L1 Sciences Fondamentales, S2
Normalisation

Jean-François Couchot

Université de Franche-Comté, UFR-ST

Plan

Introduction

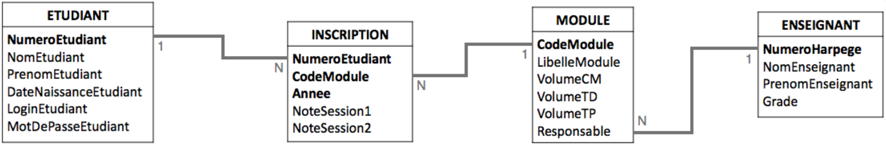
1NF, 2NF, 3NF, BCNF

4NF

Plan

Plan

Une seule relation universelle : exemple



Pourquoi pas n'utiliser qu'une seule relation ?

`UNIVERSELLE(NumeroEtudiant*, NomEtudiant, PrenomEtudiant, DateNaissanceEtudiant, LoginEtudiant, MotDePasseEtudiant, CodeModule*, LibelleModule, VolumeCM, VolumeTD, VolumeTP, NumeroHarpege, NomEnseignant,PrenomEnseignant, Grade, Annee*,NoteSession1,NoteSession2)`

NumeroEtudiant*	NomEtudiant	...	CodeModule*	LibelleModule	...	NumeroHarpege	Annee	...
23794	Dornier	...	BD_L1	Bases de données	...	7914	2009	...
23794	Dornier	...	BD_L1	Bases de données	...	7914	2010	...
23794	Dornier	...	PROG_L1	Programmation	...	7358	2009	...
32911	Martin	...	BD_L1	Bases de données	...	7914	2010	...
32911	Martin	...	PROG_L1	Programmation	...	7358	2010	...
...	...	...	...	...	...	...	...	...

Conséquences générales sur les données

- Intégrité : un libellé de module unique pour un CodeModule ?
- Redondance : (NumeroEtudiant, NomEtudiant,...) répété de nombreuses x

Une seule relation universelle : anomalies

NumeroEtudiant*	NomEtudiant	...	CodeModule*	LibelleModule	...	NumeroHarpege	Annee	...
23794	Dornier	...	BD_L1	Bases de données	...	7914	2009	...
23794	Dornier	...	BD_L1	Bases de données	...	7914	2010	...
23794	Dornier	...	PROG_L1	Programmation	...	7358	2009	...
32911	Martin	...	BD_L1	Bases de données	...	7914	2010	...
32911	Martin	...	PROG_L1	Programmation	...	7358	2010	...
...	...	...	...	...	...	...	...	...

Insertions impossibles sauf si NULL autorisé

- Etudiant inscrit dans aucun module
- Enseignant responsable d'aucun module
- Module sans inscription

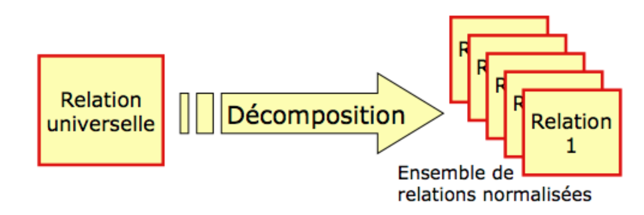
Suppressions impossibles

- Etudiant, seul inscrit dans un module : perte des informations sur le module relatives au module. . .

Normalisation : démarches possibles

Approche par décomposition : vue en L1

1. A partir de toutes les données du SI : construction d'une relation universelle
2. Décomposition de cette relation universelle en relations (1NF, 2NF...) sans anomalie jusqu'à obtention d'un schéma normalisé



Approche méthodique de construction : vue en L2

1. Construction d'un schéma relationnel avec méthode (dictionnaire des données, matrice de dépendance fonctionnelle).
2. Validation

Décomposition : intuition $\rightsquigarrow$ perte possible

Décomposition de $R(Att_1, Att_2, \dots)$

- Choix des ss-relations : $R_1(Att_{R_{1,1}}, Att_{R_{1,2}}, \dots)$, $R_2(Att_{R_{2,1}}, Att_{R_{2,2}}, \dots)$, …tq. $\{Att_{R_{1,1}}, Att_{R_{1,2}}, \dots\} \cup \{Att_{R_{2,1}}, Att_{R_{2,2}}, \dots\} \cup \dots = \{Att_1, Att_2, \dots\}$
- Projection de R, sur les attributs de R_1, de R_2 pour construire R_1, R_2,…

