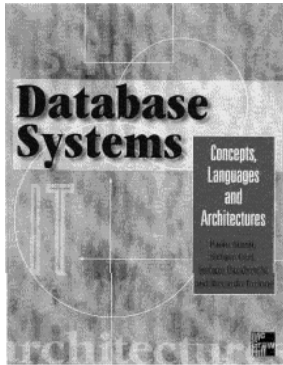
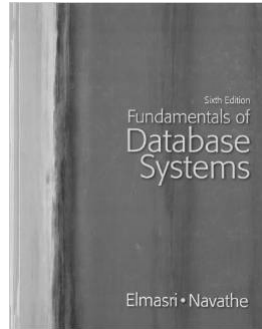


Introduction aux Bases de Données Relationnelles



Atzeni, Ceri et Paraboschi



Elmasri et Navathe

Paraboschi est l'auteur des transparents de ce cours avec fond bleu, comme celui-ci.
Les transparents d'Elmasri et Navathe sont utilisés pour 2 sujets. Traductions de l'italien et anglais, en anglais et français, faites par Kuttler. Un cours est original : passage de TRC à SQL.

Précisions : contrôle continu

TP:

- Présence et le fait d'avoir rendu: 8pt.
- Qualité du rendu, sur 2 ou 3 aléatoires: 6 pt.
- Contrôle de TP (dernière séance): 6 pt.

TD:

- Assiduité en TD: 10 pt
- Interros surprise en amphi: 10 pt

Organisation

Inscription administrative et pédagogique

Inscription sur moodle

Activation du compte FIL : techniciens M3

Activation de postgres sur votre compte FIL
nécessaire à partir du TP 2

P.R.O.F. : inscription par l'enseignant

Pour rendu de l'extérieur: configuration du
VPN de Lille 1

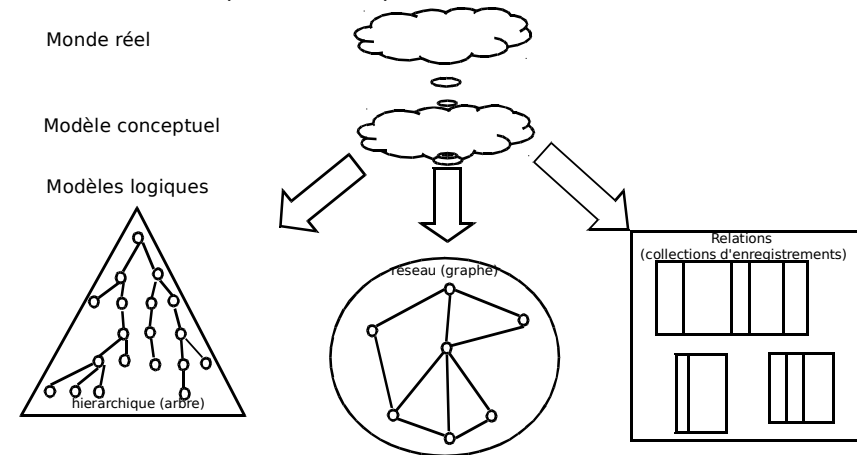
Thématiques des cours

- Modèle relationnel
- Algèbre relationnelle
- SQL :
 - Création de bases (définition, insertion)
 - Requêtes: simple au complexe
- Calcul relationnel des tuples (TRC)
- Passage de TRC à SQL
- Modélisation
- Datalog
- Normalisation
- Optimisation

Cours 1: modèle relationnel

Les modèles de données

sont des combinaisons de constructions pour représenter la réalité de manière structurée et simplifiée. Ils ne couvrent que certains aspects de la réalité, afin d'en simplifier la compréhension.

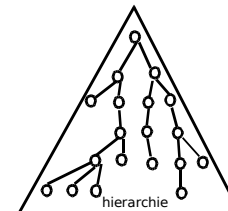


Histoire des modèles

- Modèle hiérarchique (1960)
- Modèle réseau (1970)
- Modèle relationnel (1980)
- Modèle orienté objet (1990)
- Modèle XML (2000)

Modèle hiérarchique

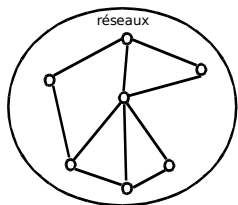
- Les données sont codées par des enregistrements
- Les associations logiques entre données sont représentées par des pointeurs dans un arbre (hiérarchie de la racine aux feuilles)



Modèle réseaux (CODASYL)

