Bases de Données (2ème partie) Normalisation des relations

Cours Deug MIAS 2ème année - I5

Version 1.0 - 7/03/02

E. Desmontils

`Emmanuel.Desmontils@irin.univ-nantes.fr`

Faculté des sciences et des techniques

Dpt-Info, Bureau 203

Plan

- Introduction
- Notion de dépendance fonctionnelle
- Formes normales

Introduction

- Soit la relation suivante : Notes(matière, nom, prenom, notecc, noteex, responsable) ;

Introduction

- Soit la relation suivante : Notes(matière, nom, prenom, notecc, noteex, responsable) ;
- Exploitation coûteuse ;
- Risque de présenter les anomalies suivantes :
  - Redondance (100 étudiants → 100 fois le nom de la matière...),
  - Anomalies d'insertion (impossible de représenter un enseignant qui n'est pas affecté, un étudiant sans notes... soit information absente, soit des «null» partout),
  - Anomalies de suppression (suppression des notes → disparition de la matière et de l'enseignant responsable !),
  - Anomalies de modification (Changer d'enseignant → 100 modification !) ;

Introduction

- *Théorie de la normalisation* : limiter ces problèmes ;
- Analyse des dépendances entre les attributs qui sont à l'origine des phénomènes redondances ;
- Poser des méthodes systématiques pour décomposer une relation en plusieurs relations plus adaptées pour la gestion du problème

  $\rightarrow$ La relation d'origine doit pouvoir être retrouvée par jointures

Notion de DF

- Idée : spécifier les relations entre les attributs (notion de dépendance entre les informations)

Notion de DF

- Idée : spécifier les relations entre les attributs (notion de dépendance entre les informations)
- Définition : Soit $R(\Delta)$ une relation, Δ ses attributs, et $X \subseteq \Delta$ et $Y \subseteq \Delta$, il existe une *dépendance fonctionnelle (DF) entre X et Y* (on dit aussi que *X détermine Y*), notée $X \rightarrow Y$, si dans la relation R chaque valeur de X détermine une et une seule valeur de Y.
- X est appelé *source* et Y *cible*

Notion de DF

- Idée : spécifier les relations entre les attributs (notion de dépendance entre les informations)
- Définition : Soit $R(\Delta)$ une relation, Δ ses attributs, et $X \subseteq \Delta$ et $Y \subseteq \Delta$, il existe une *dépendance fonctionnelle (DF) entre X et Y* (on dit aussi que *X détermine Y*), notée $X \to Y$, si dans la relation R chaque valeur de X détermine une et une seule valeur de Y.
- X est appelé *source* et Y *cible*
- i.e. $X \to Y$: si deux t-uples ont même valeur sur X alors ils ont même valeur sur Y.

Notion de DF

- Idée : spécifier les relations entre les attributs (notion de dépendance entre les informations)
- Définition : Soit $R(\Delta)$ une relation, Δ ses attributs, et $X \subseteq \Delta$ et $Y \subseteq \Delta$, il existe une *dépendance fonctionnelle (DF) entre X et Y* (on dit aussi que *X détermine Y*), notée $X \rightarrow Y$, si dans la relation R chaque valeur de X détermine une et une seule valeur de Y.
- X est appelé *source* et Y *cible*
- i.e. $X \rightarrow Y$: $\forall t_1 \in R$ et $\forall t_2 \in R$, $\Pi_X(t_1) = \Pi_X(t_2) \Rightarrow \Pi_Y(t_1) = \Pi_Y(t_2)$

UNIVERSITÉ DE NANTES

Notion de DF

- $< R(\Delta), \mathcal{F} >$, où $\mathcal{F}$ est l'ensemble des DF associées à la relation R, est appelé *schéma relationnel* (ou *schéma de relation*)
- Attention, une DF est valable sur toutes les valeurs possibles de t-uples et pas seulement sur celles présentes à un moment donné.
- Généralisation de la notion de DF : *dépendances multi-valuées*, notée $X \twoheadrightarrow Y$. A chaque valeur de X n'est pas associée une valeur de Y mais un ensemble de valeurs...

Notion de DF

- Définition : une *dépendance fonctionnelle élémentaire* $X \rightarrow A$ est une DF telle que A ne dépends pas aussi d'un sous-ensemble de X.
- Soit $< R(\Delta), \mathcal{F} >$, $X \subseteq \Delta$, $A \subseteq \Delta$, $X \rightarrow A$ DF élémentaire $\Leftrightarrow \forall X' \subset X$, $\not\exists X' \rightarrow A \in \mathcal{F}$
- Définition : une *dépendance fonctionnelle directe* : $X \rightarrow Y$, $\not\exists Z$ tel que $U \rightarrow Z$ et $Z \rightarrow Y$.
- Possibilité de visualiser les DF élémentaires de $\mathcal{F}$ par un graphe, appelé *graphe de dépendances*.