Exemples de décomposition

NumeroEtudiant	NomEtudiant	DateNaissanceEtudiant
27750	Martin	2-09-1989
38911	Martin	26-02-1998

Décomposition 1

NumeroEtudiant	NomEtudiant
27750	Martin
32911	Martin

NumeroEtudiant	DateNaissanceEtudiant
27750	2-09-1989
32911	26-02-1998

Jointure selon NumeroEtudiant $\rightsquigarrow$ R (pas de perte)

Décomposition 2

NumeroEtudiant	NomEtudiant
27750	Martin
32911	Martin

NomEtudiant	DateNaissanceEtudiant
Martin	2-09-1989
Martin	26-02-1998

Jointure selon NomEtudiant $\rightsquigarrow$ perte !

NumeroEtudiant	NomEtudiant	DateNaissanceEtudiant
27750	Martin	2-09-1989
27750	Martin	26-02-1998
32911	Martin	2-09-1989
32911	Martin	26-02-1998

Bilan des objectifs

Démarche générale de normalisation

- Pour éviter les anomalies
- Par décomposition d'une relation universelle
- Sans perte d'information

Comment décomposer sans perte ?

- En ne regroupant dans une relation que les attributs “fortement liés”
- En éliminant le maximum de redondance et ne garder que le stricte nécessaire
- Notion de dépendance fonctionnelle

Plan

Dépendance fonctionnelle : définition et exemple

Définition pour X et Y deux sous-ensembles des attributs d'une rel. R

- Y est en *dépendance fonctionnelle* avec X si pour chaque valeur de X, on a une seule valeur de Y dans R
- Notation : $X \rightarrow Y$
- X et Y : nommés *déterminant* et *ensemble dépendant* resp.

Exemple de dépendance fonctionnelle

NumeroEtudiant	NomEtudiant	DateNaissanceEtudiant
27750	Martin	2-09-1989
32911	Martin	26-02-1998

- Pour chaque numéro : un seul nom et une seule date de naissance
- Pour chaque nom : possible d'avoir plusieurs numéros d'étudiants
- Bilan : $\{\text{NumeroEtudiant}\} \rightarrow \{\text{NomEtudiant}, \text{DateNaissanceEtudiant}\}$
- Simplifiée en : $\text{NumeroEtudiant} \rightarrow \text{NomEtudiant}, \text{DateNaissanceEtudiant}$
- Déterminant : NumeroEtudiant

Dépendance fonctionnelle totale (DFT)

Définition pour X et Y deux sous-ensembles des attributs d'une rel. R

Y est en *dépendance fonctionnelle totale* (DFT) avec X si

- $X \rightarrow Y$ et
- Pour tout X' tq. $X' \subsetneq X$, Y n'est pas en dépendance fonctionnelle avec X'

Exemple

- NumeroEtudiant, CodeModule $\rightarrow$ NomEtudiant : DF non totale
- NumeroEtudiant $\rightarrow$ NomEtudiant : DFT

Fermeture des DF : motivation, définition

Motivation

A partir des dépendances proposées, comment en inférer d'autres ? Et jusqu'à combien ?

Définition : fermeture d'un ensemble de DF

L'ensemble de toutes les dépendances fonctionnelles qui sont inférées à partir d'un ensemble S de dépendances fonctionnelles est appelé *fermeture* de S et est dénoté $S+$

Comment construire $S+$?

- A partir des règles d'Armstrong ...

Plan

Règles d'inférences d'Armstrong [1]

Règles pour X, Y et Z sous-ensembles d'attributs de R

- Réflexivité : si $Y \subset X$, alors $X \rightarrow Y$
- Augmentation : si $X \rightarrow Y$ alors $X, Z \rightarrow Y, Z$
- Transitivité : si $X \rightarrow Y$ et $Y \rightarrow Z$ alors $X \rightarrow Z$
- Fermeture $S+$: obtenable à partir des 3 règles

Exemple d'application des règles d'inférence d'Armstrong

- "PrenomEtudiant, DateNaissanceEtudiant" et réflexivité :
  - PrenomEtudiant, DateNaissanceEtudiant → PrenomEtudiant...
- NoEtudiant → NomEtudiant, PrenomEtudiant et augmentation :
  - NoEtudiant, CodeModule → NomEtudiant, PrenomEtudiant, CodeModule
- CodeModule → NoHarpege, NoHarpege → NomEnseignant et Transitivité :
  - CodeModule → NomEnseignant

1. Armstrong, W. W. (1974, August). Dependency Structures of Data Base Relationships. In IFIP congress (Vol. 74, pp. 580-583).

