7. Formes normales 4NF et 5NF

i. Dépendance fonctionnelle multivaluée

Exemple : Considérons la relation ETUDIANT (nomE, cours, sport) nomE : nom d'étudiant ; cours : discipline étudiée ; sport : sport pratiqué

ETUDIANT

nomE	cours	sport
DURAND	mathématiques	foot
DUPONT	latin	volley
DUBOIS	physique	escrime
DUVAL	chimie	natation
DUBOIS	chimie	foot
DURAND	physique	volley
DUVAL	mathématiques	natation
DUPONT	mathématiques	volley
DUVAL	physique	foot

$\pi_{nomE,cours}$ (ETUDIANT)

nomE	cours
DURAND	mathématiques
DURAND	physique
DUPONT	latin
DUPONT	mathématiques
DUBOIS	physique
DUBOIS	chimie
DUVAL	chimie
DUVAL	mathématiques
DUVAL	physique

$\pi_{nomE,sport}$ (ETUDIANT)

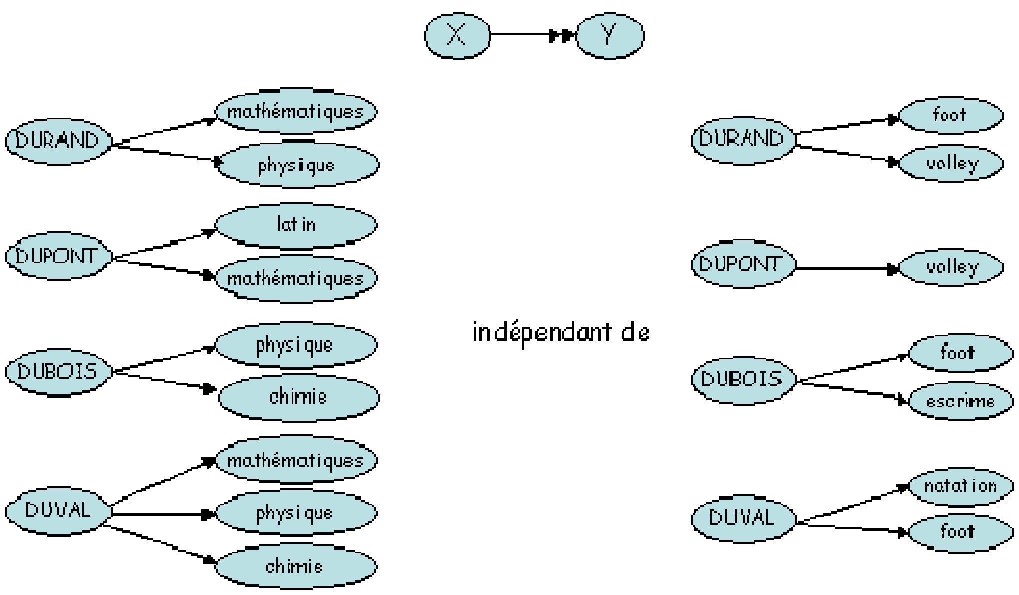
nomE	sport
DURAND	foot
DURAND	volley
DUPONT	volley
DUBOIS	escrime
DUBOIS	foot
DUVAL	natation
DUVAL	foot

nomE ->> cours et **nomE ->> sport** sont des dépendances fonctionnelles multivaluées

Etant donné une relation R et X,Y des ensembles d'attributs de R, on dit qu'il existe une dépendance multivaluée X ->> Y si

a) la connaissance d'une valeur de X détermine un unique ensemble de valeurs de Y

b) cet ensemble est indépendant des autres attributs R-{X,Y} de R



ii. Quatrième forme normale : 4NF

La quatrième forme normale concerne les dépendances multivaluées

Reprenons l'exemple précédent avec ETUDIANT (nomE, cours, sport) où nomE ->> cours, nomE ->> sport sont des dépendances fonctionnelles multivaluées. On notera que ETUDIANT est en BCNF.

Remarque : on notera que pour une relation R(A, B, C), si A ->> B, alors A ->> C. Pour cette raison, on note quelquefois la DM sous la forme A ->> B|C

Théorème de Fagin : Une relation R(A, B, C) pour laquelle A ->> B|C peut être décomposée sans perte en R1(A,B) et R2(A,C).

Appliquons le théorème de Fagin à l'exemple précédent : R1 (numE, cours) et R2 (numE, sport)

R1

nomE	cours
DUBOIS	physique
DUBOIS	chimie
DUPONT	latin
DUPONT	mathématiques
DURAND	mathématiques
DURAND	physique
DUVAL	chimie
DUVAL	mathématiques
DUVAL	physique

R2

nomE	sport
DUBOIS	escrime
DUBOIS	foot
DUPONT	volley
DURAND	foot
DURAND	volley
DUVAL	natation
DUVAL	foot

ETUDIANT n'est pas en 4NF puisque nomE ->> cours n'implique pas nomE -> sport. De même nomE ->> sport n'implique pas nomE -> cours.
R1 et R2 sont en 4NF

Remarque : On notera qu'une relation 4NF est une relation BCNF.

iii. Cinquième forme normale : 5NF

On pourrait penser qu'une relation est décomposable en deux relations. Ceci est faux, il peut arriver qu'une relation soit décomposable en trois relations.

Exemple : considérons la relation R(A, B, C) dans laquelle il n'y a pas de dépendances fonctionnelles (monovaluées ou multivaluées) non triviales.

R

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

La décomposition en deux relations R1(A,B) et R2(B,C) n'est pas une bonne décomposition :

R1

A	B
a1	b1
a1	b2
a2	b1

R2

B	C
b1	c2
b2	c1
b1	c1

R1 ⋈ R2

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

La jointure de R1 et R2 ne redonne pas R et fait apparaître un tuple parasite

Décomposons maintenant $R = \Pi_{AB}(R) \bowtie \Pi_{BC}(R) \bowtie \Pi_{CA}(R)$

Les relations R1, R2, R3 sont en 5ème forme normale. Par exemple il est impossible d'avoir $R1 = \Pi_{A}(R1) \bowtie \Pi_{B}(R1)$. Idem pour R2 et R3.

en trois relations : R1 (A, B), R2(B, C) et R3 (C, A)

R1

A	B
a1	b1
a1	b2
a2	b1

R2

B	C
b1	c2
b2	c1
b1	c1

R3

C	A
c2	a1
c1	a1
c1	a2

$R1 \bowtie R2$

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

$\bowtie$

R3

C	A
c2	a1
c1	a1
c1	a2

=

R

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

On retrouve bien ici la relation originale R.

Lorsqu'une relation R est décomposable en relations R1, R2, R3,… (projections de R sur certains groupes d'attributs), on dit qu'il existe une **dépendance de jointure**. La relation R implique donc une dépendance de jointure.

5ème forme normale :
Une relation R est en 5ème forme normale si elle n'implique pas de dépendance de jointure. La relation R précédente n'est pas en 5ème forme normale. En effet, on constate la dépendance de jointure

On démontre que la forme 5NF implique la forme 4NF et donc en conclusion :