



# Modèle relationnel

Faculté Polydisciplinaire d'Ouarzazate

Filière : Informatique et Gestion d'Entreprise

2012/2013

Mohamed Nemiche



## Table des matières

<b>1. Introduction .....</b>	<b>3</b>
<b>1.1 Constat .....</b>	<b>4</b>
<b>2. Le Modèle relationnel .....</b>	<b>7</b>
<b>2.1. Historique sur le Modèle Relationnel .....</b>	<b>7</b>
<b>2.2. Notions de base .....</b>	<b>8</b>
<b>2.3. Concepts du Modèle Relationnel .....</b>	<b>9</b>
<b>2.4. Contraintes d'intégrité .....</b>	<b>11</b>
<b>3. Dépendance fonctionnelle et normalisation .....</b>	<b>11</b>
<b>3.1 Intérêt de la normalisation .....</b>	<b>11</b>
<b>3.2 Notion de dépendance fonctionnelle .....</b>	<b>15</b>
<b>3.3 Forme normales .....</b>	<b>17</b>
<b>4. Exercices d'application .....</b>	<b>21</b>

## 1. Introduction

Au cours des premières années d'utilisation des bases de données, le modèle hiérarchique a été utilisé comme premier moyen de décrire une base de données de manière plus humaine. Ensuite, le modèle réseau a particulièrement régné dans sa norme Codasyl. Ainsi, au début des années 70, il semblait que le modèle à appliquer lors de la mise en place de bases de données serait Codasyl et le serait pendant plusieurs années.

Edgar Frank Codd définit cependant les bases du modèle relationnel à la fin des années 1960. En 1970, il publie l'article "Un modèle relationnel de données pour les grandes banques de données partagées". Actuellement, cet article est l'un des plus influents de l'histoire de l'informatique. C'est parce qu'il a défini les bases du modèle dit relationnel de bases de données. Auparavant, le seul modèle théorique normalisé était le modèle Codasyl, utilisé massivement dans les années 1970 comme paradigme du modèle réseau de base de données.

Codd s'appuie sur les travaux des mathématiciens Cantor and Childs (dont la théorie des ensembles est la véritable base du modèle relationnel). Selon Codd, les données sont regroupées dans des relations (actuellement appelées tables), qui sont une structure qui regroupe des données faisant référence à la même entité de manière organisée. De plus, les relations structurent les données indépendamment de leur stockage réel sur l'ordinateur. C'est-à-dire qu'il s'agit d'un élément conceptuel et non physique.

Codd essayait essentiellement d'empêcher aux utilisateurs de la base de données d'avoir à apprendre les tenants et les aboutissants internes du système. C'est ce qui s'est passé avec le modèle de réseau, dominant lorsque Codd a conçu le modèle relationnel, qui était plutôt physique. Son approche était révolutionnaire en évitant les concepts du monde informatique dans son modèle.

Bien qu'il ait travaillé pour IBM, cette société n'a pas bien accueilli ses théories. En fait, IBM a continué de travailler sur son système de gestion de base de données réseau IMS. Ce sont d'autres entreprises (en particulier Oracle) qui ont appliqué leurs théories.

Quelques années plus tard, le modèle a commencé à être utilisé de plus en plus, pour finalement devenir le modèle de base de données le plus populaire. Aujourd'hui, presque toutes les bases de données suivent ce modèle, bien que ces dernières années, les bases de données dites NoSQL se soient montrées de plus en plus fortes. Elles se sont révélées très efficaces dans les bases de données nécessitant une énorme quantité d'instructions.

Le modèle relationnel garantit l'indépendance physique, c'est-à-dire, la manière de stocker les données doit être absolument indépendante du modèle conceptuel. Si la manière de stocker les données change (si le schéma physique change), il n'est pas nécessaire de changer les schémas logiques. Cela permet aux utilisateurs de se concentrer sur les résultats qu'ils souhaitent obtenir de la base de données, quelle que soit la manière dont les données sont réellement stockées.

**La flexibilité** La base de données offre facilement différentes vues en fonction des utilisateurs et des applications. La vision des données est adaptée à l'utilisateur qui en a besoin.

**L'uniformité** Les structures logiques ont toujours une seule forme logique (les tables). Autrement dit, gérer le modèle relationnel, c'est gérer les tables.

**La simplicité** Facilité de gestion (quelque peu discutable, mais certainement vraie si on compare avec les systèmes de gestion de base de données antérieurs à ce modèle.

### **1.1 Constat**

Comment gérer (mémoriser et traiter) un ensemble d'étudiants (n°, nom, prénom, groupe, filière...) ?

**La Première solution** : l'utilisation des fichiers

La difficulté est comment traiter les requêtes (répondre aux questions) suivantes ?

- n° de Youssef Alami ?
- les étudiants du groupe 51 ?
- ...

A chaque nouvelle question, il faut écrire une nouvelle procédure. En plus il faut prendre en considération la structure des données stockées et du langage de programmation.

Travailler directement sur un fichier présente plusieurs inconvénients :

- Manipulation de données lourde et compliquée. Il faut être expert en programmation
- Le programmeur doit connaître la localisation physique des fichiers, la structure physique des enregistrements, le mode d'accès à ces fichiers
- Toute modification de la structure des enregistrements (ajout d'un champ par exemple) entraîne la réécriture de tous les programmes qui manipulent ces fichiers