A vous de jouer : DF

Base de données de contrats d'assurance

- Un membre peut avoir plusieurs contrats, plusieurs numeros d'enfant.

Données d'une relation universelle

NoMem.	NomMem.	PrenomMem.	NoEnfant	NoCont.	CodeCont.	TypeCont.
1	Dupont	Jean	4	011	AssV1	AssuranceVoiture
1	Dupont	Jean	3	011	AssV1	AssuranceVoiture
2	Dupont	Marie	4	124	AssVi	AssuranceVie
2	Dupont	Marie	3	124	AssVi	AssuranceVie
1	Dupont	Jean	4	123	AssVi	AssuranceVie
1	Dupont	Jean	3	123	AssVi	AssuranceVie
1	Dupont	Jean	4	012	AssV2	AssuranceVoiture
1	Dupont	Jean	3	012	AssV2	AssuranceVoiture

Questions

- Quelles semble être l'ensemble S des dépendances fonctionnelles ?
- Cloture transistive de S ?

A vous de jouer : DF

Base de données de contrats d'assurance

- Un membre peut avoir plusieurs contrats, plusieurs numeros d'enfant.

Données d'une relation universelle

NoMem.	NomMem.	PrenomMem.	NoEnfant	NoCont.	CodeCont.	TypeCont.
1	Dupont	Jean	4	011	AssV1	AssuranceVoiture
1	Dupont	Jean	3	011	AssV1	AssuranceVoiture
2	Dupont	Marie	4	124	AssVi	AssuranceVie
2	Dupont	Marie	3	124	AssVi	AssuranceVie
1	Dupont	Jean	4	123	AssVi	AssuranceVie
1	Dupont	Jean	3	123	AssVi	AssuranceVie
1	Dupont	Jean	4	012	AssV2	AssuranceVoiture
1	Dupont	Jean	3	012	AssV2	AssuranceVoiture

Questions

- Quelles semble être l'ensemble S des dépendances fonctionnelles ?
  - DF1 : NoMem $\rightarrow$ NomMem, PrenomMem
  - DF2 : NoCont $\rightarrow$ CodeCont
  - DF3 : CodeCont $\rightarrow$ TypeCont
- Cloture transistive de S ?

A vous de jouer : DF

Base de données de contrats d'assurance

- Un membre peut avoir plusieurs contrats, plusieurs numeros d'enfant.

Données d'une relation universelle

NoMem.	NomMem.	PrenomMem.	NoEnfant	NoCont.	CodeCont.	TypeCont.
1	Dupont	Jean	4	011	AssV1	AssuranceVoiture
1	Dupont	Jean	3	011	AssV1	AssuranceVoiture
2	Dupont	Marie	4	124	AssVi	AssuranceVie
2	Dupont	Marie	3	124	AssVi	AssuranceVie
1	Dupont	Jean	4	123	AssVi	AssuranceVie
1	Dupont	Jean	3	123	AssVi	AssuranceVie
1	Dupont	Jean	4	012	AssV2	AssuranceVoiture
1	Dupont	Jean	3	012	AssV2	AssuranceVoiture

Questions

- Quelles semble être l'ensemble S des dépendances fonctionnelles ?
  - DF1 : NoMem $\rightarrow$ NomMem, PrenomMem
  - DF2 : NoCont $\rightarrow$ CodeCont
  - DF3 : CodeCont $\rightarrow$ TypeCont
- Cloture transistive de S ? Ajout de DF4 : NoCont $\rightarrow$ TypeCont

Fermeture $X+$ p.r. à S l'ensemble des DF

Fermeture $X+$

Soit un ensemble d'attributs X de R et S l'ensemble des DF concernant R. La *fermeture* de X pour S est l'ensemble $X+$ des attributs qui dépendant fonctionnellement de X.

Algorithme de calcul de la fermeture $X+$

entrées : X : ensemble d'attributs, S : ensemble de dépendances fonctionnelles
sortie : $X+$: fermeture de X
$X+ := X$; changé := vrai
TANT QUE changé **et** il existe une DF non marquée **FAIRE**
  changé := faux
  POUR CHAQUE DF : $X' \rightarrow Y$ de S **FAIRE**
    SI $X' \subseteq X+$ **ALORS**
      $X+ := X+ \cup Y$
      marquer la DF comme "utilisée"
      changé := vrai
    FSI
  FIN POUR
FIN TANT QUE

Clé candidate d'une relation universelle R

Définition pour X d'un ensemble d'attributs de R

X *une clé candidate* si :

- Pour tout sous-ensemble d'attributs Y de R : $X \rightarrow Y$,
- La fermeture $X+$: égale à R
- X est irréductible : tout Y strictement inclus dans X aurait une fermeture $Y+$ strictement incluse dans R

Remarque

X est une clé candidate si X est le plus petit ensemble d'attributs qui implique tous les attributs de R.

