de dialoguer avec la base de données (Figure III.11). Cet outil possède :

- Un éditeur de schémas pour créer, modifier, conserver et consulter des schémas exprimés dans le modèle (Figure III.12).
- Des notations visuelles intuitives pour la représentation spatiale et temporelle.
- Des traducteurs capables de générer un schéma logique du schéma conceptuel et permettant de transformer un schéma MADS en son équivalent Entité-Association, INTERLIS ou relationnel.
- Un langage de manipulation associé au modèle conceptuel pour exprimer les requêtes d'interrogation dans le formalisme du modèle.
- Un éditeur de requêtes visuel permettant d'écrire une requête selon trois (03) modes :
 - Mode schéma* : l'utilisateur construit par manipulations directes la structure des objets désirés en résultat et spécifie les conditions à satisfaire.
 - Mode croquis* : l'utilisateur dessine le croquis de la configuration spatiale qu'il recherche.
 - Mode cartographique* : l'utilisateur définit les conditions spatiales que doivent satisfaire les objets sur des cartes affichées à l'écran.
- Un support d'un dictionnaire de données et un générateur de documentation HTML.

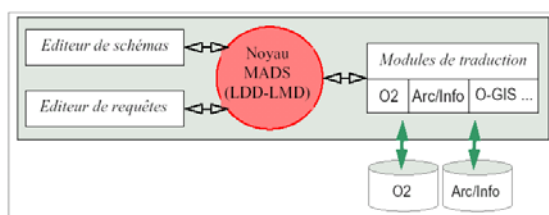


Figure III.11 : Architecture du système MADS

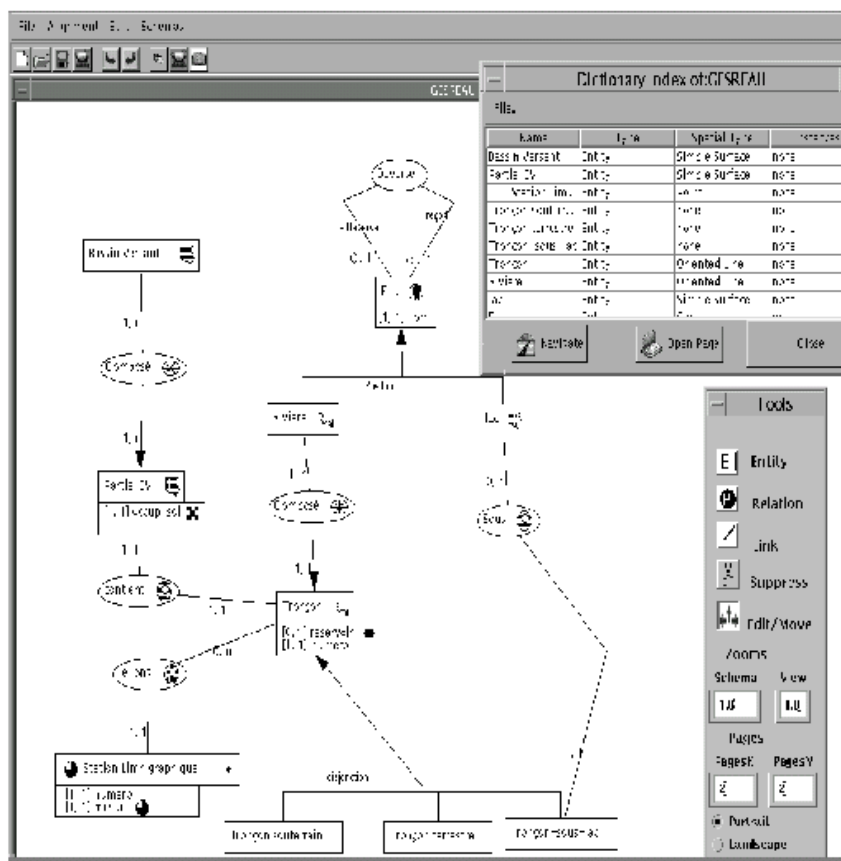


Figure III.12 : Editeur de schémas

Discussion

L'objectif qui a le plus marqué l'élaboration de la méthode MADS, est celui de l'orthogonalité entre les trois concepts : structurel, spatial et temporel. Cela veut dire qu'à une application donnée, il ne peut exister qu'une seule possibilité de représentation des objets dans le schéma conceptuel. Cette orthogonalité permet d'obtenir un modèle à la fois simple (les trois concepts sont indépendants) et puissant. Ce qui permet de regrouper dans une application des données classiques avec des données spatio-temporelles.

MADS permet aussi de décrire explicitement les relations topologiques (spatiales ou temporelles) entre les objets. La sémantique de ces relations peut être personnalisée pour satisfaire les besoins spécifiques d'une application.

Enfin, cette méthode dispose d'un outil visuel très complet, confort de son utilisation dû au fait de ne plus avoir à prendre en considération les particularités techniques des SIG ou SGBD.

III.4.3. PERSPECTORY 2000

PERSPECTORY est un outil AGL développé au centre de recherche en géomatique (Université de Laval, Canada) par l'équipe de Y.Bédard [BED00]. Il est basé dans sa modélisation, sur le formalisme UML, étendu avec des pictogrammes spatiaux et temporels afin de représenter les objets géographiques.

a- Concept spatial

PERSPECTORY utilise un langage PVL Spatial (Spatial Plug-in for Visual Language) qui offre une richesse d'expression nécessaire pour représenter la spatialité des objets géographiques dans le schéma conceptuel de UML. La description spatiale d'un objet est représentée par un pictogramme spatial à gauche de la classe associée, et si un des attributs de l'objet varie spatialement, un pictogramme spatial est défini à droite de cet attribut (Figure III.13).

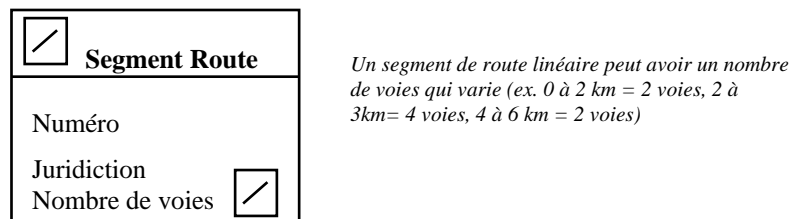


Figure III.13 : Exemple de classe spatiale de PERSPECTORY

Les différents pictogrammes spatiaux définis dans le langage PVL Spatial sont de type simple (table III.6) et composé (table III.7) :

Picto	Définition
	0 dimension (ex : une maison est représentée par un point)
	1 dimension (ex : un segment de route est représenté par une ligne)
	2 dimension (ex : un lac est représenté par un polygone)

Table III.6 : Pictogrammes simples de PVL

Picto	Définition
	complexe (ex : un réseau hydrographique est composé de rivières (ligne) et de lacs (polygone))
	alternatif (ex : un bâtiment est de forme point si la surface < 500m ² et polygone si elle est > 500m ²).
	nombre de formes d'un objet (0,1 avoir au plus une seule forme, 1,N avoir au moins une forme, 1,1 avoir une seule forme (par défaut)).
	multiple (ex : une ville est représentée par un polygone et son centre par un point)
	n'importe quelle forme
	forme décrite dans le dictionnaire de données
	dérivé (ex : une ville est représentée par une forme dérivées par d'autres informations géométriques).

Table.III.7 : Pictogrammes composés de PVL

On constate dans la modélisation de PERSPECTORY des pictogrammes sous forme italique. Cela signifie qu'un objet peut dériver une représentation spatiale par une association d'agrégation (Figure III.14).

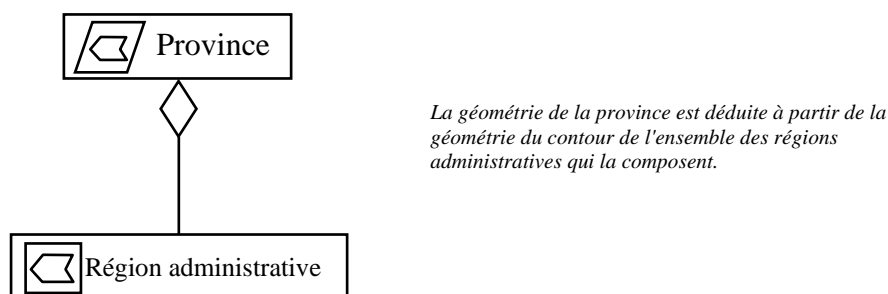


Figure III.14 : Exemple de lien d'agrégation

b- Concept temporel

La temporalité de PERSPECTORY est définie par l'évolution de l'objet dans le temps. Celle-ci peut être :

- **Descriptive** : elle survient suite au changement de valeurs des attributs de l'objet (Figure III.15-a).
- **Spatiale** : elle survient suite au changement de la position ou de la forme de l'objet (Figure III.15-b).



Figure III.15 : Exemple de l'évolution dans PERSPECTORY

Cette temporalité peut être instantanée ou durable. Elle est représentée dans le diagramme de classe de UML par une variété de pictogrammes temporels du langage PVL (table III.8).

Picto	Définition
	0Dimension - instantanée: enregistrement d'une date
	1Dimension - durable: enregistrement de deux dates
	<i>alternatif</i> : Un feu possède une <u>existence</u> durable lorsqu'il brûle durant plusieurs jours ou une, instantanée lorsqu'il ne brûle que quelques heures.
	<i>multiple</i> : Un festival a une <u>existence</u> ponctuelle si on considère le moment de son événement principal ou durable si on considère toute la durée du festival.
	tout temporalité possible
	<i>temporalité compliquée</i> (elle est décrite dans le dictionnaire)
	<i>dérivé</i> : les dates d'existence sont dérivées d'une autre classe

Table III.8 : Pictogrammes temporels de PVL

c- Exemple illustratif

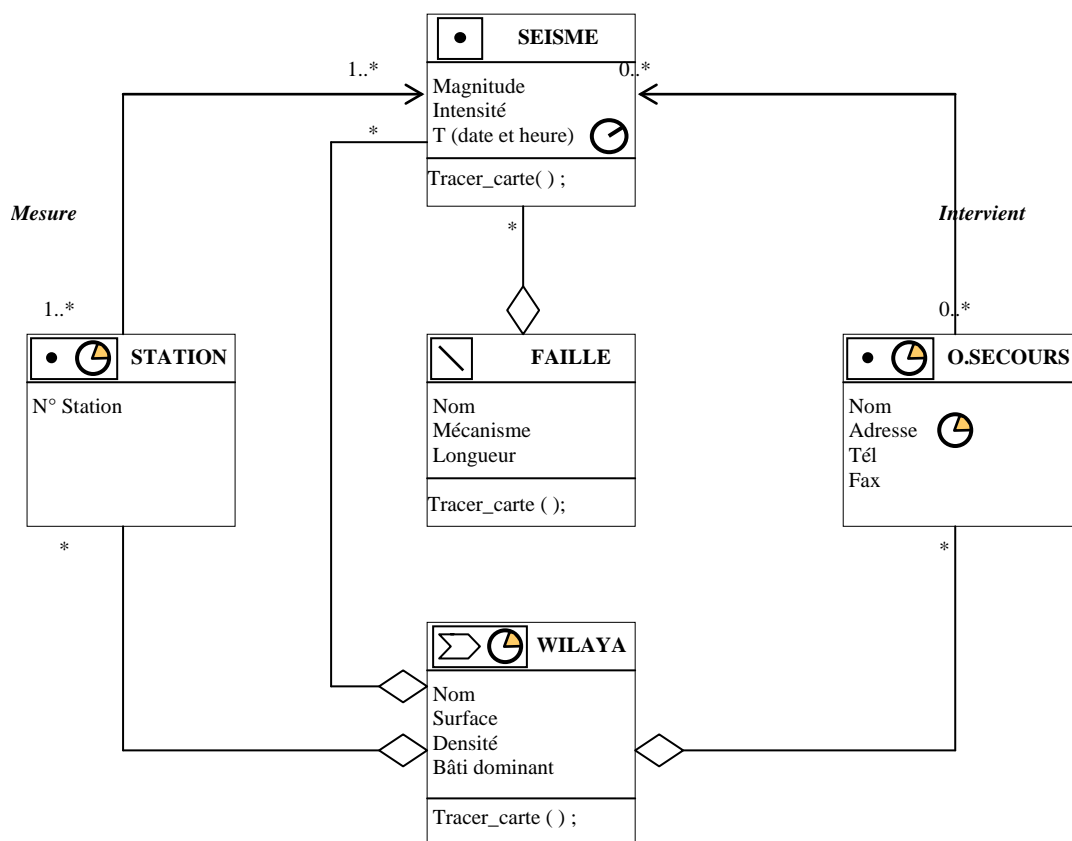


Figure III.16 : Exemple de modélisation selon PERSPECTORY

d- Développement de l'outil AGL

PERSPECTORY est un outil AGL de conception des systèmes d'information géographique. Il est développé comme un gabarit du logiciel Visio en Visual Basic et une base de données Access, inclut un générateur pour Oracle 8i spatial et bientôt pour ArcInfo, MapInfo, Géomédia, ArcView.

PERSPECTORY comporte :

- un éditeur graphique de schémas avec une boîte à outils contenant les différents composants d'un diagramme de classes du formalisme UML (classes, associations, lien d'agrégation, ...) ainsi que des pictogrammes spatiaux et temporels (Figure III.17).
- un dictionnaire de données stocké dans une base de données Access.
- un générateur de code pour structurer la base de données sur Oracle8i, bientôt sur MapInfo, ArcView, ArcInfo,
- un traducteur en langage Interlis.

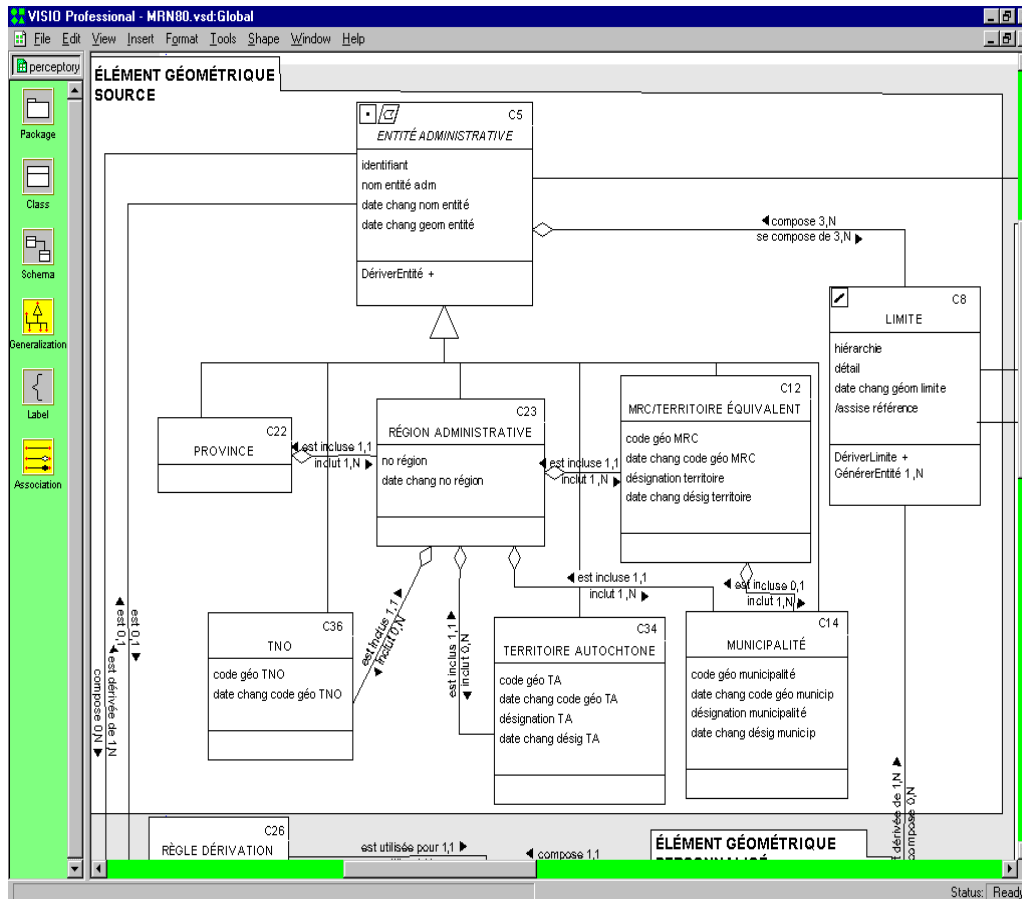


Figure III.17 : PERSPECTORY 2000

Discussion

Contrairement aux méthodes précédentes, PERSPECTORY ne possède pas son propre modèle de conception des SIG. Il se base sur le formalisme UML pour modéliser les données spatio-temporelles. Les trois concepts sont abordés et il est enrichi par le langage visuel spatial et temporel PVL.

Enfin, PERSPECTORY est orienté en tant qu'outil AGL pour la conception des SIG, très simple à utiliser mais conçu seulement pour générer le modèle de conception ainsi que le code de la base de données sous Oracle8i. Les développeurs de cet outil ont prévu le développement :

- d'un générateur XML.
- d'un générateur du squelette de code pour d'autres logiciels SIG : ArcInfo, MapInfo, ArcView, Géomedia,
- d'un utilitaire de gestion de vues, permettant de créer des sous-modèles.
- d'un module de gestion des contraintes d'intégrités spatiales entre les objets.

III.5. Conclusion

Dans ce chapitre, nous avons étudié les différentes méthodes de conception des SIG qui disposent d'un outil AGL. Toutes mettent l'accent sur l'importance de la phase de conceptualisation. Reste que ces méthodes sont limitées dans leurs concepts : représentation graphique des objets à deux dimensions et ne traitent pas le changement de la géométrie dans les objets et de leurs liens topologiques en fonction de la résolution et l'échelle cartographique.

Cette étude nous a permis de classer ces méthodes de conception.

- Des méthodes qui intègrent toutes les représentations topologiques possible (ex : MADS). Elles sont mieux placées pour modéliser les domaines d'application de type topologique.
- D'autres méthodes enrichissent leur modèle par de nouveaux concepts (classes et méthodes) permettant de modéliser les domaines d'application réseau (ex : GEO-OOA).
- Alors que d'autres se basent sur des modèles classiques étendus afin de prendre en compte les données spatio-temporelles (PERSPECTORY : UML+PVL).

Cette grande diversité reflète l'étendue du domaine SIG et rend difficile l'identification d'un ensemble minimal de spécifications pour le développement de ces systèmes. C'est la raison qui nous a mené à définir dans la partie suivante, un environnement de développement des projets SIG basé sur l'intégration des différents modèles des méthodes de conception et la réutilisation de fragments.

PARTIE II

Environnement de développement des
systèmes d'information géographique (SIG)

Dans cette partie, nous allons présenter notre approche. Elle consiste à définir un environnement de développement basé sur l'intégration des fragments de modèles. Afin d'atteindre le but souhaité, il est nécessaire de connaître certains aspects. On parlera ainsi :

- Des étapes de développement d'un projet, en particulier les projets SIG. Pour cela, nous allons introduire dans le premier chapitre de cette partie (Chapitre IV), les différentes démarches développées dans l'ingénierie des systèmes d'information pour mener un projet. Ainsi, nous allons choisir parmi ces démarches, le modèle le mieux placé pour répondre à nos besoins et qui permettra de définir une architecture de notre approche.
- De la méthodologie d'intégration des schémas. Nous allons essayer d'adapter les concepts présentés dans le deuxième chapitre de cette partie (Chapitre V) : traduction des schémas, détection et résolution des conflits, détection des correspondances inter-schémas, Tous ces concepts vont sûrement mettre en évidence l'étape d'intégration des fragments de modèles présentée dans notre démarche de développement d'un projet SIG.

CHAPITRE IV

Chapitre IV

Développement des Systèmes d'Information Géographique (SIG)

IV.1. Introduction

Les systèmes d'information, en particulier les SIG, deviennent de nos jours des systèmes stratégiques dans des entreprises et leur qualité devient également un facteur critique. Le développement des bonnes démarches de conception est pour cette raison, devenu un sujet de recherche à part entière.

Dans ce chapitre, nous allons présenter les différentes démarches conçues pour le développement des systèmes d'information en particulier les SIG.