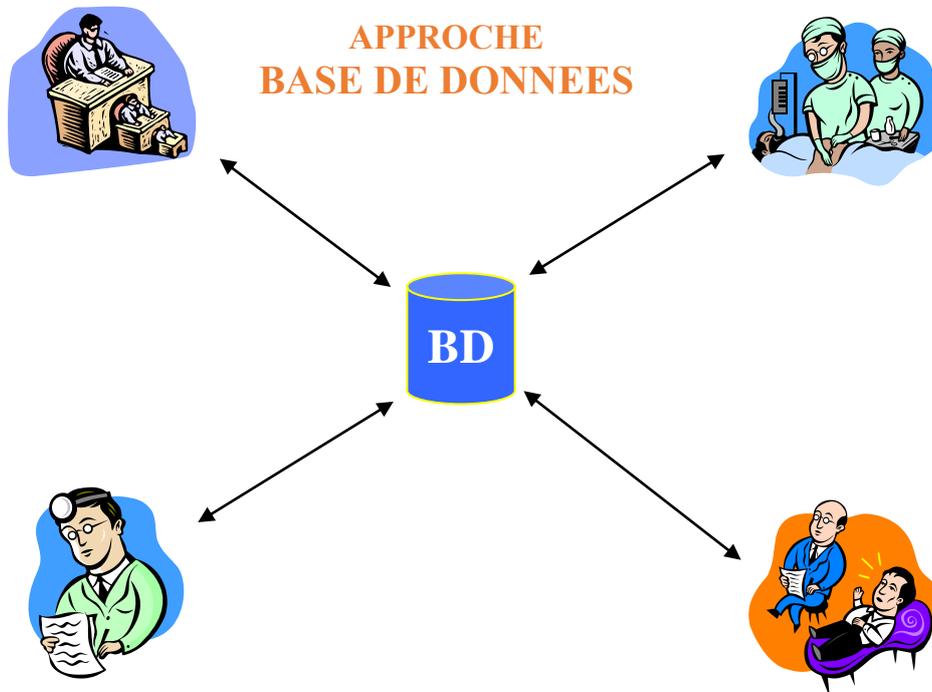
## L'APPROCHE SYSTEMES DE FICHIERS

## Problèmes

- Difficultés de gestion
- Incohérence des données
- Coûts élevés
- Maintenance difficile
- Gestion de pannes ???
- Partage des données ???
- Confidentialité ???



**Solution:** L'approche « Bases de données »

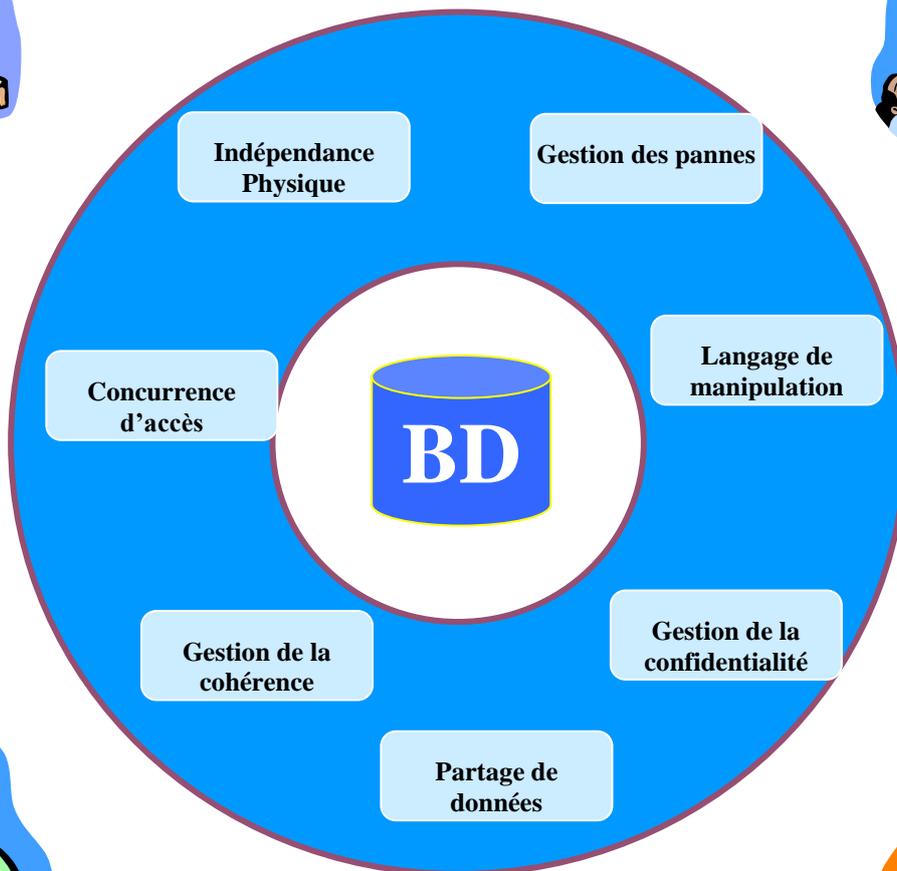


**Définition (intuitive) :** une base de données est un ensemble d'informations, (fichiers), partagé par plusieurs utilisateurs. Ces informations sont interrogées et mises à jour par l'intermédiaire d'un logiciel.

**La Modélisation des données :** consiste à **organiser** correctement les données et à éliminer la **redondance** de données.

**Système de Gestion de Bases de Données :** est un logiciel qui gère un ensemble de fonctionnalités en relation avec l'interrogation, la cohésion et le partage de données et l'administration de la base de données.

### GRANDES FONCTIONALITES D'UN SGBD



Comme exemple d'application on peut citer un système de gestion de la clientèle d'une banque et le catalogue électronique d'une bibliothèque.

**Définition** : une base de données est un ensemble structuré de données (1) enregistrées sur des supports accessibles par l'ordinateur (2) pour satisfaire simultanément plusieurs utilisateurs (3) de manière sélective (4) en un temps opportun(5).

(1) : Organisation et description de données ;

(2) : Stockage sur disque ;

(3) : Partage des données ;

(4) : Confidentialité ;

(5) : Performance.

## Modélisation et modèles

**Définition (Modélisation)** : Expression du monde réel observée en utilisant les concepts d'un *modèle* de représentation. C'est une représentation de la partie du monde « utile » à l'application.

