

# Dépendances Fonctionnelles et Normalisation

## 1. Dépendances fonctionnelles

*Définition 1 : Dépendance Fonctionnelle (DF)*

Soit  $R(U)$  une relation avec  $U$  l'ensemble de ses attributs.

Soit  $X, Y \subseteq U$ , i.e  $X$  et  $Y$  sont deux attributs ou ensemble d'attributs de  $R$ .

On dit qu'il existe une DF entre  $X$  et  $Y$  (ou que  $X$  détermine  $Y$  ou encore que  $Y$  est déterminé par  $X$ ), notée  $X \twoheadrightarrow Y$  si et seulement si :

"  $t_1$  et  $t_2$ , deux tuples de  $R$ , si  $t_1[X] = t_2[X]$  alors  $t_1[Y] = t_2[Y]$

Pour une DF  $X \twoheadrightarrow Y$ ,  $X$  est la source et  $Y$  la cible.

Exemple : Considérons l'extension suivante de la relation  $R$

R	A	B	C
	a1	b1	c1
	a1	b1	c2
	a2	b2	c2

Quelles sont les DF existant dans  $R$  ?

Dans une relation, tout attribut est en DF avec la clef primaire

Question : Pourquoi en est-on sûr ?

### Définition 2 : Clef possible ou clef candidate

Soit  $R(U)$  une relation et  $X \subseteq U$ .

$X$  est une clef possible ou candidate si et seulement si on a  $X \twoheadrightarrow Y$

où  $Y = U - X$  (i.e. le complémentaire de  $X$  dans  $U$ ).

La clef primaire d'une relation doit être choisie dans l'ensemble des clefs possibles de la relation.

S'il n'y en a aucune, tous les attributs de la relation constituent sa clef. On parle alors de relation « toute clef ».

Exemple : ETUDIANT avec NUM\_ET et NUM\_SECU.

*Propriétés des DF :  $X$  et  $Y$  sont deux ensembles d'attributs.*

<b>Réflexivité</b>	Si $Y \subseteq X$	alors $X \twoheadrightarrow Y$ (et donc $X \twoheadrightarrow X$ )
<b>Augmentation</b>	Si $X \twoheadrightarrow Y$ et $W \subseteq Z$	alors $X, Z \twoheadrightarrow Y, W$
<b>Transitivité</b>	Si $X \twoheadrightarrow Y$ et $Y \twoheadrightarrow Z$	alors $X \twoheadrightarrow Z$
<b>Pseudo-transitivité</b>	Si $X \twoheadrightarrow Y$ et $Y, Z \twoheadrightarrow W$	alors $X, Z \twoheadrightarrow W$
<b>Union</b>	Si $X \twoheadrightarrow Y$ et $X \twoheadrightarrow Z$	alors $X \twoheadrightarrow Y, Z$
<b>Décomposition</b>	Si $X \twoheadrightarrow Y, Z$	alors $X \twoheadrightarrow Y$ et $X \twoheadrightarrow Z$

### Définition 3 : Dépendance Fonctionnelle Totale (DFT)

Soit  $R(U)$  une relation,  $X, Y \subseteq U$  et  $X \twoheadrightarrow Y$ .

On a une DFT entre  $X$  et  $U$ ,  $X \twoheadrightarrow (DFT) Y$ , si et seulement si

"  $X' \subseteq X$ ,  $X' \twoheadrightarrow Y$  n'est pas vérifiée.

### Définition 4 : Clef minimale

Soit  $R(U)$  une relation,  $X \subseteq U$  et  $X$  est une clef de  $R$ .

On dit que  $X$  est une clef minimale de  $R$  si et seulement si "  $X' \subseteq X$ ,  $X'$  n'est pas une clef de  $R$ .

## 2. Théorie de la normalisation

La normalisation permet de garantir que le schéma relationnel est un bon schéma, i.e. a de bonnes propriétés.

Exemple de mauvais schéma. Considérons la relation ENSEIGNANT suivante :

ENSEIGNANT	NUMERO	NOM	CATEGORIE	CLASSE	SALAIRE
	1	Dupont	Maître Conférences	de 1	12000
	2	Martin	Maître Conférences	de 1	12000

3	Smith	Professeur	2	17000
4	Dupont	Assistant	1	10000
5	Durant	Assistant	1	10000

On a en fait la contrainte suivante : tous les enseignants de même catégorie et de même classe gagent le même salaire, i.e. :

CATEGORIE, CLASSE ® SALAIRE

Il y a dans cette relation redondance de données. Ceci conduit à des anomalies de stockage (anomalies lors des opérations de mise à jours).

· **Anomalie de modification :**

S'il y a une modification du salaire pour MC 1<sup>ere</sup> classe, on fait autant d'opérations de modification qu'il y a d'enseignant de cette catégorie et de cette classe.

