

DEPENDANCES FONCTIONNELLES (DFs)

Définition

Soit $R(U)$ une relation R d'attributs U et X, Y deux sous-ensembles de U .
 Y dépend fonctionnellement de X , noté $X \rightarrow Y$ ssi :

A chaque valeur de X est associée au maximum une valeur de Y .

Exemple

R	A	B	C
	a1	b1	c1
	a1	b1	c2
	a3	b2	c3
	a4	b3	c4

$A \rightarrow C$? **NON** car à la valeur a1 de A est associée deux valeurs c1 et c2 de C .

$A \rightarrow B$? **OUI** car à chaque valeur de A est associé une seule valeur de B
(à a1 est associée b1, à a3 est associée b2 et à a4 est associée b3)

On peut aussi appliquer la méthode par comptage pour vérifier si une DF est valide dans une relation R :

$X \rightarrow Y$ est valide (est vérifiée dans R) ssi :

$$\text{Card}(R(X)) = \text{Card}(R(X, Y))$$

$\text{Card}(R)$ est la cardinalité d'une relation (nombre de tuples distincts)

Dans notre exemple $\text{Card}(R(A)) \neq \text{Card}(R(A, C))$ $3 \neq 4$

Par contre $\text{Card}(R(A)) = \text{Card}((A, B))$ $3 = 3$

$A \not\rightarrow C$ et $A \rightarrow B$

Dépendance fonctionnelle minimale (totale)

$X \rightarrow A$ est une DF minimale (totale) ssi :

Pour tout B in X , $X - B \not\rightarrow A$ (tous les attributs de X sont nécessaires pour déterminer l'attribut A)

Axiomes d'Armstrong simplifiés

Réflexivité	: $X \rightarrow X$
Augmentation	: Si $X \rightarrow Y$ Alors $X, W \rightarrow Y$
Transitivité	: Si $X \rightarrow Y$ et $Y \rightarrow Z$ Alors $X \rightarrow Z$
Pseudo-transitivité	: Si $X \rightarrow Y$ et $Y, Z \rightarrow T$ Alors $X, Z \rightarrow T$
Union	: Si $X \rightarrow Y$ et $X \rightarrow Z$ Alors $X \rightarrow Y, Z$
Décomposition	: Si $X \rightarrow Y, Z$ Alors $X \rightarrow Y$ et $X \rightarrow Z$

Clef candidate minimale

X, un sous ensemble de U est une clé candidate minimale ssi :

- 1) $X \rightarrow U$
- 2) Pour tout A in X, $X - A \not\rightarrow U$ (tous les attributs de X sont nécessaires pour déterminer U)

Il peut exister plusieurs clef candidates minimales dans une relation

Clef primaire

Correspond au choix d'une clef parmi les clefs candidates minimales

Ce choix s'effectue selon les priorités suivantes :

- 1) Priorité aux plus petits nombres d'attributs
- 2) Priorité des types numériques sur les types chaînes de caractères

Recherche d'une clef candidate minimale (et donc d'une clef primaire)

On sait que par réflexivité, $U \rightarrow U$

On commence par cette DF $U \rightarrow U$.

Ensuite on élimine de la partie gauche de la DF précédente les attributs non nécessaires en se basant sur nos DFs et sur les axiomes d'Armstrong.

Exemple

$R(A, B, C, D) F = \{ A, B \rightarrow C, D ; B \rightarrow C, D ; C \rightarrow D \}$

$A, B, C, D \rightarrow A, B, C, D$

A est il nécessaire ? **oui** car A n'est déterminé par aucun attribut (ou un groupe d'attributs)

B est il nécessaire ? **oui** (idem que A)

C est il nécessaire ? **non** car par exemple $B \rightarrow C$. On élimine C.

D est il nécessaire ? **non** car par exemple $B \rightarrow D$ (par transitivité, ne pas raisonner avec C car cet attribut n'existe plus).

(A, B) est clef candidate minimale (et donc clef primaire possible de R)

NORMALISATION EN 3NF

On dispose d'une relation $R(A, B, C, D, \dots)$ avec F un ensemble de DFs

1NF) Une relation R est en 1NF ssi tous ses attributs sont atomiques (simple ou monovalués).

2NF) Une relation R est 2NF ssi elle est 1NF et il n'existe aucune DF interdite valide sur R . Cette DF interdite est de type $X \rightarrow A$ (X est une partie de la clef et A un attribut qui n'est pas dans une clef).