Algorithme de calcul de $X+$ sur l'exemple

Dépendances fonctionnelles de la relation universelle

- DF1 : NoEtud $\rightarrow$ NomEtud, PrenomEtud, DateNaissEtud, LoginEtud, MDPEtud
- DF2 : LoginEtud $\rightarrow$ NoEtud, NomEtud, PrenomEtud, DateNaissEtud, MDPEtud
- DF3 : NoHarpege $\rightarrow$ NomEns, PrenomEns, Grade
- DF4 : CodeModule $\rightarrow$ LibelleModule, VolCM, VolTD, VolTP, NoHarpege
- DF5 : NoEtud, CodeModule, Annee $\rightarrow$ NoteSession1, NoteSession2

Montrons que $X = \{\text{NoEtud}, \text{CodeModule}, \text{Annee}\}$ est clé candidate de la relation.

Algorithme de calcul de $X+$ sur l'exemple

Initialisation

$X+ = \{\text{NoEtud}, \text{CodeModule}, \text{Annee}\}$; changé := vrai

Algorithme de calcul de $X+$ sur l'exemple

1ère Itération : $X+ = \{\text{NoEtud}, \text{CodeModule}, \text{Annee}\}$; changé := vrai

DF1 : NoEtud $\rightarrow$ NomEtud, PrenomEtud, DateNaissEtud, LoginEtud, MDPEtud

- NoEtud $\in X+$: membre droit de DF1 ajouté dans $X+$; marquage de DF1 comme utilisée ; changé à vrai
- $X+ = \{$ NoEtud, CodeModule, Annee, NomEtud, PrenomEtud, DateNaissEtud, LoginEtud, MDPEtud $\}$

DF2 : LoginEtud $\rightarrow$ NoEtud, NomEtud, PrenomEtud, DateNaissEtud, MDPEtud

- LoginEtud $\in X+$: membre droit de DF2 ajouté dans $X+$ (mais inchangée) ; marquage de DF2 comme utilisée ;

Algorithme de calcul de $X+$ sur l'exemple

1ère Itération : $X+ =$ { NoEtud, CodeModule, Annee, NomEtud, PrenomEtud, DateNaissEtud, LoginEtud, MDPEtud } ; changé := vrai

DF3 : NoHarpege $\rightarrow$ NomEns, PrenomEns, Grade

- NoHarpege $\notin X+$: DF3 non utilisée à cette itération ; $X+$ inchangée

DF4 : CodeModule $\rightarrow$ LibelleModule, VolCM, VolTD, VolTP, NoHarpege

- CodeModule $\in X+$, membre droit de DF4 ajouté dans $X+$; marque de DF4 comme utilisée ;
- $X+ =$ { NoEtud, CodeModule, Annee, NomEtud, PrenomEtud, DateNaissEtud, LoginEtud, MDPEtud, LibelleModule, VolCM, VolTD, VolTP, NoHarpege }

DF5 : NoEtud, CodeModule, Annee $\rightarrow$ NoteSession1, NoteSession2

- {NoEtud, CodeModule, Annee} $\subseteq X+$, membre droit de DF5 ajouté dans $X+$; marque de DF5 comme utilisée ;
- $X+ =$ { NoEtud, CodeModule, Annee, NomEtud, PrenomEtud, DateNaissEtud, LoginEtud, MDPEtud, LibelleModule, VolCM, VolTD, VolTP, NoHarpege, NoteSession1, NoteSession2 }

Algorithme de calcul de $X+$ sur l'exemple

2onde itération : $X+ =$ { NoEtud, CodeModule, Annee, NomEtud, PrenomEtud, DateNaissEtud, LoginEtud, MDPEtud, LibelleModule, VolCM, VolTD, VolTP, NoHarpege, NoteSession1, NoteSession2 } ;

DF3 : NoHarpege $\rightarrow$ NomEns, PrenomEns, Grade

- NoHarpege $\in X+$: membre droit de DF3 ajouté dans $X+$; DF3 marquée comme utilisée ; changé à vrai
- $X+ =$ { NoEtud, CodeModule, Annee, NomEtud, PrenomEtud, DateNaissEtud, LoginEtud, MDPEtud, LibelleModule, VolCM, VolTD, VolTP, NoHarpege, NoteSession1, NoteSession2, NomEns, PrenomEns, Grade }

Fin de la 2^{eme} itération. Plus de DF non marquée. Fin de l'algorithme.