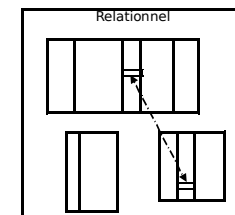
Les données sont codées par des enregistrements

Les associations logiques entre données sont représentées par des pointeurs dans un graphe



Modèle relationnel

- Les données sont représentées par des séquences de valeurs d'attributs
- Les données avec les mêmes séquences d'attributs sont regroupées en tables
- Les associations entre données sont créées en reliant des valeurs d'attributs de différentes tables



Histoire du modèle relationnel

- 1970: inventé par T. Codd (IBM Research)
Premiers projets: SYSTEM R (IBM),
Ingres (Berkeley Un.)
- 1978-80: découvertes technologiques principales
- 1981 : prix Turing pour Codd
- début 1980s: premiers systèmes commerciaux:
Oracle, IBM-SQL DS et DB2, Ingres, Informix,
Sybase
- depuis 1985: succès commercial

Définition informelle de table

Étudiant				
SID	NOM	VILLE	FORMATION	
123	Pierre	Paris	Inf	instance
107	Arnaud	Lille	Log	
415	Celine	Lille	Inf	
702	Estelle	Rome	Log	

colonne

schéma

ligne

Définition: table

- Domaine D: ensemble de valeurs quelconques
- Produit cartésien de n domaines

$$D_1 \times D_2 \times \dots \times D_n$$

Ensemble de n-uplets

$\langle d_1, d_2, \dots, d_n \rangle$, où $d_i \in D_i, 1 \leq i \leq n$

- Relation R sur $D_1 \times D_2 \times \dots \times D_n$: sous-ensemble quelconque de $D_1 \times D_2 \times \dots \times D_n$

Exemple : quatre relations

- $R_1 = (\langle a,1 \rangle, \langle b,3 \rangle)$
- $R_2 = (\langle a,1 \rangle, \langle b,3 \rangle, \langle a,2 \rangle)$
- $R_3 = ()$
- $R_4 = (\langle a,1 \rangle, \langle b,1 \rangle, \langle a,2 \rangle, \langle b,2 \rangle, \langle a,3 \rangle, \langle b,3 \rangle)$

Exemple

- Deux domaines :
 $D_1 = (a,b)$
 $D_2 = (1,2,3)$
- Produit cartésien :
 $D_1 \times D_2 =$
 $(\langle a,1 \rangle, \langle b,1 \rangle, \langle a,2 \rangle, \langle b,2 \rangle, \langle a,3 \rangle, \langle b,3 \rangle)$

Propriétés

- Arité d'une relation: nombre de domaines(n)
- Cardinalité d'une relation: nombre de tuples
- Attribut: nom donné à un domaine dans une relation

[Les noms des attributs d'une relation doivent être uniques]

Comparaison de la terminologie

Propriétés

Schema (d'une relation):

table (attribut1,... attributN)

Les noms des attributs dans une relation doivent être distincts!

R1(A,B) R2(C,D)

A	B
a	1
b	3

C	D
c	1
b	3
a	2

Définition formelle	Exemple
relation	table
attribut	colonne
tuple	ligne
domaine	type de donnée
cardinalité	nombre de lignes
arité	nombre de colonnes

Différence importante

Définition formelle :
absence de doublons
Exemples :
doublons possibles

Exemple 1: gestion des examens d'une université

cours

CID	TITRE	PROF
1	Maths	Laplace
2	CS	Dupuis

Exemple 1

examens

SID	CID	DATE	Note
123	1	7-9-13	10
123	2	8-1-13	8
702	2	7-9-13	5

Exemple 1: gestion des examens

étudiant

SID	NOM	VILLE	FORMATION
123	Pierre	Paris	Inf
415	Celine	Lille	Inf
702	Estelle	Rome	Log

examens

SID	CID	DATE	NOTE
123	1	7-9-13	10
123	2	8-1-13	8
702	2	7-9-13	5

cours

CID	TITRE	PROF
1	Maths	Laplace
2	CS	Dupuis

Exemple 1 : requête A

Quels étudiants ont obtenu la note 10 en Maths?

étudiant

SID	NOM	VILLE	FORMATION
123	Pierre	Paris	Inf
415	Celine	Lille	Inf
702	Estelle	Rome	Log

examens

SID	CID	DATE	NOTE
123	1	7-9-13	10
123	2	8-1-13	8
702	2	7-9-13	5

cours

CID	TITRE	PROF
1	Maths	Laplace
2	CS	Dupuis

Exemple 1 : requête B

Quels profs ont donné des notes à Pierre?

