

***Plus loin que la troisième forme normale :
Pourquoi et comment ?***

Rappels

- Une relation est en 3FN si :
 - pas de DF d'une partie d'une clé vers un attribut non clé (2FN)
 - pas de DF entre attributs non clés.

- Toute relation a au moins une décomposition en 3FN qui :
 - préserve les dépendances fonctionnelles (DF)
 - et est sans perte

□ **Une relation en 3FN peut comporter des redondances**

Exemple : R(VILLE, DEPARTEMENT, CODE POSTAL)

avec les DF : VILLE, DEPARTEMENT -> CODE
 CODE -> DEPARTEMENT

et la clé : VILLE, DEPARTEMENT

les redondances : pour toutes les villes ayant le même code postal, on répète le département.

Problème : une DF d'un attribut non clé vers une partie de la clé.

□ **Forme normale de BOYCE-CODD :**

Les seules DF autorisées sont celles dans lesquelles une clé détermine un attribut.

□ **Exemple :**

R (VILLE, DEPARTEMENT, CODE POSTAL)
n'est pas en BCNF

R peut être décomposée en :

R1 (VILLE, CODE POSTAL)
R2 (DEPARTEMENT, CODE POSTAL)

= Décomposition sans perte

Mais qui ne préserve pas la DF :

VILLE, DEPARTEMENT -> CODE.

R1 et R2 sont en BCNF.

□ Une relation en BCNF peut comporter des redondances

"L'étudiant de numéro NUMETUD pratique le sport SPORT et suit le COURS."

- Pas de DF
- CLE = {NUMETUD, SPORT, COURS}
- R est en 3FN et en BCNF
- Cependant R contient des redondances :

R	NUMETUD	SPORT	COURS
	100	Foot	Math
	100	Foot	Anglais
	200	Foot	Math
	200	Tennis	Anglais
	200	Foot	Anglais
	200	Tennis	Math

□ La notion de *dépendance fonctionnelle* ne suffit à définir toutes les dépendances entre les données.

Dépendance multivaluée (DM) :

R (A1, A2, ..., An)

X sous-ensemble de {A1, A2, ..., An}

Y sous-ensemble de {A1, A2, ..., An}

X \twoheadrightarrow Y : "X multidétermine Y" si, soit Z = R - X - Y,

{(xyz) et (xy'z')} ∈ R ⇒ (xy'z) et (xyz') ∈ R }

"A chaque valeur de X, il y a un ensemble de valeurs de Y associées et cet ensemble est indépendant des autres attributs."

□ **Exemple** : R (NUMETUD, SPORT, COURS)

NUMETUD->>SPORT

NUMETUD->>COURS

<=> Pas de lien entre les cours suivis et les sports pratiqués.

□ **Contre-exemple** : R (NUMEMP, LANGUE, PRODUIT)

"L'employé NUMEMP parle LANGUE et vend PRODUIT".

Si pour vendre un produit, il faut parler la langue du pays où il est distribué, il n'y a pas de dm (dépend. multiv.).

□ **4ème Forme Normale** : une relation est en 4FN si les seules DM sont celles dans lesquelles une clé multidétermine un attribut.

□ **Remarques** :

- une dépendance fonctionnelle est un cas particulier de dépendance multivaluée
- 4FN => 3FN et BCNF
- en fait, on ne considère que les DM élémentaires (parties gauche et droite minimale).

□ Exemple : VINS (BUVEUR, CRU, PRODUCTEUR)

VINS	BUVEUR	CRU	PRODUCTEUR
	Pierre	Chablis	Claude
	Pierre	Chablis	Nicolas
	Pierre	Volnay	Nicolas
	Paul	Chablis	Nicolas

VINS est en 4FN (aucune DM) :

BUVEUR \rightsquigarrow CRU Pierre n'achète pas de Volnay chez Claude

CRU \rightsquigarrow PRODUCTEUR

PRODUCTEUR \rightsquigarrow BUVEUR

VINS	BUVEUR	CRU	PRODUCTEUR
	Pierre	Chablis	Claude
	Pierre	Chablis	Nicolas
	Pierre	Volnay	Nicolas
	Paul	Chablis	Nicolas

Si on décompose VINS en 2 relations à 2 attributs chacune, on ne peut plus retrouver la relation initiale, et ce quelle que soit la décomposition choisie :

