

THESE

présentée par

José Valdeni DE LIMA

pour obtenir le titre de DOCTEUR

de L'UNIVERSITE Joseph FOURIER (GRENOBLE I)
(arrêté ministériel du 5 juillet 1984)

Spécialité : Informatique

**GESTION D'OBJETS COMPOSES DANS UN SGBD :
CAS PARTICULIER DES DOCUMENTS
STRUCTURES**

Thèse soutenue le 20 mars 1990 devant la commission d'examen.

Président	:	Y. CHIARAMELLA
Rapporteur	:	J. KOULOUMDJIAN
Rapporteur	:	C. CHRISMENT
Examineurs	:	M. LOPEZ
		M. ADIBA
		J.P. GIRAUDIN
		V. QUINT

Thèse préparée au sein du Laboratoire de Génie Informatique

BIBLIOTHEQUE DU CERIST

89/1752

RESUME

Cette thèse traite du problème de la gestion des documents structurés multimédia dans un SGBD. Par gestion, nous entendons la modélisation, la manipulation, le stockage et l'accès aux documents. Nous présentons un modèle de Documents Structurés de Bureau (DSB) et une algèbre associée pour réaliser la spécification précise des aspects fonctionnels : opérateurs de construction et restructuration des objets manipulés et fonctions d'accès. Le stockage et l'accès sont implémentés au niveau fonctionnel sous forme d'opérations sur des documents en prenant en considération leurs structures logiques. Le couplage du modèle standard ODA au modèle DSB et l'intégration au niveau fonctionnel des opérations implémentées ont permis la mise en place d'un gestionnaire autonome de documents utilisable à partir d'un SGBD relationnel. Ce gestionnaire de documents permet la spécialisation des documents et l'utilisation de valeurs nulles. Une grande partie de ce travail a été réalisée dans le cadre du projet ESPRIT DOEOIS et un prototype expérimental a été développé sur ORACLE.

MOTS CLES

Bases de Données, Modèles de Données, Méthodes de Stockage et d'Accès, Objets Complexes, Documents Structurés, Structure Logique, Spécialisation, Valeurs Nulles, Documents Textuels.

TABLE DES MATIERES

INTRODUCTION	9
CHAPITRE 1	21
OBJETS COMPLEXES : ETAT DE L'ART	23
1.1 Bases de données et Objets Complexes	24
1.2 Concepts de base	29
1.3 Modélisation d'objets complexes documents	33
1.4 Caractérisation des modèles	42
1.5 Classement des modèles	45
1.6 Systèmes de Gestion de Bases de Données Documents (SGBDD)	47
CHAPITRE 2	51
MODELE ET ALGEBRE POUR LES DOCUMENTS	
STRUCTURES DE BUREAU	53
2.1 Introduction	53
2.2 Concepts fondamentaux	55
2.3 Les opérateurs primitifs	57
2.4 Objets de l'algèbre	59
2.5 Différences entre l'algèbre formelle et l'algèbre au niveau utilisateur	60
2.6 Caractéristiques	62
2.7 Modèle descriptif associé à l'algèbre	66
2.8 Opérateurs applicables aux documents	68
2.9 Traitement des valeurs nulles	79
2.10 Association de la sémantique aux valeurs nulles	81
2.11 Représentation machine adéquate	83
2.12 Extension de l'algèbre associée au modèle	84
2.13 Conclusions	85

CHAPITRE 3	87
STRUCTURES LOGIQUES DES DOCUMENTS	89
3.1 Introduction	89
3.2 Choix des unités documentaires	92
3.3 Mécanismes de typage des documents	96
3.4 Contraintes sur la structure générique des documents	97
3.5 Conclusions	117
CHAPITRE 4	119
STOCKAGE ET ACCES AUX OBJETS COMPLEXES DOCUMENTS	121
4.1 Introduction	121
4.2 Caractéristiques des traitements des documents	123
4.3 Eléments clefs de la conception physique d'une base de documents	132
4.4 Méthodes d'accès aux documents structurés	134
4.5 Méthodes de stockage des documents structurés	137
4.6 Conclusions	156
CHAPITRE 5	159
REALISATION	161
5.1 Introduction	161
5.2 Serveur d'Information Bureautique	163
5.3 Gestionnaire de Documents Autonome (GDA)	166
5.4 Services offerts	169
5.5 Le gestionnaire relationnel GESDOC	174
5.6 Conclusions	175
CONCLUSIONS	177
BIBLIOGRAPHIE	187
ANNEXE 1 : Couplage des modèles ODA et FM	209

ANNEXE 2 : Relations de stockage 219

ANNEXE 3 : Interface fonctionnelle du gestionnaire de
documents autonome 229