· **Anomalie d'insertion :**

Pour pouvoir stocker le salaire des MC 2eme classe, il faut qu'il y ait au moins un enseignant dans cette catégorie et de cette classe.

· **Anomalie de suppression :**

Si l'unique professeur de 2eme classe part, on perd l'information sur le salaire des professeurs de classe 2.

L'objectif de la normalisation est d'éliminer les anomalies de stockage, i.e. d'éliminer la redondance.

La normalisation consiste à éclater une relation en plusieurs. Elle comprend plusieurs étapes :1 NF, 2 NF, 3NF et 3 BCNF (...)

**Première Forme Normale (1 NF)**

Une relation est en 1 NF, si tous ses attributs sont atomiques (mono-valués).

Exemples de relation non 1 NF :

ETUDIANT	NUM	NOM	PRENOM
	1	Dupont	Pierre
			Jean
	2	Durand	Marie
	3	Dupré	Sylvie
			Claudine
			Claire

LIVRE	CODE	TITRE	AUTEUR
	100	L'art des BD	Miranda
			Busta

Comment normaliser en 1 NF ?

Solutions :

1<sup>ère</sup> solution : Créer autant d'attributs que le nombre maximum de valeurs de l'attribut multi-valué (stockage horizontal).

Exemple :

ETUDIANT	NUM	NOM	PRENOM1	PRENOM2	PRENOM3
	1	Dupont	Pierre	Jean	<i>NULL</i>
	2	Durand	Marie	<i>NULL</i>	<i>NULL</i>
	3	Dupré	Sylvie	Claudine	Claire

2<sup>ème</sup> solution : Créer une nouvelle relation comportant la clef de la relation initiale et l'attribut multi-valué puis éliminer l'attributs multi-valué de la relation initiale (stockage vertical).

Exemple :

LIVRE	CODE	TITRE	AUTEURS	CODE	AUTEUR
	100	L'art des BD		100	Miranda
				100	Busta

Remarque : on vient de faire apparaître un couple clef primaire/clef étrangère

Inconvénients :

Solution 1 : stockage de valeurs nulles, impossibilité de stocker plus de valeur que prévu

Solution 2 : opération de jointure, lourdeur des auto-jointures

Avantages :

Solution 1 : tout est dans la même relation (pas de jointure)

Solution 2 : pas de valeur nulle, pas de limite au stockage

Préférer :

Solution 1 : quand le nombre de valeurs de l'attributs multi-valué est constant ou que le nombre max de valeur est faible (très peu de valeurs nulles)

Solution 2 : dans les autres cas

## Deuxième Forme Normale (2 NF)

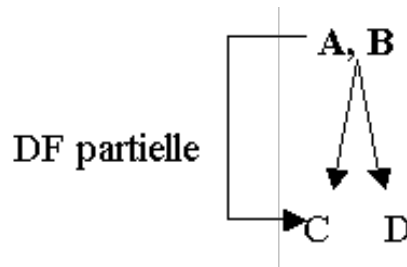
Une relation est en 2 NF si et seulement si :

- Elle est en 1 NF
- Tout attribut n'appartenant pas à la clef primaire est en DF totale avec la clef primaire.

Exemple de relation non 2 NF

ENSEIGNEMENT(NUM, **CODE\_MATIERE**, NOM, VOLUME\_HORAIRE) avec NUM ® NOM

Graphiquement :



Comment normaliser en 2 NF

- Isoler la DF partielle dans une nouvelle relation
- Eliminer l'attribut cible de la DF de la relation initiale

Exemple :

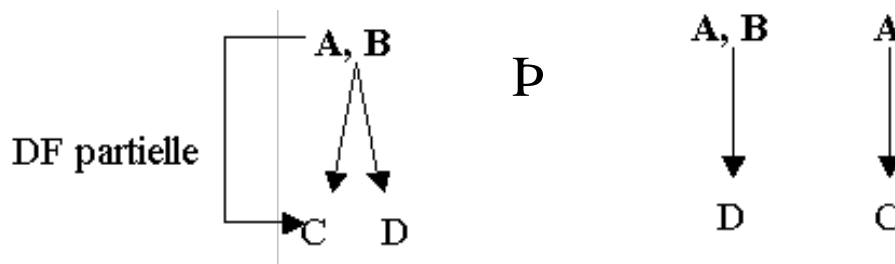
ENSEIGNEMENT( **NUM**, **CODE**, VOLUME\_HORAIRE)

ENSEIGANT( **NUM**, NOM )

Remarque : un couple clef primaire/clef étrangère est apparu

Une relation en 1 NF dont la clef est mono-attribut est forcément en 2 NF

Graphiquement :



### Troisième Forme Normale (3 NF)

Une relation est en 3 NF si et seulement si :