IV.2. Cycle de développement d'un projet

Un projet est un ensemble de processus, une organisation et une logistique spécifique pour répondre aux besoins d'un groupe d'utilisateur ou d'une clientèle. Ainsi, un projet se doit de suivre une démarche spécifique, généralement présenté par une série d'étapes. Ces étapes constituent le cycle de vie du projet.

Quelles sont les démarches de conduite d'un projet ?

IV.2.1. La démarche en cascade [ROY70]

Le principe de la démarche en cascade est de découper le projet en phases distinctes sur le principe du non-retour. Lorsque une phase est achevée, son résultat sert de point d'entrée à la phase suivante (Figure VI.1). Cette démarche, développée dans les années 1970 par W. ROYCE a servi pendant des années de modèle de référence. Elle est adaptée aux projets de courte durée. Pour des grands systèmes, elle présente l'inconvénient de ne pas permettre de mener, en parallèle, le développement de modules d'applications.

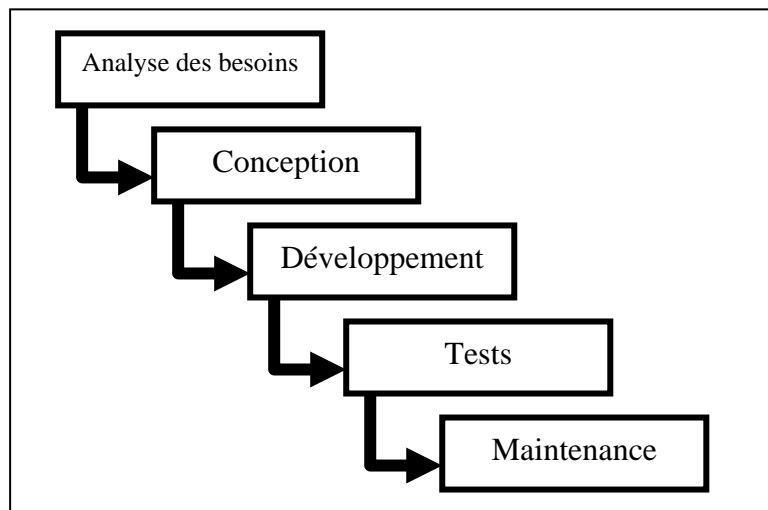


Figure IV.1 : La démarche en cascade

IV.2.2. La démarche en V [SOM89]

La démarche en V est une variante de la démarche en cascade. Elle confronte les différents niveaux de test avec les phases de projet de même niveau (Figure IV.2). Ceci permet à chaque étape de définir non seulement les fonctions, mais également les critères de validation. La cohérence entre les deux éléments permet de vérifier en continu que le projet progresse vers un produit répondant aux besoins initiaux.

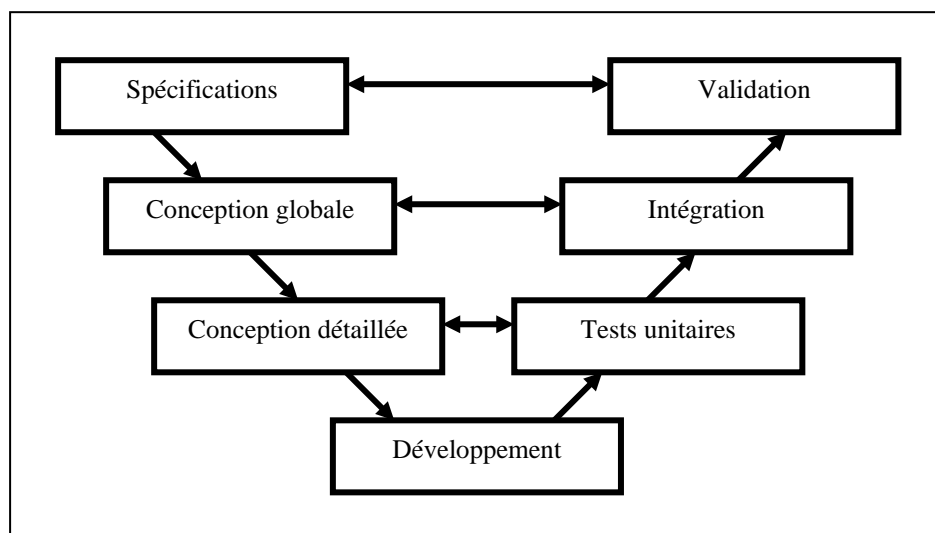


Figure IV.2 : La démarche en V

IV.2.3. La démarche par prototypage [PIG84]

Le prototypage est une technique qui permet d'obtenir assez rapidement une première version implémentable d'un système à partir de spécifications d'un sous-système représentatif du système projeté. En procédant à des essais sur le prototype, le développeur peut, en détectant les erreurs de fonctionnement, les réajuster. Il peut aussi le compléter de différentes manières, en fonction de l'extension du sous-système représentatif qui a été pris pour cible, jusqu'à obtenir le système complet.

Le prototypage a été fréquemment utilisé dans les ateliers de génie logiciel et en base de données. Il a séduit les concepteurs surtout comme moyen de validation dans la mise en œuvre de systèmes complexes.

VI.2.4. La démarche en spirale [BOH88]

Proposée par Barry W BOHEM en 1984, la démarche en spirale est une méthode de conduite d'un projet, composée de boucles dont chacune d'elles serait un sous projet présenté sous la forme d'une démarche en cascade (Figure IV.4). Chaque boucle de spirale permet :

- D'identifier les objectifs propres de la boucle.
- Les différentes alternatives pour atteindre les objectifs.
- Les contraintes de chaque alternative.

La démarche en spirale est adaptée aux grands projets complexes.

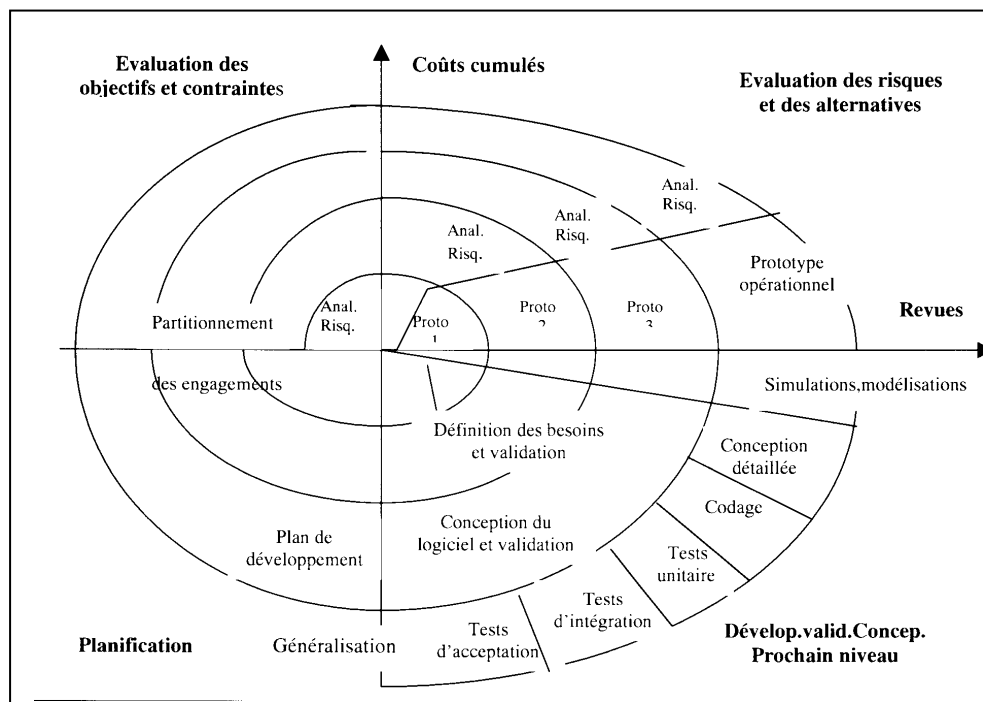


Figure IV.4 : La démarche en spirale

IV.2.5. La démarche tridimensionnelle [TAR83]

Induite par la méthode Merise, cette démarche positionne le développement d'un système par rapport à trois axes (Figure IV.5) : un axe qui décrit le cycle de vie du système, un axe qui décrit les niveaux d'abstractions et le troisième décrit le cycle de développement correspondant au cycle de vie du projet (similaire au cycle en cascade).

Ce sont les deux premiers axes qui constituent l'originalité de cette démarche. Le premier permet de planifier l'évolution du système et d'organiser son changement. Le second introduit une indépendance entre la solution conceptuelle et la solution technique.

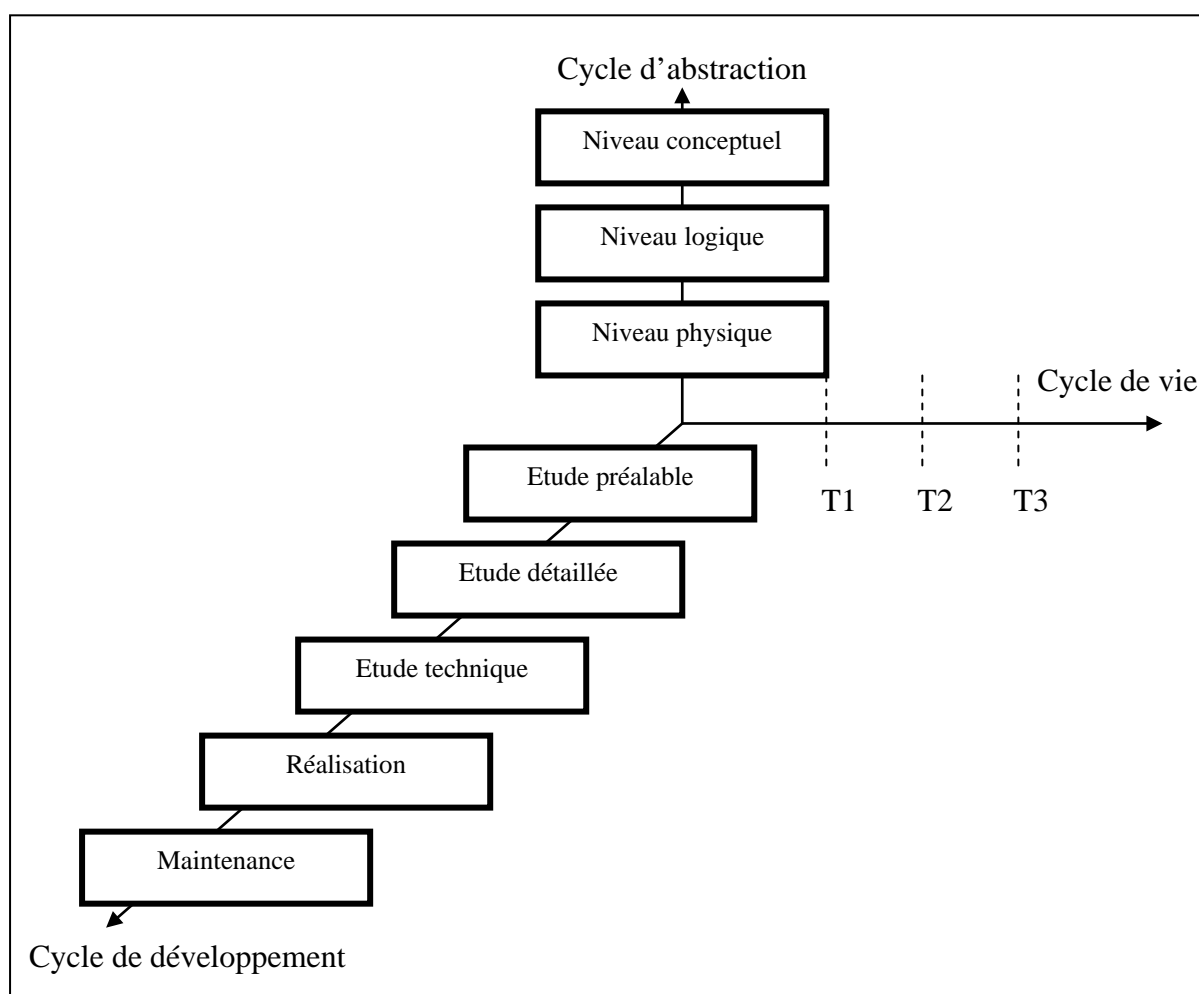


Figure IV.5 : La démarche tridimensionnelle

VI.2.6. La démarche en fontaine [HEN90]

La démarche en fontaine a été définie pour les environnements de développement orienté objet. Elle consiste en un processus incrémental qui permet de représenter non seulement le chevauchement des étapes de développement mais aussi les itérations entre les différentes étapes (Figure IV.6).

Le processus débute par une phase d'analyse des besoins et se poursuit par l'exécution des mêmes phases que celles définies dans la démarche « en cascade ». Ce processus est répété jusqu'à ce que toutes les parties du système en cours de développement aient été développées.

En fait dans cette démarche, la technique d'itération et celle du parallélisme sont combinées en même temps.

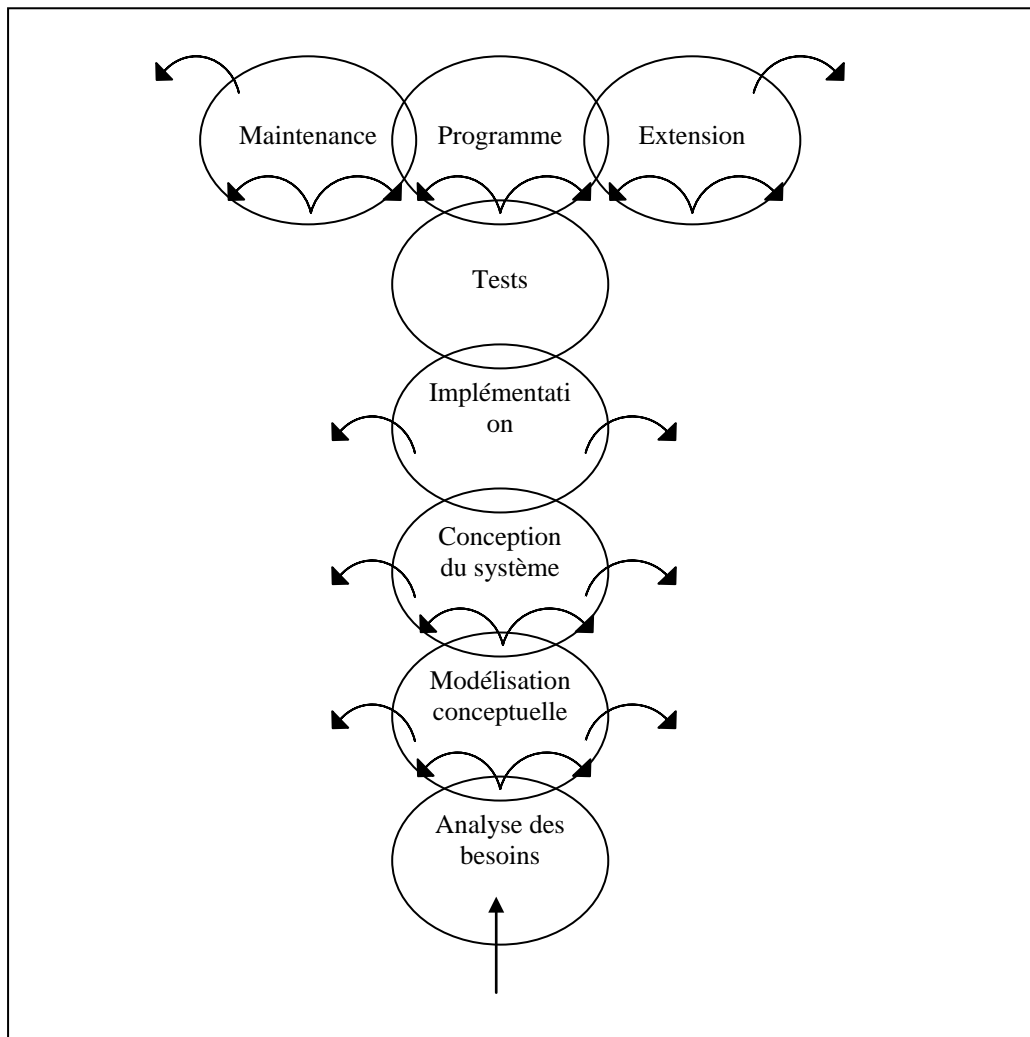


Figure IV.6 : La démarche en fontaine

IV.3. Les conditions d'un SIG efficace [THE01]

Le développement d'un projet SIG est une activité complexe qui fait intervenir de nombreuses ressources humaines, matérielles et technologiques et par conséquent, doit être menée de façon rigoureuse.

Afin d'aboutir à un système efficace, il est nécessaire de prendre en compte certaines mesures lors de la planification et la réalisation du projet. Pour cela, le concepteur doit :

- Élaborer un SIG rentable ou nécessaire pour accomplir l'objectif de l'organisme.
- Concevoir un système pertinent qui satisfait les besoins de production.
- Bâtir un système flexible capable d'évoluer au gré des changements de contexte et des besoins.
- Organiser le développement de manière sécuritaire pour éviter les choix inadéquats et les erreurs coûteuses.
- Établir des procédures rigoureuses pour contrôler la qualité des données et mesurer la performance des traitements réalisés par le SIG.
- Suivre une démarche de conception indépendante des logiciels et du matériel pour assurer un maximum d'indépendance technologique.
- Dans un projet complexe, il faut prévoir une approche multidisciplinaire impliquant plusieurs spécialistes, ce qui repose sur l'élaboration d'un vocabulaire et d'un schéma conceptuel communs.
- Définir les rôles et les responsabilités des individus et des services impliqués dans le développement pour établir une attitude de coopération.
- Impliquer les utilisateurs à toutes les étapes de la conception afin d'intégrer l'ensemble des besoins et maintenir la motivation.
- Éviter les achats massifs de matériel et de logiciel au début du processus de planification et limiter la saisie de données au strict minimum avant la phase d'implantation, afin de produire un retour sur l'investissement dans des délais acceptables.
- Rédiger la documentation durant le développement du système afin d'assurer un maximum d'indépendance face aux individus qui l'élaborent.

Pour ce qui est de la modélisation, le concepteur doit aussi :

- Élaborer une architecture globale.
- Modéliser les phénomènes, les processus, les données et les traitements afin de répartir les difficultés sur plusieurs niveaux d'intervention.
- Effectuer les choix stratégiques à trois niveaux successifs: conceptuel, logique, physique.
- Prévoir des mécanismes de contrôle de qualité, de fiabilité et d'efficacité.
- Impliquer le personnel régulier dans l'élaboration du SIG.
- Identifier les chevauchements de juridictions et réduire les duplications de rôles.

IV.5. Les sept phases de développement d'un projet SIG [THE01]

Thériault propose dans sa démarche, un découpage du projet SIG en sept phases allant du général au particulier (approche systémique) (Figure IV.7) :

IV.5.1. Étude d'opportunité

Objectif :

- a- Définir la problématique :
 - Identifier les **requérants** et les **besoins**.
 - Produire une **description sommaire** du projet.
 - Situer le projet par rapport au **mandat de l'organisme**.

- b- Dégager les orientations stratégiques :
 - Délimiter le **champ d'activité du système** envisagé (questions traitées et objectifs généraux).
 - Formuler un **projet** de SIG.
 - Rechercher des **systèmes** existants qui remplissent un rôle analogue.
 - Justifier l'**approche géomatique** pour les applications visées

- c- Évaluer la pertinence du projet :
 - Évaluer sommairement la **faisabilité** du projet (au plan technique).
 - Établir l'ordre de grandeur des **coûts**.
 - Apprécier le potentiel de **rentabilité**.
 - Recommander, s'il y a lieu, d'**entreprendre une étude de faisabilité**.

