

Modélisation des Aspects Temporels dans les Bases de Données Spatiales

Mohammed MINOUT

Promoteur: Prof. Esteban ZIMÁNYI



Résumé

L'introduction du temps dans les bases de données classiques et spatiales apparaît de plus en plus, aujourd'hui, comme une nécessité pour une gestion optimale de l'historicité. En effet, les applications de bases de données spatio-temporelles sont présentes dans un grand nombre d'applications. Le besoin, par exemple, est de sauvegarder l'historique des géométries des parcelles dans le système d'information d'un plan cadastral, la prévention d'incendie dans le système de gestion forestière, le système de navigation des véhicules, etc. Cette historique des phénomènes permet de mieux comprendre ce qui s'est produit dans le passé de manière à éventuellement anticiper certaines évolutions futures.

Etant donné ces nouveaux besoins, cette thèse se focalise sur la modélisation et l'implantation des aspects temporels dans bases de données. En effet, la conception d'une application de base de données se fait par un enchaînement de trois phases (conceptuelle, logique et physique). Au niveau conceptuel, plusieurs modèles conceptuels ont été proposés intégrant les caractéristiques temporelles et spatiales.

Malheureusement, au niveau logique, les modèles de données des SGBD actuels n'offrent pas les concepts nécessaires pour implanter le modèle conceptuel spatio-temporel. Nous proposons donc de nouvelles règles de traductions d'un schéma conceptuel, basé sur le modèle MADS (Modélisation des Applications à des données spatio-temporelles), en un schéma logique MADSLog pour les modèles cibles à savoir : relationnel et relationnel-objet. Chaque règle transforme un concept structurel, temporel et spatial du modèle MADS en un ou plusieurs concepts supportés par la cible. Par exemple, la propriété spatiale définissant la géométrie d'un type d'objet est traduite par la création d'un nouvel attribut de type spatial dans ce type d'objet. Un outil CASE (Computer-Aided Software Engineering) appelé Schema Translateur est développé dans cette thèse implémentant toutes les règles de traductions.

La traduction de schémas conceptuels en schémas logiques peut impliquer une perte sémantique en raison de la différence de la puissance d'expression entre le modèle conceptuel et le modèle de données des SGBD existants. D'où la nécessité de générer un ensemble de contraintes d'intégrité afin de préserver la sémantique définie dans le schéma conceptuel. Ces contraintes sont exprimées à ce niveau par des formules logiques.

Avec l'apparition de GML (Geographic Markup Language) qui est conçu pour la modélisation, le transport et le stockage d'informations géographiques. Nous transformons également le schéma conceptuel MADS en GML. De nouveaux schémas GML temporel et spatial sont définis qui peuvent être employés par n'importe quelle application de base de données spatio-temporelle.

Au niveau physique, nous proposons une méthode d'adaptation du schéma logique en schéma physique pour le modèle relationnel-objet. Elle permet de définir les tables, les types abstraits, les types d'objets, les domaines, etc. Notre proposition permet aussi la génération des contraintes d'intégrité au niveau physique. En effet, chaque contrainte d'intégrité (struc-

turelle, temporelle ou spatiale) qui est définie en calcul logique est exprimée soit directement par des contraintes déclaratives ou soit par des déclencheurs du SGBD choisi. Les déclencheurs spatiaux sont fondés sur les fonctionnalités prédéfinies dans Oracle, alors que les déclencheurs temporels sont basés sur les opérateurs et méthodes appliquées sur les types temporels.

Enfin, la traduction de requêtes est une deuxième clef de cette recherche. Le but de la traduction de requêtes, exprimées en algèbre, étant de reconstituer l'information au sens MADS à partir de la base de données stockées dans le SGDB cible. Elle permet de traduire les expressions algébriques MADS, qui sont définies sur le schéma conceptuel et non sur le schéma physique, en requêtes opérationnelles qui peuvent être exécutées sur une base de données spatiale et temporelle sous un SGBD ou un SIG.

Remerciement

L'accomplissement d'une thèse n'est évidemment pas le fruit d'un travail individuel mais constitue le résultat de nombreuses collaborations et discussions. Je profite donc de la fin de ce doctorat, et du début de ce texte, pour remercier ceux qui y ont contribué.