Exemple

A	B	C	D	E
a1	b1	c1	d3	e2
a1	b1	c3	d4	e3
a2	b2	c4	d2	e1
a3	b1	c1	d3	e2
a2	b2	c4	d2	e1

	$\to$ ou $\not\to$?	
A	?	B
A	?	C
B,C	?	D
A,C	?	D
B	?	D
D	?	E
A,C	?	E

Exemple

A	B	C	D	E
a1	b1	c1	d3	e2
a1	b1	c3	d4	e3
a2	b2	c4	d2	e1
a3	b1	c1	d3	e2
a2	b2	c4	d2	e1

$\rightarrow$ ou $\not\rightarrow$?

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

Exemple

A	B	C	D	E
a1	b1	c1	d3	e2
a1	b1	c3	d4	e3
a2	b2	c4	d2	e1
a3	b1	c1	d3	e2
a2	b2	c4	d2	e1

	$\rightarrow$ ou $\not\rightarrow$?	
A	$\rightarrow$	B
A	$\not\rightarrow$	C
B,C	?	D
A,C	?	D
B	?	D
D	?	E
A,C	?	E

Exemple

A	B	C	D	E
a1	b1	c1	d3	e2
a1	b1	c3	d4	e3
a2	b2	c4	d2	e1
a3	b1	c1	d3	e2
a2	b2	c4	d2	e1

	$\rightarrow$ ou $\not\rightarrow$?	
A	$\rightarrow$	B
A	$\not\rightarrow$	C
B,C	$\rightarrow$	D
A,C	?	D
B	?	D
D	?	E
A,C	?	E

Exemple

A	B	C	D	E
a1	b1	c1	d3	e2
a1	b1	c3	d4	e3
a2	b2	c4	d2	e1
a3	b1	c1	d3	e2
a2	b2	c4	d2	e1

$\rightarrow$ ou $\not\rightarrow$?

A	$\rightarrow$	B
A	$\not\rightarrow$	C
B,C	$\rightarrow$	D
A,C	$\rightarrow$	D
B	?	D
D	?	E
A,C	?	E

Exemple

A	B	C	D	E
a1	b1	c1	d3	e2
a1	b1	c3	d4	e3
a2	b2	c4	d2	e1
a3	b1	c1	d3	e2
a2	b2	c4	d2	e1

$\rightarrow$ ou $\not\rightarrow$?

A	$\rightarrow$	B
A	$\not\rightarrow$	C
B,C	$\rightarrow$	D
A,C	$\rightarrow$	D
B	$\not\rightarrow$	D
D	?	E
A,C	?	E

UNIVERSITÉ DE NANTES

Exemple

A	B	C	D	E
a1	b1	c1	d3	e2
a1	b1	c3	d4	e3
a2	b2	c4	d2	e1
a3	b1	c1	d3	e2
a2	b2	c4	d2	e1

$\rightarrow$ ou $\not\rightarrow$?

A	$\rightarrow$	B
A	$\not\rightarrow$	C
B,C	$\rightarrow$	D
A,C	$\rightarrow$	D
B	$\not\rightarrow$	D
D	$\rightarrow$	E
A,C	?	E

Exemple

A	B	C	D	E
a1	b1	c1	d3	e2
a1	b1	c3	d4	e3
a2	b2	c4	d2	e1
a3	b1	c1	d3	e2
a2	b2	c4	d2	e1

$\rightarrow$ ou $\not\rightarrow$?

$$A \rightarrow B$$
$$A \not\rightarrow C$$
$$B,C \rightarrow D$$
$$A,C \rightarrow D$$
$$B \not\rightarrow D$$
$$D \rightarrow E$$
$$A,C \rightarrow E$$

Règles de Armstrong

- 3 propriétés qui permettent la déduction de nouvelles DF et des preuves.
- Soit $R(\Delta)$ une relation et $W, X, Y, Z \subseteq \Delta$

Règles de Armstrong

- 3 propriétés qui permettent la déduction de nouvelles DF et des preuves.
- Soit $R(\Delta)$ une relation et $W, X, Y, Z \subseteq \Delta$
- *Reflexivité* : $Y \subseteq X \Rightarrow X \rightarrow Y$ (donc $X \rightarrow X$)
- NB Une telle DF est appelée *dépendance fonctionnelle triviale*.