R1	BUVEUR	CRU	R2	BUVEUR	PRODUCTEUR
	Pierre	Chablis		Pierre	Claude
	Pierre	Volnay		Pierre	Nicolas
	Paul	Chablis		Paul	Nicolas

R3	CRU	PRODUCTEUR
	Chablis	Claude
	Chablis	Nicolas
	Volnay	Nicolas

VINS \circ R1 \otimes R2
 VINS \circ R1 \otimes R3
 VINS \circ R2 \otimes R3

Pourtant, VINS contient des redondances :

2 fois Pierre Chablis

2 fois Chablis Nicolas

=> Il existe des relations non décomposables en 2 relations, mais en 3 ou plus.

□ **Dépendance de jointure :**

$R(A_1, A_2, \dots, A_n)$

X_1, X_2, \dots, X_m sous-ensembles de $\{A_1, A_2, \dots, A_n\}$

Il existe une dépendance de jointure $\{X_1, X_2, \dots, X_m\}$

si $R = \prod_{X_1}(R) \bowtie \prod_{X_2}(R) \bowtie \dots \prod_{X_m}(R)$

* **Exemple : Relation VINS**

*{BUVEUR CRU, BUVEUR PRODUCTEUR, CRU PRODUCTEUR}

□ S'il existe une telle dépendance de jointure, alors R est décomposable en m relations X_1, X_2, \dots, X_m .

* **Exemple : VINS = R1 \bowtie R2 \bowtie R3**

□ **Remarques :**

– une DM est un cas particulier de DJ (dépend. de jointure) :

$(X \twoheadrightarrow Y \Rightarrow X \twoheadrightarrow Z) \Rightarrow \{XY, XZ\}$

– si R (A_1, A_2, A_3, A_4) a 2 clés A_1 et A_2 , on a les DJ

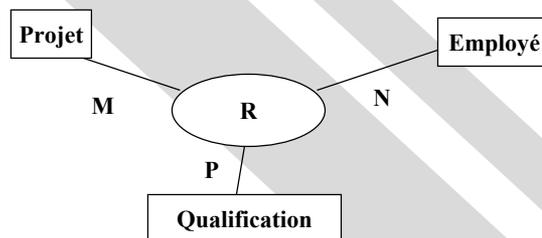
*{ A_1A_2, A_1A_3, A_1A_4 } et *{ A_1A_2, A_2A_3, A_2A_4 }

- **Cinquième Forme Normale** : Une relation est en 5FN si toutes les DJ sont impliquées par les clés.
 - * *Exemple* : VINS n'est pas en 5FN

- **Autre définition** de la 5FN (équivalente) : Une relation est en 5FN s'il n'est pas possible de créer la relation par jointure de relations plus simples avec des clés différentes.
 - * *Problème* : trouver les DJ (elles n'ont pas d'interprétation sémantique simple)
 - * Certaines tables n'ont pas de décomposition en 5FN.

Un autre exemple

Soit la relation ternaire : Projet – Employé - Qualification



1ère interprétation - Qualifications requises

numemp	numproj	qualif
101	3	A
101	3	B
101	4	A
101	4	C
102	3	A
102	3	B
103	5	D

numemp	numproj
101	3
101	4
102	3
103	5

numemp	qualif
101	A
101	B
101	C
102	A
102	B
103	D

numproj	qualif
3	A
3	B
4	A
4	C
5	D

2ème interprétation – Qualifications communes

numemp	numproj	qualif
101	3	A
101	3	B
101	4	A
101	4	B
102	3	A
102	3	B
103	3	A
103	4	A
103	5	A
103	5	C

numemp	numproj
101	3
101	4
102	3
103	3
103	4
103	5

numemp	qualif
101	A
101	B
102	A
102	B
103	A
103	C

numproj	qualif
3	A
3	B
4	A
4	B
5	A
5	C

3ème interprétation - Qualifications utilisées

numemp	numproj	qualif
101	3	A
101	3	B
101	4	A
101	4	C
102	3	A
102	3	B
102	4	A
102	4	B

numemp	numproj
101	3
101	4
102	3
102	4

numemp	qualif
101	A
101	B
101	C
102	A
102	B

numproj	qualif
3	A
3	B
4	A
4	C
4	B