$X =$ {NoEtud, CodeModule, Annee} est une clé candidate car

- $X+ = R$
- X est irréductible : retirer un des attributs de $X \rightsquigarrow X+ \neq R$

A vous de jouer : clé candidate

Données

- *UNIVERSELLE*(NoMem, NomMem, PrenomMem, NoEnfant, NoCont, CodeCont, TypeCont)
- DF1 : NoMem $\rightarrow$ NomMem, PrenomMem ; DF2 : NoCont $\rightarrow$ CodeCont
- DF3 : CodeCont $\rightarrow$ TypeCont DF4 : NoCont $\rightarrow$ TypeCont

Clé candidate d'une telle relation ?

A vous de jouer : clé candidate

Données

- *UNIVERSELLE*(NoMem, NomMem, PrenomMem, NoEnfant, NoCont, CodeCont, TypeCont)
- DF1 : NoMem $\rightarrow$ NomMem, PrenomMem ; DF2 : NoCont $\rightarrow$ CodeCont
- DF3 : CodeCont $\rightarrow$ TypeCont DF4 : NoCont $\rightarrow$ TypeCont

Clé candidate d'une telle relation ?

- Montrons que {NoMem, NoEnfant, NoCont} est une clé candidate

Plan

Décomposition sans perte d'information

Définition pour $R(T)$

- Soit B_1, B_2 deux ensembles non vides d'attributs tels que $B_1 \cup B_2 = T$ et
- $R_1 = [B_1]R$ et $R_1 = [B_2]R$ forment une décomposition de R
- Si la jointure naturelle $R_1[B_1 \cap B_2]R_2 = R$: décomposition sans perte d'information

Exemples de décomposition

NoEtud	NomEtud	DateNaissEtud
27750	Martin	2-09-1989
38911	Martin	26-02-1998

$B_1 = \{$NoEtud, NomEtud$\}$
$B_2 = \{$NoEtud, DateNaissEtud$\}$

NoEtud	NomEtud
27750	Martin
32911	Martin

NoEtud	DateNaissEtud
27750	2-09-1989
32911	26-02-1998

Jointure selon NoEtud
$\rightsquigarrow R$ (pas de perte d'info)

$B_1 = \{$NoEtud, NomEtud$\}$
$B_2 = \{$NomEtud, DateNaissEtud$\}$

NoEtud	NomEtud
27750	Martin
32911	Martin

NomEtud	DateNaissEtud
Martin	2-09-1989
Martin	26-02-1998

Jointure selon NomEtud
$\rightsquigarrow \neq R$ (perte d'info !)

Décomposition sans perte d'information

Théorème de Heath pour $R(T)$

Soit X, Y et Z une partition de T avec $X \rightarrow Y$. La décomposition de R en deux relations R_1 et R_2 définies ci-dessous est sans perte d'information.

- $R_1 = [X \cup Y]R$
- $R_2 = [X \cup Z]R$

Preuve en posant $B_1 = X \cup Y$ et $B_2 = X \cup Z$

Soit $R_1 = [B_1]R$ et $R_2 = [B_2]R$; on a $B_1 \cap B_2 = X$; montrons que $R_1[X]R_2 = R$.

- $R_1[X]R_2 \supseteq R$? Soit $r \in R$. On a $[X \cup Y]r \in [X \cup Y]R = R_1$ et $[X \cup Z]r \in [X \cup Z]R = R_2$. $[X \cup Y]r$ et $[X \cup Z]r$ ont la même valeur pour l'attribut X et donc la jointure $[X \cup Y]r[X][X \cup Z]r$ appartient à $R_1[X]R_2$.
- $R_1[X]R_2 \subseteq R$? Soit $r \in R_1[X]R_2$. Alors $[X \cup Y]r \in [X \cup Y]R = R_1$ et donc il existe $s_1 \in R$ tq $[X \cup Y]s_1 = [X \cup Y]r$. Similairement $[X \cup Z]r \in [X \cup Z]R = R_2$ et donc il existe $s_2 \in R$ tq $[X \cup Z]s_2 = [X \cup Z]r$. Donc $[X]r = [X]s_1 = [X]s_2$. Comme $X \rightarrow Y$, $[Y]s_1 = [Y]s_2$ et donc $[X \cup Y \cup Z]s_2 = r \in R$.