Règles de Armstrong

- 3 propriétés qui permettent la déduction de nouvelles DF et des preuves.
- Soit $R(\Delta)$ une relation et $W, X, Y, Z \subseteq \Delta$
- *Reflexivité* : $Y \subseteq X \Rightarrow X \rightarrow Y$ (donc $X \rightarrow X$)
- NB Une telle DF est appelée *dépendance fonctionnelle triviale*.
- *Augmentation* : $X \rightarrow Y \Rightarrow WX \rightarrow WY$ (NB $WX \equiv W \cup X$)

Règles de Armstrong

- 3 propriétés qui permettent la déduction de nouvelles DF et des preuves.
- Soit $R(\Delta)$ une relation et $W, X, Y, Z \subseteq \Delta$
- *Reflexivité* : $Y \subseteq X \Rightarrow X \rightarrow Y$ (donc $X \rightarrow X$)
- NB Une telle DF est appelée *dépendance fonctionnelle triviale*.
- *Augmentation* : $X \rightarrow Y \Rightarrow WX \rightarrow WY$ (NB $WX \equiv W \cup X$)
- *Transitivité* : $X \rightarrow Y$ et $Y \rightarrow Z \Rightarrow X \rightarrow Z$

Exemple

- Démontrer que $AD \to BE$ en ayant les dépendances suivantes :
  - $A \to B$
  - $B, C \to D$
  - $A, C \to D$
  - $D \to E$
  - $A, C \to E$
- Dem. :

Exemple

- Démontrer que $AD \to BE$ en ayant les dépendances suivantes :
  - $A \to B$
  - $B, C \to D$
  - $A, C \to D$
  - $D \to E$
  - $A, C \to E$
- Dem. :

(1) $A \to B \stackrel{Aug.D}{\Longrightarrow} AD \to BD$

Exemple

- Démontrer que $AD \rightarrow BE$ en ayant les dépendances suivantes :
  - $A \rightarrow B$
  - $B, C \rightarrow D$
  - $A, C \rightarrow D$
  - $D \rightarrow E$
  - $A, C \rightarrow E$
- Dem. :

$$(1)\quad A \rightarrow B \stackrel{Aug.D}{\Longrightarrow} AD \rightarrow BD$$

$$(2)\quad D \rightarrow E \stackrel{Aug.B}{\Longrightarrow} BD \rightarrow BE$$

Exemple

- Démontrer que $AD \rightarrow BE$ en ayant les dépendances suivantes :
  - $A \rightarrow B$
  - $B, C \rightarrow D$
  - $A, C \rightarrow D$
  - $D \rightarrow E$
  - $A, C \rightarrow E$
- Dem. :

$$(1) \quad A \rightarrow B \overset{Aug.D}{\Longrightarrow} AD \rightarrow BD$$

$$(2) \quad D \rightarrow E \overset{Aug.B}{\Longrightarrow} BD \rightarrow BE$$

$$(3) \quad (1) + (2) \overset{Trans.}{\Longrightarrow} AD \rightarrow BE$$

UNIVERSITÉ DE NANTES

Règles de Armstrong (suite)

- Pour simplifier, 3 autres règles qui peuvent se déduire des 3 premières.

Règles de Armstrong (suite)

- Pour simplifier, 3 autres règles qui peuvent se déduire des 3 premières.
- *Union* : $X \rightarrow Y$ et $X \rightarrow Z \Rightarrow X \rightarrow YZ$

Règles de Armstrong (suite)

- Pour simplifier, 3 autres règles qui peuvent se déduire des 3 premières.
- *Union* : $X \rightarrow Y$ et $X \rightarrow Z \Rightarrow X \rightarrow YZ$
- *Décomposition* :
  $X \rightarrow YZ \Rightarrow X \rightarrow Y$ et $X \rightarrow Z$

Règles de Armstrong (suite)

- Pour simplifier, 3 autres règles qui peuvent se déduire des 3 premières.
- *Union* : $X \rightarrow Y$ et $X \rightarrow Z \Rightarrow X \rightarrow YZ$
- *Décomposition* :
  $X \rightarrow YZ \Rightarrow X \rightarrow Y$ et $X \rightarrow Z$
- *Pseudo-transitivité* :
  $X \rightarrow Y$ et $WY \rightarrow Z \Rightarrow WX \rightarrow Z$

Règles de Armstrong (suite)

- Pour simplifier, 3 autres règles qui peuvent se déduire des 3 premières.
- *Union* : $X \rightarrow Y$ et $X \rightarrow Z \Rightarrow X \rightarrow YZ$
- *Décomposition* : $X \rightarrow YZ \Rightarrow X \rightarrow Y$ et $X \rightarrow Z$
- *Pseudo-transitivité* : $X \rightarrow Y$ et $WY \rightarrow Z \Rightarrow WX \rightarrow Z$
- Tout ensemble de DF peut être représenté sous *forme canonique* i.e. où la partie droite des DF n'est formée que d'un seul attribut (conséquence de l'union et la décomposition).