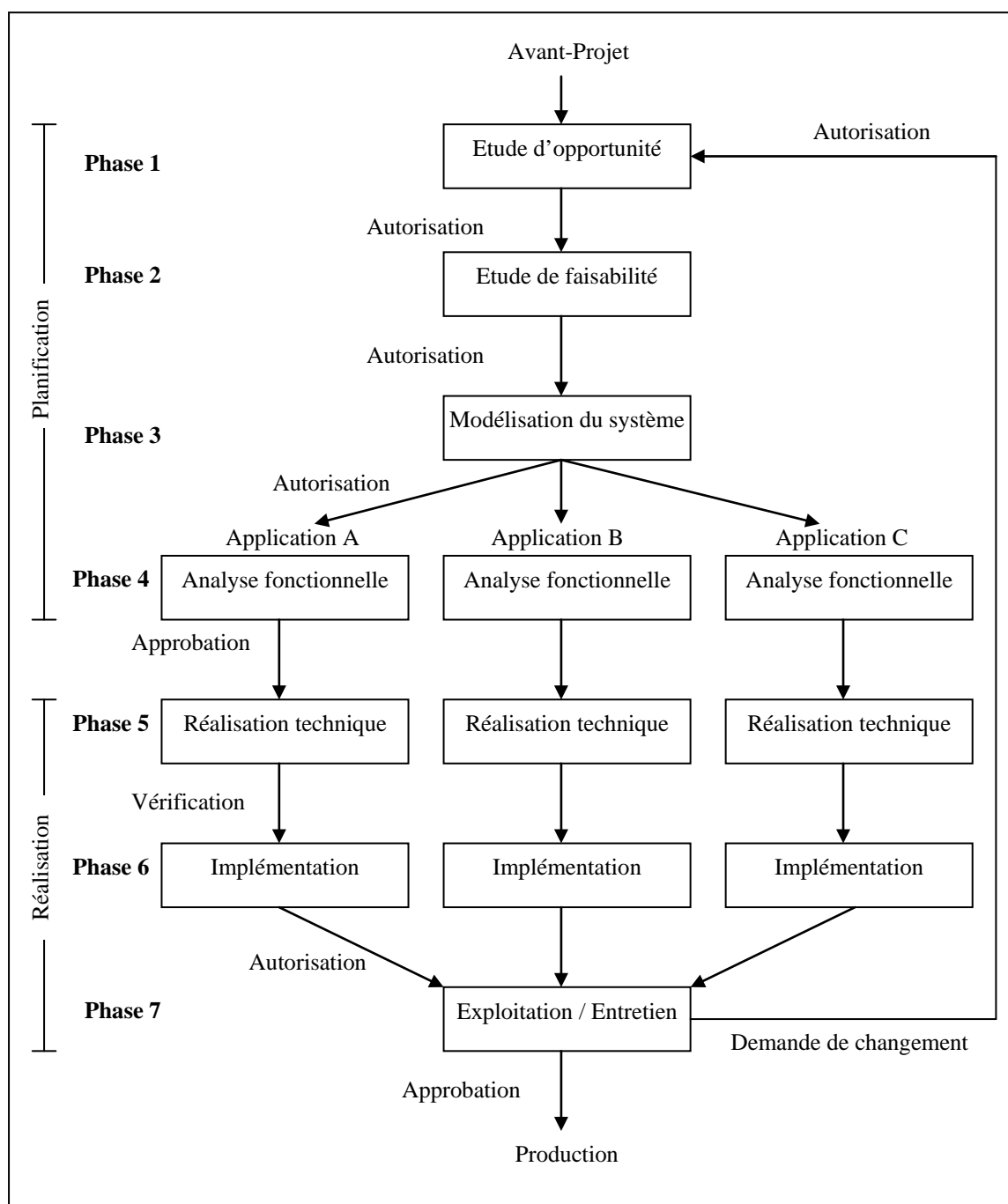


Figure IV.7 : Les sept phases de développement d'un projet SIG

IV.5.2. Étude de faisabilité

Objectif :

- a- Faire un **inventaire détaillé des besoins** :
 - Axer l'étude sur les **besoins en information**, sans se restreindre au fonctionnement actuel de l'organisme (flux d'information, nature et séquence des opérations administratives, etc.).
 - Énumérer les **problèmes** rencontrés et les **besoins** exprimés.

- Évaluer la **pertinence des SIG** pour répondre à ces besoins.
- Proposer des **objectifs prioritaires**.

- b- Élaborer des **scénarios d'implantation** :
 - Préciser les **objectifs spécifiques** du système.
 - Identifier le potentiel de **diffusion** interne et externe des **données**.
 - Réaliser l'**inventaire des entités, des relations et des processus** représentés dans le SIG.
 - Élaborer les **modèles conceptuels** des données et des traitements.
 - Identifier les **applications et les sous-systèmes requis**.
 - Spécifier les **modalités de fonctionnement** du SIG (interface avec l'utilisateur, respect des normes administratives, etc.).
 - Vérifier la **faisabilité technique**.
 - Étudier la **qualité des données** et évaluer le **potentiel d'intégration**.
- c- Faire l'**évaluation économique** du projet.
- d- Recommander, s'il y a lieu, d'entreprendre la réalisation du SIG.

IV.5.3. Modélisation du système

Objectif :

- a- Définir l'**architecture générale** du système :
 - Terminer les **modèles conceptuels** (définir les propriétés et les attributs).
 - Réaliser les **modèles logiques** des données (identifier les fiches et les liens relationnels) et des traitements (identifier les fonctions du SIG et établir les séquences de réalisation).
 - Prévoir les mécanismes de **conversion** nécessaires pour **intégrer les données existantes**.
 - Spécifier les **critères de rendement** du système.
 - Identifier les **composantes requises** (inventaire, comparaison et sélection des logiciels adéquats; matériel informatique nécessaire).
 - Planifier la réalisation du SIG avec une **séquence d'applications**.
 - Ébaucher les **modèles physiques** (échanciers de production, contenu des guides d'opération, essais fonctionnels, etc.).
- b- Réaliser des **prototypes**.
- c- Établir des **priorités de développement**.

IV.5.4. Analyse fonctionnelle

Objectif :

- a- Planifier la **réalisation de chaque application** :
 - Élaborer une **stratégie de traitement** adéquate et efficace.
 - Valider les **processus d'acquisition** des données.
 - **Synchroniser** le développement du système avec la saisie des données.
 - Vérifier la **compatibilité des modèles** de traitement avec les lois, les règlements et les normes en vigueur.
 - **Prévoir les modifications** nécessaires pour adapter le système aux changements de réglementation.

- b- Déterminer les **composantes requises** pour chaque application :
 - **Comparer les logiciels** capables d'accomplir les tâches prévues.
 - Dresser un **inventaire comparatif des appareils** requis pour exploiter les logiciels retenus.
 - Préparer les données nécessaires pour effectuer les **tests de fiabilité et de performance**.
 - Identifier les **ressources humaines** nécessaires pour effectuer les tests et obtenir les autorisations requises.
 - Vérifier l'existence d'un **service local** (représentant, vendeur, consultant, etc.) pour accompagner la solution informatique retenue.
 - Déterminer les besoins en formation pour **amorcer le recyclage du personnel**

IV.5.5. Réalisation technique

Objectif :

- a- Réaliser l'**évaluation technique** du système et des composantes :
 - Faire une **évaluation qualitative**: disponibilité des fonctions requises et aptitude du logiciel pour effectuer les tâches prévues.
 - Faire une **évaluation quantitative**: mesure de performance de chaque combinaison logiciel-matériel pour accomplir un ensemble de tâches typiques.
 - Identifier les **facteurs de risques** pour chaque option technologique.

- b- Sélectionner les **logiciels** et le **matériel** :
 - Prévoir l'**envergure maximale** du SIG.
 - Étudier la **compatibilité de chaque élément** avec les autres composantes du système.
 - Analyser les **mécanismes d'échange** de données avec les partenaires (formats, réseaux, contrôle des accès, partage des frais, etc.).
 - Prévoir les fréquences d'**entretien** et de **remplacement** du logiciel et du matériel.
 - Entreprendre la **formation spécifique du personnel**.
- c- Produire les **modèles physiques** des données et des traitements.
- d- **Préparer le matériel** requis pour l'implantation.

IV.5.6. *Implantation*

Objectif :

- a- Constituer la **base de données** :
 - Choisir les **modes de saisie** des données.
 - Valider la **qualité des données**: précision, exactitude, exhaustivité, pertinence, consistance logique, etc.
 - Produire et diffuser les **métadonnées**.
- b- Réaliser la **programmation** des fonctions de traitement :
 - Programmer et vérifier les **fonctions primaires** requises par l'application.
 - Intégrer les **fonctions élémentaires** pour produire les fonctions de système.
 - Rédiger la **documentation** de système et les modes d'emploi.
- c- Vérifier la **performance** et la **fiabilité** du système :
 - Vérifier le **comportement** et la **performance** des fonctions primitives, intégrées et de système.
 - Effectuer des **tests de fiabilité** du système pour accomplir les tâches planifiées.
- d- Réaliser des **projets pilotes**.
- e- Compléter la **formation des utilisateurs**.
- f- Effectuer les corrections nécessaires pour **optimiser le système**.

IV.5.7. Exploitation et entretien du SIG

Objectif :

- a- Utiliser le SIG pour accomplir les opérations de gestion et de planification visées par l'ensemble des applications complétées.
- b- Perfectionner les mécanismes de sécurité.
- c- Poursuivre la saisie et la validation des données.
- d- Effectuer les contrôles de qualité.
- e- Suggérer des modifications ou des ajouts dans le système.

Discussion

La démarche proposée, met en évidence et d'une manière générale les étapes à suivre pour développer un projet SIG.

Les remarques que l'on peut tirer de cette démarche sont résumées en deux grands points :

- la démarche repose sur une approche systémique qui décompose tout projet SIG en sous-systèmes et sépare les données des traitements. Elle aura ainsi, les mêmes retombés (avantages et inconvénients) que les méthodes systémiques des SI.
- Elle assure toutes les phases du cycle de vie d'un projet allant de l'étude des besoins jusqu'à l'exploitation et l'entretien du SIG, sauf que l'on ne retrouve pas les processus et les modèles pour mener à bien chaque étape de développement du projet. Par exemple dans la modélisation conceptuelle des données et des traitements, assuré par l'étape 3 de la démarche, les modèles ne sont pas définis. Le choix est laissé au concepteur du système.

Cette démarche est considérée comme une référence pour le développement des projets dans le domaine des SIG.

IV.6. Conclusion

L'étude réalisée dans ce chapitre, résume les différentes démarches développées pour mener un projet SI de sa planification à sa réalisation et l'entretien. Sans oublier que les mêmes démarches ont été enrichies pour prendre en considération les projets SIG. C'est le cas pour la démarche proposée par [THE01].

Cette étude nous a permis de définir une architecture globale du développement d'un projet SIG prévue pour notre approche.

CHAPITRE V

Chapitre V

Intégration de schémas : Concepts et méthodologie

V.1. Introduction

L'importance de pouvoir intégrer différentes sources d'information (et notamment des bases de données) au sein d'une organisation n'est plus à démontrer. En effet, les besoins en informations de plus en plus nombreux, les applications de plus en plus complexes nécessitent un partage de l'information, devenue une ressource stratégique dont l'utilisation doit être optimisée.

L'hétérogénéité des systèmes locaux, en termes de méthodes, de modèles, de systèmes et d'organisation rend très difficile toute tentative de partage de données et de services. De nombreux travaux de recherche ont ainsi été entrepris afin de faciliter la mise en commun de ressources informatiques. Ainsi parle-t-on d'*interopérabilité* et d'*intégration* dans le domaine des bases de données. Les travaux concernant l'intégration (aussi bien de schémas que de bases de données) ont donné lieu à une multitude d'approches, dont celle de Batini[BAT86], et plus récemment celle de Martin[COM98].

Dans ce chapitre, nous allons présenter les différents concepts liés aux approches d'intégration des bases de données en vue de les mettre en évidence dans le cas des SIG.

V.2. Concepts d'intégration de schémas

Avant de présenter les techniques d'intégration des schémas, nous allons tout d'abord, cerner les concepts liés à ce processus afin de comprendre la problématique posée.

V.2.1. Diversité de schémas

Dans la plupart des cas, les bases de données à intégrer sont développées indépendamment. Elles seront par conséquent, hétérogènes. Cela induit un nombre non négligeable de conflits.

Batini [BAT86] et Spaccapietra [SPA92] ont distingué plusieurs classes de causes de diversité des schémas :

- Perspectives différentes ;
- Descriptions différentes des propriétés d'objets ;
- Hétérogénéité des modèles de données ;
- Equivalence à travers les constructions du modèle ;
- Incompatibilité des spécifications conceptuelles.

- **Perspectives différentes**

Durant le processus de conception d'une Base de données, différents groupes de concepteurs adoptent leurs propres points de vues pour la modélisation d'un même objet. Ces perceptions peuvent être complètement différentes ou seulement intersectées.

- **Descriptions différentes des propriétés d'objets**

Lorsque deux concepteurs conçoivent la même réalité, ils peuvent lui assigner des modélisations différentes. Leur rassemblement au sein d'une même base de données conduit à une situation de type : redondance des informations, incohérence (le même attribut n'a pas la même valeur dans chaque modèle) ou niveaux de précision différents pour un même objet.

Prenons par exemple les points de vue de deux concepteurs différents pour la modélisation d'un réseau routier : un cartographe et un ingénieur des ponts et chaussées. Le premier aura pour principal souci d'obtenir une cartographie cohérente du réseau routier. Il sera attentif aux positions relatives des différentes routes, à leur parcours, ainsi qu'aux attributs géographiques permettant d'identifier et de différencier les catégories. L'ingénieur des ponts et chaussées sera intéressé par le mode de construction de la route, le matériau utilisé, le mode de drainage, les possibilités d'élargissement, ... etc.

- **hétérogénéité des modèles de données**

Les concepteurs utilisent des modèles de données différents dans leur modélisation (relationnel, CODASYL, orienté objet, ...). Ce qui conduit à des conflits appelés conflits d'hétérogénéité et sont en général traités tout à fait au début du processus d'intégration par une traduction des différents modèles en un modèle canonique.

- **Equivalence à travers les constructions du modèle**

Les schémas exprimés par le même modèle de données, et qui représentent d'une façon équivalente le même domaine d'une application conduisent à des conflits connus sous le nom de conflits structurels. Par conséquent, le schéma le plus riche, permet une grande variété de possibilités pour représenter la même situation, mais le processus d'intégration sera plus complexe, alors que les schémas pauvres peuvent servir à réduire les conflits structurels. Par exemple, dans la Figure V.1, « Autoroute » et « Route communale » sont distinguées par une spécialisation dans le premier schéma, et par une différence de la valeur de l'attribut Type dans le deuxième.

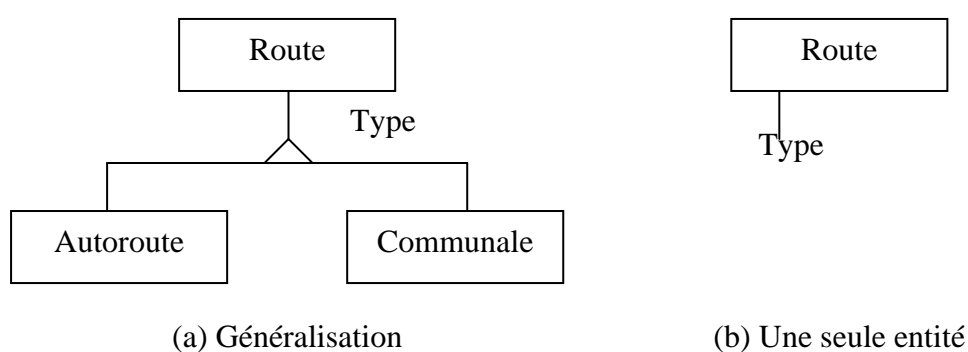


Figure V.1 : Exemple d'équivalence de schémas

- **Spécifications conceptuelles incompatibles**

Le choix des concepts non compatibles nécessite des informations complémentaires qui peuvent être fournies par une analyse sémantique des différents schémas. Dans la Figure V.2, un employé est concerné par un seul projet avec une cardinalité (1 :n). Alors qu'un employé peut participer à plusieurs projets.

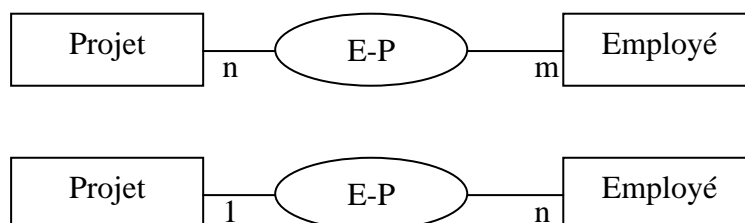


Figure V.2 : Incompatibilité des spécifications

V.2.2. Propriétés inter-schémas

Des relations entre les concepts appartenant aux différents schémas peuvent exister d'une façon implicite. Ces relations sont liées aux propriétés sémantiques de chaque schéma et sont appelées propriétés inter-schémas.

Dans la figure suivante (Figure V.3), l'entité « Wilaya » appartient au schéma 1, et l'entité « commune » au schéma 2. Il est à déduire qu'entre les deux schémas, il existe une relation d'agrégation, telle que « wilaya » est un groupe de « commune ».

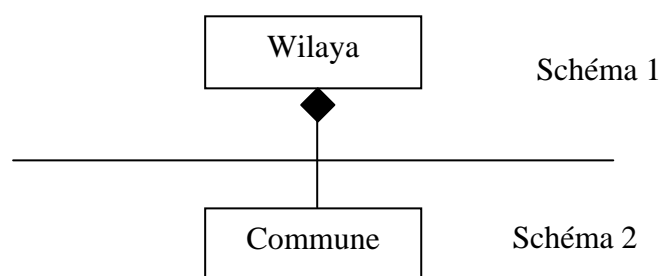


Figure V.3 : Exemple de relation inter-schémas

Dans un modèle orienté objet, les relations inter-schémas peuvent être distinguées d'une façon plus claire, car l'abstraction du modèle préconise des relations de types généralisation/spécialisation et l'agrégation entre les classes d'objets.

Les propriétés inter-schémas, une fois localisées, sont exploitées pour la restructuration et la résolution des conflits en utilisant des opérateurs de réorganisation.

V.3. Méthodologie d'intégration de schémas

Une méthode d'intégration est définie dans [COM98] comme suit : « *une méthode d'intégration est un ensemble d'opérations dotant le concepteur d'une méthode d'aide à l'intégration par ordinateur (computer aided integration method). Ce concepteur est appelé à choisir la meilleure opération pour une paire de concepts données dans le but de réaliser une intégration correcte* ».

Le cœur du problème est le processus d'intégration des bases de données. Il consiste à prendre en entrée un ensemble de bases de données (schémas), et à produire en sortie une description unifiée des schémas initiaux et les règles de traduction (mapping) qui vont permettre l'accès aux données existantes à partir du schéma intégré.

De nombreuses études se sont intéressées à l'intégration des schémas. Il en résulte deux types de systèmes d'intégration appelés aussi « systèmes fédérés » [VER01]:

- **Fortement couplé** : il comporte un schéma global unifié constituant le schéma de la base de données. Il est basé sur un modèle spécifique appelé modèle de données commun ou modèle canonique. Cependant, une étape de traduction des schémas à intégrer devient un pré-requis indispensable et est traitée comme un problème à part.
- **Faiblement couplé** : contrairement au précédent, le système de fédération faiblement couplé n'a pas recours à un schéma unique global, mais fournit un langage de requêtes multi-bases de données qui est une abstraction des langages de requêtes des différentes bases de données et qui permet de cacher les problèmes d'hétérogénéité. Seul le langage est basé sur le modèle de données commun pour accéder aux différents sources de données.