Par exemple, la modélisation d'une « personne » sera différente dans une application gérant l'inscription d'étudiants à des diplômes, et dans la BDD de la sécurité sociale.

Existe plusieurs modèles de représentation.

- 1960 Modèle hiérarchique, modèle réseau.
- 1970 Modèle relationnel
- 1990 Modèle objet

## 2. Le Modèle relationnel

### *2.1. Historique sur le Modèle Relationnel*

Conçu en 1970 par E.F CODD (alors chercheur chez IBM-SAN JOSE- CALIFORNIE), Ce modèle qui se développe le plus actuellement pour des raisons de simplicités de représentation.

Basé sur la théorie des ensembles et les mathématiques relationnelles, il permet d'appliquer aux données tous les opérateurs ensemblistes et relationnels. Le grand apport de ce modèle est la normalisation des données. On peut citer aussi :

- Modèle de description des informations
- Modèle très utilisé, quoiqu'ancien
- Fondement mathématique
- Modèle associé à l'algèbre relationnelle

Le modèle relationnel repose sur quatre notions principales : le domaine, la relation, les dépendances fonctionnelles et la clé.

## 2.2. Notions de base

### 2.2.1. Notions de domaine, d'attribut

Un attribut est une information qui contribue à la description de l'objet que l'on veut modéliser. Exemple nom, prénom, adresse ...

Ces attributs peuvent être :

- calculés à partir d'autres attributs,
- non calculables
- + élémentaires,
- + structurés

Dans ce cas-là, en général, ils ne sont pas mémorisés. On garde les attributs élémentaires.

Le modèle relationnel, en soi, ne fait pas la différence entre les 3 !!!

Dans ce cas-là, en général, ils ne sont pas mémorisés. On garde la formule de calcul et les attributs élémentaires

$$\text{Taille\_moyenne} = (\text{mini} + \text{maxi}) / 2$$

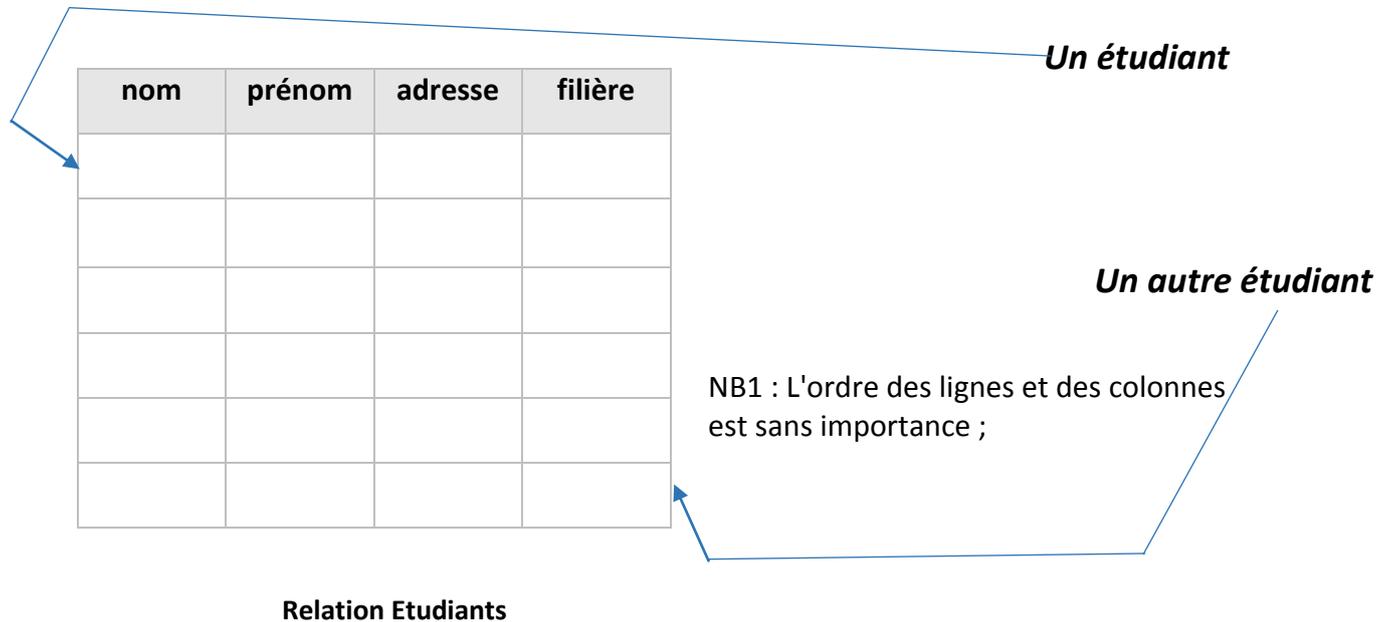
Un domaine : ensemble fini ou infini de valeurs distinctes que peut prendre un attribut (entiers, booléens...). Un domaine peut être partagé par plusieurs attributs.

### 2.2.2 Notion de relation

Il y a deux façons de « voir » les choses :

- 1 Sous l'angle mathématique : Une **relation** est un sous-ensemble du produit cartésien des **domaines**.
- 2 Sous l'angle pragmatique : Une relation, c'est une table dans laquelle les colonnes correspondent aux **attributs** et les lignes aux occurrences

Une relation, c'est une table dans laquelle les colonnes correspondent aux attributs et les lignes aux occurrences :



### 2.2.3 Notion de clé

Clé : attribut ou ensemble d'attributs dont les valeurs identifient de manière unique chaque tuple de la relation.