J'aimerais premièrement témoigner de ma reconnaissance à Esteban Zimányi, mon directeur de thèse, pour m'avoir donné l'opportunité de travailler sur un sujet passionnant dans des conditions matérielles et intellectuelles très favorables. Son ouverture d'esprit, son soutien, sa vision et sa détermination sont des caractéristiques qui m'ont non seulement marquée mais qui ont été très profitables à l'accomplissement de cette recherche.

J'aimerais deuxièmement remercier les Professeurs Roel Wuyts et Gianluca Bontempi pour leur présence au sein de mon comité d'accompagnement et pour les nombreuses suggestions et remarques fournies. En outre, je remercie grandement les Professeurs Christelle Vangenot et Yves De Smet d'avoir accepté de participer à mon jury de thèse.

Ma reconnaissance s'adresse aussi aux membres de mon service plus particulièrement aux J.M. Dricot, S. Boucher, J.P. Norguet, F. Servais, B. Verahegen, N. Vander Heyden, E. Malinowski et J.T. shele. Nos discussions, échanges d'idées et exposés m'ont apporté bon nombre de suggestions. A toutes les personnes rencontrées au cours de ce travail dont les questions, les remarques et les conseils furent une grande source d'inspiration et de découvertes.

Enfin, j'aimerais remercier ma famille qui m'a supporté durant la dernière ligne droite de cette thèse et tout particulièrement à Dounia qui m'a fortement conseillé et soutenu depuis la naissance de ce travail.

Table des matières

1	Introduction	1
1.1	Introduction	1
1.2	Contexte	2
1.2.1	Phases principales de la modélisation de données	2
1.2.2	Aspects spatiaux	4
1.2.3	Aspects temporels	5
1.2.4	Interrogation et visualisation des bases de données spatio-temporelles	6
1.3	Motivations et objectifs de la thèse	6
1.3.1	Motivations	6
1.3.2	Objectifs	8
1.4	Principales contributions	9
1.4.1	Traduction de schémas conceptuels	9
1.4.2	Représentation de schémas logiques en relationnel-objet	10
1.4.3	Représentation de schémas conceptuels en GML	10
1.4.4	Traduction de requêtes algébriques MADS en SQL	11
1.5	Organisation de la thèse	12
2	Préliminaires & état de l’art	13
2.1	Introduction	13
2.2	Modélisation conceptuelle de données spatiales et temporelles	13
2.2.1	Approche entité-association	14
2.2.2	Approche orienté objet	15
2.2.3	Approche XML	16
2.2.4	Modèle MADS	17
2.2.5	Comparaison entre MADS, STER et Perceptory	30
2.3	Implémentation de la temporalité et de la spatialité	31
2.3.1	Temporalité	31
2.3.2	Spatialité	34
2.4	Conclusion	37
3	Du modèle MADS vers un modèle MADSLog	39
3.1	Introduction	39
3.2	Processus de traduction de schémas	40
3.2.1	Définition d’une transformation	40
3.2.2	Sémantique d’une transformation	43
3.3	Règles de transformations structurelles	43

3.3.1	Attribut multivalué	44
3.3.2	Attribut complexe	45
3.3.3	Domaine énuméré ou intervalle	46
3.3.4	Génération d'identifiant	47
3.3.5	Type d'association	48
3.3.6	Lien de généralisation	54
3.4	Règles de transformations temporelles	55
3.4.1	Attributs temporels	55
3.4.2	Attribut variable dans le temps	59
3.4.3	Cycle de vie d'un type d'objet	61
3.4.4	Cardinalités temporelles	63
3.4.5	Type d'association de synchronisation	65
3.5	Règles de transformations spatiales	68
3.5.1	Attribut spatial	68
3.5.2	Type d'objet spatial	71
3.5.3	Type de données spatial spécialisé	71
3.5.4	Attribut variable dans l'espace	72
3.5.5	Attribut variable dans l'espace et dans le temps	74
3.5.6	Type d'objet spatial variable dans le temps	76
3.5.7	Type d'association topologique	78
3.6	Conclusion	80
4	Traduction du schéma MADSLLog en relationnel-objet	83
4.1	Introduction	83
4.2	Modèle relationnel-objet	84
4.2.1	SQL:2003	84
4.2.2	Oracle 10g	85
4.3	Directives de réécriture	86
4.3.1	Domaine structuré	87
4.3.2	Type d'objet	88
4.3.3	Type d'association	89
4.3.4	Application d'une partie de réseau fluvial	92
4.4	Préservation sémantique	98
4.4.1	Contraintes d'intégrité structurelles	98
4.4.2	Contraintes d'intégrité temporelles	105
4.4.3	Contraintes d'intégrité spatiales	115
4.5	Conclusion	120
5	Représentation du schéma conceptuel MADSL en GML	123
5.1	Introduction	123
5.2	Concepts de base en GML	124
5.2.1	Entité géographique	124
5.2.2	Propriétés des entités	124
5.2.3	Schéma XML	125
5.2.4	Noyau des schémas GML	126
5.3	Conversion du schéma MADSL en schéma GML	126
5.3.1	Schéma GML pour la dimension thématique	128