Plan

Introduction

1NF, 2NF, 3NF, BCNF

- 1NF
- 2NF
- 3NF
- BCNF (3.5 NF)

4NF

Les premières formes normales

Procédure de normalisation

- Réalisée par réductions successives réversibles d'un ensemble de relations en une forme plus "satisfaisante"
- Construite à partir du concept de formes normales

Forme normale : qu'est-ce et inclusions

Une relation est dans une *forme normale* particulière si elle satisfait un ensemble de conditions prédéfinies.

$$1NF \supset 2NF \supset 3NF \supset BCNF \supset 4NF \supset 5NF$$

$1NF \supset 2NF \supset 3NF$:

- introduites par Codd
- permettent la décomposition d'un ensemble de relations sans perte d'information, en utilisant les DF

Plan

Introduction

1NF, 2NF, 3NF, BCNF

- 1NF
- 2NF
- 3NF
- BCNF (3.5 NF)

4NF

Première forme normale

Définition

- Une relation est en première forme normale (1NF) si tous ses attributs sont atomiques.
- Un schéma relationnel en 2NF : si toutes ses relations en 2NF

Exemple : une relation ETUDIANT pas 1NF

NoEtu*	NomEtu	PrenomEtu	DateNaissEtu	Login	MDPEtu
23794	Dornier	Arnaud	18-08-1998	adornier	adornier
32911	Martin	Maxime	26-02-1998	mmartin	mmartin
33818	Schatt	Bastien, Charles	06-05-1998	bschatt	bschatt

Prénom pouvant contenir 2 valeurs : pas atomique, donc pas en 1NF.

Décomposition pour se conformer à 1NF

Décomposition des attributs problématiques.

A vous de jouer : 1NF ?

Données

- *UNIVERSELLE*(NoMem, NomMem, PrenomMem, NoEnfant, NoCont, CodeCont, TypeCont)

NoMem.	NomMem.	PrenomMem.	NoEnfant	NoCont.	CodeCont.	TypeCont.
1	Dupont	Jean	4	011	AssV1	AssuranceVoiture
1	Dupont	Jean	3	011	AssV1	AssuranceVoiture
2	Dupont	Marie	4	124	AssVi	AssuranceVie
2	Dupont	Marie	3	124	AssVi	AssuranceVie
1	Dupont	Jean	4	123	AssVi	AssuranceVie
1	Dupont	Jean	3	123	AssVi	AssuranceVie
1	Dupont	Jean	4	012	AssV2	AssuranceVoiture
1	Dupont	Jean	3	012	AssV2	AssuranceVoiture

Est-ce 1NF ?

A vous de jouer : 1NF ?

Données

- *UNIVERSELLE*(NoMem, NomMem, PrenomMem, NoEnfant, NoCont, CodeCont, TypeCont)

NoMem.	NomMem.	PrenomMem.	NoEnfant	NoCont.	CodeCont.	TypeCont.
1	Dupont	Jean	4	011	AssV1	AssuranceVoiture
1	Dupont	Jean	3	011	AssV1	AssuranceVoiture
2	Dupont	Marie	4	124	AssVi	AssuranceVie
2	Dupont	Marie	3	124	AssVi	AssuranceVie
1	Dupont	Jean	4	123	AssVi	AssuranceVie
1	Dupont	Jean	3	123	AssVi	AssuranceVie
1	Dupont	Jean	4	012	AssV2	AssuranceVoiture
1	Dupont	Jean	3	012	AssV2	AssuranceVoiture

Est-ce 1NF ?

- Tous ses attributs sont atomiques $\rightsquigarrow$ Oui

Plan

Deuxième forme normale

Deuxième forme normale

Une relation R est en deuxième forme normale (2NF) si :

- elle est en 1NF
- tout attribut n'appartenant pas à la clé : en dépendance fonctionnelle totale (DFT) avec la clé

Un schéma relationnel en 2NF : si toutes ses relations en 2NF

Décomposition d'une relation R pour se conformer à 2NF

$R(X*, Y*, Z, T)$	décomposée comme suit $\Rightarrow$	$R_1 = [Y*, Z]R$
$+\ Y* \rightarrow Z \in S+$	jusqu'à plus applicable	$R_2 = [X*, Y*, T]R$
Z le $+$ grand possible		