V.3.2. Les étapes d'intégration

La plupart des approches [BAT86] et [COM98], décomposent le processus d'intégration de schémas en plusieurs phases, dont les principales sont (Figure V.5) :

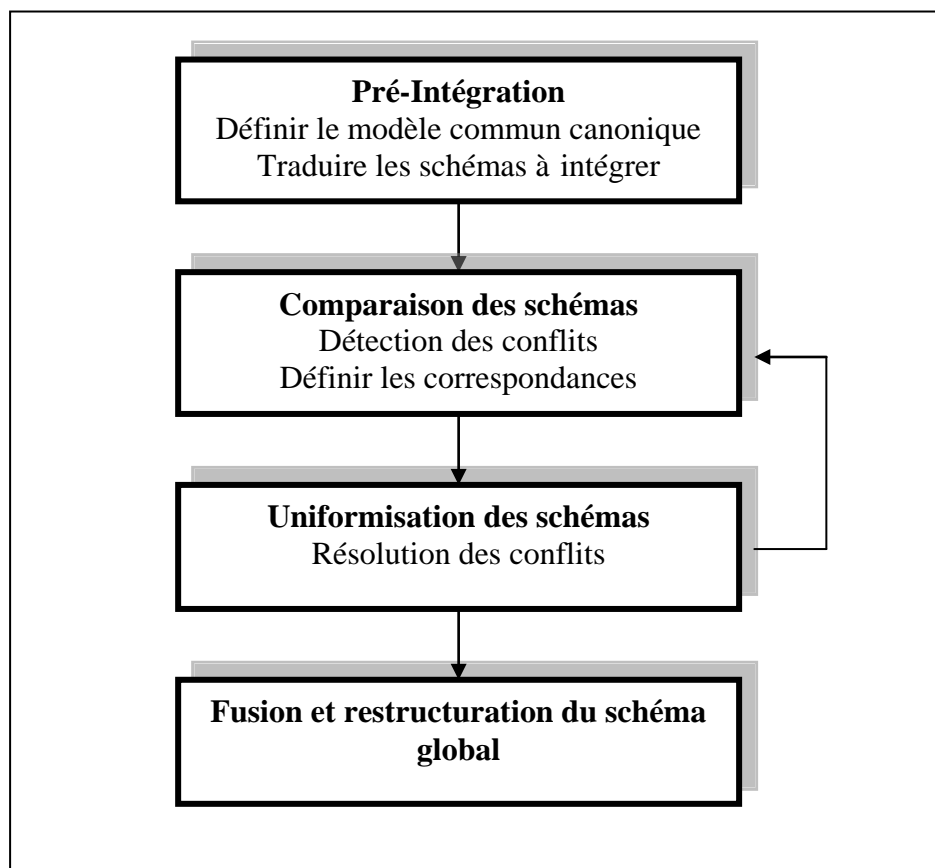


Figure V.4 : Les étapes d'intégration

- **Pré-Intégration** : le but de cette étape est la traduction des différents schémas à intégrer (schémas composants) dans un modèle de données commun appelé modèle canonique, pour extraire le maximum d'informations concernant les éléments (attributs, classes, méthodes, ...), les relations ainsi que les règles entre ces derniers. La pré-intégration joue un rôle d'homogénéisation des schémas vis à vis des différents modèles de données utilisés pour leur expression. Le choix du modèle commun utilisé pour la pré-intégration est important. Celui-ci doit permettre d'exprimer d'une manière simple les structures pouvant être définies par les modèles utilisés pour les schémas.
- **Comparaison des schémas** : les schémas une fois traduits, sont ensuite analysés et comparés pour déterminer les correspondances entre les concepts, et détecter les cas de conflits. Les propriétés inter-schémas peuvent être détectées à ce niveau.
- **Uniformisation des schémas** : une fois les conflits détectés, un effort de résolution est appliqué. Cette étape ainsi que la précédente travaillent de façon cyclique, jusqu'à élimination de tous les conflits.
- **Fusion et restructuration du schéma global** : il s'agit de la dernière étape. Elle permet la vérification et la validation du schéma global obtenu après la fusion des schémas composants. Une analyse du schéma final permet de déduire de nouvelles propriétés (relation redondantes ou nouvelles, généralisation, ...). Pour cela, une restructuration du schéma global sera indispensable pour atteindre les objectifs.

V.3.1. Caractéristique du schéma global

Le schéma global obtenu par intégration doit vérifier certaines contraintes [BAT86][TIA93], parmi lesquelles, on retrouve :

- **Complétude et correction** : le schéma intégré doit contenir tous les concepts présents dans n'importe quel schéma composant. En d'autre terme, le schéma intégré doit être une représentation fidèle de l'union de tous les domaines de l'application associés aux schémas composants.
- **Minimal** : un même concept représenté dans plusieurs schémas, doit apparaître une seule fois dans le schéma global. Cela élimine tout type de redondance.

- **Compréhensible** : le schéma doit être facile à comprendre par les concepteurs et les utilisateurs. Ceci implique, que si le modèle permet plusieurs représentations, la plus riche sera utilisée (utilisation d'une hiérarchie de généralisation par exemple).

V.3.3. Les conflits lors de l'intégration [PAR96]

Les conflits de données sont issus des différences de modélisation et de l'interprétation des entités par les concepteurs des systèmes d'information.

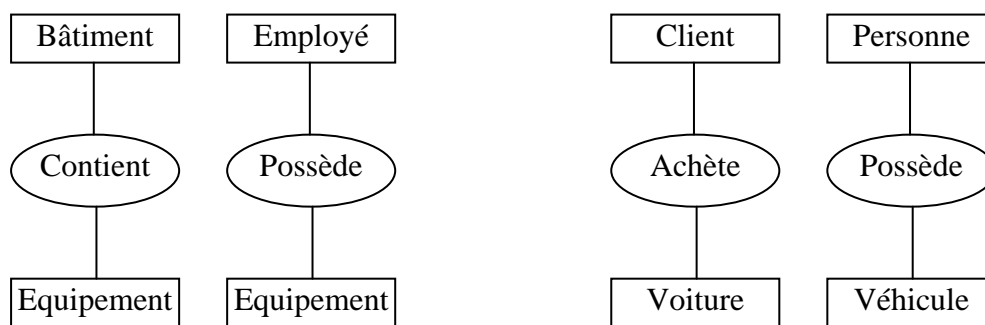
Plusieurs travaux ont proposé des taxonomies de conflits de données [PAR96]. Elles sont regroupés en deux grandes catégories : conflits de nomination et conflits structurels.

VI.3.3.1. Conflits de nomination

Les objets représentés dans les différents schémas, sont identifiés par des noms distincts. Les concepteurs de chaque schéma adoptent leurs propres terminologie et noms. Cela conduit à une prolifération d'une inconsistance des noms à travers les composants du schéma global.

Les conflits dus à la nomination sont de deux types :

- Homonyme** : ce type de conflit existe lorsque le même nom est utilisé pour deux ou plusieurs concepts différents. Par exemple, l'entité « équipement » peut être définie comme des équipements de bureau d'un bâtiment dans un schéma, et comme des pièces de fournitures d'un employé dans un autre schéma (Figure V.5-a).



(a) Exemple d'homonyme

(b) Exemple de synonyme

Figure V.5 : Les conflits de nomination

- Synonyme** : le même concept est décrit par deux noms différents ou plus. Par exemple, l'entité « véhicule » représente le même objet « voiture » dans deux schémas différents (Figure V.5-b).

La solution traditionnelle proposée pour éliminer les conflits de nomination, consiste à changer le nom de l'un des concepts dans le cas homonyme et de renommer l'un des concepts en utilisant le nom de l'autre dans le cas synonyme.

VI.3.3.2. Conflits structurels

Un conflit structurel survient lorsque les entités en correspondance sont décrites par des concepts de niveaux de représentation différents ou soumis à des contraintes différentes [PAR96].

Ces conflits peuvent être classés en plusieurs catégories :

- a- **Conflit de type de données** : résultant de la situation où un même objet est modélisé par des concepts différents dans les schémas composants. C'est le cas, par exemple, lorsqu'un objet est représenté comme entité dans un schéma et comme attribut dans un autre. Les solutions proposées [PAR96] pour résoudre ce type de conflits ne sont pas généralisées pour tout type de concepts.
- b- **Conflit de dépendance** : dû au fait que plusieurs entités sont reliées entre elles par des dépendances différentes dans les schémas composants. Par exemple, la relation Projet-Employé peut être 1 :n dans un schéma et n :m dans un autre.
- c- **Conflit de comportement** : ceci est rencontré lorsque différentes politiques d'insertions/éliminations sont associées à la même classe d'objet dans les différents schémas. Par exemple, dans un schéma, un département peut être autorisé à exister sans employés. Tandis que dans l'autre, l'élimination du dernier employé conduira à l'élimination de département.
- d- **Conflit de clé** : ce cas de conflits est perçu lorsque des clés différentes sont assignées au même concept dans les différents schémas. Par exemple, la clé d'un participant à un projet est « employé.mat » dans un schéma, et « Département.Nomemp » dans l'autre.
- e- **Conflit d'unité** : dû au fait que les attributs sémantiquement égaux, utilisent des unités différentes dans des schémas composants différents. Par exemple, une profondeur utilise l'unité mètre dans un schéma et le mile dans un autre.
- f- **Conflit de précision** : les attributs des différents schémas qui sont équivalents, demandent plus de précision pour les distinguer. Cela conduit à un conflit de précision. Par exemple, l'attribut « phone » signifie le numéro de poste d'un employé

dans un schéma et dans un autre, il est utilisé pour définir son numéro de téléphone personnel.

- g- **Conflit de type d'incompatibilité au niveau d'abstraction** : lorsque deux objets sémantiquement similaires sont décrits dans des niveaux d'abstraction différents. Par exemple, dans un schéma traitant la participation d'agents dans des projets, on peut trouver l'objet membre pour décrire les participants. Par contre dans un autre schéma, on trouvera le concept « Etudiant » et « Enseignant » séparément pour désigner les participants.

Un autre cas de conflit peut exister dans les modèles OO. Une même information est définie comme propriété statique (attribut) dans un schéma et dynamique (méthode) dans un autre. Par exemple, l'information « age » d'une personne est désignée comme un attribut « age » dans un schéma, et une méthode de calcul basée sur la date de naissance dans un autre.

Les solutions proposées [PAR96] pour résoudre les différents types de conflits diffèrent d'une méthode à une autre et dépendent des modèles de données utilisées dans l'intégration.

V.3.3.3. Cas d'un SIG [BEN99]

- a- **Conflit géométrique** : la géométrie d'un objet est représenté par une classe dans un schéma et par un attribut dans un autre. Par exemple, le modèle de MADS modélise la géométrie d'un objet comme étant un attribut alors que le modèle POLLEN représente explicitement la géométrie par un ensemble de classes géométriques (Point, Ligne, Polygone) (Figure V.6). Dans ce cas, les sources d'informations sont en conflit pour représenter les relations topologiques.

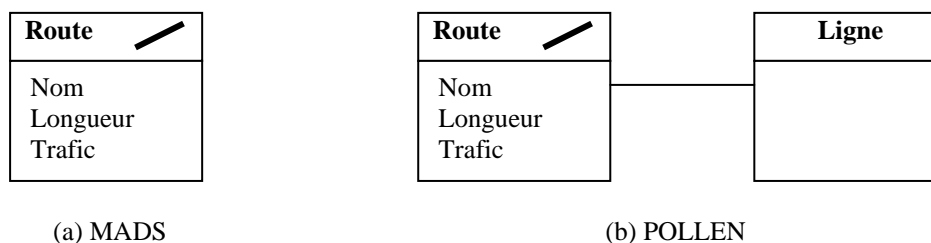


Figure V.6 : Exemple de conflit géométrique

- b- **Conflit structurel** : dans ce type de conflit, le même objet est perçu avec deux structures différentes. Les objets peuvent être regroupés en tant que relation ou en tant que objet complexe. Ce cas de conflit figure suite au changement de résolution

- spatiale. Par exemple, dans un schéma, une zone bâtie est représentée par un ensemble d'objets « bâtiments » et dans un autre comme étant un objet complexe « zone bâtie ».
- c- **Conflit de généralisation** : il est produit par des différences dans les liens de généralisation/spécialisation entre objets. Par exemple, un schéma ne peut utiliser qu'un seul concept pour représenter tous les types de routes. Tandis qu'un autre emploiera deux concepts pour distinguer les objets autoroute et route départementale liés à un concept plus général route.
 - d- **Conflit sémantique** : les conflits sémantiques sont issus de différences dans la localisation et le découpage de l'information géographique ainsi que dans les caractéristiques associées à chaque objet. Par exemple, des conflits de localisation peuvent provenir du choix du système de projection, point de référence, unité de mesure,
 - e- **Conflit d'agrégation spatiale** : apparaît lorsque la représentation d'un objet géographique varie selon l'échelle choisie et selon le niveau de détail souhaité. Par exemple, une route est représentée sous forme surfacique dans une représentation à grande échelle, et sous forme linéaire dans une représentation utilisant une échelle plus réduite.
 - f- **Conflit topologique** : dû à la différence dans la façon de modéliser la relation topologique. Celle-ci est représentée dans un schéma, par des associations topologiques et dans un autre, par des contraintes d'intégrités spatiales. Par exemple, la relation topologique « adjacence » est modélisée par une association « procheDe » dans un schéma (c'est le cas de MADS). Alors que dans un autre (cas de CONGOO), elle est définie par une contrainte d'intégrité « se trouve à une distance inférieure à 300m ».
 - g- **Conflit temporel** : la temporalité d'un objet est représentée par une entité dans un schéma. Tandis que dans un autre, elle est représentée par un attribut temporel.

V.3.4. Correspondances inter-schémas

Pour construire le schéma global, la relation entre les concepts doit être connue. Des efforts ont été faits pour définir des outils pour l'identification automatique des correspondances inter-schémas [PAR96] en recherchant les caractéristiques identiques ou similaires [LAR02] : noms, identifiants, composants, propriétés, attributs, méthodes,

a- Correspondances entre classes :

- **Objets équivalents** : deux objets sont équivalents si et seulement si tous leurs états sont équivalents (attributs et comportements). Par exemple, dans la modélisation d'un réseau routier, les objets « carrefour » et « rond-point » sont des objets équivalents.
- **Classes équivalentes** : soit P l'ensemble de toute instance d'une classe C . deux classes $C1$ et $C2$ sont équivalentes ssi :

$$\forall o1 \in P1 / \exists o2 \in P2 : o1 \equiv o2 \text{ et } \forall o2 \in P2 / \exists o1 \in P1 : o1 \equiv o2$$

- **Classes compatibles** : deux classes $C1$ et $C2$ sont compatibles si on peut trouver une super-classe avec une sémantique significative.

On peut distinguer trois types :

- **class-containment** : une classe $C1$ contient une autre classe $C2$ ssi : $(P2 \subset P1)$;
- **class-overlapping** : $C1$ et $C2$ sont deux classes overlapping ssi : $P1 \cap P2 \neq \emptyset$;
- **class-disjointness** : $C1$ et $C2$ sont deux classes disjointes ssi : $P1 \cap P2 = \emptyset$;
- **Classes sémantiquement disjointes** : $\neg \exists P$ tel que : $P1 \subset P$ ou $P2 \subset P$. d'une autre manière, deux classes sont sémantiquement disjointes si elles ne sont pas compatibles.

b- Correspondances entre attributs

- **Attributs identiques** : soit D un domaine donné, S une sémantique liée à un concept donné. Deux attributs $a1, a2$ sont identiques ssi

$$(D(a1) = D(a2)) \wedge (S(a1) = S(a2))$$

- **Attributs équivalents** : deux attributs $a1$ et $a2$ sont dits équivalents ssi :

$$(D(a1) \neq D(a2)) \wedge (S(a1) \equiv S(a2))$$

- **Ensemble-attributs équivalents** : soit $E1 = \{a1, a2, \dots, an\}$ et $E2 = \{b1, b2, \dots, bn\}$ deux ensembles d'attributs des classes $C1$ et $C2$ respectivement :

$$E1 \Leftrightarrow E2 \text{ ssi : } \forall ai \in E1, \exists bj \in E2 : ai \equiv bj \text{ et } \forall ai \in E2, \exists bj \in E1 : ai \equiv bj$$

Par exemple, dans la figure V.7, l'ensemble $E1$ est donné par $\{NameOs, VersionOs\}$, et $E2$ est donné par $\{Name, Version\}$ alors que $NameOs \equiv Name$ et $VersionOs \equiv Version$, donc on considère que $E1 \equiv E2$.

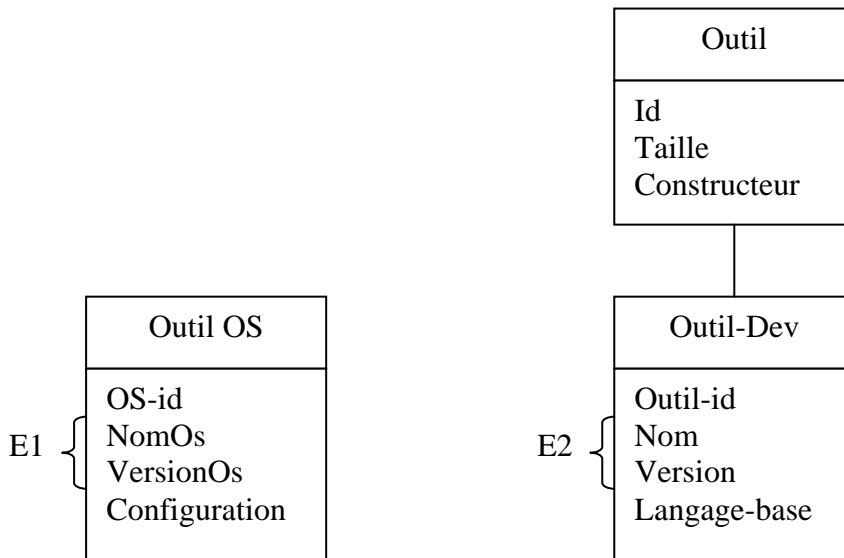


Figure V.7 : Exemple d'ensemble-attributs équivalents

- Attributs-classe équivalents : soit C une classe et $E = \{a_1, a_2, \dots, a_n\}$ un ensemble d'attributs d'une autre classe C' : $E \Leftrightarrow C$ ssi : $\forall a_i \in E, \exists b_j \in C : a_i \equiv b_j$. Par exemple, dans la Figure V.8, l'ensemble $E = \{\text{Diplôme, Responsabilité}\}$ est équivalent à la classe Rôle.

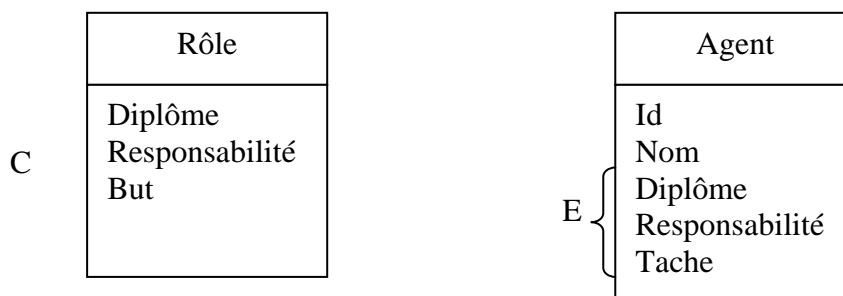


Figure V.8 : Exemple attributs-classe équivalents

c- Correspondances entre les méthodes

- **Méthodes équivalentes :**

soient $M_1 \in C_1$ et $M_2 \in C_2$, $M_1 \Leftrightarrow M_2$ ssi : $(C_1 \Leftrightarrow C_2) \wedge (S(M_1) \Leftrightarrow S(M_2))$;

- **Méthodes identiques :** M_1 est exactement la même que M_2 ;
- **Méthodes imbriquées :** M_1 est imbriquée dans M_2 ssi : $(S(M_1) \subset S(M_2))$;

V.4. Quelques méthodes d'intégration

V.4.1. Méthode écrite par C.Batini, M.Lenzerini & S.B Navathe [BAT86]

Cette méthode propose une approche d'intégration des schémas dans un modèle E/A. Elle est présentée en trois phases : analyse des conflits, fusion des schémas, enrichissement et restructuration du schéma global.

- a- **analyses des conflits** : durant cette étape, les conflits de nomination et structurels sont détectés. Le premier est résolu en renommant les entités en conflit et le deuxième est résolu en plusieurs étapes de restructuration des schémas composants (entité-relation, relation-entité et attribut en relation). Par contre, les conflits dus aux contraintes d'intégrité sont résolus, en laissant le choix au concepteur.
- b- **Fusion** : la fusion est réalisée par une simple superposition des concepts communs.
- c- **Enrichissement et restructuration du schéma global** : de nouvelles relations sont utilisées pour accroître la clarté du nouveau schéma. Les redondances dans le schéma global sont éliminées pour le rendre minimal.

V.4.2. Méthode décrite par Johan Eder & Heinz Frank [FRA94]

La méthode est basée sur le modèle orienté objet (OO). Elle est divisée en trois grandes étapes : Pré-Intégration, intégration des classes, et enfin intégration des types.

- a- **Pré-Intégration** : dans cette étape, une analyse générale des schémas composants permet de détecter les conflits. Le résultat de cette phase inclut des informations pour la poursuite du processus.
- b- **Intégration des classes** : consiste à faire une analyse sémantique des classes (partiellement parallèle avec les types) pour établir les relations entre ces dernières. Les classes peuvent être :
 - *Non disjointes* : intégration des superclasses, intégration des types et intégration des classes ainsi que les sous-classes (généralisation).
 - *Sémantiquement équivalentes mais disjointes* : intégration des superclasses et intégration des types (agrégation).
 - *Redondantes* : intégration des superclasses, types, classes et sous-classes.
- d- **Intégration des types** : elle consiste en une analyse structurelle dans le but de :
 - Résoudre les conflits structurels,
 - Intégrer les super-types,
 - Intégrer les types d'attributs,

- Intégrer les méthodes,
- Intégration statique/dynamique,
- Intégration des sous-types,
- Intégrer les classe correspondantes.