N°	Nom	Prénom	Adresse	Filière
a12	Alami	Yousef	xx1	IGE
a13	toto	Mohamed	yy1	Tourisme
a14	titi	Ali	zz1	Cinema

- Jamais de tuple en double !
- Il existe donc toujours un identifiant : dans le cas le pire, c'est l'ensemble des attributs de la relation !
- Il peut exister plusieurs identifiants

### 2.3. Concepts du Modèle Relationnel

Une **relation** est caractérisée par un nom R et un ensemble d'attributs A1, A2, ... An. Si une relation a n attributs, on dit que la relation R est de **arité** n et on note **R (A1, A2, ... An)**

Ex. : Produit (numProd, libellé, pu)

Un **attribut** est caractérisé par un nom Ai et un domaine noté **dom(Ai)**, ensemble des valeurs possibles de Ai.

Ex. : dom (pu) =] 0, 10000]

- Valeur **nulle** (notée **NULL**) : valeur particulière indiquant que la valeur d'un attribut n'est pas connue ou que l'attribut ne s'applique pas.

ex1. Cas un client dont on ignore la ddn.

ex2. Cas d'un employé ne possédant pas de téléphone.

- Un tuple d'une relation R (A1, A2, ... An) est un ensemble de valeurs **<v1, v2 , ...vn>** telles que  $v_i \in \text{dom}(A_i)$  ou  $v_i = \text{NULL}$

*<36, nokia, 500> est un tuple de la relation Produit.*

**EXEMPLE :**

Soit la relation « Ouvrage » dans sa représentation tabulaire, elle se représente ainsi :

**OUVRAGE :**  
 ↑  
 Nom de la relation  
 Ou table

N°Ouvrage	Titre	Auteur	Qte_ stock
5412	Le rêve Mexicain	xxxx	120
7318	Le Parfum	xxxx	57
8417	De près de loin	xxxx	140
9003	Jour de silence à Tanger	xxxx	95
9075	L'avenir radieux	xxxx	110

←  
 Tuple ou  
 Occurrence de la  
 relation

→  
 Valeur de l'attribut  
 « Qte\_ stock »

Dans sa représentation en extension, la relation « Ouvrage » se représente ainsi : Ouvrage ( N°Ouvrage,Titre, Auteur, Qte\_stock).

Degré : Correspond aux nombres d'attributs de la relation. (La relation « Ouvrage » est de degré 4).

**Cardinalité :** Elle représente le nombre de tuples ou d'occurrences ou de lignes de la relation. (La cardinalité de « Ouvrage » est 5).

L'**intention** d'une relation est définie par son **schéma** :

- nom de la relation
- liste des attributs + domaines
- contraintes d'intégrité

**Exemple :**

Produit (numProd : nombre entier, libellé : chaîne de caractères, pu : nombre réel)

**Deux contraintes :**

- 1) clé primaire : numProd
- 2)  $0 < pu \leq 10\,000$

## 2.4. Contraintes d'intégrité

**Clé d'une relation :** Groupe d'attributs **minimum** qui identifie de manière **unique** un tuple dans une relation. Toute relation doit avoir au moins une clé documentée, c'est la **clé primaire**.

- ex. numProd : clé primaire de Produit.

Dans ce support la clé primaire est **soulignée** dans le schéma.

- ex. Produit (numProd, libellé, pu)

**Clé étrangère d'une relation :** attribut(s) constituant la clé primaire d'une autre relation. Les clés étrangères définissent les CI référentielles.

- Notation : la clé étrangère est en **italique** dans le schéma.

**Contraintes liées au domaine :** les données doivent vérifier certaines conditions pour être cohérentes.

- ex.  $pu > 0$  ET  $pu \leq 10000$

## 3. Dépendance fonctionnelle et normalisation

Un des buts de la normalisation est de passer d'une « grosse » relation à un ensemble de plus petites, plus « propres ».

### 3.1 Intérêt de la normalisation

Si l'on fait tout ce travail, c'est parce qu'on va y gagner ! Normaliser une relation (ou un ensemble de relations), c'est éviter des problèmes :

- de redondance
  - de stockage.
- }
- ajout conditionnel d'informations,



- suppression superflue d'informations,
- modification répétitive d'informations

Considérons une extension de la relation Étudiants

redondance

Nom	Prénom	Code Filière	Filière
Alami	Ali	IGE	Informatique et gestion d'entreprise
Slaoui	Rachid	IGE	Informatique et gestion d'entreprise
Nasri	Hind	IGE	Informatique et gestion d'entreprise
Belhaj	Amina	IGE	Informatique et gestion d'entreprise
Ouhbi	Fouad	IGE	Informatique et gestion d'entreprise

Dans ce tableau (bien entendu, nous l'avons choisi pour cela !), il y a 5 fois le fait que la filière informatique et gestion d'entreprise, le code de la filière IGE.

Relation Etudiants 1

Clé étrangère

Nom	Prénom	Code Filière
Alami	Ali	IGE
Slaoui	Rachid	IGE
Nasri	Hind	IGE
Belhaj	Amina	IGE
Ouhbi	Fouad	IGE

Relation Filière

Code Filière	Filière
IGE	Informatique et gestion d'entreprise

Après normalisation, le système de tables n'aura plus que l'attribut **Code Filière** comme redondance. C'est ce que l'on pourrait appeler les coûts fixes. Le prix de la non-redondance, en quelque sorte.

### Ajout conditionnel

Reprenons l'extension précédente de la relation Étudiants

Nom	Prénom	Code Filière	Filière
Alami	Ali	IGE	Informatique et gestion d'entreprise
Slaoui	Rachid	IGE	Informatique et gestion d'entreprise
Nasri	Hind	IGE	Informatique et gestion d'entreprise

<b>Belhaj</b>	Amina	IGE	Informatique et gestion d'entreprise
<b>Ouhbi</b>	Fouad	IGE	Informatique et gestion d'entreprise

Supposons que l'on vienne de connaître l'existence d'un autre étudiant, par exemple Ait Taleb Samir de la filière Techniques cinématographiques (TECC).

### **Etape n°1**

*Relation Filière*

Code Filière	Filière
<b>IGE</b>	Informatique et gestion d'entreprise
<b>TECC</b>	Techniques Cinématographiques

Après normalisation, la nouvelle filière sera enregistrée avant même le premier étudiant ...

