

Introduction aux Bases de Données Relationnelles

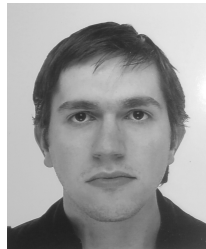
G2 et 3:
Celine Kuttler



G4 :Maude Pupin

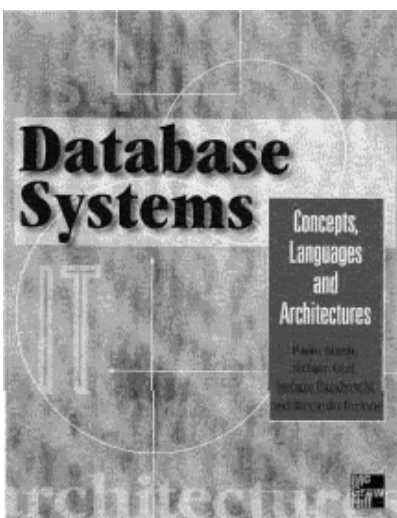


G1 :Alexandre Temperville

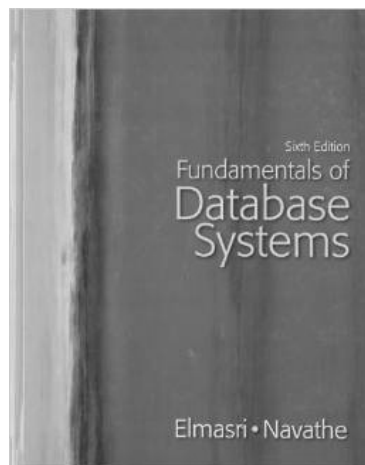


G5 :Aymeric Blot

Littérature



Atzeni, Ceri et Paraboschi



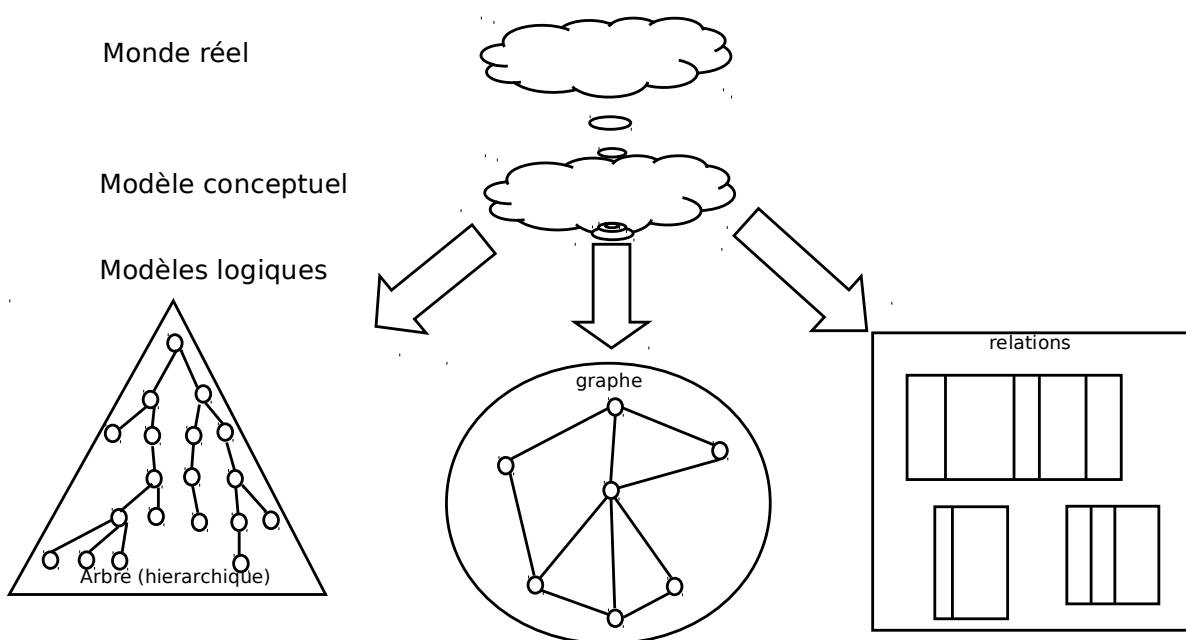
Elmasri et Navathe

Paraboschi est l'auteur des transparents de ce cours.
Traductions de l'italien, en français et anglais, faites par Kuttler.