V.4.3. Méthode décrite par W.Gothard, A.Newfeld & P.C.Lockemann [GOT 92]

Cette méthode propose un processus d'intégration basé sur des prédicats (de similitude et de confirmation) en utilisant un modèle sémantique.

Au début, la méthode génère le premier type de prédicat entre les objets, attributs ou relations, dont une similitude est possible (grâce à une analyse sémantique). Par la suite, le processus doit vérifier les prédicats un par un pour le confirmer ou l'infirmier.

De plus, la méthode présente plusieurs contraintes et se concentre essentiellement sur les notions d'inclusion, d'association, rôle, ... entre les domaines.

Les étapes d'intégration sont :

- **Comparaison des schémas** : consiste à comparer les différents schéma pour détecter les conflits de représentation (noms, structures, ...).
- **Conformité des schémas** : le but de cette étape est de préparer les différents schémas pour l'intégration en résolvant les conflits.
- **Fusion et restructuration** : consiste en une superposition des différents schémas résultant de l'étape précédente, puis une amélioration pour l'atteinte des objectifs d'un schéma global.

V.4.4. Méthode décrite par P.Martin & W.Powley [MAR 95]

La méthode offre un processus d'intégration pour des schémas relationnels, décomposé en cinq étapes :

- **Exportation des schémas locaux** : permet de fournir les différents schémas en entrée du système et de les traduire en un modèle de données canonique. Elle permet aussi de fournir des contextes (sémantique) pour l'interprétation des données.
- **Résolution des conflits au niveau attribut** : consiste à identifier les attributs qui sont en conflits (noms, types et scalaires). Ils seront par la suite résolus pour les intégrer dans le schéma.

- **Résolution des conflits au niveau relation** : les relations entre les différents schémas exportés et qui représentent des concepts équivalents sont identifiées. Ainsi, les conflits détectés seront résolus en utilisant des fonctions de transformation et de concaténation de chaînes pour former un attribut.
- **Résolution des conflits au niveau schéma** : consiste à identifier certains types de conflits structurels. Par exemple, une relation dans un schéma est représentée par plusieurs relations dans un autre schéma. Ce type de conflit est résolu en utilisant les opérateurs de jointure, union,
- **Fusion** : consiste à fusionner les différents schémas en schéma global unifié.

V.4.5. Méthode de K.Jia Ling & L.P. Chen [JIA93]

Cette méthode est basée sur les correspondances entre classes-classes, attributs-attributs et les correspondances à travers les compositions hiérarchiques dans les différents schémas.

La méthode définit plusieurs règles pour la restructuration des classes ainsi que pour leur intégration. L'auteur ne spécifie pas la méthode en étapes mais plutôt sous forme de plusieurs algorithmes, chacun dédié à la résolution d'une correspondance donnée. Ainsi, la méthode s'occupe plutôt de l'aspect modélisation et non pas de l'aspect structurel, et préconise des restructurations des schémas en utilisant des règles (inversion, construire classes, détruire attributs, ...) avant l'intégration pour une simplification.

V.5. Conclusion

La méthodologie d'intégration, comme elle est définie dans ce chapitre, consiste à prendre en entrée un ensemble de bases de données (schémas) et à produire en sortie une description unifiée des schémas initiaux (schémas intégrés) tout en prenant en compte les points en communs (correspondances, conflits, ...).

Ces techniques d'intégration sont aussi capables de répondre aux besoins des concepteurs SIG pour intégrer des bases de données géographiques de différentes sources.

Pour ce qui est de notre cas, ces techniques vont nous servir à résoudre la problématique posée qui est : « intégration des fragments de modèles spatio-temporels de divers méthodes de conception des SIG en un schéma global ». Nous allons essayer de les adapter dans notre approche définie dans le prochain chapitre.

CHAPITRE VI

Chapitre VI

Environnement de développement d'un projet SIG :

Une approche basée sur l'intégration des fragments de modèles

VI.1. Introduction

Le développement d'un projet SIG fait le plus souvent appel à des applications déjà élaborées et à des données déjà mémorisées que ce soit dans des fichiers ou dans des bases de données indépendantes. C'est le cas pour les grandes entreprises où le projet se traduit par un développement indépendant de plusieurs bases de données pour des applications spécifiques. Par exemple, dans le cadre d'une application de gestion du risque sismique, le SIG fait appel à des différents services pour permettre aux responsables de gérer les interventions et de coordonner les travaux. Plus précisément, le service d'intervention peut être amené à combiner des informations provenant du service relatif à la gestion du réseau routier.

Le développement indépendant d'un projet SIG consiste à retrouver les sous-systèmes qui le composent. Ainsi, une fois définis et après leur modélisation, il est évident de développer une architecture pour les faire collaborer d'une manière transparente pour l'utilisateur.

Dans ce chapitre, nous proposons une approche de développement basée sur l'intégration des fragments de modèles. Ces derniers représentent les schémas conceptuels des sous-systèmes.

Notre approche est conçue suivant deux points :

- Assurer les deux phases de développement d'un projet : analyse des besoins et la modélisation du système, suivi d'une phase d'intégration des fragments de modèles.
- S'inspirer des méthodes d'intégration des schémas et proposer une architecture d'intégration spécifique à notre cas.

VI.2. Présentation générale de notre démarche

Pour tout espace géographique, il existe une représentation spatiale en couches d'information géographique (voir Chapitre I). Le développement d'un SIG n'est que la modélisation des couches qui le composent [ADN01].

C'est dans ce contexte que nous avons conçu notre démarche. Elle propose une modélisation basée sur l'intégration des fragments de modèles des couches spatiales. Afin d'atteindre l'objectif souhaité, nous avons défini un environnement de développement basé sur le processus itératif incrémental (modèle en fontaine définie dans le Chapitre IV). Cet environnement est composé de trois (03) phases (Figure VI.1):

- Analyse des besoins
- Modélisation
- Construction du schéma fédéré du projet

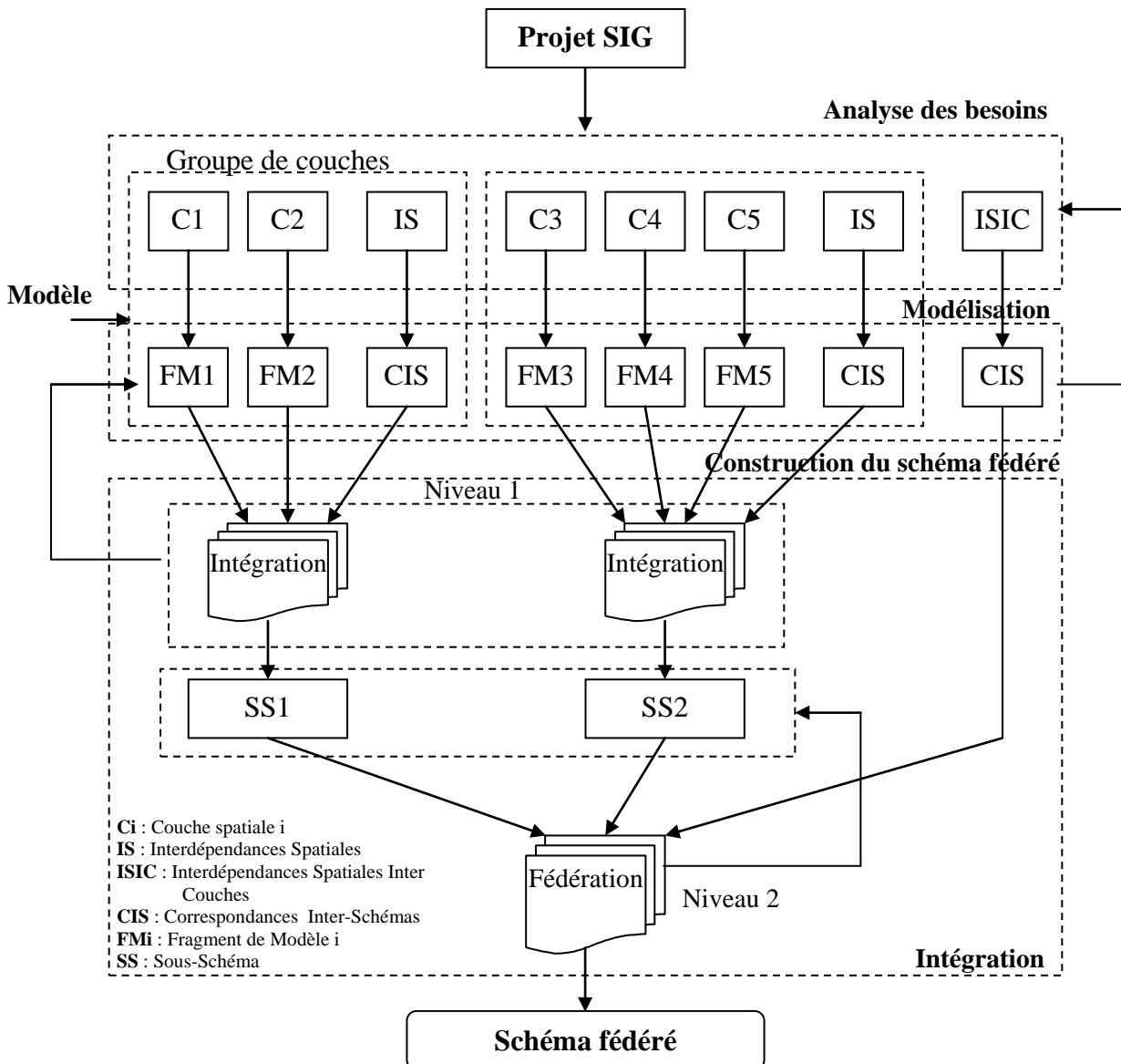


Figure VI.1 : schéma général du développement d'un projet SIG

Avant de procéder à la description de notre démarche, nous présentons dans ce qui suit quelques définitions des concepts utilisés.

a- Couche spatiale

Une couche spatiale est un ensemble d'objets géographiques qui appartiennent au même thème du domaine d'application. Par exemple : réseau routier, découpage administratif, occupation du sol, Sur une même couche, on peut avoir plusieurs objets géographiques de différentes formes géométriques (point, ligne, polygone). Par exemple, dans un réseau routier, on trouvera des carrefours, des ponts représentés par des points et des segments de route représentés par des lignes.

Règle 1 : une couche définit un thème

b- Géométrie d'une couche

La géométrie d'une couche spatiale dépend de la forme de ses objets géographiques. Elle peut être :

- *Simple* : la couche spatiale regroupe des objets de même type (Point, Ligne ou Polygone). Par exemple, la couche « découpage administratif » composée de wilayas, daïras et communes est représentée par la géométrie polygone héritée de la forme géométrique de la wilaya, daïra et commune.

Règle 2 : une couche simple est composée d' un seul type d'objets géographiques

- *Composée* : la couche spatiale regroupe un ensemble d'objets géographiques de différentes formes géométriques. par exemple, le réseau routier est représenté par une géométrie composée de Point et Ligne dont les points représentent les carrefours, les ponts, ..., et les lignes représentent les segments de route.

Règle 3 : une couche composée contient plusieurs types d'objets géographiques

c- Modèle géométrique :

Une couche spatiale est représentée selon un modèle géométrique. Il est déterminé à partir des relations qui existent entre les objets de la couche. On peut avoir trois (03) types de modèle : spaghetti, réseau et topologie (Voir Chapitre I)

Règle 4 : une couche est caractérisée par un modèle géométrique

d- Groupe de couches spatiales :

Un groupe de couches spatiales contient un ensemble de couches qui partagent les mêmes concepts spatiaux (géométrie, relations spatiales, méthodes, ...). Généralement, ces couches sont regroupées selon leur modèle géométrique (réseau, spaghetti, topologie). Par exemple, le réseau routier, réseau téléphonique et réseau électrique appartiennent au même groupe de couches appelé « réseau ».

e- Interdépendances spatiales :

Une interdépendance spatiale est une relation topologique (intersection, adjacence, ...) qui relie deux objets géographiques appartenant à des couches différentes. Cette relation est obtenue par la superposition de ces couches.

Dans notre cas, on peut avoir deux types d'interdépendances spatiales :

- *Intra-couches* : elle relie les couches spatiales d'un même groupe.
- *Inter-couches* : elle relie les couches spatiales de groupes différents.

f- Fragment de modèle :

Un fragment de modèle est le résultat de la modélisation d'une couche spatiale.

g- Intégration :

C'est l'étape de fusion des fragments de modèles et de correspondance des sous-schémas pour former le schéma fédéré. Dans notre cas, on peut avoir deux niveaux d'intégration :

- *Intégration des fragments de modèles «1^{er} niveau* » : les fragments de modèles du même groupe sont regroupés pour former un sous-schéma.
- *Fédération des sous-schémas «2^{ième} niveau* » : les sous-schémas obtenus au premier niveau d'intégration sont inter reliés pour construire le schéma fédéré.

h- Sous-schéma :

C'est le résultat de l'intégration des fragments de modèles du même groupe de couches spatiales réalisée dans le premier niveau d'intégration.

i- Schéma fédéré :

Le schéma fédéré est une vision virtuelle du schéma global du système. Il repose sur la définition des correspondances entre les sous-schémas. Cela est prévue dans le deuxième niveau de fédération des sous-schémas.

VI.2.1. Analyse des besoins

Cette phase consiste à retrouver l'ensemble des couches spatiales qui composent le système. Elles seront par la suite regroupées en groupe de couches selon leur modèle géométrique. Une fois déterminées, une étape de reconnaissance des interdépendances spatiales est définie pour retrouver les correspondances entre les couches que ce soit du même groupe (interdépendances spatiales intra-couches) ou de groupes différents (interdépendances spatiales inter-couches).

VI.2.2. Modélisation

Les couches spatiales obtenues dans la phase d'analyse des besoins seront modélisées indépendamment. Le choix du modèle dépend des caractéristiques géométriques du groupe associé à la couche. Ainsi, le modèle choisi sera appliqué sur l'ensemble des couches du même groupe.

A la fin de cette phase, nous obtiendrons des fragments de modèles. Ils sont au nombre de couches spatiales.

VI.2.3. Construction du schéma fédéré

Les fragments de modèles obtenus seront intégrés dans un schéma fédéré en deux niveaux d'intégration :

- **Premier niveau** : Pour chaque groupe de couches spatiales, Les fragments de modèles sont intégrés pour former le sous-schéma. Les interdépendances spatiales intra-couches sont introduites pour faire correspondre les fragments de modèles. La réutilisation des fragments existants peut se faire à ce niveau.
- **Deuxième niveau** : les sous-schémas obtenus dans le premier niveau d'intégration sont inter reliés pour former le schéma fédéré du système. Les interdépendances spatiales inter-couches sont introduites pour faire correspondre les sous-schémas.

VI.3. Description de la démarche

VI.3.1. Analyse des besoins

Le but de cette phase est de retrouver l'ensemble des couches spatiales qui composent le système, leurs caractéristiques géométriques et les interdépendances spatiales qui les relient. Avant de procéder à l'acquisition de ces concepts, il est évident de précéder cette étape par une étude de l'existant. Une étude qui permet de rassembler les informations nécessaires et

b- Scénario

Un scénario est un exemple de déroulement d'opérations qui représentent des réactions à des types de situations auxquelles sont confrontés les acteurs pour répondre à des objectifs précis [ALI99]. Un exemple de scénario est présenté dans le Chapitre III.4 de cette thèse. Il va nous servir à concevoir le SIG pour la gestion du risque sismique.

VI.3.1.1. Processus d'extraction des couches spatiales

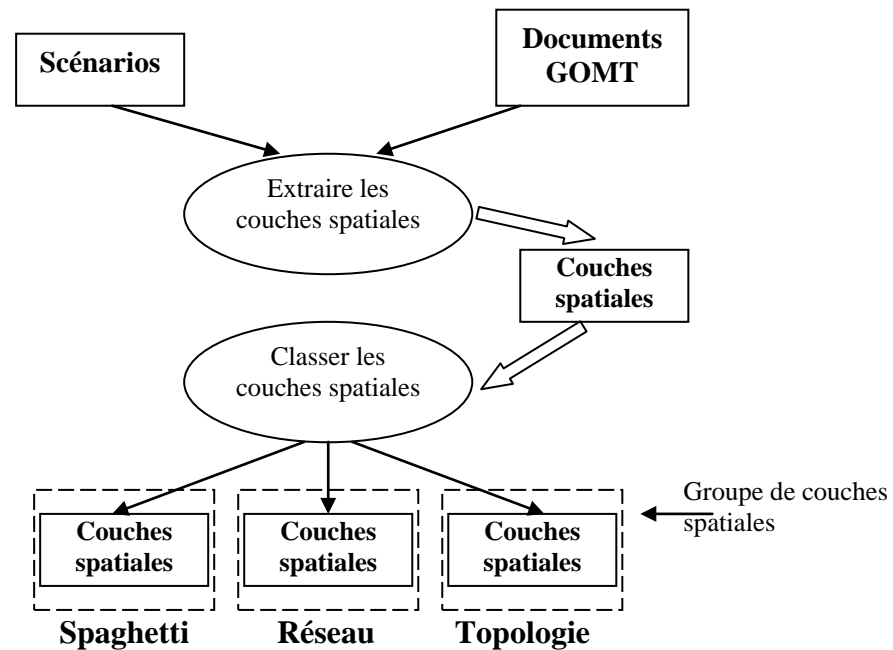


Figure VI.3 : Processus d'extraction des couches spatiales

A partir des scénarios obtenues dans l'étude de l'existant et des documents GOMT :

- Extraire l'ensemble des couches spatiales qui composent le système. Pour cela, nous avons défini une structure de données qui contient des propriétés descriptives et géométriques de la couche (figure VI.4).
- Classifier les couches spatiales une fois obtenues selon leur modèle géométrique. Nous obtiendrons ainsi au maximum trois (03) catégories de groupes de couches spatiales : réseau, spaghetti et topologie (figure VI.3)

Nom de la couche	Géométrie
- Description	
- {objets géographiques}	
- Documents cartographiques	
- Modèle géométrique	

Figure VI.4 : Structure de la couche spatiale

Exemple :

R.routier	Point , ligne
<ul style="list-style-type: none"> - Réseau routier - {Route, Pont, carrefour, ...} - Carte touristique 1 :50000 - Réseau 	

Structure de la couche « réseau routier »

VI.3.1.2. Processus de reconnaissance des interdépendances spatiales :

Les interdépendances spatiales sont obtenues en superposant les couches spatiales. Pour cela, nous allons :

- Extraire toutes les relations topologiques possibles entre les couches spatiales. Le tableau VI.1 présente les différentes relations topologiques qui peuvent exister entre deux couches spatiales.

Couche i	Couche j	Relations topologiques
Polygone	Polygone	Disjonction, Adjacence, recouvrement, Inclusion, Egalité
Polygone	Ligne	Disjonction, Adjacence, Croisement, Inclusion
Polygone	Point	Disjonction, Adjacence, Inclusion
Ligne	Ligne	Disjonction, croisement, Egalité
Ligne	Point	Inclusion, disjonction, adjacence
Point	Point	Disjonction, Adjacence, Egalité

Table VI.1 : Relations topologiques entre deux couches spatiales

- Eliminer les relations inutiles et garder celles qui répondent aux objectifs.
- Eliminer les relations ne vérifiant pas les contraintes d'intégrité spatiale.
- Construire un graphe de topologie dont les nœuds représentent les couches et les arcs représentent les relations topologiques (figure VI.5).

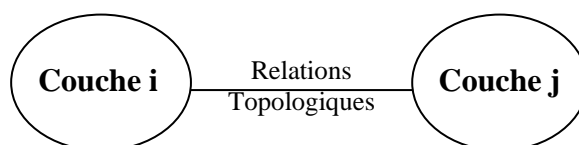
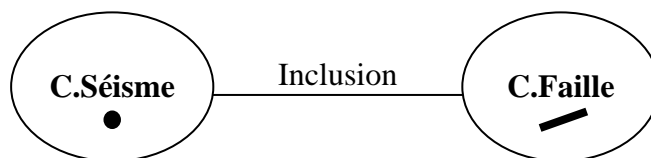


Figure VI.5 : Relations topologiques entre deux couches spatiales

Exemple :



A la fin de ce processus, nous obtiendrons un graphe de topologie dont les arcs reliant les couches du même groupe représentent les interdépendances spatiales intra-couches. Tandis que les arcs reliant les couches de groupes différents représentent les interdépendances spatiales inter-couches.