Exemple Scolarité : mise en 2NF

Relation Universelle et $S+$, cloture transitive des DF

R(NoEtu*, CodeModule*, Annee*, NomEtu, PrenomEtu, DateNaissEtu, LoginEtu, MDPEtu, LibelleModule, VolCM, VolTD, VolTP, NoHarpege, NoteSession1, NoteSession2, NomEns, PrenomEns, Grade)
DF1 : NoEtu $\rightarrow$ NomEtu, PrenomEtu, DateNaissEtu, LoginEtu, MDPEtu
DF2 : LoginEtu $\rightarrow$ NoEtu, NomEtu, PrenomEtu, DateNaissEtu, MDPEtu
DF3 : NoHarpege $\rightarrow$ NomEns, PrenomEns, Grade
DF4 : CodeModule $\rightarrow$ LibelleModule, VolCM, VolTD, VolTP, NoHarpege
DF4b : CodeModule $\rightarrow$ LibelleModule, VolCM, VolTD, VolTP, NoHarpege, NomEns, PrenomEns, Grade (transitivité)
DF5 : NoEtu, CodeModule, Annee $\rightarrow$ NoteSession1, NoteSession2

Décompositions successives

R + DF1 :
R_1(NoEtu*, NomEtu, PrenomEtu, DateNaissEtu, LoginEtu, MDPEtu)
R_2(NoEtu*, CodeModule*, Annee*, LibelleModule, VolCM, VolTD, VolTP, NoHarpege, NoteSession1, NoteSession2, NomEns, PrenomEns, Grade) : pb DF4/DF4b

Exemple Scolarité : mise en 2NF

Relation Universelle et $S+$, cloture transitive des DF

R(NoEtu*, CodeModule*, Annee*, NomEtu, PrenomEtu, DateNaissEtu, LoginEtu, MDPEtu, LibelleModule, VolCM, VolTD, VolTP, NoHarpege, NoteSession1, NoteSession2, NomEns, PrenomEns, Grade)

DF1 : NoEtu → NomEtu, PrenomEtu, DateNaissEtu, LoginEtu, MDPEtu

DF2 : LoginEtu → NoEtu, NomEtu, PrenomEtu, DateNaissEtu, MDPEtu

DF3 : NoHarpege → NomEns, PrenomEns, Grade

DF4 : CodeModule → LibelleModule, VolCM, VolTD, VolTP, NoHarpege

DF4b : CodeModule → LibelleModule, VolCM, VolTD, VolTP, NoHarpege, NomEns, PrenomEns, Grade (transitivité)

DF5 : NoEtu, CodeModule, Annee → NoteSession1, NoteSession2

Décompositions successives

R_1(NoEtu*, NomEtu, PrenomEtu, DateNaissEtu, LoginEtu, MDPEtu)

R_2(NoEtu*, CodeModule*, Annee*, LibelleModule, VolCM, VolTD, VolTP, NoHarpege, NoteSession1, NoteSession2, NomEns, PrenomEns, Grade) + DF4b :

R_3(CodeModule*, LibelleModule, VolCM, VolTD, VolTP, NoHarpege, NomEns, PrenomEns, Grade)

R_4(NoEtu*, CodeModule*, Annee*, NoteSession1, NoteSession2)

Exemple Scolarité : mise en 2NF

Relation Universelle et $S+$, cloture transitive des DF

R(NoEtu*, CodeModule*, Annee*, NomEtu, PrenomEtu, DateNaissEtu, LoginEtu, MDPEtu, LibelleModule, VolCM, VolTD, VolTP, NoHarpege, NoteSession1, NoteSession2, NomEns, PrenomEns, Grade)

DF1 : NoEtu → NomEtu, PrenomEtu, DateNaissEtu, LoginEtu, MDPEtu

DF2 : LoginEtu → NoEtu, NomEtu, PrenomEtu, DateNaissEtu, MDPEtu

DF3 : NoHarpege → NomEns, PrenomEns, Grade

DF4 : CodeModule → LibelleModule, VolCM, VolTD, VolTP, NoHarpege

DF4b : CodeModule → LibelleModule, VolCM, VolTD, VolTP, NoHarpege, NomEns, PrenomEns, Grade (transitivité)

DF5 : NoEtu, CodeModule, Annee → NoteSession1, NoteSession2

Décompositions successives

R_1(NoEtu*, NomEtu, PrenomEtu, DateNaissEtu, LoginEtu, MDPEtu)

R_3(CodeModule*, LibelleModule, VolCM, VolTD, VolTP, NoHarpege, NomEns, PrenomEns, Grade)

R_4(NoEtu*, CodeModule*, Annee*, NoteSession1, NoteSession2)