Cours 1: modèle relationnel

Les modèles de données

représentent la réalité de manière structurée et simplifiée. Ils n'en couvrent que certains aspects, afin d'en simplifier la compréhension.

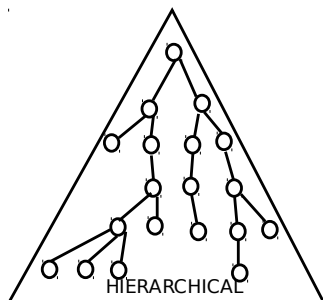


Histoire des modèles

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

Modèle hiérarchique

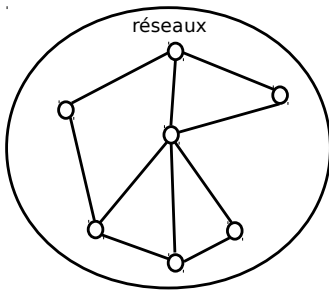
- Les données sont codées par des enregistrements
- Les associations logiques entre données sont représentées par des pointeurs dans une structure arborescente



Modèle réseaux (CODASYN)

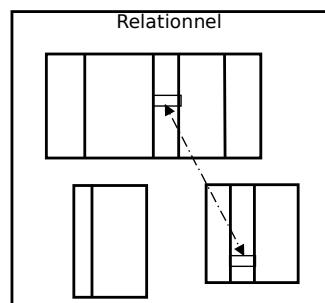
Les données sont codées par des enregistrements

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

Définition informelle de table

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

Definition: table

- Domaine D: ensemble de valeurs quelconques
- Produit cartésien de n domaines
 $D_1 \times D_2 \times \dots \times D_n$
 $\langle d_1, d_2, \dots, d_n \rangle$, 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

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

Exemple : quatre relations

- $R1 = (\langle a,1 \rangle, \langle b,3 \rangle)$
- $R2 = (\langle a,1 \rangle, \langle b,3 \rangle, \langle a,2 \rangle)$
- $R3 = ()$
- $R4 = (\langle a,1 \rangle, \langle b,1 \rangle, \langle a,2 \rangle, \langle b,2 \rangle, \langle a,3 \rangle, \langle b,3 \rangle)$

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]

Propriétés

Schema (d'une relation):

table (attribut1,... attributN)

Les noms des attributs dans une relation doivent être distincts!

R1(A,B)

A	B
a	1
b	3

R2(C,D)

C	D
c	1
b	3
a	2

Comparaison de la terminologie

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: université

cours

CID	TITRE	PROF
1	Maths	Leguichet
2	CS	Duchat

Exemple : université

examens

SID	CID	DATE	Note
123	1	7-9-03	10
123	2	8-1-03	8
702	2	7-9-03	5

Exemple : université

étudiant

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

exams

SID	CID	DATE	NOTE
123	1	7-9-03	10
123	2	8-1-03	8
702	2	7-9-03	5

cours

CID	TITRE	PROF
1	Maths	Leguichet
2	CS	Duchat

Requête 1

- 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

ID	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	Leguichet
2	CS	Duchat

Requête 2

- 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

ID	CID	DATE	NOTE
123	1	7-9-03	10
123	2	8-1-03	8
702	2	7-9-03	5

cours

CID	TITRE	PROF
1	Maths	Leguichet
2	CS	Duchat

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.

2
3

Sémantiques possibles des NULLs

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

2
4

Contraintes d'intégrité

- Il peut exister des instances de bases de données, qui sont correctes au niveau syntaxique, mais qui ne sont pas en accord avec la réalité représentée par la base.
- Il faut ajouter des contraintes aux instances, afin de garantir l'intégrité des données.

Clés

- Ensemble d'attributs permettant d'identifier les tuples d'une relation r , sans ambiguïté
- Formellement:
 - Un ensemble d'attributs K est une super-clé pour r , si r ne contient pas deux tuples distincts t_1 et t_2 , tels que $t_1[K] = t_2[K]$
 - K est une clé pour r si c'est une super-clé minimale pour r
(cad K ne contient pas d'autre super-clé comme vrai sous-ensemble)

Exemple de clés

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

- SID est une clé:
 - Est une super-clé
 - Est minimal puisqu'il ne contient qu'un seul attribut

Une clé ne peut pas être NULL

- Si elle était NULL, elle ne pourrait pas être une super-clé
- Un but du modèle relationnel est de limiter le nombre de valeurs NULL et de fournir de l'information sans ambiguïté

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 ?
 - Quoi sont des clés et super-clés dans une relation ?