Une fois les couches spatiales et leurs interdépendances sont obtenues, nous allons procéder à la phase modélisation pour obtenir les fragments de modèles.

VI.3.2. Modélisation

L'une des raisons qui nous a mené à définir une démarche de développement à base d'intégration est la multiplicité des concepts de modélisation des SIG qui n'a fait qu'accroître l'embarras du développeur pour choisir le modèle le mieux adapté à ses besoins.

VI.3.2.1. Choix du modèle

Un modèle conceptuel doit offrir aux développeurs une grande richesse d'expression pour répondre à la diversité des besoins et en même temps, leur permettre de déduire un schéma de données lisible et facile à appréhender. Or, cela n'est pas le cas pour certaines méthodes de conception des SIG. Leur modèle ne dispose pas de cette richesse de concepts pour répondre à la diversité de domaines des SIG. On distingue des méthodes qui intègrent tous les concepts de modélisation liées à la gestion des réseaux (ex : GEO-OOA[KPS 97], OMT-G [KAR00]). Par contre, d'autres méthodes offrent une grande richesse dans ces concepts pour gérer la topologie (ex : MADS[PAR97]). Cette diversité est due au fait que ces méthodes de conception ont été conçues dans un contexte applicatif.

Afin d'aider le développeur du projet SIG à choisir le modèle conceptuel adéquat, nous présentons un tableau comparatif des méthodes de conception des SIG basé sur les trois (03) modèles géométriques d'une couche spatiale : réseau, spaghetti et topologie (tableau VI.2).

Modèles	Modèles géométriques		
	Réseau	Spaghetti	Topologie
GEO-OOA	- Classes réseau : Node et Link	Oui	- Agrégation spatiale
MADS	- Opérations	Oui	- Agrégation spatiale - Généralisation - Relations topologiques
OMT-G	- Classes réseau : unidirectional line Bidirectional line Node - Associations : Link-Link Link-Node	Oui	- Contraintes d'intégrités spatiales
PERSPECTORY	-	Oui	- Agrégation spatiale
POLLEN	-	Oui	- Relations topologiques
GEO-OM	-	Oui	- Relations topologiques
GEO2	-	Oui	- Agrégation spatiale
CONGOO	-	Oui	- Contraintes d'intégrités spatiales

Table VI.2 : Comparaison des modèles conceptuels selon le modèle géométrique

VI.3.2.2. Les étapes de modélisation

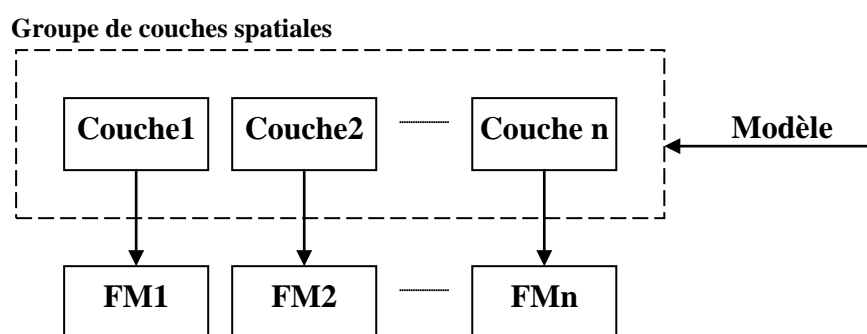


Figure VI.6 : Modélisation des couches spatiales

Les couches spatiales obtenues dans la phase d'analyse des besoins seront modélisées indépendamment. Le développeur choisira le modèle qui convient pour modéliser les caractéristiques géométriques d'une couche. Ainsi, le modèle choisi sera appliqué sur l'ensemble des couches du même groupe (Figure VI.6).

A la fin de cette phase, nous obtiendrons des fragments de modèles qui seront par la suite intégrés pour construire le schéma conceptuel global du projet SIG.

VI.3.3. Construction du schéma fédéré

La modélisation indépendante des couches spatiales proposée dans notre démarche, nous a conduit à une situation d'intégration spécifique :

- Les fragments de modèles obtenus sont disjoints puisque l'objet géographique appartient à une seule couche spatiale.
- Les fragments de modèles sont homogènes pour le même groupe de couches et hétérogènes entre les différents groupes.
- Les interdépendances spatiales relient les objets géographiques appartenant à des couches différentes. leur modélisation permettra de relier les fragments de modèles pour en construire le schéma fédéré.

Afin d'arriver à construire le schéma fédéré du projet SIG, nous proposons deux niveaux d'intégration :

VI.3.3.1. Intégration des fragments de modèles « Niveau 1 »

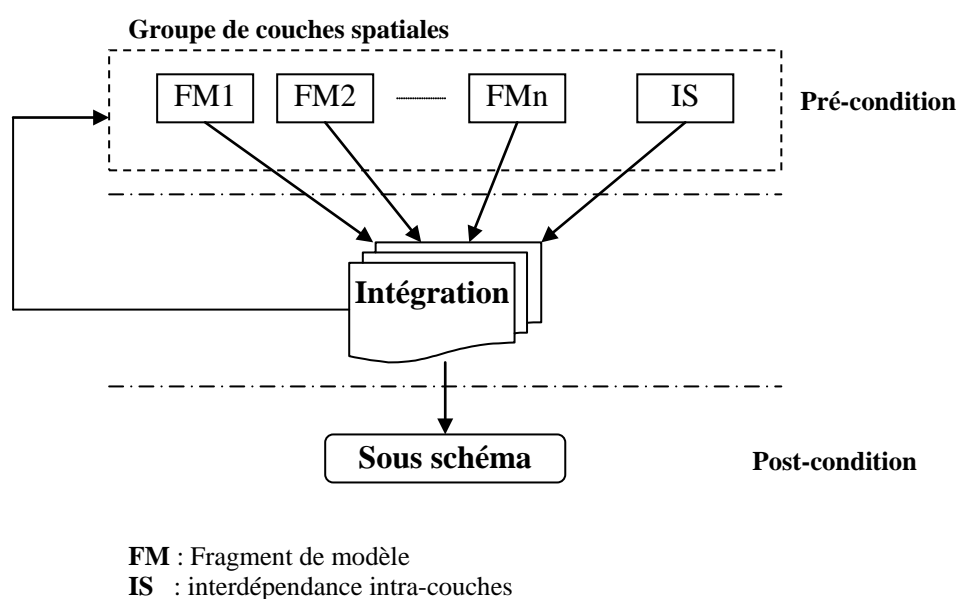


Figure VI.7 : Intégration des fragments de modèles

a- Pré-condition

- Les fragments de modèles à intégrer sont homogènes et disjoints. Ils appartiennent au même groupe de couches spatiales et ils sont modélisés par le même modèle.
- Les interdépendances spatiales sont définies par des associations topologiques pour relier les fragments de modèles.

b- Post-condition

- Le sous-schéma obtenu après intégration des fragments de modèles doit vérifier les caractéristiques d'un schéma complet, minimal et compréhensible.

c- Les étapes d'intégration**c.1. Pré-Intégration**

Cette étape consiste à traduire les différents schémas à intégrer en un modèle de données commun (modèle canonique). Or que dans notre cas, les fragments de modèles sont obtenus en appliquant le même modèle pour les couches spatiales du même groupe. Ces fragments sont alors homogènes et prêts à être intégrés.

c.2. Comparaison des fragments de modèles et détection des conflits**c.2.1. Détection des conflits**

La disjonction et l'homogénéité des fragments de modèles n'empêchent pas d'avoir certains types de conflits.

Le tableau suivant (tableau VI.3) synthétise les différents conflits qui peuvent exister entre les fragments de modèles.

	Conflits	
	SI	SIG
Fragments de modèles non-disjoints	<ul style="list-style-type: none"> - Synonyme - Homonyme - Type de données - Dépendance - Unité - Précision - Clé - Comportement 	<ul style="list-style-type: none"> - Géométrique - Structurel - Sémantique - Topologique - Agrégation spatiale - Généralisation - Temporel
Fragments de modèles disjoints	<ul style="list-style-type: none"> - Homonyme 	<ul style="list-style-type: none"> - Sémantique

Table VI.3 : classification des conflits

* Conflit d'homonyme

Ce type de conflit peut se produire lors de l'intégration. Un même nom peut être utilisé pour deux objets géographiques différents. Par exemple, l'objet collecteur réfère aux transformateurs dans un réseau électrique (figure VI.8-a). Tandis que dans un réseau routier, désigne des carrefours (figure VI.8-b).

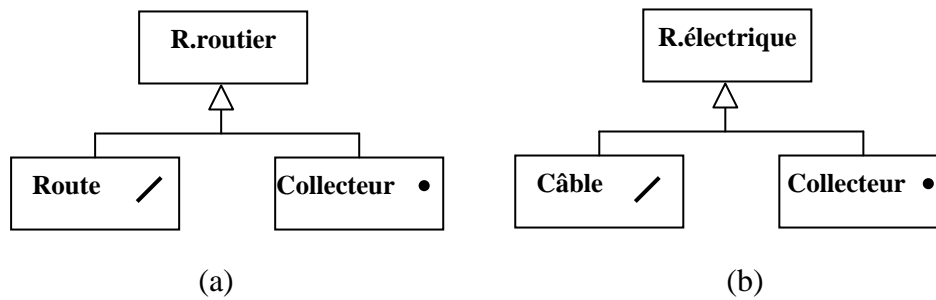


Figure VI.8 : Exemple d'homonyme

* Conflit sémantique spatiale

Le choix de type de positionnement, du point de référence et du système de projection diffère d'une couche spatiale à une autre. Cela peut induire à des conflits de type sémantique.

c.2.3. Correspondances inter-schémas

Pour construire le schéma conceptuel, il est nécessaire de retrouver les équivalences entre les fragments de modèles à intégrer. Ces équivalences se présentent au niveau des concepts : classes-classes, classes-attributs, ... (voir Chapitre V).

Pour ce qui est de notre cas, les équivalences ne sont pas définies puisque les fragments de modèles sont disjoints. Elles seront remplacées par les interdépendances spatiales déduites en phase d'analyse des besoins pour correspondre les fragments de modèles. Ces relations seront intégrées dans le sous-schéma comme étant des associations topologiques pour relier les fragments de modèles. Pour cela, nous proposons d'utiliser les associations définies dans le modèle de MADS. Celui-ci représente toutes les relations topologiques qui peuvent exister.

c.3. Uniformisation des schémas

Dans cette étape, nous proposons une solution pour résoudre les conflits détectés en comparant les fragments de modèles.

c.3.1. Résolution des conflits

* Homonyme

La solution proposée consiste à changer le nom de l'un des concepts qui sont en conflit.

* Sémantique spatiale

La résolution des conflits sémantiques spatiaux consiste à choisir des caractéristiques de projection communes (positionnement, point de référence, système de projection). Ainsi, les différentes couches spatiales seront converties en utilisant des opérations de conversion.

c.4. Fusion et restructuration des sous-schémas

La fusion des fragments de modèles en un sous-schéma nécessite d'introduire de nouveaux concepts.

c.4.1. Classe Pivot

C'est une super-classe d'un ensemble d'entités ayant un même type géométrique (Point, ligne ou polygone) de même couche spatiale. Par exemple, dans un réseau routier, la classe Pivot « Point-réseau » est la super-classe des sous-classes « carrefour » et « pont » dont la géométrie est de type point (figure VI.9).

La structure de la classe pivot est définie comme suit :

Nom : nom de la classe

Attributs : des attributs géométriques hérités des sous classes ou attributs spécifiques.

Méthodes : Méthodes géométriques hérités des sous classes ou méthodes spécifiques.

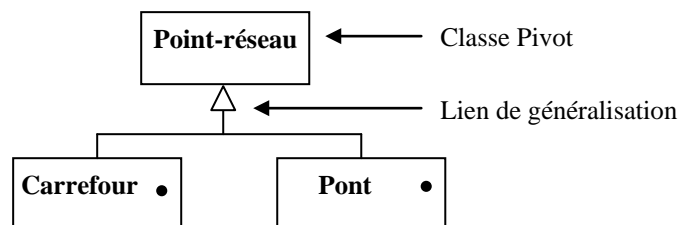


Figure VI.9 : Exemple de classe Pivot dans un réseau routier

c.4.2. lien de généralisation

C'est une association qui permet d'établir le lien entre la super-classe « Pivot » et l'ensemble des sous-classes dérivant de la même forme géométrique (figure VI.8).

La fusion des fragments de modèles en un sous-schéma consiste à :

- Définir les classes Pivots de chaque fragment de modèle.
- Relier les classes ainsi obtenues en utilisant les correspondances inter-schémas définies dans l'étape précédente.
- Analyser le sous-schéma pour déduire de nouvelles propriétés (relation redondantes, nouvelles, ...).

Nous obtiendrons à la fin de cette étape, pour chaque groupe de couches le sous-schéma correspondant. Il sera par la suite utilisé dans le deuxième niveau d'intégration pour obtenir le schéma fédéré.

VI.3.3.2. Fédération des sous-schémas « Niveau 2 »

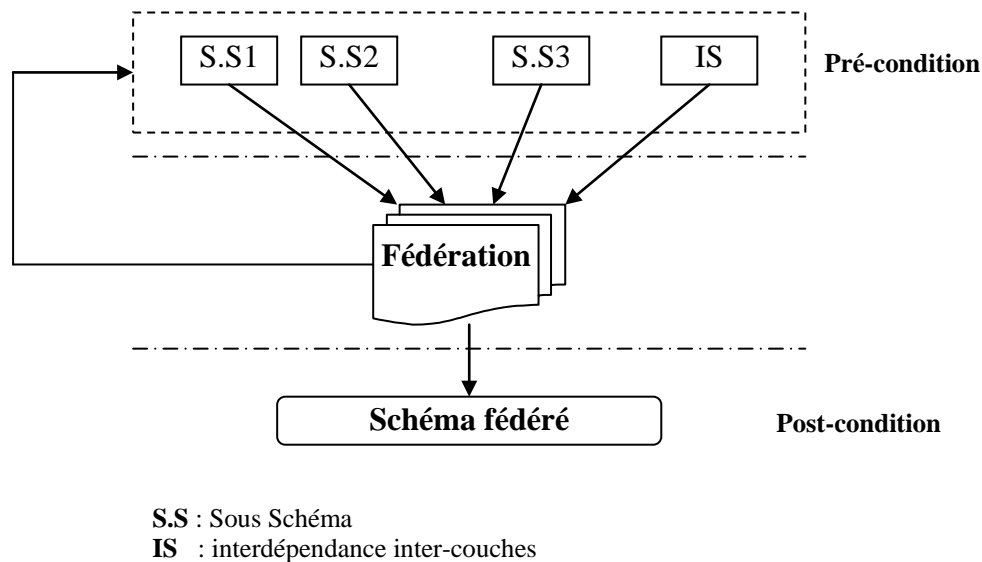


Figure VI.10 : Fédération des sous-schémas

a- Pré-condition :

- Les sous-schémas sont le résultat de l'intégration des fragments de modèles de chaque groupe de couches spatiales.
- Les sous-schémas sont hétérogènes et disjoints.

Pourquoi hétérogène ?

La modélisation des couches spatiales diffère d'un groupe de couches à un autre. Chacun utilise son propre modèle (voir la phase de modélisation). Par exemple, les couches réseau routier et réseau électrique qui appartiennent au même groupe de couches spatiales « réseau », seront modélisés par la méthode GEO-OOA. Tandis que la couche découpage administratif qui appartient au groupe « topologie », sera modélisée par la méthode MADS.

Pourquoi disjoints ?

Les fragments de modèles sont disjoints. Après leur intégration, les sous-schémas obtenus seront aussi disjoints.

- Le nombre de sous-schéma à intégrer sera inférieur ou égal à 3. Ce nombre est lié au modèle géométrique associé aux groupes de couches spatiales : réseau, spaghetti et topologie.

b- Post-condition :

- Le schéma fédéré obtenu après la correspondance des sous-schémas est multi modèles. Il donne une vision virtuelle du schéma global du système.

c- Les étapes de fédération :

Contrairement au premier niveau d'intégration où les sous-schémas obtenus sont des schémas unifiés, l'étape de fédération des sous-schémas n'a pas recours à un schéma global unifié. Le schéma fédéré donne une vue globale du schéma conceptuel de la base de données.

Le système fédéré adopté dans notre démarche, ne donne pas l'importance aux étapes de (1) traduction des sous-schémas en un modèle commun, (2) détection et résolution des conflits. Par contre, il est indispensable que les correspondances soient définies entre les sous-schémas.

c.1. Comparaison des sous-schémas

c.1.1. Correspondances inter-schémas

Dans ce niveau de fédération, les correspondances entre les sous-schémas sont représentées par les interdépendances spatiales inter-couches. Afin de les intégrer dans le schéma global, nous proposons d'utiliser les associations topologiques définies dans MADS.

c.2. Restructuration du schéma fédéré

Pour relier les sous-schémas en un schéma fédéré, nous allons introduire les concepts utilisés dans le premier niveau d'intégration (classe Pivot, lien de généralisation). Mais cette fois-ci, la classe Pivot sera la super classe des sous-classes Pivot du même groupe de couches spatiales dont elles possèdent la même géométrie (Point, Ligne ou Polygone) (figure VI.11). Le lien de généralisation permettra ainsi de relier les classes Pivot du deuxième niveau d'intégration avec les sous-classes pivots du premier niveau.

Ainsi, la correspondance des sous-schémas sera définie comme suit :

- Etablir un deuxième niveau d'hierarchie contenant les classes pivots de chaque sous-schéma.
- Relier les classes pivots ainsi obtenues en utilisant les correspondances inter-schémas définies dans l'étape (c.1.2).
- Analyser le schéma fédéré une fois construit pour déduire de nouvelles propriétés (relations redondantes, nouvelles, ...).

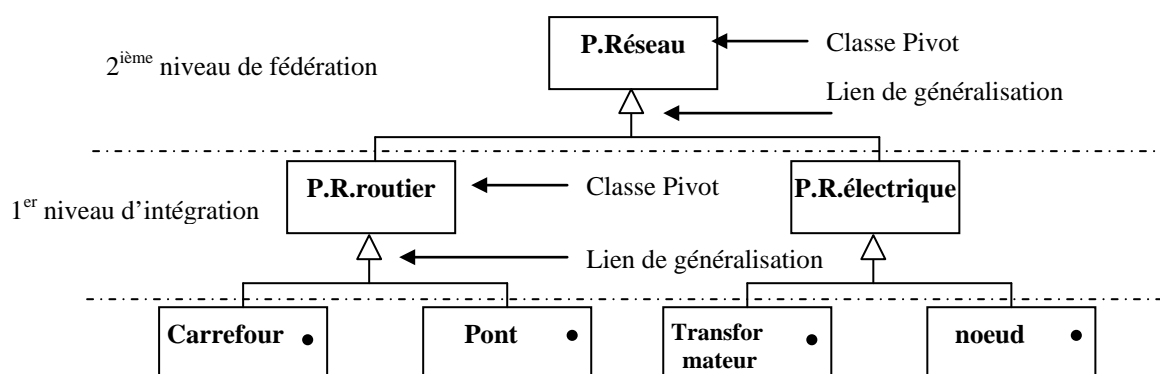


Figure VI.11 : Classe Pivot et lien de généralisation

A la fin de cette étape, nous obtiendrons le schéma fédéré du projet SIG.

VI.3.4. Réutilisation des fragments de modèles

L'architecture à base d'intégration proposée dans cette démarche, est très ouverte pour intégrer des fragments de modèles existants. Cela peut se faire au premier niveau d'intégration. Mais il n'est possible qu'après la traduction du fragment (figure VI.12). Pour cela, nous allons définir un processus qui résume les étapes de pré-intégration des fragments de modèles existants (figure VI.13).