### **Etape n°2**

*Relation Etudiants 1*

Nom	Prénom	Code Filière
<b>Alami</b>	Ali	IGE
<b>Slaoui</b>	Rachid	IGE
<b>Nasri</b>	Hind	IGE
<b>Belhaj</b>	Amina	IGE
<b>Ouhbi</b>	Fouad	IGE
<b>Ait Taleb</b>	Samir	TECC

### **Suppression superflue**

Reprenons notre extension favorite de la relation Étudiants ajoutons-y l'information concernant Ait Taleb.

Nom	Prénom	Code Filière	Filière
<b>Alami</b>	Ali	IGE	Informatique et gestion d'entreprise
<b>Slaoui</b>	Rachid	IGE	Informatique et gestion d'entreprise
<b>Nasri</b>	Hind	IGE	Informatique et gestion d'entreprise
<b>Belhaj</b>	Amina	IGE	Informatique et gestion d'entreprise
<b>Ouhbi</b>	Fouad	IGE	Informatique et gestion d'entreprise
<del><b>Ait Taleb</b></del>	<del>Ali</del>	<del>TECC</del>	<del>Techniques Cinématographiques</del>

Supposons maintenant que cet étudiant soit le seul de la filière TECC dans la base de données. Supprimer les informations concernant le Ait Taleb nous fait OBLIGATOIREMENT perdre celles sur la filière

**suppression de l'étudiant → suppression de la filière !!**

Relation Etudiants 1

Nom	Prénom	Code Filière
Alami	Ali	IGE
Slaoui	Rachid	IGE
Nasri	Hind	IGE
Belhaj	Amina	IGE
Ouhbi	Fouad	IGE
<del>Ait Taleb</del>	<del>Samir</del>	<del>TECC</del>

Relation Filière

Code Filière	Filière
IGE	Informatique et gestion d'entreprise
TECC	Techniques Cinématographiques

Après normalisation, on pourra conserver les références de la filière TECC tout en ayant enlevé celles de l'étudiant. On n'aura supprimé que ce qui était strictement nécessaire.

### **Modification répétitive**

Reprenons encore une fois (la dernière ???) l'extension de la relation Étudiants:

Nom	Prénom	Code Filière	Filière
Alami	Ali	IGE	Informatique et gestion d'organisation
Slaoui	Rachid	IGE	Informatique et gestion d'organisation
Nasri	Hind	IGE	Informatique et gestion d'organisation
Belhaj	Amina	IGE	Informatique et gestion d'organisation
Ouhbi	Fouad	IGE	Informatique et gestion d'organisation
Ait Taleb	Ali	TECC	Techniques Cinématographiques

Supposons maintenant que nous ayons mal écrit le libelle de la filière informatique. Rectifier cette erreur va nécessiter la modification de TOUTES les occurrences de la relation (de toutes les lignes de la table) :

Nom	Prénom	Code Filière	Filière
Alami	Ali	IGE	Informatique et gestion d'entreprise
Slaoui	Rachid	IGE	Informatique et gestion d'entreprise

<b>Nasri</b>	Hind	IGE	Informatique et gestion d'entreprise
<b>Belhaj</b>	Amina	IGE	Informatique et gestion d'entreprise
<b>Ouhbi</b>	Fouad	IGE	Informatique et gestion d'entreprise
<b>Ait Taleb</b>	Ali	TECC	Techniques Cinématographiques

Après la normalisation, s'il y a une erreur, on pourra la rectifier en une seule écriture modificative. Le gain de temps peut être appréciable.

*Relation Filière*

Code Filière	Filière
IGE	Informatique et gestion d'entreprise
TECC	Techniques Cinématographiques

Théorie de la normalisation : limiter ces problèmes en analysant les dépendances entre les attributs qui sont à l'origine des phénomènes redondances.

Poser des méthodes systématiques pour décomposer une relation en plusieurs relations plus adaptées pour la gestion du problème

◊ La relation d'origine doit pouvoir être retrouvée par jointures

Idée : spécifier les relations entre les attributs (notion de dépendance entre les informations).

Définition : Soit  $R(\Delta)$  une relation,  $\Delta$  ses attributs, et  $X \subseteq \Delta$  et  $Y \subseteq \Delta$ , il existe une dépendance fonctionnelle (DF) entre X et Y (on dit aussi que X détermine Y), notée  $X \rightarrow Y$ , si dans la relation chaque valeur de X détermine une et une seule valeur de Y.

- X est appelé source et Y cible.
- i.e.  $X \rightarrow Y$  : si deux tuples ont même valeur sur X alors ils ont même valeur sur Y.

**3.2 Notion de dépendance fonctionnelle**

A	B	C	D	E
a1	b1	c1	d3	e2
a1	b1	c3	d4	e3
a2	b2	c4	d2	e1
a3	b1	c1	d3	e2
a2	b2	c4	d2	e1

		$\rightarrow$ ou $\nrightarrow$ ?	
	A	$\rightarrow$	B
	A	$\nrightarrow$	C
	B,C	$\rightarrow$	D
	A,C	$\rightarrow$	D
	B	$\nrightarrow$	D
	D	$\rightarrow$	E
	A,C	?	E

1

## Règles d'Armstrong

Tous les attributs sont pris dans un ensemble  $\Delta$  :

- Si  $Y \subseteq X \subseteq \Delta$ , alors  $X \rightarrow Y$  (réflexivité)
- Si  $X \rightarrow Y$  et  $Z \subseteq \Delta$ , alors  $XZ \rightarrow YZ$  (augmentation)
  - NB ( $XZ = X \cup Z$ )
- Si  $X \rightarrow Y$  et  $Y \rightarrow Z$ , alors  $X \rightarrow Z$  (transitivité)

### Exemple :

Démontrer que  $AD \rightarrow BE$  en ayant les dépendances fonctionnelles suivantes :

$A \rightarrow B$

$B, C \rightarrow D$

$A, C \rightarrow E$

$D \rightarrow E$

$A, C \rightarrow E$

Dem:

(1)  $A \rightarrow B \xrightarrow{\text{(Augmentation)}} AD \rightarrow BD$

(2)  $D \rightarrow E \xrightarrow{\text{(Augmentation)}} BD \rightarrow BE$

(3) (1) + (2)  $\xrightarrow{\text{(Augmentation)}} AD \rightarrow BE$

Pour simplifier, 3 autres règles qui peuvent se déduire des 3 premières.

