

Saïd Benaïchou

Bases de données et systèmes d'information

Modèle relationnel, SQL,
optimisation des requêtes,
transactions, modélisation des données



La côte de l'ouvrage : 2-005-916

Table des matières

PARTIE 1 LE MODELE RELATIONNEL	19
Chapitre 1 Les bases de données	21
1 Introduction	21
2 La notion de base de données	22
3 Le SGBD	24
4 Modèle de données et type de SGBD	26
5 Les modèles de la première génération	27
5.1 Le modèle hiérarchique	27
5.2 Le modèle réseau	30
6 L'architecture ANSI/X3/SPARC	34
7 Les objectifs d'un SGBD	36
8 Architectures matérielles pour les bases de données	37
8.1 L'architecture de première génération	37
8.2 L'architecture client/serveur à deux niveaux	38
8.3 L'architecture client/serveur à trois niveaux	39
8.4 L'architecture distribuée	40
9 La base de données support	41
Exercices	42
Chapitre 2 Notions de base sur les bases de données relationnelles	51
1 Introduction	51
2 Le modèle relationnel	52
2.1 Schéma de relation, attribut et domaine	52
2.2 Tuple et relation	53
2.3 Clés candidates et clé primaire	54
2.4 Les clés étrangères	55
3 La notion de base de données relationnelle	56
3.1 Le schéma de base de données relationnelle	56

3.2 La base de données relationnelle	57
4 Les contraintes d'intégrité	58
4.1 Contrainte sur un attribut	59
4.2 Contrainte avec plusieurs attributs	59
4.3 La contrainte d'entité	60
4.4 La contrainte d'intégrité référentielle	60
4.5 Contraintes dynamiques	62
5 L'algèbre relationnelle	62
5.1 Les opérateurs classiques	62
<i>a) L'union, l'intersection et la différence</i>	62
<i>b) Le produit cartésien</i>	64
5.2 Les opérateurs unaires	65
<i>a) La projection</i>	65
<i>b) La restriction</i>	65
5.3 La jointure	66
5.4 La jointure naturelle	67
5.5 Jointure de plusieurs relations	68
5.6 Jointure externe	69
5.7 Semi-jointure	70
5.8 La division	71
6 Formulation des requêtes avec l'algèbre relationnelle	71
6.1 Requêtes avec simple extraction de données	71
6.2 Requêtes avec calculs	72
<i>a) Calculs en lignes</i>	73
<i>b) Calculs en colonnes</i>	73
<i>c) Combinaison des calculs en lignes et en colonnes</i>	74
<i>d) Calculs par groupe de lignes</i>	75
7 Le calcul relationnel	76
7.1 Le calcul des prédicats du premier ordre (rappel)	76
7.2 Le calcul de domaines	78
7.3 Le calcul des tuples	79
7.4 Le langage QBE	80
Exercices	83
Chapitre 3 Les dépendances	87
1 Introduction	87
2 Les dépendances fonctionnelles	88
2.1 La notion de dépendance fonctionnelle	88
2.2 Fermeture d'un ensemble de dépendances fonctionnelles	90

Table des matières	9
2.3 Règles d'inférence	91
2.4 Fermeture d'un sous-ensemble d'attributs	94
2.5 La recherche des clés candidates	95
<i>a) Caractérisation des clés</i>	95
<i>b) Recherche d'une clé candidate</i>	96
<i>c) Recherche de toutes les clés candidates</i>	97
2.6 Couverture minimale	98
2.7 Graphe des dépendances fonctionnelles	100
3 Les dépendances multivaluées	100
3.1 La notion de dépendance multivaluée	100
3.2 Règles d'inférence et fermeture	103
3.3 Dépendance multivaluée imbriquée	104
4 Dépendance de jointure	106
5 Fermeture d'un ensemble de dépendances	109
Exercices	112
Chapitre 4 La normalisation	117
1 Introduction	117
2 Décomposition sans perte d'information	118
2.1 Décomposition d'un schéma de relation	119
2.2 La notion de décomposition sans perte	119
3 Méthodes de décomposition d'un schéma de relation	121
3.1 Décomposition avec une dépendance fonctionnelle	122
3.2 Décomposition avec une dépendance multivaluée	123
3.3 Décomposition avec une dépendance de jointure	124
4 Objectif de la décomposition	126
5 Décomposition avec préservation des dépendances	127
6 Formes normales et algorithmes de décomposition	130
6.1 La première forme normale	131
6.2 Deuxième et troisième forme normale	132
6.3 Forme normale de Boyce/Codd	134
6.4 Les algorithmes de décomposition en BCNF et en 3NF avec préservation des dépendances	136
<i>a) L'algorithme de décomposition</i>	137
<i>b) L'algorithme de synthèse</i>	138
6.5 Quatrième forme normale	140
6.6 Cinquième forme normale	142
Exercices	144