Exemple

- Démontrer l'union :
  $X \to Y$ et $X \to Z \Rightarrow X \to YZ$
- Hypothèse : $X \to Y$ et $X \to Z$

Exemple

- Démontrer l'union :
  $X \to Y$ et $X \to Z \Rightarrow X \to YZ$
- Hypothèse : $X \to Y$ et $X \to Z$

(1) $X \to Y \stackrel{Aug.X}{\Longrightarrow} XX \to XY$

Exemple

- Démontrer l'union :
  $X \to Y$ et $X \to Z \Rightarrow X \to YZ$
- Hypothèse : $X \to Y$ et $X \to Z$

$$(1)\ \ X \to Y \stackrel{Aug.X}{\Longrightarrow} XX \to XY \equiv X \to XY$$

Exemple

- Démontrer l'union :
  $X \rightarrow Y$ et $X \rightarrow Z \Rightarrow X \rightarrow YZ$
- Hypothèse : $X \rightarrow Y$ et $X \rightarrow Z$

$$(1)\quad X \rightarrow Y \stackrel{Aug.X}{\Longrightarrow} XX \rightarrow XY \equiv X \rightarrow XY$$

$$(2)\quad X \rightarrow Z \stackrel{Aug.Y}{\Longrightarrow} XY \rightarrow ZY$$

Exemple

- Démontrer l'union :
  $X \to Y$ et $X \to Z \Rightarrow X \to YZ$
- Hypothèse : $X \to Y$ et $X \to Z$

$$(1)\ \ X \to Y \stackrel{Aug.X}{\Longrightarrow} XX \to XY \equiv X \to XY$$

$$(2)\ \ X \to Z \stackrel{Aug.Y}{\Longrightarrow} XY \to ZY$$

$$(3)\ \ (1) + (2) \stackrel{Trans.}{\Longrightarrow} X \to ZY$$

Exemple

- Démontrer l'union :
  $X \to Y$ et $X \to Z \Rightarrow X \to YZ$
- Hypothèse : $X \to Y$ et $X \to Z$

$$(1)\ \ X \to Y \overset{Aug.X}{\Longrightarrow} XX \to XY \equiv X \to XY$$

$$(2)\ \ X \to Z \overset{Aug.Y}{\Longrightarrow} XY \to ZY$$

$$(3)\ \ (1) + (2) \overset{Trans.}{\Longrightarrow} X \to ZY \equiv X \to YZ$$

Autre exemple

- Soit la table :
  notes(matiere, nom, prenom,
  noetu, notecc, noteex, responsable)
- $noetu \rightarrow nom$
- $noetu \rightarrow prenom$
- $matiere \rightarrow responsable$
- $matiere, noetu \rightarrow notecc$...
- Mais attention
  $matiere, nom, prenom \not\rightarrow notecc$!
- En effet, il faudrait $nom, prenom \rightarrow noetu$
  (ce qui n'est pas vrai à cause des homonymes)

Fermeture

- Idée : identifier tous les attributs qui dépendent d'autres attributs

Fermeture

- Idée : identifier tous les attributs qui dépendent d'autres attributs
- Définition : la *fermeture* d'un ensemble $\mathcal{F}$ de DF (ou *clôture*), notée $\mathcal{F}^+$, est l'ensemble de toutes les DF qui sont conséquences de $\mathcal{F}$ par l'application des règles d'Armstrong.

Fermeture

- Idée : identifier tous les attributs qui dépendent d'autres attributs
- Définition : la *fermeture* d'un ensemble $\mathcal{F}$ de DF (ou *clôture*), notée $\mathcal{F}^+$, est l'ensemble de toutes les DF qui sont conséquences de $\mathcal{F}$ par l'application des règles d'Armstrong.
- $\forall f \in \mathcal{F}^+,\ \mathcal{F} \models f$

Fermeture

- Idée : identifier tous les attributs qui dépendent d'autres attributs
- Définition : la *fermeture* d'un ensemble $\mathcal{F}$ de DF (ou *clôture*), notée $\mathcal{F}^+$, est l'ensemble de toutes les DF qui sont conséquences de $\mathcal{F}$ par l'application des règles d'Armstrong.
- $\forall f \in \mathcal{F}^+,\ \mathcal{F} \models f$
- NB : souvent lourd à calculer !

Fermeture

- Deux ensembles de DF, df_1 et df_2, sont *équivalents* si leurs fermetures sont égales :
  $df_1 \equiv df_2 \Leftrightarrow df_1^+ = df_2^+$
- $\{A \rightarrow B, B \rightarrow C, B \rightarrow D\}$
  $\equiv \{A \rightarrow B, A \rightarrow D, B \rightarrow CD\}$

Fermeture

- Deux ensembles de DF, df_1 et df_2, sont *équivalents* si leurs fermetures sont égales : $df_1 \equiv df_2 \Leftrightarrow df_1^+ = df_2^+$
- $\{A \to B, B \to C, B \to D\}$ $\equiv \{A \to B, A \to D, B \to CD\}$
- Pour tout ensemble de DF, il existe un ensemble de DF élémentaires non triviales équivalent.