5.3.2	Schéma GML pour la dimension temporelle	133
5.3.3	Schéma GML pour la dimension spatiale	136
5.4	Conclusion	140
6	Interrogation de bases de données	141
6.1	Introduction	141
6.2	Architecture de l'éditeur de requêtes	141
6.3	Algèbre MADS	142
6.3.1	Opérations algébriques unaires	144
6.3.2	Opérations algébriques binaires	146
6.4	Règles de traductions de requêtes	146
6.4.1	Formalisation théorique	148
6.4.2	Attribut multivalué	149
6.4.3	Attribut complexe	150
6.4.4	Type d'association binaire	151
6.4.5	Cycle de vie	153
6.4.6	Attribut variable dans le temps	154
6.4.7	Mécanisme de transformation	155
6.5	Transcripteur GML	156
6.5.1	Architecture du transcripteur GML	156
6.5.2	Conception du document GML	157
6.6	Conclusion	158
7	Implémentation	161
7.1	Architecture de l'éditeur de schémas MADS	161
7.2	Technologies	164
7.2.1	XML	164
7.2.2	Définition de type de données	164
7.2.3	Modèle des objets des documents	164
7.2.4	Langage de programmation	165
7.2.5	Oracle	165
7.3	Traducteur de schémas	166
7.3.1	Architecture	166
7.3.2	Schéma conceptuel MADS en XML	166
7.3.3	Schéma logique MADSLog d'Oracle en XML	172
7.3.4	Schéma logique MADSLog d'Oracle en XML	173
7.3.5	Packages du traducteur	176
7.3.6	Algorithme de traduction	178
7.4	Conclusion	179
8	Conclusion et perspectives	181
A	DTD pour le modèle conceptuel MADS et logique MADSLog	185
A.1	DTD pour le modèle conceptuel	185
A.2	DTD pour le modèle logique	190
A.2.1	DTD pour Oracle	190
A.2.2	DTD pour Access-ArcView	192

B	Implémentation des contraintes d'intégrité	195
B.1	Déclencheurs pour type d'association plusieurs à plusieurs	195
B.2	Déclencheurs pour cycle de vie	195
B.3	Déclencheurs pour attribut variable dans le temps	196
B.4	Déclencheurs pour types d'association de synchronisation	197
B.4.1	Opérateur précède	197
B.4.2	Opérateur succède	198
B.4.3	Opérateur égale	198
B.4.4	Opérateur rencontre	199
B.4.5	Opérateur pendant	199
B.4.6	Opérateur commence	200
B.4.7	Opérateur finit	200
C	Implémentation des méthodes temporelles du TDT	203
C.1	Définition des méthodes sur les types ttemporels MADS	203
C.2	Définition des méthodes définies sur le cycle de vie	205
C.3	Implémentation dans Oracle	205
C.4	Type variable dans le temps	207
C.5	Cycle de vie du type d'objet ou d'association	209
D	Validation du document GML	211
E	DTD pour l'algèbre MADS	215