PARTIE 2 LE LANGAGE SQL	147
Chapitre 5 Notions de base du langage SQL2	149
1 Introduction	149
2 Les types de données	150
2.1 Types pour les chaînes de caractères	151
2.2 Types numériques	151
2.3 Types pour les données temporelles	152
<i>a) Les dates et heures</i>	152
<i>b) Les durées</i>	153
<i>c) Opérations autorisées</i>	154
2.4 Les chaînes de bits	154
2.5 Les domaines	155
2.6 Types de données supplémentaires	156
3 Création d'une table	156
3.1 L'instruction de base	157
3.2 Contraintes sur les colonnes	159
<i>a) Valeur NULL et contrainte NOT NULL</i>	159
<i>b) Valeur par défaut</i>	161
<i>c) La clé primaire</i>	161
3.3 Modification de la structure d'une table	162
4 Les requêtes	163
4.1 Requêtes élémentaires	163
4.2 Requêtes avec critère de sélection	165
5 Les prédicats IN, BETWEEN, LIKE et IS NULL	168
6 Les noms de corrélation	171
7 Le tri	172
8 Traitement des expressions	173
8.1 Expressions dans la liste de projection	173
8.2 Evaluation des expressions en présence des valeurs NULL	175
<i>a) Expressions de type prédéfini</i>	175
<i>b) Expressions logiques</i>	176
<i>c) Les prédicats IS [NOT] TRUE, IS [NOT] FALSE et IS [NOT] UNKNOWN</i>	177
9 Fonctions SQL2	178
9.1 Fonctions pour les chaînes de caractères	178
9.2 Les données temporelles	180
9.3 Fonctions pour les valeurs NULL	182
10 Les opérateurs de conversion et de choix	183
10.1 L'opérateur CAST	184

Table des matières	11
10.2 L'opérateur CASE	185
11 Comparaison de lignes	186
11.1 Les opérateurs relationnels	186
11.2 Le prédicat IS NULL	188
12 Les vues	189
13 La pratique avec le système	190
13.1 Le schéma	190
13.2 La connexion	191
13.3 Catalogue et dictionnaire des données	192
Exercices	194
Chapitre 6 Requêtes élaborées	201
1 Introduction	201
2 Requêtes avec jointure	202
2.1 Jointure SQL/89	202
<i>a) Jointure de deux tables</i>	202
<i>b) Jointure de plusieurs tables</i>	204
<i>c) Autojointure</i>	205
<i>d) Le produit cartésien</i>	206
<i>e) La jointure naturelle</i>	206
2.2 Jointure dans SQL2	207
<i>a) La θ-jointure</i>	207
<i>b) La jointure naturelle</i>	208
<i>c) La jointure externe</i>	209
3 Les opérateurs union, intersection et différence	210
4 Calcul statistique	212
4.1 Fonctions d'agrégat	213
4.2 GROUP BY, HAVING et SELECT complet	214
4.3 Fonctionnalités OLAP	218
5 Sous-requête	221
5.1 Sous-requête retournant une seule valeur	221
5.2 Sous-requête retournant une table	222
5.3 Sous-requête corrélée	222
5.4 Imbrication des sous-requêtes	224
6 Les prédicats ALL, ANY, SOME, IN, MATCH, EXISTS et UNIQUE	224
7 Formulation avancée	228
8 Requête intermédiaire et récursivité	229
Exercices	230

Chapitre 7 Mises à jour, intégrité et sécurité des données	237
1 Introduction	237
2 Les mises à jour	238
2.1 L'insertion	238
<i>a) Option VALUES</i>	238
<i>b) Option VALUES pour plusieurs lignes</i>	241
<i>c) Option requête</i>	241
2.2 La modification	242
2.3 La suppression	243
2.4 La mise à jour à travers les vues	244
3 Les contraintes d'intégrité	246
3.1 Contraintes dans les tables	246
<i>a) La contrainte d'entité</i>	247
<i>b) Les clés candidates</i>	248
<i>c) La contrainte d'intégrité référentielle</i>	248
<i>d) Contrainte sur les lignes de tables</i>	251
<i>e) Cas d'une table existante</i>	252
3.2 Les assertions	253
3.3 Contraintes de domaines	255
4 La sécurité des données	257
4.1 Concepts de base	258
<i>a) Techniques disponibles</i>	258
<i>b) Contrôle d'accès au système</i>	258
4.2 Privilèges dans SQL2	259
4.3 Octroi de privilèges	260
4.4 Révocation des privilèges	262
Exercices	263
 PARTIE 3 TECHNIQUES AVANCÉES	 267
 Chapitre 8 Méthodes pour une évaluation optimisée des requêtes	 269
1 Introduction	269
2 Arbre de requête	271
3 Principe d'évaluation du coût d'un arbre	272
4 Coût en écriture d'un opérateur	274
5 Implémentation des opérateurs	276