Fermeture

- Deux ensembles de DF, df_1 et df_2, sont *équivalents* si leurs fermetures sont égales : $df_1 \equiv df_2 \Leftrightarrow df_1^+ = df_2^+$
- $\{A \rightarrow B, B \rightarrow C, B \rightarrow D\}$ $\equiv \{A \rightarrow B, A \rightarrow D, B \rightarrow CD\}$
- Pour tout ensemble de DF, il existe un ensemble de DF élémentaires non triviales équivalent.
- La *fermeture d'un ensemble d'attributs* X est l'ensemble des attributs déterminés par X relativement à l'ensemble des DF.

Notions de clé

- Définition : une *clé* (ou *superclé*) de $R(\Delta)$ munie de l'ensemble de DF $\mathcal{F}$ est un groupe d'attributs X tel que : $X \to \Delta \in \mathcal{F}^+$.

Notions de clé

- Définition : une *clé* (ou *superclé*) de $R(\Delta)$ munie de l'ensemble de DF $\mathcal{F}$ est un groupe d'attributs X tel que : $X \rightarrow \Delta \in \mathcal{F}^+$.
- Définition : une *clé minimale* (ou *clé candidate*) X est une clé sur $< R(\Delta), \mathcal{F} >$ telle que : $X \rightarrow \Delta \in \mathcal{F}^+$ et $\nexists Y \subset X$ tel que $Y \rightarrow \Delta \in \mathcal{F}^+$
- NB : une relation peut comporter plusieurs clés candidates

Notions de clé

- Définition : une *clé* (ou *superclé*) de $R(\Delta)$ munie de l'ensemble de DF $\mathcal{F}$ est un groupe d'attributs X tel que : $X \rightarrow \Delta \in \mathcal{F}^+$.
- Définition : une *clé minimale* (ou *clé candidate*) X est une clé sur $< R(\Delta), \mathcal{F} >$ telle que : $X \rightarrow \Delta \in \mathcal{F}^+$ et $\nexists Y \subset X$ tel que $Y \rightarrow \Delta \in \mathcal{F}^+$
- NB : une relation peut comporter plusieurs clés candidates
- Définition : une *clé primaire* X de $< R(\Delta), \mathcal{F} >$ est la clé candidate privilégiée pour la relation R.

Couverture irredondante

- Objectif : gérer les DF, i.e. simplifier, supprimer les redondances... $\Rightarrow$ ensemble équivalent mais plus pertinent de DF
- Définition : une *dépendance fonctionnelle redondante* est une DF élémentaire obtenue par transitivité à partir des autres DF.

Couverture irredondante

- Objectif : gérer les DF, i.e. simplifier, supprimer les redondances... $\Rightarrow$ ensemble équivalent mais plus pertinent de DF
- Définition : une *dépendance fonctionnelle redondante* est une DF élémentaire obtenue par transitivité à partir des autres DF.
- Définition : une *couverture irredondante* (ou *couverture minimale*) $Irr(\mathcal{F})$ est une partie de l'ensemble $\mathcal{F}$ de DF élémentaires telle que :
  - Toutes les dépendances élémentaires sont dans $Irr(\mathcal{F})^+$
  - Aucune DF de $Irr(\mathcal{F})$ n'est redondante

Couverture irredondante

- Objectif : gérer les DF, i.e. simplifier, supprimer les redondances... $\Rightarrow$ ensemble équivalent mais plus pertinent de DF
- Définition : une *dépendance fonctionnelle redondante* est une DF élémentaire obtenue par transitivité à partir des autres DF.
- Définition : une *couverture irredondante* (ou *couverture minimale*) $Irr(\mathcal{F})$ est une partie de l'ensemble $\mathcal{F}$ de DF élémentaires telle que :
  - Toutes les dépendances élémentaires sont dans $Irr(\mathcal{F})^+$
  - Aucune DF de $Irr(\mathcal{F})$ n'est redondante
- Cette couverture irredondante n'est pas unique

Couverture irredondante

- Objectif : gérer les DF, i.e. simplifier, supprimer les redondances... $\Rightarrow$ ensemble équivalent mais plus pertinent de DF
- Définition : une *dépendance fonctionnelle redondante* est une DF élémentaire obtenue par transitivité à partir des autres DF.
- Définition : une *couverture irredondante* (ou *couverture minimale*) $Irr(\mathcal{F})$ est une partie de l'ensemble $\mathcal{F}$ de DF élémentaires telle que :
  - Toutes les dépendances élémentaires sont dans $Irr(\mathcal{F})^+$
  - Aucune DF de $Irr(\mathcal{F})$ n'est redondante
- $Irr(\mathcal{F})^+ = \mathcal{F}^+$ ou $Irr(\mathcal{F}) \equiv \mathcal{F}$