- **Union** :  $X \rightarrow Y$  et  $X \rightarrow Z \implies X \rightarrow YZ$
- **Décomposition** :  $X \rightarrow Y$  et  $Z \subseteq Y$  alors  $X \rightarrow Z$
- **Pseudo-transitivité** :  $X \rightarrow Y$  et  $WY \rightarrow Z \implies WX \rightarrow Z$

## Typologie des dépendances fonctionnelles

Une dépendance  $X \xrightarrow{R} Y$  est :

- Triviale si  $Y \subseteq X$
- Élémentaire si pour tout  $X' \subset X$ , la dépendance fonctionnelle  $X' \xrightarrow{R} Y$  n'est pas vraie ( $Y$  ne dépend pas fonctionnellement d'une partie de  $X$ )
- Canonique si sa partie droite ne comporte qu'un seul attribut
- Directe si (i) elle est élémentaire et si (ii)  $Y$  ne dépend pas transitivement de  $X$ .



Soit la relation :

notes	matiere	etudiant	notes
	15	Albert Paul	8, 12.5
	15	Duplo Bertrand	2.5, 0, 18

Cette relation n'est pas en 1NF car l'attribut etudiant n'est pas atomique. La décomposition en première forme normale consiste à remplacer l'attribut etudiant par nom et prénom

Notes(matiere, nom, prenom, note1, note2, note3) ou ...

### 3.3.2 Deuxième forme normale

Définition : une relation est dite en **deuxième forme normale** (2FN) si elle est en 1FN et qu'un attribut n'appartenant à aucune clé soit en DF élémentaire avec toutes les clés.

Soit R la relation suivante :  $\langle R(\underline{A}, \underline{B}, C), \{AB \rightarrow C, B \rightarrow C\} \rangle$ . Cette relation n'est pas en deuxième forme normale car l'attribut non clé C dépend seulement d'une partie de la clé

Pour passer en 2FN, il suffit d'appliquer une décomposition de la relation basée sur  $B \rightarrow C$ .

D'où  $\langle R_1(\underline{A}, \underline{B}), \{AB \rightarrow B\} \rangle$  et  $\langle R_2(\underline{B}, C), \{B \rightarrow C\} \rangle$

### 3.3.3 Troisième forme normale

Définition : une relation est dite en **troisième forme normale** (3FN) lorsqu'elle est en 2FN et que tout attribut n'appartenant pas à une clé ne dépend pas d'un attribut non clé.

- Toutes les DF sont directes.
- Tout attribut n'appartenant pas à une clé est en DF élémentaire directe avec la clé.
- Dès que deux attributs sont en DF et qu'ils ne font pas partie d'une clé alors ils ne sont pas en 3FN.

Soit R la relation suivante :  $\langle R(\underline{A}, B, C), \{A \rightarrow B ; B \rightarrow C\} \rangle$ . R n'est pas en 3FN car A détermine C indirectement (par transitivité).

Pour passer en 3FN, il suffit de décomposer la relation avec  $B \rightarrow C$

D'où  $\langle R_1(\underline{A}, B), \{A \rightarrow B\} \rangle$  et  $\langle R_2(\underline{B}, C), \{B \rightarrow C\} \rangle$

Toute relation admet une décomposition en 3FN à jonction conservative (sans perte) et avec préservation des DF.

#### Exemple

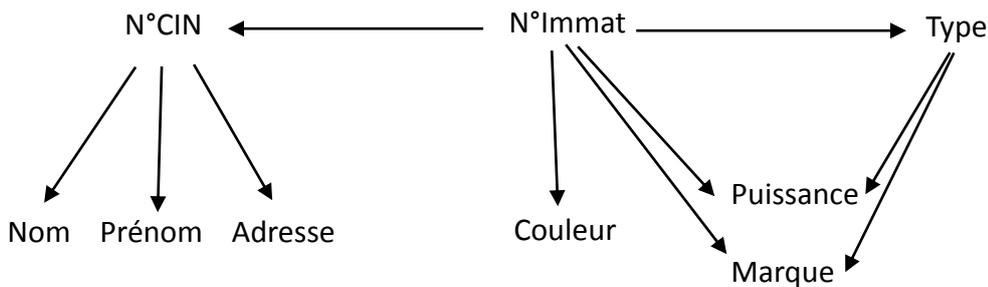
Service d'immatriculation de voitures dans une préfecture

- Soient les DF suivantes :

## Modèle relationnel

- N°Immat -> Couleur, Type, Puissance, Marque
- N°CIN -> Nom, Prénom, Adresse
- N°Immat -> N°CIN
- Type -> Marque, Puissance

On crée le graphe :



On supprime les transitivités

On obtient :

Personne (N°CIN, Nom, Prénom, Adresse)

Voiture (N°Immat, Couleur, Type\*, N°CIN\*)

Types (Type, Puissance, Marque)

### 3.3.4 Troisième forme normale de Boyce-Codd

Une relation est en **troisième forme normale de Boyce-Codd-Kent** si :