3NF) Une relation R est 3NF ssi elle est 2NF et il n'existe aucune DF interdite valide sur R . Cette DF interdite est de type $X \rightarrow A$ (X n'est pas une clef ou ne contient pas une clef et A un attribut qui n'est pas dans une clef)

Remarque importante

Pour simplifier la normalisation on suppose que la notion de clef dans les définitions correspond à la clef primaire (une clef candidate minimale)

Résumé 2NF, 3NF

2NF

DF interdite de type :

$X \rightarrow A$

- a) X est une partie de la clef primaire
- b) A n'appartient pas à la clef primaire

Exemples de DFs interdites en 2NF

$R(\underline{A}, B, C)$; $A \rightarrow C$ est interdite et $B \rightarrow C$ est interdite

3NF

DF interdite de type :

$X \rightarrow A$

- a) X n'est pas la clef ou ne contient pas la clef primaire
- b) A n'appartient pas à la clef primaire

Exemples de DFs interdites en 3NF

$R(\underline{A}, B, C, D)$

$C \rightarrow D$ est interdite, $D \rightarrow C$ est interdite,

$A, C \rightarrow D$ est interdite, $B, C \rightarrow D$ est interdite,

$A, D \rightarrow C$ est interdite, $B, D \rightarrow C$ est interdite

$B, C \rightarrow D$ est interdite

L'algorithme de normalisation en 3NF

Entrée :

$R(A, B, C, D, \dots)$ et F , un ensemble de DFs

Sortie :

$R_1, R_2, \dots$ Chaque relation est en 3NF

Méthode :

Pour chaque relation R_i N3NF faire

Trouver une DF interdite $X \rightarrow A$;

Maximiser la DF $X \rightarrow A$ (ajouter à A tous les attributs déterminés par X)

Isoler la DF interdite dans une nouvelle relation

$R_j(X, Y)$ avec X clef primaire de R_j

Enlever de la relation R_i le(s) attribut(s) Y

Fin Pour

Sélectionner les relations en 3NF

Indiquer les clefs étrangères

Exemple

$R(A, B, C, D)$ $F = \{ A, B \rightarrow C, D$
 $B \rightarrow C, D$
 $C \rightarrow D \}$

$R(\underline{A}, B, C, D)$ n'est pas en 2NF car $B \rightarrow C$

Maximiser $B \rightarrow C$: $B \rightarrow C, D$

Isoler la DF interdite : $R_1(\underline{B}, C, D)$

Enlever de R les attributs C, D : $R_2(\underline{A}, B)$

$R_2(\underline{A}, B)$ est en 3NF car F ne comporte aucune DF interdite ni en 2NF ni en 3NF.

$R_1(\underline{B}, C, D)$ est 2 NF la clef comporte un seul attribut

R_1 n'est pas en 3NF car $C \rightarrow D$ est interdite en 3NF.

En appliquant le même processus on aura pour finir (les clefs étrangères en italique gras) :

$R_2(\underline{A}, \underline{B})$

$R_{11}(\underline{C}, D)$

$R_{12}(\underline{B}, C)$

Normalisation en 1NF

Soit la relation $R(\underline{A}, B, C^*)$, l'attribut C est multivalué. R est donc en 1NF (Non First Normal Form).

Pour normaliser R en 1NF, il existe deux cas :

Cas 1 : La cardinalité de l'attribut multivalué C est petite. Par exemple C est de type énuméré $\{c1, c2, c3\}$.

Remplacer $R(\underline{A}, B, C^*)$ par $R(\underline{A}, B, C1, C2, C3)$

Cas2 : La cardinalité de l'attribut multivalué C est grande.

Remplacer $R(\underline{A}, B, C^*)$ par deux relations : $R1(\underline{A}, B)$ et $R2(\underline{A}, C)$

Exemple

Cas 1

R	<u>A</u>	B	C*
	a1	b1	{c1, c2, c3}
	a2	b1	{c1}
	a3	b2	
	a4	b3	{c1, c2}

R	<u>A</u>	B	C1	C2	C3
	a1	b1	c1	c2	c3
	a1	b1	c1		
	a3	b2			
	a4	b3	c1	c2	

Cas 2

R	<u>A</u>	B
	a1	b1
	a1	b1
	a3	b2
	a4	b3

R	<u>A</u>	<u>C</u>
	a1	c1
	a1	c2
	a1	c3
	a2	c1
	a4	c1
	a4	c2