Couverture irredondante

- Objectif : gérer les DF, i.e. simplifier, supprimer les redondances... $\Rightarrow$ ensemble équivalent mais plus pertinent de DF
- Définition : une *dépendance fonctionnelle redondante* est une DF élémentaire obtenue par transitivité à partir des autres DF.
- Définition : une *couverture irredondante* (ou *couverture minimale*) $Irr(\mathcal{F})$ est une partie de l'ensemble $\mathcal{F}$ de DF élémentaires telle que :
  - Toutes les dépendances élémentaires sont dans $Irr(\mathcal{F})^+$
  - Aucune DF de $Irr(\mathcal{F})$ n'est redondante
- $\nexists \mathcal{F}' \subset Irr(\mathcal{F}),\ \mathcal{F}' \equiv \mathcal{F}$

Exemple

- Soit l'ensemble
  $\mathcal{F} = \{A \rightarrow B, BC \rightarrow D, D \rightarrow E, AC \rightarrow DE\}$

Exemple

- Soit l'ensemble $\mathcal{F} = \{A \to B, BC \to D, D \to E, AC \to DE\}$
- $\mathcal{F} \equiv \mathcal{F}' = \{A \to B, BC \to D, D \to E, AC \to D, AC \to E\}$ (forme canonique de $\mathcal{F}$)

Exemple

- Soit l'ensemble $\mathcal{F} = \{A \rightarrow B, BC \rightarrow D, D \rightarrow E, AC \rightarrow DE\}$
- $\mathcal{F} \equiv \mathcal{F}' = \{A \rightarrow B, BC \rightarrow D, D \rightarrow E, AC \rightarrow D, AC \rightarrow E\}$ (forme canonique de $\mathcal{F}$)
- $AC \rightarrow D : A \rightarrow B \stackrel{Aug.C}{\Longrightarrow} AC \rightarrow BC$

Exemple

- Soit l'ensemble $\mathcal{F} = \{A \rightarrow B, BC \rightarrow D, D \rightarrow E, AC \rightarrow DE\}$
- $\mathcal{F} \equiv \mathcal{F}' = \{A \rightarrow B, BC \rightarrow D, D \rightarrow E, AC \rightarrow D, AC \rightarrow E\}$ (forme canonique de $\mathcal{F}$)
- $AC \rightarrow D : A \rightarrow B \stackrel{Aug.C}{\Longrightarrow} AC \rightarrow BC$
- $AC \rightarrow BC + BC \rightarrow D \stackrel{Trans.}{\Longrightarrow} AC \rightarrow D$

Exemple

- Soit l'ensemble $\mathcal{F} = \{A \rightarrow B, BC \rightarrow D, D \rightarrow E, AC \rightarrow DE\}$
- $\mathcal{F} \equiv \mathcal{F}' = \{A \rightarrow B, BC \rightarrow D, D \rightarrow E, AC \rightarrow D, AC \rightarrow E\}$ (forme canonique de $\mathcal{F}$)
- $AC \rightarrow D : A \rightarrow B \stackrel{Aug.C}{\Longrightarrow} AC \rightarrow BC$
- $AC \rightarrow BC + BC \rightarrow D \stackrel{Trans.}{\Longrightarrow} AC \rightarrow D$
- donc cette dépendance est redondante

Exemple

- Soit l'ensemble $\mathcal{F} = \{A \rightarrow B, BC \rightarrow D, D \rightarrow E, AC \rightarrow DE\}$
- $\mathcal{F} \equiv \mathcal{F}' = \{A \rightarrow B, BC \rightarrow D, D \rightarrow E, AC \rightarrow D, AC \rightarrow E\}$ (forme canonique de $\mathcal{F}$)
- $AC \rightarrow D : A \rightarrow B \stackrel{Aug.C}{\Longrightarrow} AC \rightarrow BC$
- $AC \rightarrow BC + BC \rightarrow D \stackrel{Trans.}{\Longrightarrow} AC \rightarrow D$
- donc cette dépendance est redondante
- $AC \rightarrow D + D \rightarrow E \stackrel{Trans.}{\Longrightarrow} AC \rightarrow E$

Exemple

- Soit l'ensemble $\mathcal{F} = \{A \rightarrow B, BC \rightarrow D, D \rightarrow E, AC \rightarrow DE\}$
- $\mathcal{F} \equiv \mathcal{F}' = \{A \rightarrow B, BC \rightarrow D, D \rightarrow E, AC \rightarrow D, AC \rightarrow E\}$ (forme canonique de $\mathcal{F}$)
- $AC \rightarrow D : A \rightarrow B \stackrel{Aug.C}{\Longrightarrow} AC \rightarrow BC$
- $AC \rightarrow BC + BC \rightarrow D \stackrel{Trans.}{\Longrightarrow} AC \rightarrow D$
- donc cette dépendance est redondante
- $AC \rightarrow D + D \rightarrow E \stackrel{Trans.}{\Longrightarrow} AC \rightarrow E$
- donc cette dépendance est aussi redondante