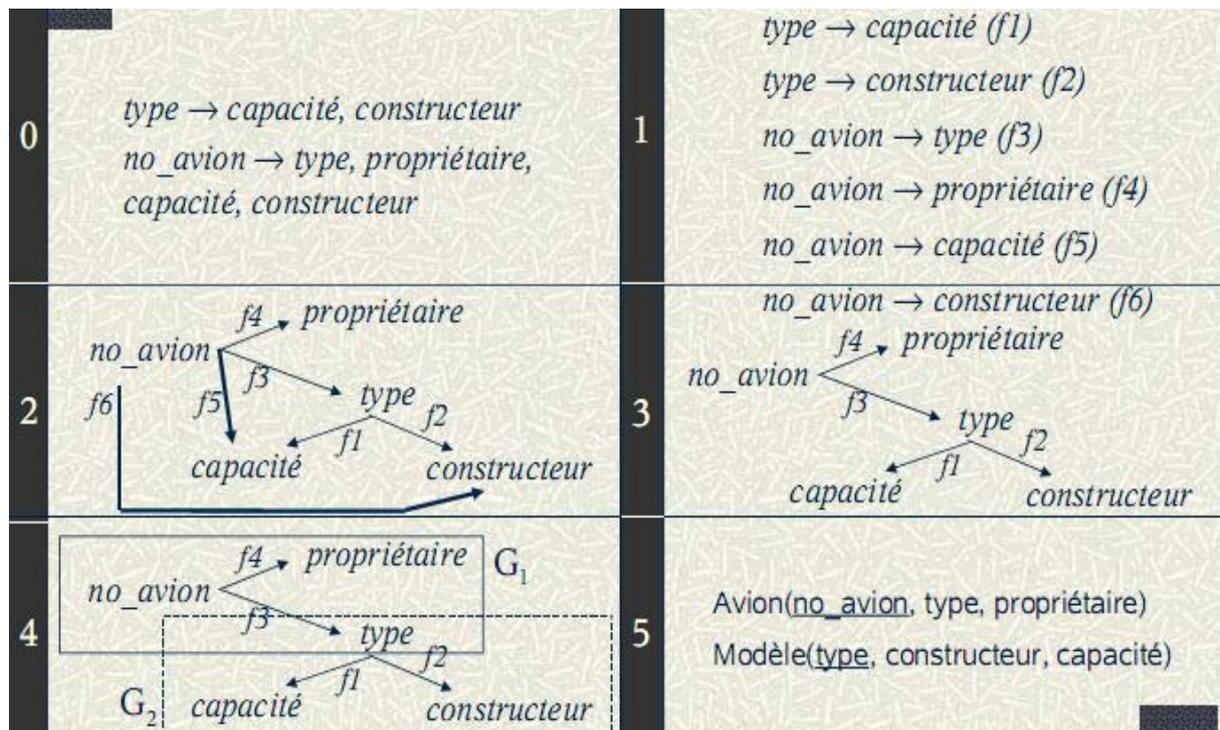
- 3FNBCK
- elle est en troisième forme normale,
  - chaque fois qu'une dépendance fonctionnelle non triviale de la relation R est vérifiée, alors sa source contient une clé de R.

### Algorithme de normalisation par synthèse simplifié

- 0 Ecrire les dépendances fonctionnelles (df)
- 1 Rendre canonique et élémentaires les df qui ne le sont pas
- 2 Représenter les nouvelles df sous forme d'un graphe dont les nœuds sont les attributs impliqués dans les df et les arcs les df elles-mêmes
- 3 Eliminer les df non directes

- 4 Partitionner les df en  $G_1, \dots, G_n$  de façon à ce que toutes les df d'une partition aient la même partie gauche
- 5 Constituer une relation  $R_i$  par  $G_i$ , la partie gauche de  $G_i$  étant clé de  $R_i$  : chaque  $R_i$  est alors en 3NF

### Exemple



## 4. Exercices d'application

### Exercice 1

Soit les deux relations suivantes R1 (A, B, C) et R2 (D, E, A). Soit les extensions suivantes des deux relations :

A	B	C
1	..	..
2	..	..
3	..	..
4	..	..

D	E	A
..	..	2
..	..	2
..	..	1
..	..	4

A	B	C
1	..	..
2	..	..
3	..	..
4	..	..

D	E	A
..	..	2
..	..	5
..	..	1
..	..	4

Expliquer si la contrainte d'intégrité référentielle entre A de R1 et A de R2 est respectée dans les deux extensions.

### Exercice 2

Soit la relation R(A,B,C) avec l'extension suivante :

A	B	C
a1	b1	c1
a2	b1	c2
a2	b2	c3
a1	b2	c2
a2	b3	c3
a1	b3	c3

Quelle est la clé primaire de la relation R.

### Exercice 3

Soit le schéma de la base de données Bibliothèque suivante :

Etudiant(NumEtd,NomEtd,AdresseEtd)

Livre(NumLivre,TitreLivre,NumAuteur,NumEditeur,NumTheme,AnneeEdition)

Auteur(NumAuteur,NomAuteur,AdresseAuteur)

Editeur(NumEditeur,NomEditeur,AdresseEditeur)

Theme(NumTheme,IntituléTheme)

Prêt(NumEtd,NumLivre,DatePret,DateRetour)

Un étudiant peut emprunter plusieurs livres à la fois. Chaque fois qu'un livre est retourné, la date de retour est mise à jour. Sinon sa valeur reste null. Pour des raisons de statistique, on conserve dans la table Prêt tous les tuples les livres retournés.

Un livre ne peut pas être emprunté le même jour ou il est retourné !

- 1) Donner les clés de ces relations. Justifier.
- 2) Donner toutes les contraintes d'intégrités référentielles qui apparaissent dans ce schéma.

**Exercice 4**

On considère la relation R(A,B,C,D,E, F) sur laquelle sont définies les dépendances fonctionnelles suivantes :

$A, B \twoheadrightarrow C$

$D \twoheadrightarrow C$

$D \twoheadrightarrow E$

$C, E \twoheadrightarrow F$

$E \twoheadrightarrow A$

1. Compléter les cases vides du tableau :

<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>
	1		110		54
<i>x</i>	2	<i>j</i>	100	<i>n</i>	52
<i>w</i>	1	<i>i</i>	110	<i>m</i>	
	2		100		52

**Exercice 5**

On considère la relation R (A, B, C) avec l'ensemble de DF {  $A \rightarrow B ; B \rightarrow C$  }. Par exemple, R pourrait être la relation FILM (No\_exploitation, Titre, Realisateur).