Table des matières	13
5.1 La jointure	276
<i>a) Algorithme des boucles imbriquées</i>	277
<i>b) Algorithme de tri/fusion</i>	277
<i>c) Algorithme de tri/fusion avec clé candidate</i>	279
<i>d) Jointure avec index</i>	279
5.2 La restriction	280
<i>a) Restriction sans index</i>	281
<i>b) Restriction avec index</i>	282
5.3 La projection	282
6 Méthodes pour l'optimisation	283
7 Optimisation algébrique	284
8 L'ordonnancement des jointures	287
9 Exemple de recherche d'un plan optimal	289
10 Techniques pour la performance	294
10.1 Augmentation de la mémoire tampon	294
10.2 La méthode pipeline	294
10.3 Exécution en parallèle des opérateurs	295
11 Optimisation dans les systèmes distribués	296
11.1 La fragmentation des données	297
<i>a) La fragmentation verticale</i>	297
<i>b) La fragmentation horizontale</i>	298
<i>c) La fragmentation mixte</i>	299
11.2 La réplication des données	299
Exercices	300
Chapitre 9 Les transactions	307
1 Introduction	307
2 La notion de transaction	308
3 Le problème des transactions concurrentes	309
3.1 La perte d'une mise à jour	310
3.2 La lecture non reproductible	310
3.3 La lecture impropre	311
3.4 Le problème de la contrainte déjouée	311
4 Mise en œuvre des transactions	312
5 Implémentation de l'atomicité et la durabilité : le journal	313
6 Gestion des pannes	314
7 La procédure de reprise à chaud	315
8 Le protocole de validation à deux phases	317
9 Gestion de la concurrence	318

10 La technique du verrouillage	319
11 Le problème de l'interblocage	320
12 Verrouillage à granularité fixe et variable	322
13 La sérialisabilité avec l'estampillage	324
14 Niveaux d'isolation	326
15 SQL et les transactions	327
15.1 Types d'interférence	327
15.2 Les instructions SQL	327
<i>a) Validation, annulation et initialisation</i>	328
<i>b) Paramètres d'une transaction</i>	329
<i>c) Initialisation explicite et points de sauvegarde d'une transaction</i>	330
15.3 Contrainte immédiate et différée	330
Exercices	333
Chapitre 10 Programmation SQL	337
1 Introduction	337
2 SQL intégré	338
2.1 Les instructions	338
2.2 Les variables hôtes	339
2.3 La gestion des exceptions et des erreurs	341
<i>a) SQLCODE et SQLSTATE</i>	341
<i>b) L'instruction déclarative WHENEVER</i>	343
2.4 L'instruction SELECT	344
2.5 La notion de curseur	347
<i>a) Principe de fonctionnement</i>	347
<i>b) Options d'un curseur</i>	350
3 SQL Dynamique	353
3.1 Evaluation d'instructions autre que SELECT et non paramétrées	353
3.2 Evaluation d'instructions autre que SELECT avec paramètres connus	355
3.3 SELECT avec liste de projection et paramètres connus	356
3.4 Evaluation d'instructions quelconques	357
<i>a) Les DESCRIPTOR AREA</i>	358
<i>b) Manipulation des DESCRIPTOR AREA</i>	358
<i>c) Exécution des instructions SQL</i>	360
4 Le langage PSM	362
4.1 Concepts de base	363
<i>a) Variable et affectation</i>	363
<i>b) Instructions de branchement conditionnel</i>	363
<i>c) Les boucles</i>	364