Exemple

- Soit l'ensemble $\mathcal{F} = \{A \rightarrow B, BC \rightarrow D, D \rightarrow E, AC \rightarrow DE\}$
- $\mathcal{F} \equiv \mathcal{F}' = \{A \rightarrow B, BC \rightarrow D, D \rightarrow E, AC \rightarrow D, AC \rightarrow E\}$ (forme canonique de $\mathcal{F}$)
- $AC \rightarrow D : A \rightarrow B \overset{Aug.C}{\Longrightarrow} AC \rightarrow BC$
- $AC \rightarrow BC + BC \rightarrow D \overset{Trans.}{\Longrightarrow} AC \rightarrow D$
- donc cette dépendance est redondante
- $AC \rightarrow D + D \rightarrow E \overset{Trans.}{\Longrightarrow} AC \rightarrow E$
- donc cette dépendance est aussi redondante
- Couverture minimale $= \{A \rightarrow B, BC \rightarrow D, D \rightarrow E\}$

Décomposition d’une relation

- On peut toujours décomposer une relation suivant une DF

Décomposition d'une relation

- On peut toujours décomposer une relation suivant une DF
- On ne peut pas décomposer une relation sans DF

Décomposition d'une relation

- On peut toujours décomposer une relation suivant une DF
- On ne peut pas décomposer une relation sans DF
- La décomposition suivant une DF ne perd pas d'information

Décomposition d'une relation

- On peut toujours décomposer une relation suivant une DF
- On ne peut pas décomposer une relation sans DF
- La décomposition suivant une DF ne perd pas d'information
- Une table peut être décomposée en deux tables selon une DF telles que l'une des tables contient tous les attributs de la DF et l'autre, tous les attributs de la table d'origine sauf ceux à droite de la DF.

Décomposition d'une relation

- On peut toujours décomposer une relation suivant une DF
- On ne peut pas décomposer une relation sans DF
- La décomposition suivant une DF ne perd pas d'information
- Une table peut être décomposée en deux tables selon une DF telles que l'une des tables contient tous les attributs de la DF et l'autre, tous les attributs de la table d'origine sauf ceux à droite de la DF.
- Si on a $< R(X, Y, Z), \{X \to Y\} >$ alors $R(X, Y, Z) = R_1(X, Y) \bowtie R_2(X, Z)$

Formes Normales

- Déterminer la «qualité» d'une relation (par rapport à la redondance...) en utilisant les dépendances fonctionnelles.
- 6 formes normales : contraintes de plus en plus fortes sur la structure des relations
- Dans le cadre de ce cours : les 3 premières

Première forme normale

- Définition : une relation est dite en *première forme normale* (1NF) si tous ses attributs sont atomiques.
- Les attributs ne peuvent pas être décomposés du point de vue du contexte dans lequel est envisagé la relation.
- Soit la relation :

notes	matiere	etudiant	notes
	I5	Albert Paul	8, 12.5
	I5	Duplo Bertrand	2.5, 0, 18

- Pas un 1NF ! : notes(matiere,nom,prenom,note1,note2,note3) ou ...

Deuxième forme normale

- Définition : une relation est dite en *deuxième forme normale* (2FN) si elle est en 1FN et qu'un attribut n'appartenant à aucune clé soit en DF élémentaire avec toutes les clés.
- i.e. : $< R(A, B, C), \{AB \rightarrow C, B \rightarrow C\} >$ pas en 2FN.
- Pour passer en 2FN, il suffit d'appliquer une décomposition de la relation basée sur $B \rightarrow C$.
- D'où $< R_1(A, B), \{AB \rightarrow B\} >$ et $< R_2(B, C), \{B \rightarrow C\} >$

Troisième forme normale

- Définition : une relation est dite en *troisième forme normale* (3FN) lorsqu'elle est en 2FN et que tout attribut n'appartenant pas à une clé ne dépend pas d'un attribut non clé.
- Toutes les DF sont directes.
- Toutes attribut n'appartenant pas à une clé est en DF élémentaire directe avec la clé.
- Dès que deux attributs sont en DF et qu'ils ne font pas partie d'une clé alors ils ne sont pas en 3FN.

Troisième forme normale

- i.e. : $< R(A, B, C), \{A \rightarrow BC, B \rightarrow C\} >$ pas en 3FN.
- Pour passer en 3FN, il suffit de décomposer la relation avec $B \rightarrow C$
- D'où $< R_1(A, B), \{A \rightarrow B\} >$ et $< R_2(B, C), \{B \rightarrow C\} >$
- Toute relation admet une décomposition en 3FN à jonction conservative (sans perte) et avec préservation des DF.