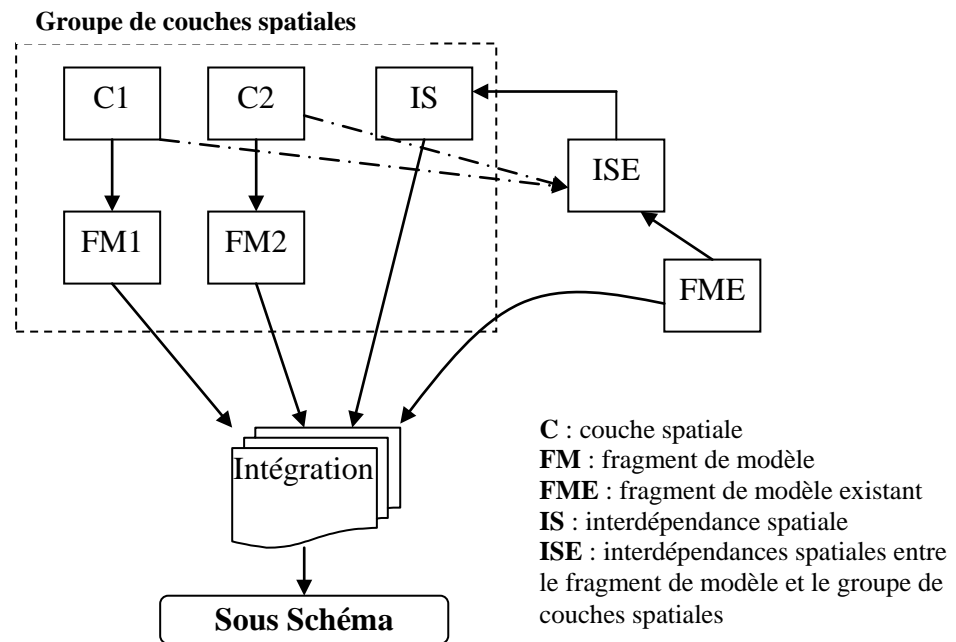


Figure VI.12 : intégration des fragments de modèles existants

VI.3.4.1. Processus de pré-intégration des fragments de modèles existants (figure VI.13)

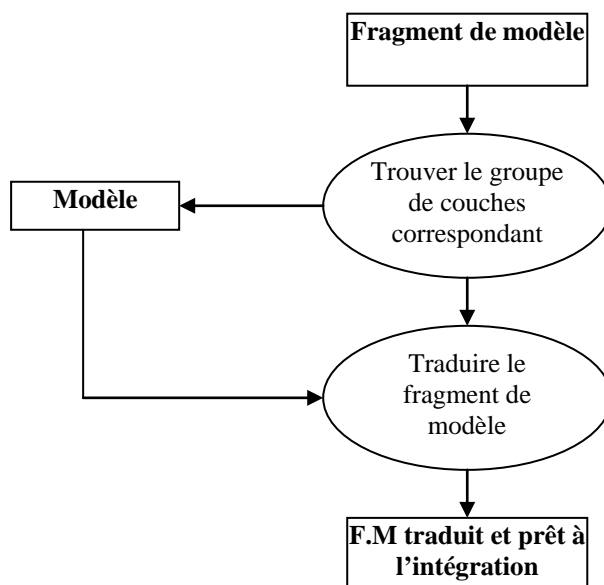


Figure VI.13 : Processus de pré-intégration des fragments de modèles existants

- Trouver le groupe de couches spatiales correspondant au fragment de modèle. Cela permettra de connaître le modèle commun pour traduire le fragment.
- Une fois le modèle commun connu, traduire le fragment de modèle. Dans cette étape, il est fortement possible de trouver les liens de correspondances entre le fragment de modèle et le modèle commun puisque ce dernier permet de modéliser l'ensemble des couches spatiales du même groupe. Il sera alors nécessaire de définir des techniques de traduction pour passer du fragment de modèle au modèle commun.

A la fin de cette étape, nous obtiendrons des fragments de modèles traduits et prêts à être intégré dans les sous-schémas du système.

VI.3.4.2. Interdépendances spatiales

Avant de procéder à l'intégration des fragments de modèle existants, on doit compléter l'ensemble des interdépendances spatiales inter et intra couches par des nouvelles relations topologiques. Elles seront déduites en faisant correspondre le fragment de modèle à intégrer avec l'ensemble des couches spatiales du système. Pour cela, nous allons appliquer le processus de reconnaissance des interdépendances spatiales proposé en phase d'analyse des besoins afin de retrouver ces nouvelles relations. Mais cela nécessite de connaître la forme géométrique modélisée dans le fragment de modèle à intégrer.

VI.4. Application : Un SIG pour la gestion du risque sismique

VI.4.1. Analyse des besoins

a- Extraction des couches

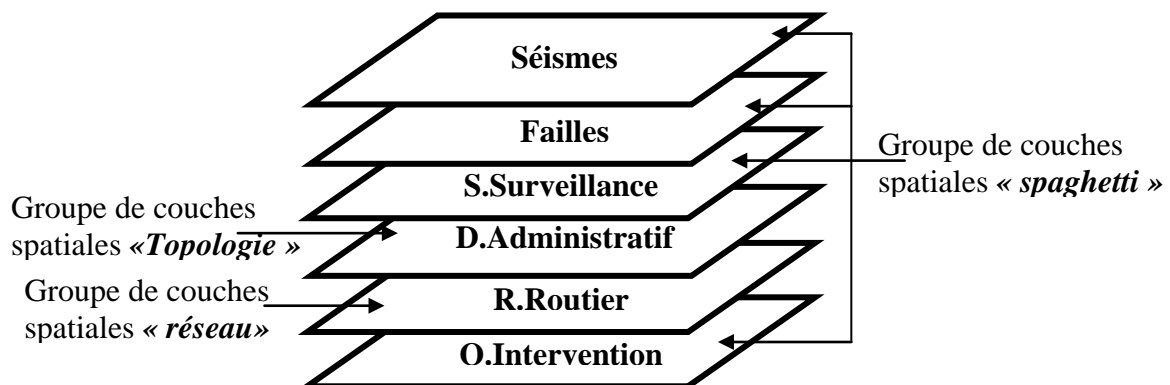


Figure VI.14 : Couches spatiales du système pour la gestion du risque sismique

b- Structures des couches

Séisme	Point
Séisme {Séisme (intensité, magnitude, région, date et heure)} Catalogue de la sismicité historique. Spaghetti	

Couche « Séisme »

S.Surveillance	Point
Réseau des Stations de surveillances {S-surveillance(Nom-station, adresse, type (mondiale, régionale, mobile)} Carte du réseau des stations de surveillances Spaghetti	

Couche « Station de surveillance »

R.Failles	Ligne
Réseau de failles {Faille(Nom-faille, longueur, Mécanisme-rupture(normale, inverse,décrochement)} Carte de réseau de Failles. Spaghetti	

Couche « Réseau de failles »

O.intervension	Point
Organismes d'interventions et de secours {Hopital (Nom, adresse, téléphone, fax), Protection-civile, Gendarme) Carte topographique. Spaghetti	

Couche « Organisme d'intervention et de secours »

R.routier	Ligne, Point
Réseau routier {Route (Nom, type (communale, autoroute, nationale), Voie-urbaine (Nom, type (avenue, boulevard)), Ponts, Carrefours} Carte routière Réseau	

Couche « Réseau routier »

D.Administratif	Polygone
Découpage administratif {Wilaya (Nom, surface, densité, type-construction), Daïra, Commune} Carte de découpage administratif de l'Algérie. Topologie	

Couche « Découpage administratif »

c- Interdépendances spatiales

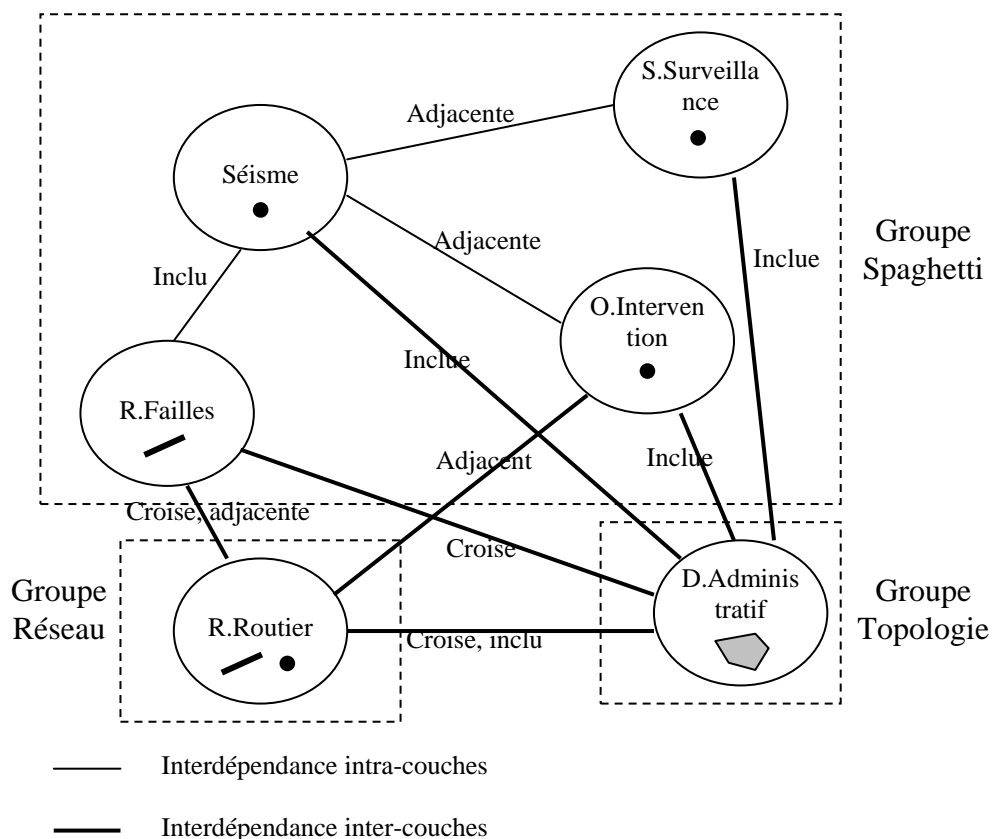


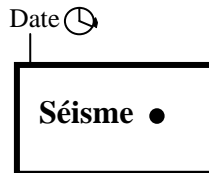
Figure VI.15 : Graphe de topologie

VI.4.2. Modélisation

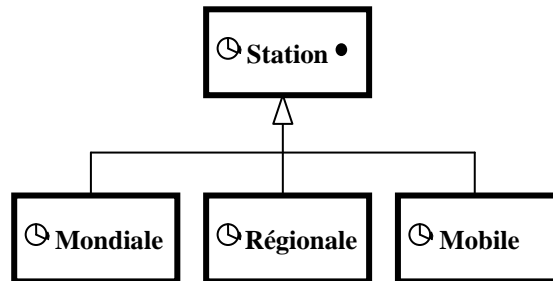
Groupe de couches spatiales « spaghetti »

Dans ce groupe, nous avons choisi le modèle de MADS pour modéliser l'ensemble des couches spatiales. Ainsi, nous obtiendrons les schémas suivants :

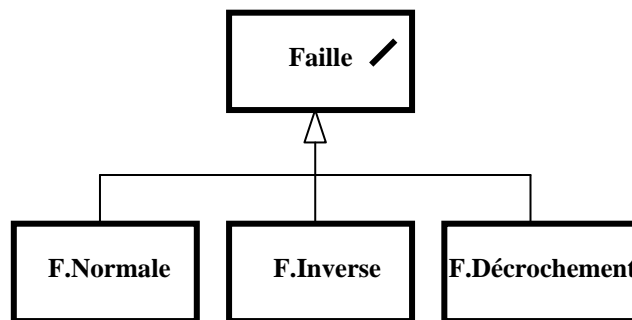
a- Couche « séisme »



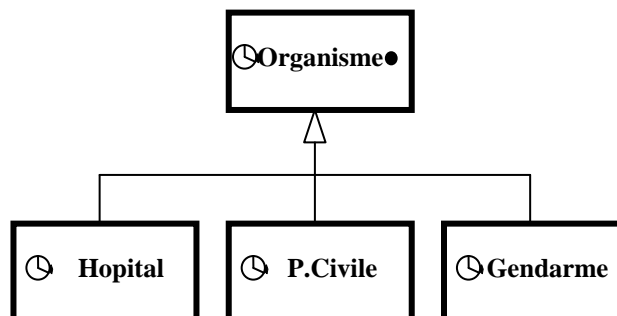
b- Couche «Station de surveillance »



c- Couche « Réseau de failles »



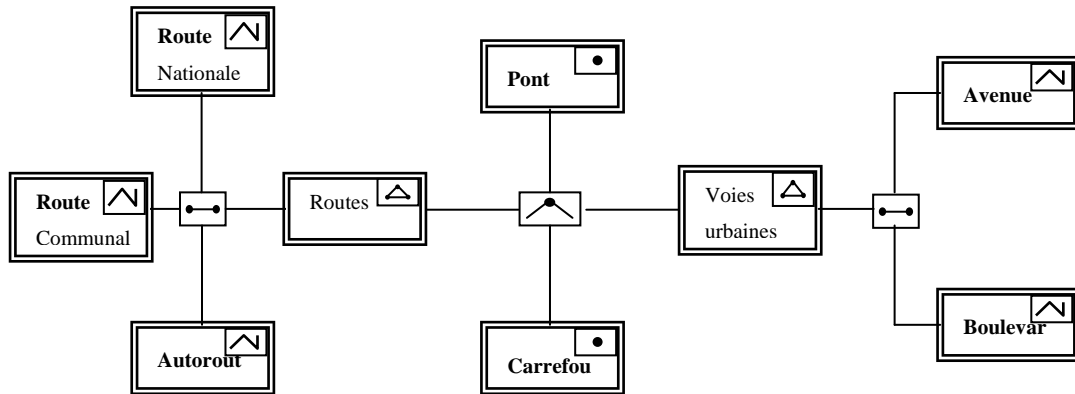
d- Couche « Organisme d'intervention et de secours »



Groupe de couches spatiales « Réseau » :

Une seule couche est définie dans ce groupe, elle est modélisée par le modèle de la méthode GEO-OOA.

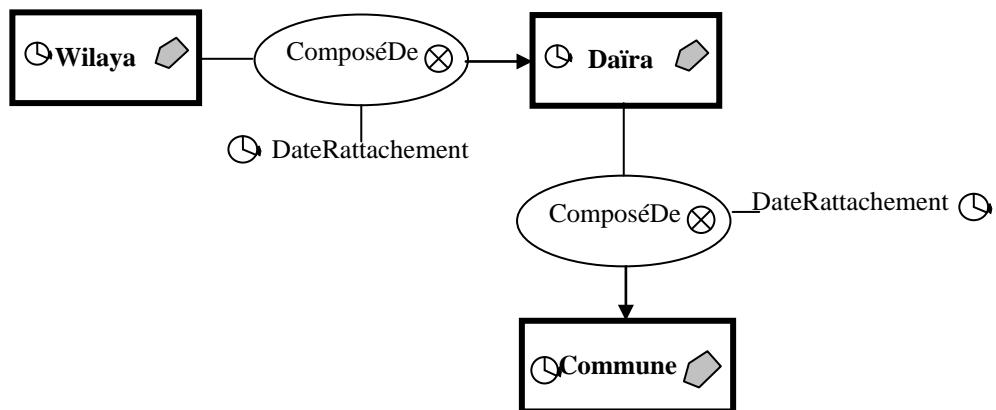
e- Couche « Réseau routier »



Groupe de couches spatiales « Topologie » :

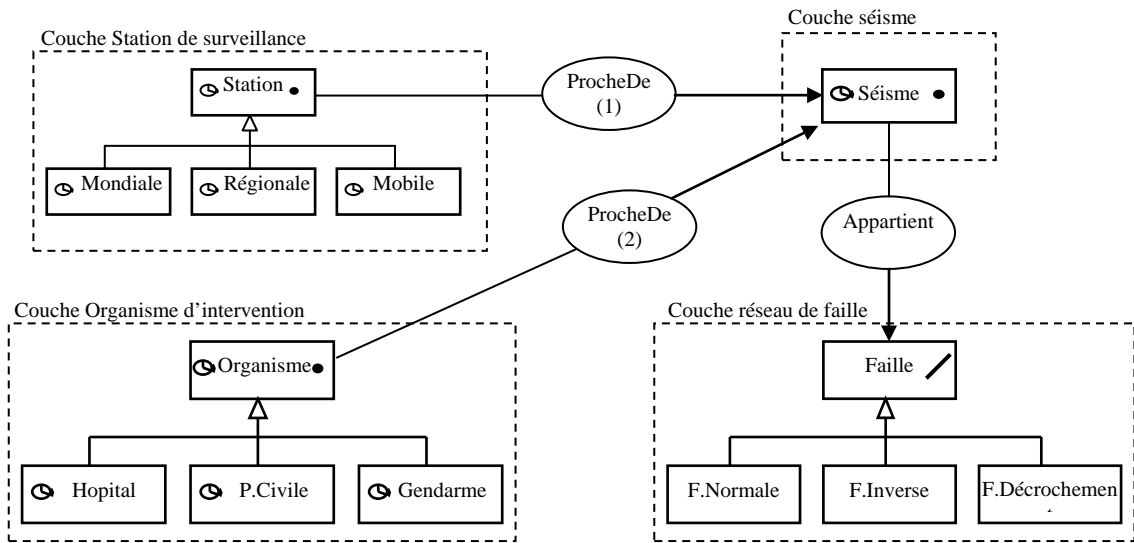
Ce groupe est composé par une seule couche « découpage administratif ». Pour sa modélisation, nous avons choisi le modèle de MADS.

f- Couche « Découpage administratif »

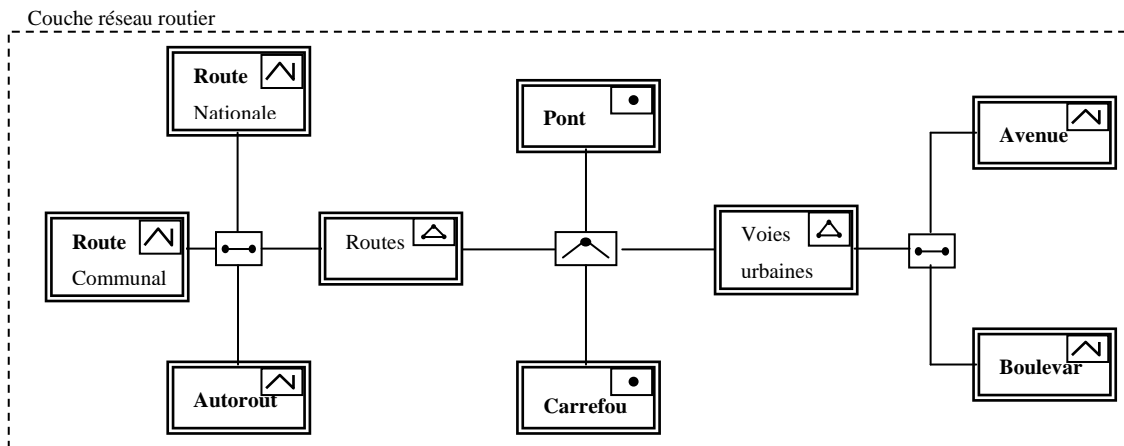


VI.4.3. Construction du schéma fédéré

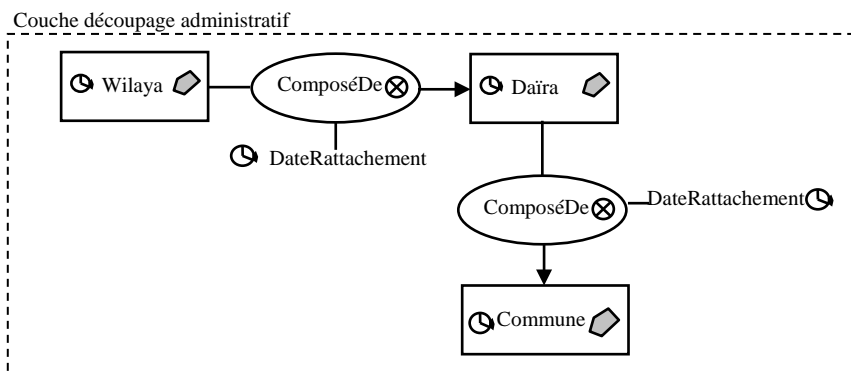
a- Intégration des fragments de modèles « Premier niveau »