Table des figures

1.1	Les phases principales de conception de bases de données.	3
2.1	Représentation temporelle non groupée de l'historique des employés	16
2.2	Représentation temporelle groupée de l'historique des employés	17
2.3	Hierarchie de types de données temporels.	20
2.4	Attributs temporels.	21
2.5	Types d'objets temporels.	22
2.6	Trois exemples généralisation temporelle.	22
2.7	Type d'association temporel.	23
2.8	Hierarchie des types de données spatiaux.	25
2.9	Attribut profondeur modélisé avec et sans attribut variable.	25
2.10	Deux exemples de généralisation spatiale.	26
2.11	Raffinement et redéfinition du type spatial.	27
2.12	Spécialisation multiple.	27
2.13	Exemple de type d'association topologique.	28
2.14	Exemple de type d'association d'aggrégation spatiale.	29
2.15	Le modèle spatial d'Oracle.	36
3.1	Transformation de l'attribut autresNoms en modèle relationnel.	44
3.2	Transformation de l'attribut autresNoms en modèle relationnel-objet.	44
3.3	Transformation de l'attribut adresse en modèle relationnel.	45
3.4	Transformation des attributs complexes et multivalués Ports et Bateaux en modèle relationnel.	46
3.5	Transformation de l'attribut adresse en modèle relationnel-objet.	46
3.6	Transformation des attributs complexes et multivalués Ports et Bateaux en modèle relationnel-objet.	46
3.7	Transformation du domaine énuméré typeOccupation.	47
3.8	Génération de l'identifiant en relationnel.	47
3.9	Génération de l'identifiant en relationnel-objet.	48
3.10	Transformation de l'association ContrôléPar en modèle relationnel.	49
3.11	Transformation de l'association ContrôléPar en modèle relationnel-objet.	49
3.12	Transformation de l'association binaire plusieurs à plusieurs Déborde en modèle relationnel.	50
3.13	Transformation de l'association binaire plusieurs à plusieurs Déborde en modèle relationnel-objet.	51
3.14	Transformation de l'association Achète en modèle relationnel.	52

3.15	Transformation de l'association Achète en modèle relationnel-objet.	53
3.16	Transformation du lien de généralisation.	54
3.17	Transformation de l'attribut débit de type Interval en modèle relationnel. . . .	55
3.18	Transformation de l'attribut débit de type Interval en modèle relationnel-objet.	56
3.19	Transformation de l'attribut débit de type IntervalBag et InstantBag en modèle relationnel.	56
3.20	Transformation de l'attribut débit de type IntervalBag et InstantBag en modèle relationnel-objet.	57
3.21	Transformation de l'attribut débit de type ComplexTime(a) SimpleTime (b). . .	57
3.22	Transformation de l'attribut débit de type Time.	58
3.23	Transformation de l'attribut mesure variable dans le temps en modèle relationnel.	59
3.24	Transformation de l'attribut mesure variable dans le temps en modèle relationnel-objet.	60
3.25	Transformation du cycle de vie du type d'objet Station en modèle relationnel.	61
3.26	Transformation du cycle de vie de type d'objet Station en modèle relationnel-objet.	62
3.27	Transformation du type d'association temporelle Entretien en modèle relationnel.	64
3.28	Transformation du type d'association temporelle Entretien en modèle relationnel-objet.	64
3.29	Transformation du type d'association de synchronisation Détruit en modèle relationnel.	66
3.30	Transformation du type d'association de synchronisation Détruit en modèle relationnel-objet.	66
3.31	Transformation de l'attribut spatial geometry de type spatial SurfaceBag. . . .	68
3.32	Transformation de l'attribut spatial geometry de type PointBag.	69
3.33	Transformation de l'attribut spatial geometry de type LineBag.	69
3.34	Transformation de l'attribut spatial geometry de type Surface.	69
3.35	Transformation de l'attribut spatial geometry de type spatial ComplexGeo. . . .	70
3.36	Transformation de l'attribut spatial geometry de type spatial SimpleGeo). . . .	70
3.37	Transformation de l'attribut spatial geometry de type spatial Geo.	70
3.38	Transformation du type d'objet spatial Lac en un type d'objet non spatial. . .	71
3.39	Transformation de la spatialité spécialisée du type d'objet Lac en type générique Geo.	71
3.40	Transformation de l'attribut profondeur variable dans l'espace en modèle relationnel.	72
3.41	Transformation de l'attribut profondeur variable dans l'espace en modèle relationnel-objet.	73
3.42	Transformation de l'attribut profondeur variable dans l'espace et dans le temps en modèle relationnel.	74
3.43	Transformation de l'attribut profondeur variable dans l'espace et dans le temps en modèle relationnel-objet.	75
3.44	Transformation du type d'objet spatial Rivière variable dans le temps en modèle relationnel.	77
3.45	Transformation du type d'objet spatial Rivière variable dans le temps en modèle relationnel-objet.	77
3.46	Transformation de l'association topologique Déborde en modèle relationnel. . .	78
3.47	Transformation de l'association topologique Déborde en modèle relationnel-objet.	79

