



Chapitre 3: Dépendances et normalisation

Enseignant:

Nour H. BEN SLIMEN ATTAOUI

Ecole Supérieure de la Statistiques et de l'Analyse de l'Information (ESSAI)

Besoin de normalisation

La relation: **R (produit,client,adresse,quantité_com,prix_unit,montant)**

produit	client	adresse	Quantité_com	Prix_unit	montant
Shampooing	Dupuis	Lille	10	5	50
Crème	Dupuis	Lille	150	3	450
Mousse à raser	Dupuis	Lille	20	4	80
Shampooing	Martin	Marseille	15	5	75
Mousse à raser	Dubois	Paris	50	4	200
Shampooing	Dupont	Lyon	30	5	150

Remarques

Quand un client fait plusieurs commandes, il y a redondance du couple (client, adresse)

Si le client change d'adresse, il faut changer plusieurs lignes

Augmentation du temps de mise à jour

Augmentation des risques d'incohérence → Anomalie de modification

Si un produit n'a pas été commandé, on ne connaît pas son prix

→ Anomalie d'insertion

Si Martin annule sa commande, on perd son nom et son adresse

Si un produit ne figure que dans une seule ligne et qu'on supprime la ligne

Perte d'informations → Anomalie de suppression

Si on insère la ligne **Shampooing, Dubois, Dijon, 5, 5, 25**, le système ne détectera pas que Dubois a 2 adresses différentes

Une représentation cohérente

4

R1(code_client,client,adresse)

code_client	client	Adresse
CL001	Dupuis	Lille
CL002	Martin	Marseille
CL003	Dubois	Paris
CL004	Dupont	Lyon

R2(code_produit,produit,prix_unit)

code_produit	produit	prix_unit
P0001	Shampooing	5
P0002	Crème	3
P0003	Mousse à raser	4

R3(code_commande,code_client,code_produit,quantité_com)

code_commande	code_client	code_produit	quantité_com
C0001	CL001	P0001	10
C0002	CL001	P0002	150
C0003	CL001	P0003	20
C0004	CL002	P0001	15
C0005	CL003	P0003	50
C0006	CL004	P0001	30

La normalisation

- **But** : fixer des règles permettant de passer de la 1ère représentation à la 2ème
- **Normaliser une relation** : représenter la relation sous une forme qui respecte des critères assurant l'intégrité des données (évitant ainsi au concepteur de la base un certain nombre d'erreurs graves)
- **Dépendance** : contrainte exprimée par une règle que doivent vérifier les données pour que la base de données soit dans un état cohérent
- **Théorie de la normalisation** : repose sur l'analyse des dépendances entre attributs. Elle propose des méthodes systématiques visant à décomposer les relations afin qu'il soit toujours possible de reconstituer la relation originale par l'opération de jointure
- **3 types de dépendances** : dépendances fonctionnelles, dépendances multivaluées et dépendances jonctionnelles

Dépendances fonctionnelles DF

Définition 1

○ Soit R une relation, A l'ensemble des attributs et X, Y des sous-ensembles de A . Par définition, il existe une **dépendance fonctionnelle** entre X et Y (notée $X \rightarrow Y$) si dans la relation R , chaque valeur de X détermine une et une seule valeur de Y

○ **Exemple:** client \rightarrow adresse
(prix_unit, quantité_com) \rightarrow montant

3 relations :

- vol(numVol, depart, arrivee, numav, numpil, jdep, hdep, jarr, harr)
- pilote(numpil, nom, prenom)
- avion(numav, type)

numVol \rightarrow numpil

- numVol est appelé « source de la DF »
- numPilote est appelé « but de la DF »

Dépendances fonctionnelles DF

Définition 2 : clé d'une relation

- Quand dans une relation, un attribut ou un groupe d'attributs est source de dépendances fonctionnelles ayant respectivement pour but chacun des autres attributs de la relation, cet attribut ou groupe d'attributs est appelé **Clé de la relation**

Exemple: vol(numvol,depart,arrivee,numav,numpil,jdep,hdep,jarr,harr)

8 DF de source numvol :

numvol \rightarrow depart
numvol \rightarrow arrivee
numvol \rightarrow numav
numvol \rightarrow numpil

numvol \rightarrow jdep
numvol \rightarrow hdep
numvol \rightarrow jarr
numvol \rightarrow harr

numvol est la clé de la relation vol

- Un attribut est appelé **attribut clé** s'il appartient à au moins une clé de R
- Un attribut est appelé **attribut non clé** s'il n'appartient à aucune clé de R

Dépendances fonctionnelles DF

Définition 3 : DF élémentaire

- Soit X un groupe d'attributs et Y un attribut unique n'appartenant pas à X . La dépendance fonctionnelle $X \rightarrow Y$ est dite **élémentaire** si Y ne dépend pas aussi d'un sous-ensemble de X
- X est la plus petite quantité d'information permettant de déduire Y
- Si une clé est composée de plusieurs attributs, les DF dont la clé est la source sont dites élémentaires si aucun sous-ensemble de la clé ne suffit pour être source de la DF