- 1) Quelle est la clé primaire de R ? Dans quelle forme normale se trouve cette relation ?
- 2) L'extension de la relation R' suivante est-elle une extension de R ?

R'	A	B	C
	A1	B1	C1
	A2	B1	C2
	A3	B2	C1
	A4	B3	C3

3) Trouver une extension R'' conforme à R, à partir de R'.

4) Proposer une décomposition en 3FN de R sans perte d'information.

### **Exercice 6**

Nous considérons ce schéma relationnel résultant d'une première enquête :

CLIENT(NumClient,RaisonSociale,NumRepresentant,Tauxrepresentant)

D'une part chaque CLIENT n'est affecté qu'à un seul REPRESENTANT

- 1) Représenter les dépendances fonctionnelles
- 2) Quelle est la forme normale ?
- 3) Définir le schéma équivalent en 3<sup>e</sup> forme normale.

### **Exercice 7**

Soient :

La relation R1(NumMatriculeOuvrier,NomOuvrier,NumRéparation,NumMachine,TempsPassé, Dateréparation,NomMachine,NumAtelier,NomAtelier)

Les dépendances fonctionnelles suivantes :

NumMatriculeOuvrier  $\longrightarrow$  NomOuvrier

NumRéparation  $\longrightarrow$  Dateréparation

NumMatriculeOuvrier, NumRéparation  $\longrightarrow$  TempsPassé

NumRéparation  $\longrightarrow$  NumMachine

NumMachine  $\longrightarrow$  NomMachine

NumMachine  $\longrightarrow$  NomAtelier

NumAtelier  $\longrightarrow$  NomAtelier

NumMachine  $\longrightarrow$  NumAtelier

- 1) Clé primaire de R1 ?
- 2) Trouver le graphe des dépendances fonctionnelles
- 3) Définir le schéma équivalent en 3<sup>e</sup> forme normale.

**Exercice 8**

On considère l'ensemble des dépendances fonctionnelles suivantes :

e,f,g →t

a →v

a →b

b →c

b →w

c →d

c →z

d →c

d →u

- 1) Trouver le graphe des dépendances fonctionnelles.
- 2) Définir le schéma équivalent en 3<sup>e</sup> forme normale (indiquer les clés primaires et étrangères).

**Exercice 9**

Une entreprise comprend différents services, chacun étant caractérisé par un numéro unique (NOSER), un nom (NOMS) de service, le numéro (NORES) et le nom (NOMR) de son responsable. Un budget (BUSER) est attribué à un service.

Chaque service gère un ou plusieurs projets, mais un projet est géré par un seul service. Un projet est caractérisé par un numéro (NOPRO) supposé unique et un nom (NOMP). Un budget (BUPRO) est attribué à un projet.

Les employés de l'entreprise sont affectés à un instant donné à un seul projet. Un employé est caractérisé par un numéro (NOEMP) supposé unique et un nom (NOME). Chaque employé peut être joint par l'intermédiaire d'un numéro de téléphone (NOTEL). Un numéro de téléphone peut être partagé entre plusieurs employés.

Un employé est installé dans un bureau caractérisé par un numéro unique (NOBUR). Un bureau peut accueillir plusieurs employés et plusieurs appareils téléphoniques. La localisation d'un bureau est repérée par le nom de son bâtiment (NOMB). Un bureau est rattaché pour gestion à un seul service.

- 1) Déterminer les dépendances fonctionnelles qui existent sur l'ensemble des attributs.
- 2) Donner une représentation de la base sous forme d'un ensemble de relations en 3FN.

## Références Bibliographiques

A.MESGUISH B.NORMIER. COMPRENDRE LES BASES DE DONNEES THEORIE ET PRATIQUE MASSON 1981

C.DELOBEL M.ADIBA. BASES DE DONNEES ET SYSTEMES RELATIONNELS DUNOD 1982

G.GARDARIN BASES DE DONNEES : LES SYSTEMES ET LEURS LANGAGES EYROLLES 1983

G.GARDARIN P.VALDURIEZ. RELATIONAL DATABASES AND KNOWLEDGE BASES ADDISON WESLEY 1989

G.GARDARIN P.VALDURIEZ. SGBD AVANCES BASES DE DONNEES OBJETS, DEDUCTIVES, REPARTIES EYROLLES 1991

LES BASES DE DONNEES AVANCEES. DU MODELE RELATIONNEL AU IMAD SALEH HERMES MODELE ORIENTE OBJET OUVRAGE COLLECTIF COORDONNATEUR 1992

J. AKOKA, I. COMYN WATTIAU : CONCEPTION DES BASES DE DONNEES RELATIONNELLES : CONCEPTS, METHODES ET CAS CORRIGES. VUIBERT, 2009.

MEIER. INTRODUCTION PRATIQUE AUX BASES DE DONNEES RELATIONNELLES, 2009

N. BOUDJLIDA : BASES DE DONNEES ET SYSTEMES D 'INFORMATION. LE MODELE RELATIONNEL : LANGAGES, SYSTEMES ET METHODES. DUNOD, 1999

MEIER . INTRODUCTION PRATIQUE AUX BASES DE DONNEES RELATIONNELLES, 2002.

BIZOI : SQL POUR ORACLE 10G,EYROLLES, 2006

CONCEPTIRAMEZ ELMASRI ET SHAMKANT NAVATHE : CONCEPTION ET ARCHITECTURE BASE DE DONNEES 4IEME EDITION, PEARSON EDUCATION FRANCE, 2009