Sous schéma1 : Intégration des fragments de modèles du groupe « spaghetti »



Sous schéma2 : Intégration des fragments de modèles du groupe « réseau »



Sous schéma3 : Intégration des fragments de modèles du groupe « topologie »

b- Fédération des sous-schémas « deuxième niveau » (Figure VI.16)

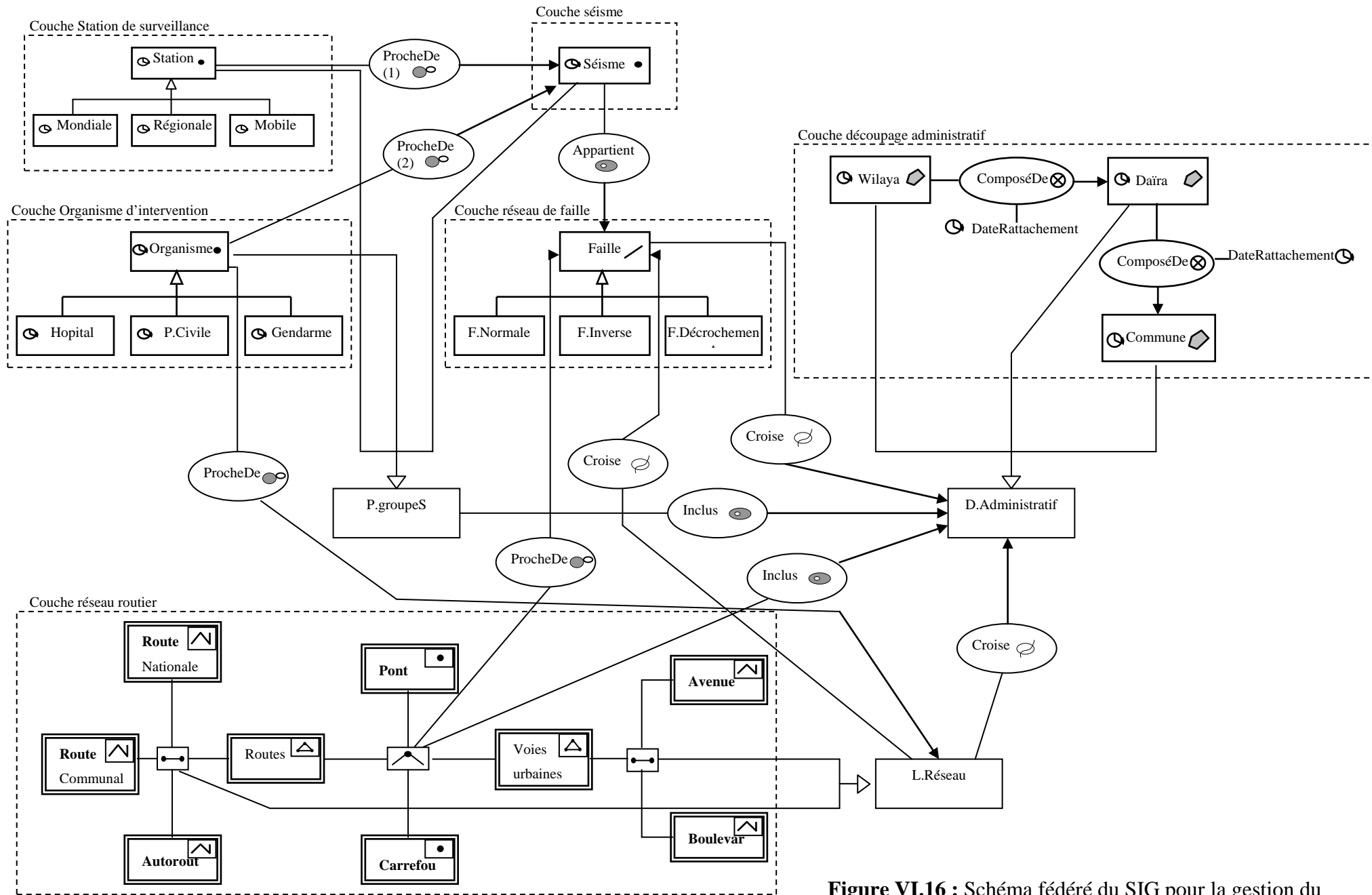


Figure VI.16 : Schéma fédéré du SIG pour la gestion du risque sismique

VI.5. Conclusion

Dans ce travail, nous avons présenté une approche de développement des systèmes d'information géographique (SIG) basée sur l'intégration des fragments de modèles. L'environnement de développement proposé, est composé de (03) grandes étapes :

- Une étape de décomposition du système en couches spatiales.
- La modélisation indépendante des couches spatiales en utilisant une variété de méthodes de conception des SIG. Ainsi, nous obtiendrons des fragments de modèles.
- L'intégration des fragments de modèles pour obtenir le schéma fédéré du système.

Cela est réalisée en deux niveaux :

- Intégration des fragments de modèles du même groupe de couches spatiales (spaghetti, réseau et topologie) pour former les sous-schémas. Pour cela, nous avons adapté des techniques développées dans l'ingénierie des systèmes d'information. Plus précisément, les méthodologies d'intégration, traduction des schémas, résolution des conflits, détection des correspondances inter-schémas,
- Fédération des sous-schémas obtenus pour former le schéma fédéré. Nous avons adopté une approche de fédération des systèmes hétérogènes sans schéma global. Le schémas obtenu donne une vision virtuel du schéma conceptuel global du système. Ce qui donne une ouverture à plusieurs travaux de recherches cités en perspectives.

L'intérêt de l'approche proposée repose sur le plan conceptuel :

- La décomposition du domaine d'application en couches spatiales, simplifie la modélisation du système.
- La méthodologie d'intégration proposée dans la troisième phase, permet d'intégrer dans le même schéma conceptuel des fragments de modèles déjà conçus.

**CONCLUSION
GENERALE**

CONCLUSION GENERALE

Les différents aspects proposés dans cette thèse montre que le développement des systèmes d'information géographique (SIG) est une tâche complexe. Toutes les méthodes de développement analysées, ne répondent pas de façon globale aux besoins des concepteurs des SIG.

Pour faciliter la tâche, nous avons proposé, dans ce travail, un environnement de développement des SIG qui guidera le concepteur dans sa phase d'analyse et de conception à schématiser une vue globale du schéma conceptuel d'une base de données.

L'approche proposée, consiste à retrouver les sous-systèmes qui composent un projet SIG. Généralement, ils sont définis par les couches spatiales du domaine d'application. La modélisation indépendante de ces couches utilisant diverses méthodes de conception, permettra d'obtenir des fragments de modèles. Ces derniers sont par la suite, intégrés pour former un schéma fédéré du système, en adoptant les approches de fédération des systèmes homogènes et hétérogènes.

Notre objectif reste à définir un environnement riche de techniques et de méthodes de conception des SI et SIG afin de faciliter le développement de tout projet SIG.

Les perspectives de notre travail sont nombreuses. Nous cherchons à :

- Développer des techniques de traduction des schémas permettant de faire intégrer des fragments de modèles existants.
- Améliorer la solution proposée après l'avoir appliqué sur divers domaines SIG.
- Rapprocher le travail effectué de l'approche Data Warehouse.
- Passer au-delà du conceptuel en proposant une architecture logique et physique spécifique à notre approche.
- Joindre à cette approche, un langage de requêtes spatio-temporel spécifique.
- Développer des outils AGLs assurant les (03) phases proposées dans notre approche.

BIBLIOGRAPHIE

BIBLIOGRAPHIE

- [1] « cours SIG vectoriels », http://62.241.132.4/ge/prog01-02/sigsenghor/htm/Cours%2004_%20SIG%20Vectoriels.htm
- [2] Cours, Chapitre VI, « Introduction aux systèmes d'information géographique », <http://www.cnerta.educagri.fr/infogeo/sigcarto/cours/courprot/courfr6.htm>
- [3] Cours, Chapitre II, « PRINCIPES GÉNÉRAUX », <http://www.cnerta.educagri.fr/infogeo/sigcarto/cours/courprot/courfr2.htm>
- [ABD01] N.Abdat, « GOMT : une technique de modélisation des SIG basée sur le modèle objet », Thèse de Magister Informatique, 2001
- [ADN01] M.El Adnani, A.Tikniouine : « Couches fonctionnelles et rôles : nouveaux concepts pour une modélisation cohérente et intégrale des informations géographiques », CYBERGEO n° 198, 2001.
- [ALI99] Z.Alimazighi : « Environnement de développement des systèmes d'information : une approche méthodologique basé sur l'analyse de documents et scénarios », Thèse de Doctorat d'état, USTHB, Mars 1999
- [ALL02] T.Allili, H.Beldjoudi, A.K. Yelles-Chaouche, "Conceptualisation d'un système d'information géographique pour le risque naturel : Application aux risques sismiques », Journées scientifiques de la technologie de l'information, Université de Mentouri Constantine, 13-18 Avril 2002.
- [ALL01] T.Allili, Z.Alimazighi, « Etude comparative des outils de développement des SIG », Mini-projet , Labo LSI, Institut Informatique, USTHB, Alger, 2001.
- [BAT86] C.Batini, M.Lenzerini, S.B.Narathe : « A Comparative Analysis of methodologies for database schema integration », ACM Computing Surviving, Vol 18, n°4, December 1986.
- [BEL00] H.Beldjoudi, A.K Yelles-Chaouche, A.About : « Le système d'information géographique au service de la réduction du risque sismique », CRAAG 2000.
- [BED00] Bédard.Y : « Modélisation de base de données spatiales avec UML étendu et Perspectory 2000 », Journée CASSINI du réseau français de chercheurs en SIG, La Rochelle, France, 8-10 Septembre 2000. <http://sirs.scg.ulaval.ca/yvanbedard/slideshow/publication/229/OGIS>.
-

- [BED96] Y.Bédard, Y.Cardon, Z.Maamar, B.Moulin, D.Valiere : « Adapting data models for the design of spatio-temporel databases, Computer, Environment and Urban systems », an International Journal, Vol 20, n°1, 1996.
- [BENA02] R.D.Benhacène, « Une démarche de développement de systèmes d'information par fragments de modèles », Thèse de Magister, 2002.
- [BEN99] Benslimane.D, Jouanot.F, Laurini.R, Yétongnon.K, Cullot.N, Savonnet.M : « Interopérabilité de SIG : un état de l'art », Revue Internationale de géomatique, 1999.
- [CAR 93] Caron.C, Bédard.Y, Gagnon.P : « MODUL-R : un formalisme adapté pour les SIRS », Revue de géomatique, vol3, n°3, 1993, p 283-306.
- [CAR91] C.Caron : “nouveau formalisme de modélisation conceptuelle adapté aux SIRS”, Mémoire de M.Sc, Département des sciences géomatique, Université Laval, Quebec, Canada, 1999.
- [CER01] « Modèle des relations géométriques »
[http:// www.certu.fr/sitart/geomat/minisio1/signatur/dossier16/dossier16.htm](http://www.certu.fr/sitart/geomat/minisio1/signatur/dossier16/dossier16.htm), 2001
- [CHA95] J-L.Chatelain, B.Guillier, M.Souris, E.Dupérier, H.Yepes, « SIG et évaluation des risques naturels : Application aux risque sismiques de QUITO », revue Mappmonde, n°3, 1995.
- [CLA98] C.Claramunt, « Un Modèle de Vue Spatiale pour une Représentation Flexible de Données Géographiques », Université de Bourgogne, 1998
http://etna.int-evry.fr/~adm_bd/Cigales/Publications/ Theses/
- [CLA97] Claramunt.C , Coulondre.S, Libourel.T : « Autour des méthodes orientées objet pour la conception des SIG », Revue internationale de géomatique, vol1,n°1,1997.
- [COA91] P.Coad, E.Yourdon : « Object-oriented analysis », Prentice Hall, 2nd edition 1991.
- [CON01] “Geographic Guideline”
[http:// www.conabio.gob.mx/sig_ingles/geographic_guideline.htm](http://www.conabio.gob.mx/sig_ingles/geographic_guideline.htm)
- [COS] « MODELISATION DE DONNEES GEOGRAPHIQUES : Utilisation d'un logiciel AGL pour la modélisation et la documentation d'applications SIG », COSIG, c/o Office fédéral de topographie, <http://www.cosig.ch>
- [ENG01] ”Introduction to ArcView”
[http:// www.engineering.usu.edu/classe/6400/hw4/hw4.htm](http://www.engineering.usu.edu/classe/6400/hw4/hw4.htm)
- [ESRI01] « Documentation technique »
[http:// www.esri.com](http://www.esri.com) / [http:// www.esrifrance.fr](http://www.esrifrance.fr)
- [ESRI98] “ERSI ShapeFile Technical Description”, An ERSI White paper July 1998
-

- [FRA97] H.Frank, J.Eder : “Integration of behaviour Models”, in Proceedings of ER’97, Workshop on behavioural Modelling and Design Transformation.
- [GAY97] O.Gayte, J-P.Cheylan, T.Libourel, S.Lardon :”conception de SI sur l’environnement”, collection géomatique, Ed Hermès, Paris, 1997.
- [GEO01] [http:// www.geoconcept.fr](http://www.geoconcept.fr)
- [GOT92] W.Gothard, P.C.Lackemann, A.Newfeld : “System guided view integration for object oriented database”, in IEEE Transaction on Knowledge and DataEngenning, 1992.
- [Heb96] B.S.Henderson, A.Bultuis : “The COMMA project : first draft for limited circulation”, Cottar Technical Report Sydney, 1996.
- [HOW] H.Howard, M.Hwang, H.Lin : “GIS Mapping of seismic hazards and seismic risk in Shelby country, Tennessee, USA”.
- [JIA93] Jia-Ling Koh, Arbee L.P.Chen : “Integration of heterogeneous object schemas”, In Proc. International Conference on Conceptual Modelling / the Entity Relationship Approach, 1993.
- [KAR00] Karla.A.V.Borges, Clodoveu A. Davis Jr, Alberto H. F. Laender : « OMT-G: An Object-Oriented Data Model for Geographic Applications », <http://www.pbh.gov.br/prodabel/cde/omt-g.html>, <http://www.omtg.hpg.com.br>, 2000.
- [KPS 97] ÖSTERS G., PAGEL B.U., SIX H.W: « GIS-application development with GeoOOA », International Journal of GIS, vol. 11(4), p. 307-335, 1997.
- [LAR02] N.Lardjane, « Intégration de modèles de procédés logiciels », Thèse de Magister Informatique, 2002.
- [LAU93] Laurini et Milleret-Raffort, « La structure d’un SIG », 1993
- [LAU93] R.Laurini, F.Milleret-Raffort : « Les bases de données en géomatique », Edition HERMES, Paris, 1993.
- [MAP01] [http:// www.mapinfo.com](http://www.mapinfo.com)
- [MAR95] P.Martin, W.Powley : “CORDS schema integration environment”, External technical report, dept of computing and information science, Queen’s University, Kingstom, Canada, 1995.
- [MIF01] “The MapInfo Interchange File (MIF) Format specification”, MapInfo Corporation, [http:// www.mapinfo.com](http://www.mapinfo.com)
- [PAN96] D.Pantazis, J-P.Donnay : “La conception de SIG- méthode et formalisme”, collection géomatique, Ed Hermès, Paris 1996.
-

- [PAR97] Parent.C, Spaccapietra.S, Zimanyi.E, Donini.P, Plazanet.C, Vangenot.C, Rognon.N, Grausaz.P.A : « MADS, un modèle conceptuel pour des applications spatio-temporelles », Revue Internationale de géomatique, vol7,n°3-4, 1997.
- [PAR96] Parent.C, Spaccapietra.S : « Intégration de base de données : Panorama des problèmes et des approches », Ingénierie des systèmes d'information Vol.4, N°3, 1996.
- [PAR 95] C.Parent, C.Claramunt, N.Rognon : « Concepts pour un modèle conceptuel spatial », rapport interne, Projet MADS, EPFL, 1995.
- [PER01] « Fondement des systèmes d'information géographique »
[http:// perso.club-internet.fr/giza/www/commentaires/fondsig.htm](http://perso.club-internet.fr/giza/www/commentaires/fondsig.htm),2001
- [PER01] « Nouvelle tendance dans les technologies de l'information spatiale »
[http:// www.perso.respublica.fr/gerssat/sig/gereravecunsig.htm](http://www.perso.respublica.fr/gerssat/sig/gereravecunsig.htm)
- [POU97] J.Pouliot, N.Rognon, Y.Bédard, F.Golay : « De la sélection à la mise en œuvre d'outils de conception pour les SIRS », Revue internationale de géomatique, Vol 7, n° 3-4/1997, p 259-277.
- [PRI01] « Base de données géographiques »
[http:// www.prism.uvsq.fr](http://www.prism.uvsq.fr), 2001
- [RES01] Cours : « Enjeux de la gestion de base de données par un SIG »,
<http://perso.respublica.fr/gerssat/gerera/gereravecensig.htm>, 2001.
- [SBG01] « Idrisi Tutorial on www »
[http:// www.sbg.ac.at/geo/idrisi/wwwtutor/tuthome.htm](http://www.sbg.ac.at/geo/idrisi/wwwtutor/tuthome.htm)
- [SCH96] Scholl.M, Voisard.A, Peloux.J.P, Raynal.L, Rigaux.P : « SGBD Géographiques, Spécificités », Ed International Thomson Publishing, France 1996.
- [SMA01] “An Object Oriented GIS, Issues and Solution
Discrete Geometry with seamless topology in a GIS.
AM/FM Data modelling for utilities
An overview of SmallWorld Magik
GIS Database are different
CASE Tool for GIS customisation”
[http:// www.SmallWorld.uk.com](http://www.SmallWorld.uk.com)
- [SPA92] S.Spaccapietra, C.Parent, Y.Dupont : “Model independent assertions for integration of heterogeneous schemas”, in VLDB Journal , p 81-126, 1992.
- [THE01] M.Thériault, E.Girard, F.Bégin : “La gestion d'un projet de SIG : aspects informatiques et institutionnels”, Département d'aménagement et de géographie, Université de Laval, Canada.
<http://62.241.132.4/ge/prog01-02/sigsenghor/htm/syllabus.htm>
-

- [THE01] M.Thériault, E.Girard, F.Bégin : “Concepts fondamentaux des SIG”, Département d’aménagement et de géographie, Université de Laval, Canada.
<http://62.241.132.4/ge/prog01-02/sigsenghor/htm/syllabus.htm>
- [THE99] M.Thériault, E.Girard, F.Bégin, « Les SGBD et la modélisation du territoire », Département de géographie, Université de laval, Canada, 1999
- [TIK01] A.Tikniouine, M.El Adnani : « Essai d’intégration des SIG à représentation multiple et des méthodes multicritères d’aides à la décision pour l’aménagement du territoire ».
- [UA01] “GIS in Fisheries management”
[http:// www.ua.es/copemed/reports/gis/maltacourse/day2.pdf](http://www.ua.es/copemed/reports/gis/maltacourse/day2.pdf)
- [UNI01] « Structure de données appliquées aux SIG »
[http:// www.unites.uqam.ca/dgeo/geo7530/stb2000.htm](http://www.unites.uqam.ca/dgeo/geo7530/stb2000.htm), 2001
- [UNI00] « La modélisation : représenter la réalité dans un système informatisé »,
<http://www.unites.uqam.ca/dgeo/geo7530/stb2000.htm>
- [VER01] H.Verjus, « Conception et construction de fédérations de progiciels », thèse de doctorat en Informatique, Université de SAVOIE (ESLA), 2001.
-

RESUME

Pour faciliter le développement des systèmes d'information géographique, les concepteurs proposent une approche de modélisation indépendante. Elle consiste à décomposer tout projet SIG en sous systèmes. Elle permettra ainsi, la réutilisation des données déjà existantes que ce soit des fichiers ou des bases de données indépendantes. Le développement de ces nouveaux systèmes repose désormais sur la capacité du système informatique à réaliser l'interopérabilité entre bases de données.

Dans cette étude, nous présentons une approche de développement d'un système d'information géographique basée sur l'intégration des fragments de modèles. La décomposition du système en couches spatiales permettra de retrouver l'ensemble des sous systèmes à développer. Nous aurons ainsi, la possibilité d'utiliser de diverses méthodes de conception et de réintégrer des fragments déjà existants.

Mots clés. Objet géographique, méthodes de conception, couche spatiale, interdépendances spatiales, fragment de modèle, Intégration, SIG.