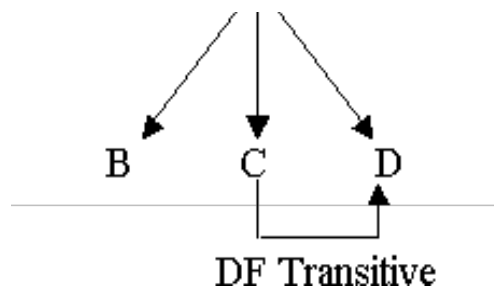
- Elle est en 2 NF
- Elle ne contient pas de DF Transitive entre attributs non clef

Exemple :

ENSEIGNANT(NUM, NOM, CATEGORIE, CLASSE, SALAIRE) avec CATEGORIE, CLASSE ® SALAIRE

Graphiquement :





Comment normaliser en 3 NF :

- Isoler la DF transitive dans la nouvelle relation
- Eliminer la cible de la DF de la relation initiale

Exemple :

ENSEIGNEMENT( **NUM**, NOM, CATEGORIE, CLASSE )

SALAIRE( **CATEGORIE**, **CLASSE**, SALAIRE )

Opération de mise à jour :

- **Modification** : si le salaire d'un MC 1<sup>ère</sup> classe change, une seule modification
- **Insertion** : on peut stocker le salaire des MC 2<sup>ème</sup> classe sans avoir d'enseignant dans cette catégorie
- **Suppression** : si le Professeur en classe 2 part, on conserve le salaire des professeurs en classe 2.

## Forme Normale de Boyce et Codd (BCNF)

Une relation est en BCNF si et seulement si :

- Elle est en 2 NF
- Toute source de DF est une clef primaire minimale

Remarques :

1. si R est en BCNF alors elle est en 3 NF

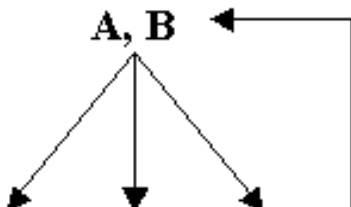
Preuve : Si R est en BCNF alors " X tel que X ® U, X est une clef minimale.

Supposons qu'il existe Y, Z Ì U tel que Y ® Z, mais dans ce cas Y est une clef minimale. On ne peut donc jamais avoir de transitivité avec deux attributs non clef

2. si R est en 3 NF alors elle n'est pas forcément en BCNF

Preuve : cf. contre exemple ci-dessous.

Graphiquement : relation non BCNF



- Attributs atomiques → 1 NF
- Toutes les DF de source AB sont minimales → 2 NF
- Pas de transitivité entre attribut non clef → 3 NF

C̄    D̄    Ē

Comme  $E \twoheadrightarrow B$  et que  $E$  n'est pas une clef minimale, cette relation n'est pas en BCNF.

Exemple de relation non BCNF

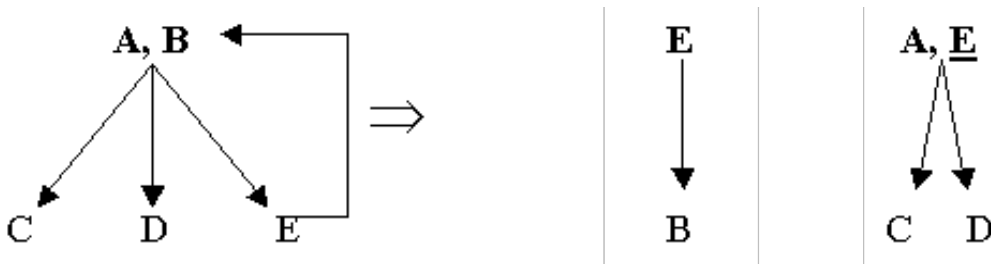
Soit la relation ADRESSE( **Rue**, **Lieu-dit**, **Bureau-Postal**, Code Postal )

- Quelles sont les DF existant dans cette relation ?
- En quelle forme normale est la relation ?

Comment normaliser en BCNF

- Isoler la DF problématique dans une nouvelle relation
- Eliminer la cible de la DF problématique et la remplacer par sa source dans la relation initiale

Graphiquement : relation non BCNF



Exemple :

### **Théorème de décomposition de Casey-Delobel**

A partir de la 2 NF, on normalise en appliquant le théorème suivant.

*Théorème de décomposition :*

Soit  $R(X, Y, Z)$  une relation où  $X, Y, Z$  sont des ensembles d'attributs.

Soit  $X \twoheadrightarrow Y$  une DF vérifiée dans  $R$ . Alors, il existe  $R_1$  et  $R_2$  deux relations telle que  $R_1(X, Y)$  et  $R_2(X, Z)$  et  $R = \text{Jointure}(R_1, R_2 / R_1.X = R_2.X)$

La décomposition de R dans les deux relations R1 et R2 est garantie sans perte d'information.