Dépendances fonctionnelles DF

Définition 4 : DF directe

- Une DF de source A et de but B est dite **directe** s'il n'existe pas d'attribut C tel que

$$A \rightarrow C$$

Et

$$C \rightarrow B$$

Dépendances fonctionnelles DF

Propriétés

Axiomes d'Amstrong

- Réflexivité Si $X \subseteq Y$ alors $Y \rightarrow X$
- Augmentation Si $X \rightarrow Y$ alors $X, W \rightarrow Y, W$
- Transitivité Si $X \rightarrow Y$ et $Y \rightarrow Z$ alors $X \rightarrow Z$

A partir des axiomes, on peut déduire les règles

- 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 $X \rightarrow Z$
- Pseudo-transitivité Si $X \rightarrow Y$ et $W, Y \rightarrow Z$ alors $W, X \rightarrow Z$

Dépendances fonctionnelles DF

fermeture transitive

On appelle **fermeture transitive F^+** d'un ensemble F de DFE, l'ensemble de toutes les DFE qui peuvent être composées par transitivité à partir des DFE de F .

Exemple:

- Soit l'ensemble $F = \{A \rightarrow B, B \rightarrow C, B \rightarrow D, A \rightarrow E\}$.
- La fermeture transitive de F est $F^+ = \{A \rightarrow B, B \rightarrow C, B \rightarrow D, A \rightarrow E, A \rightarrow C, A \rightarrow D\}$

Normalisation des relations

Définition 5- relation normalisée

Définition 6- 1ère Forme Normale

Une relation est **normalisée** si un même attribut n'est pas représenté plusieurs fois (même avec des valeurs différentes) et si un même attribut n'est pas décomposable en d'autres attributs

Toute relation normalisée est dite en **1ère Forme Normale**

Relation non normalisée - exemple

nom	diplôme	enfants	
		prénom	date naissance
Dupuis	bac	Luc	1986
	licence	André	1989
Dupont	bac	Jeanne	1991
	ingénieur	Lucie	1995

La relation n'est pas normalisée

- L'attribut **enfants** est décomposable en 2 attributs
- L'attribut **diplôme** est représenté plusieurs fois pour une même personne

Relation 1ère Forme Normale - exemple

La relation: **vente**(num_client,num_article,nom_client,nom_article) où le couple (**num_client,num_article**) est sa clé

Les DF : (num_client,num_article) \rightarrow nom_client
(num_client,num_article) \rightarrow nom_article

Inconvénients de la 1ère Forme Normale

Certaines dépendances fonctionnelles élémentaires qui devraient exister n'existent pas :

num_client \rightarrow nom_client
num_article \rightarrow nom_article

Normalisation des relations

définition 7 : relation en 2ème Forme Normale

Une relation est en **2ème Forme Normale** si et seulement si elle est en 1FN et si tous les attributs non clés sont pleinement dépendants des clés (toutes les dépendances entre une clé et un attribut non clé sont élémentaires)

ou

Une relation est en **2ème Forme Normale** si toutes les DF issues de la clé sont élémentaires (aucun sous-ensemble de la clé ne suffit pour être source de la DF)

Relation qui n'est pas en 2ème Forme Normale - exemple

- o La relation

vente(num_client,num_article,nom_client,nom_article)
est en 1ère Forme Normale mais pas en 2ème Forme Normale

- o Des DF : $(\underline{\text{num_client}}, \underline{\text{num_article}}) \rightarrow \text{nom_client}$
 $(\underline{\text{num_client}}, \underline{\text{num_article}}) \rightarrow \text{nom_article}$

- o On peut extraire les DF élémentaires

$\text{num_client} \rightarrow \text{nom_client}$
 $\text{num_article} \rightarrow \text{nom_article}$

Relation qui est en 2ème Forme Normale - exemple

num-client → nom-client
num-client → num-représentant
num-client → nom-représentant

- o La relation:

représentant(num_client, nom_client, num_représentant, nom_représentant)

est en 2ème Forme Normale car les 3 DF qui en découlent sont élémentaires

- o Inconvénients de la 2ème forme normale

La DF **num_client → nom_représentant** n'est pas directe puisqu'on a

Num_client → num_représentant et qu'il existe une DF naturelle

Num_représentant → nom_représentant

Normalisation

définition 8 : relation en 3ème Forme Normale

Une relation est en **3ème Forme Normale** si elle est déjà en 2^{ème} Forme Normale et si toutes les DF issues de la clé sont **directes**

- ⇒ Aucune DF n'est issue d'un sous-ensemble de la clé
- ⇒ Aucune DF n'est issue d'un attribut non clé vers un attribut non clé

Relation en 3ème Forme Normale

Exemple

o Les relations:

client(num_client, nom_client, adresse)

vente(num_commande, num_article, quantité)

sont en 3ème Forme Normale

o On a les DF directes et élémentaires:

Num_client → nom_client

Num_client → adresse

Num_commande, num_article → quantité



MERCI