étudiant

SID	NOM	VILLE	FORMATION
123	Pierre	Paris	Inf
415	Celine	Lille	Inf
702	Estelle	Rome	Log

examens

SID	CID	DATE	NOTE
123	1	7-9-13	10
123	2	8-1-13	8
702	2	7-9-13	5

cours

CID	TITRE	PROF
1	Maths	Laplace
2	CS	Dupuis

Exemple 2 : gestion du personnel

employe

MATR	NOM	DATE-EMB	SALAIRE	MATR-MGR
1	Paul	1-1-15	3K	2
2	Georges	1-1-07	2,5K	null
3	Jean	1-7-06	2K	2

affectation

MATR	NUM-PROJ	POURC
1	3	50
1	4	50
2	3	100
3	4	100

projet

NUM-PROJ	TITRE	TYPE
3	Idea	Esprit
4	Wide	Esprit

Requêtes

- Qui est le manager de Paul?

employe

MATR	Nom	DATE-EMB	SALAIRE	MATR-MGR
1	Paul	1-1-15	3K	2
2	Georges	1-1-07	2,5K	null
3	Jean	1-7-06	2K	2

affectation

MATR	NUM-PROJ	POURC
1	3	50
1	4	50
2	3	100
3	4	100

projet

NUM-PROJ	TITRE	TYPE
3	Idea	Esprit
4	Wide	Esprit

Requêtes

Dans quels types de projets travaille Jean?

employe

MATR	Nom	DATA-ASS	SALAIRE	MATR-MGR
1	Paul	1-1-15	3K	2
2	Georges	1-1-07	2,5K	null
3	Jean	1-7-06	2K	2

affectation

MATR	NUM-PROJ	POURC
1	3	50
1	4	50
2	3	100
3	4	100

projet

NUM-PROJ	TITRE	TYPE
3	Idea	Esprit
4	Wide	Esprit

L'information incomplète

Firstname	Middle name	Lastname
Franklin	Delano	Roosevelt
Winston		Churchill
Charles		De Gaulle
Josip		Stalin

Traitement d'information incomplète dans le modèle relationnel

- Méthode naïve mais effective:
 - La valeur nulle (NULL): indique l'absence de valeur dans un domaine (NULL n'appartient à aucun domaine!)
- Chaque attribut a soit une valeur du domaine, ou alors au lieu de cela, la valeur NULL.

Sémantiques possibles des NULLs

- (au moins) trois cas différents
 - valeur inconnue
 - valeur inexistante
 - valeur sans information
- Les systèmes de bases de données ne distinguent pas entre ces trois cas.

Clés

- Sous-ensemble des attributs du schéma, avec les propriétés d'unicité et de minimalité
- Unicité: il n'y a jamais deux tuples dans la même relation, avec la même clé
- Minimalité: si on enlève un attribut de la clé, la propriété d'unicité est perdue

Comment améliorer le schéma ?

- Contraintes d'intégrité: excluent certaines instances de la base, qui ne représentent pas correctement le monde applicatif.
- Pour garantir l'intégrité des données, on utilise:
 - des clés
 - l'intégrité référentielle
 - des contraintes sur les valeurs
 - des contraintes génériques

Exemple de clés

SID	NOM	VILLE	FORMATION
123	Pierre	Paris	Inf
107	Arnaud	Lille	Log
415	Celine	Lille	Inf
702	Estelle	Rome	Log

Clés pour l'exemple 1: gestion des examens d'une université

etudiant

SID	NOM	VILLE	FORMATION

examens

SID	CID	DATE	NOTE

cours

CID	TITRE	PROF

Clés pour l'exemple 2: gestion du personnel

employes

MATR	NOM	DATE-EMB	SALAIRE	MATR-MGR

affectation

MATR	NUM-PROJ	POURC

projet

NUM-PROJ	NOM	TYPE

Exemple 3 : gestion de commandes

client

ID-CLI	Nom	TAX-ID

commande

ID-COM	ID-CLI	DATE	VALEUR

detail

ID-COM	ID-PROD	QTE

produit

ID-PROD	NOM	PRIX

Requêtes

- Quelles sont les commandes de Paul?
- Combien de commandes pour Paul?
- Combien de feux d'artifices ont été commandés le 15/11/2014?
- Calculer, par client, la somme des valeurs de toutes ses commandes
- Donner la commande de la valeur la plus haute

Résumé

- Nous avons vu:
 - Qu'est-ce qu'une relation ?
 - Qu'est un attribut ?
 - C'est quoi, un domaine d'attribut ?
 - C'est quoi, NULL ?
 - Quel est le rôle des clés?