Table des matières	15
4.2 Fonctions et procédures stockées	368
5 Les déclencheurs	371
Exercices	374
PARTIE LA MODELISATION DES DONNEES	381
Chapitre 11 La modélisation conceptuelle des données avec UML	383
1 Introduction	383
2 Concepts de base	385
2.1 La classe	385
2.2 Les attributs	386
2.3 L'identifiant	388
2.4 L'association	389
<i>a) Le concept d'association</i>	389
<i>b) Les associations réflexives</i>	392
<i>c) Les associations n-aires</i>	394
<i>d) Les attributs d'une association</i>	396
2.5 La multiplicité	399
3 Autres relations	401
3.1 L'agrégation et la composition	401
3.2 La généralisation	403
<i>a) Le concept de généralisation</i>	403
<i>b) Recherche des généralisations</i>	405
<i>c) L'héritage multiple</i>	409
4 Les contraintes d'intégrité	410
5 Contraintes sur les relations	412
5.1 Contraintes entre associations	412
<i>a) La contrainte de totalité</i>	413
<i>b) La contrainte d'exclusion</i>	413
<i>c) La contrainte de partition</i>	414
<i>d) La contrainte de simultanéité</i>	414
<i>e) La contrainte d'inclusion</i>	415
5.2 Contraintes sur la généralisation	415
<i>a) La contrainte de totalité</i>	415
<i>b) La contrainte d'exclusion</i>	416
<i>c) La contrainte de partition</i>	416
6 Expression des contraintes d'intégrité avec le langage OCL	418

6.1 Notions de base	419
6.2 Types de données et opérateurs	420
6.3 Expression de plusieurs contraintes dans un contexte	422
6.4 Navigation à travers les associations	423
6.5 Accès aux attributs	425
6.6 Opérations sur les collections	426
Exercices	429
Chapitre 12 Le modèle Entité/Association	433
1 Introduction	433
2 Concepts de base	434
2.1 L'entité, les propriétés et l'identifiant	434
2.2 L'association	435
<i>a) Le concept d'association</i>	435
<i>b) Les associations réflexives</i>	437
<i>c) Les associations n-aires</i>	438
<i>d) Les propriétés d'une association</i>	439
2.3 Les cardinalités	440
3 Identifiant relatif	441
4 La généralisation	442
4.1 Le concept de généralisation	443
4.2 Extension à l'héritage multiple	445
5 Contraintes sur les relations	446
5.1 Contraintes entre associations	446
<i>a) La contrainte de totalité</i>	446
<i>b) La contrainte d'exclusion</i>	447
<i>c) La contrainte de partition</i>	447
<i>d) La contrainte de simultanéité</i>	447
<i>e) La contrainte d'inclusion</i>	448
5.2 Contraintes sur la généralisation	449
<i>a) La contrainte de totalité</i>	450
<i>b) La contrainte d'exclusion</i>	450
<i>c) La contrainte de partition</i>	450
6 Les dépendances fonctionnelles	450
6.1 Les dépendances fonctionnelles entre propriétés	451
6.2 Les dépendances fonctionnelles inter-entités	452
7 La normalisation	454
7.1 La normalisation des entités	456
<i>a) La première forme normale (1NF)</i>	456

Table des matières	17
<i>b) Les formes normales supérieures (2NF, 3NF et BCNF)</i>	457
7.2 La normalisation des associations	458
<i>a) Association non en 2NF</i>	458
<i>b) Association non en BCNF</i>	460
Exercices	463
Chapitre 13 La modélisation logique et physique des données	469
1 Introduction	469
2 La modélisation logique des données	470
2.1 La classe et l'entité	470
2.2 Association sans multiplicité 1 ou 0..1 (resp. sans cardinalité maximale 1)	470
2.3 Association binaire avec multiplicité 1 ou 0..1 (resp. avec cardinalité maximale 1)	472
<i>a) Cas de la multiplicité 1 (resp. des cardinalités - 1,1 -)</i>	472
<i>b) Cas de la multiplicité 0..1 (resp. des cardinalités - 0,1 -)</i>	473
2.4 L'association réflexive	475
2.5 L'agrégation et la composition	476
2.6 La généralisation	477
2.7 La généralisation multiple	478
2.8 L'identification relative (modèle E/A)	478
3 La modélisation physique des données	479
3.1 Description de l'étape	479
3.2 Transformation des schémas de relation	482
3.3 Expression des contraintes non graphiques	482
3.4 Contraintes entre associations	484
<i>a) La contrainte de totalité</i>	485
<i>b) La contrainte d'exclusion</i>	486
<i>c) La contrainte de partition</i>	486
<i>d) La contrainte de simultanéité</i>	487
<i>e) La contrainte d'inclusion</i>	488
3.5 Contraintes sur la généralisation	489
<i>a) La contrainte de totalité</i>	489
<i>b) La contrainte d'exclusion</i>	490
<i>c) La contrainte de partition</i>	490
3.6 Dépendance fonctionnelle inter-entités (modèle E/A)	490
Exercices	491

REFERENCES

495

INDEX

497