4.1	Type structuré utilisé comme valeur et comme type d'objet en SQL:2003. . .	85
4.2	Type structuré utilisé comme valeur et comme type d'objet en Oracle 10g. . .	86
4.3	Transformation de l'attribut autresNoms MADSLLog en SQL:2003 et en Oracle 10g.	87
4.4	Transformation de l'attribut adresse MADSLLog en SQL:2003 et en Oracle 10g. . .	88
4.5	Transformation d'un type d'objet MADSLLog en SQL:2003 et en Oracle 10g. .	89
4.6	Transformation d'un type d'association un à plusieurs MADSLLog en SQL:2003 et Oracle 10g.	90
4.7	Transformation d'un type d'association plusieurs à plusieurs MADSLLog en SQL:2003 et Oracle 10g.	91
4.8	Transformation d'une généralisation en SQL:2003 et Oracle 10g.	92
4.9	Schéma MADS d'une application de gestion d'un réseau fluvial.	93
4.10	Transcription du MADSLLog en modèle relationnel-objet.	95
4.11	Transformation de l'attribut autresNoms en modèle relationnel-objet.	98
4.12	Transformation du domaine énuméré typeOccupation	99
4.13	Génération de l'identifiant en relationnel-objet.	100
4.14	Transformation de l'association ContrôléPar en modèle relationnel-objet. . . .	100
4.15	Transformation de l'attribut débit de type Interval en modèle relationnel-objet.	106
4.16	Transformation de l'attribut débit de type ComplexTime (a) SimpleTime (b). .	107
4.17	Transformation finale de l'attribut débit de type Time	108
4.18	Transformation de l'attribut mesure variable dans le temps en modèle relationnel-objet.	109
4.19	Transformation du cycle de vie de type d'objet Station	111
4.20	Transformation du type d'association temporelle Entretient en modèle relationnel-objet.	112
4.21	Transformation du type d'association de synchronisation Détruit en modèle relationnel-objet.	114
4.22	Transformation de l'association topologique Déborde en modèle relationnel-objet.	120
5.1	Hierarchie de différents schémas xsd en GML.	127
5.2	Schéma MADS d'une application de gestion d'un réseau fluvial.	128
5.3	Types de données de données temporels (TDT) en MADS.	134
6.1	Architecture de l'éditeur de requêtes.	142
6.2	Schéma MADS d'application de gestion fluviale.	144
6.3	Traduction de MADSLLog en modèle relationnel-objet.	147
6.4	Architecture du transcripateur GML.	156
6.5	Résultat de la requête.	157
7.1	Architecture de l'éditeur de schémas MADS	162
7.2	Interface graphique de l'éditeur de schémas	163
7.3	Architecture du traducteur de schémas	167
7.4	Schéma MADS d'application de gestion fluvial.	168
7.5	La hiérarchie des packages du traducteur de schémas	176
7.6	Interface graphique du traducteur de schémas	177
7.7	Diagramme de classes simplifié du package multirepresentationrules	178
7.8	Transformation de l'arbre pour l'élément période	179

C.1 Modèle logique d’un type variable dans le temps 207

Liste des tableaux

2.1	Types de relation topologique.	29
2.2	Comparaison de sémantiques spatiales et temporelles.	30
3.1	Table d'équivalence entre MADS et MADSLog (Dimension thématique). . . .	41
3.2	Table d'équivalence entre MADS et MADSLog (Dimension temporelle). . . .	42
3.3	Table d'équivalence entre MADS et MADSLog (Dimension spatiale).	43
4.1	Directives pour les bases de données relationnel-objet.	87
5.1	Table de correspondance entre MADS et GML.	137