Troisième forme normale de Boyce-Codd

- Définition : une relation est dite en *troisième forme normale de Boyce-Codd* si elle est en 3FN et que les seules DF élémentaires sont celles dans lesquelles une clé candidate détermine un attribut (la source de chaque DF).
- i.e. : $< R(A, B, C), \{AB \rightarrow C, C \rightarrow B\} >$ pas en 3FNBC.
- Pour obtenir la 3FNBC, il peut y avoir perte de DF lors des décompositions.

Exemple

- Notes(*matiere*,nom, prenom, *noe*, notecc, noteex, responsable, nobureau)
- 1FN
- (1) $matiere \rightarrow responsable$
  (2) $noe \rightarrow nom, prenom$
  (3) $matiere, noe \rightarrow notecc, noteex$
  (4) $matiere, noe \rightarrow nom, prenom$
  (5) $matiere, noe \rightarrow responsable$
  (6) $responsable \rightarrow nobureau$
- (1 et 5) ou (2 et 4) $\Rightarrow$ pas en 2FN !

Exemple

- Notes(*matiere*,notecc, noteex,*noe*),
- Etudiant(nom, prenom, *noe*)
- Matiere(*matiere*, responsable, nobureau)
- 2FN
- (1) $matiere \rightarrow responsable$
  (2) $noe \rightarrow nom, prenom$
  (3) $matiere, noe \rightarrow notecc, noteex$
  (4) $matiere, noe \rightarrow nom, prenom$
  (5) $matiere, noe \rightarrow responsable$
  (6) $responsable \rightarrow nobureau$
- (1 et 6) $\Rightarrow$ pas en 3FN !

Exemple

- Notes(*matiere*,notecc, noteex,*noe*),
- Etudiant(nom, prenom, *noe*)
- Matiere(*matiere*, noresp)
- Responsable(*noresp*,nobureau)
- 3FN
- (1) $matiere \rightarrow responsable$
  (2) $noe \rightarrow nom, prenom$
  (3) $matiere, noe \rightarrow notecc, noteex$
  (4) $matiere, noe \rightarrow nom, prenom$
  (5) $matiere, noe \rightarrow responsable$
  (6) $responsable \rightarrow nobureau$
- Bien de normaliser par parfois préférable de ne pas le faire pour des raisons de performance (efficacité)

Décomposition des relations

- Idée : Découper les relations qui ne sont pas en 3FN pour n'obtenir que des relations en 3FN.

Décomposition des relations

- Idée : Découper les relations qui ne sont pas en 3FN pour n'obtenir que des relations en 3FN.
- Solution 1 : algorithmes de décomposition (obtention de tables en 3FNBC-4FN) :
  - décomposition progressive des relations en fonctions de DF.
  - La fin des décompositions est détectée lorsque soit toutes les DF ont été utilisées soit toutes les relations sont en 3FN.

Décomposition des relations

- Idée : Découper les relations qui ne sont pas en 3FN pour n'obtenir que des relations en 3FN.
- Solution 2 : algorithmes de synthèse (obtention de tables en 3FN). Soit $< R(\Delta), \mathcal{F} >$,
  - Calculer $\mathcal{I} = Irr(\mathcal{F})$.
  - Partitionner $\mathcal{I}$ en $\mathcal{I}_1$ ...$\mathcal{I}_n$ où chaque $\mathcal{I}_j$ possède des DF ayant la même partie gauche.
  - Pour chaque $\mathcal{I}_j$, construire $< R_j(\Delta_j), \mathcal{I}_j >$ où Δ_j est l'ensemble des attributs apparaissant dans $\mathcal{I}_j$.

Décomposition des relations

- Idée : Découper les relations qui ne sont pas en 3FN pour n'obtenir que des relations en 3FN.
- Attention, les algorithmes par décomposition peuvent perdre des DF et ceux par synthèse, de l'information.

Remarques sur la base du cours

- Récupérer sur «`http://www/~desmonti/login_sql`».
- Faire "source login_sql"
- Se connecter à sql "sqlplus"
- Donner un nom d'utilisateur entre "i501" et "i518"
- Entrer le mot de passe : le même que le nom d'utilisateur
- Prompt "SQL>"
- Entrer les requêtes... le nom des tables est préfixé par "i5su". Exemple : "i5su.etudiants".

Remarques sur la base du cours

- Possibilité de faire un fichier texte contenant les requêtes. Cela permet de les faire ensemble et de les corriger.
- Éditer le fichier.
- Sous SQL taper : "start nom_fichier.sql;" pour exécuter les requêtes contenues dans le fichier "mon_fichier.sql".
- Pour enregistrer les résultats dans un fichier texte faire :
  "spool res.lst;"
  "start mon_fichier.sql;"
  "spool off;"
- Terminer SQL : "exit;"