A vous de jouer : 2NF

Données

- *UNIVERSELLE*(NoMem, NomMem, PrenomMem, NoEnfant, NoCont, CodeCont, TypeCont)
- Clé candidate : {NoMem, NoEnfant, NoCont}
- DF1 : NoMem $\rightarrow$ NomMem, PrenomMem ; DF2 : NoCont $\rightarrow$ CodeCont
- DF3 : CodeCont $\rightarrow$ TypeCont DF4 : NoCont $\rightarrow$ TypeCont

Mise en 2NF

- Transformer cette relation pour avoir un schéma 2NF

Plan

Introduction

1NF, 2NF, 3NF, BCNF

- 1NF
- 2NF
- 3NF
- BCNF (3.5 NF)

4NF

Troisième forme normale

Définition

Une relation est en troisième forme normale (3NF) si :

- elle est en 2NF
- Tout attribut n'appartenant pas à la clé : ne dépend pas d'un attribut non-clé

Un schéma relationnel en 3NF : si toutes ses relations en 3NF

Décomposition d'une relation R pour se conformer à 3NF

$R(X*, Y*, Z, T)$	décomposée comme suit $\Rightarrow$	$R_1 = [Z*, T]R$
$+ Z \rightarrow T \in S+$	jusqu'à plus applicable	$R_2 = [X*, Y*, Z]R$
T le $+$ grand possible		

Exemple Scolarité : mise en 3NF

A partir du schéma 2NF et de $S+$, cloture transitive des DF

R_1(NoEtu*, NomEtu, PrenomEtu, DateNaissEtu, LoginEtu, MDPEtu)
R_3(CodeModule*, LibelleModule, VolCM, VolTD, VolTP, NoHarpege, NomEns, PrenomEns, Grade)
R_4(NoEtu*, CodeModule*, Annee*, NoteSession1, NoteSession2)
DF1 : NoEtu $\rightarrow$ NomEtu, PrenomEtu, DateNaissEtu, LoginEtu, MDPEtu
DF2 : LoginEtu $\rightarrow$ NoEtu, NomEtu, PrenomEtu, DateNaissEtu, MDPEtu
DF3 : NoHarpege $\rightarrow$ NomEns, PrenomEns, Grade
DF4 : CodeModule $\rightarrow$ LibelleModule, VolCM, VolTD, VolTP, NoHarpege
DF4b : CodeModule $\rightarrow$ LibelleModule, VolCM, VolTD, VolTP, NoHarpege, NomEns, PrenomEns, Grade (transitivité)
DF5 : NoEtu, CodeModule, Annee $\rightarrow$ NoteSession1, NoteSession2

Décompositions de R_3 selon DF3

R_1(NoEtu*, NomEtu, PrenomEtu, DateNaissEtu, LoginEtu, MDPEtu)
R_4(NoEtu*, CodeModule*, Annee*, NoteSession1, NoteSession2)
R_5(NoHarpege*,NomEns, PrenomEns, Grade)
R_6(CodeModule*, LibelleModule, VolCM, VolTD, VolTP, NoHarpege)

A vous de jouer : 3NF

Données 2NF

- M2 (NoMembre*, NomMembre, PrenomMembre)
- M3 (NoMembre*, NoEnfant*, NoContrat*)
- M4 (NoContrat*, CodeContrat, TypeContrat)
- DF1 : NoMem $\rightarrow$ NomMem, PrenomMem ; DF2 : NoCont $\rightarrow$ CodeCont
- DF3 : CodeCont $\rightarrow$ TypeCont DF4 : NoCont $\rightarrow$ TypeCont

Mise en 3NF

- Transformer cette relation pour avoir un schéma 3NF

Plan

Introduction

1NF, 2NF, 3NF, BCNF

- 1NF
- 2NF
- 3NF
- BCNF (3.5 NF)

4NF

Forme normale de Boyce-Codd

Définition

Une relation est en forme normale de Boyce-Codd (BCNF) si :

- elle est en 3NF et pour chaque dépendance $X \rightarrow Y$, au moins une des conditions suivantes est vraie :
  - $X \rightarrow Y$ est une dépendance functionnelle triviale ($Y \subseteq X$)
  - X : une super-clé

Un schéma relationnel en BCNF : si toutes ses relations en BCNF

Décomposition pour se conformer à BCNF

$R(X*, Y*, Z, T)$	Passage en BCNF	$R_1 = [Z*, Y]R$
avec $Z \rightarrow Y$	$\Longrightarrow$	$R_2 = [X*, Z*, T]R$

Attention, cette décomposition entraîne la perte d'une DF ($X, Y \rightarrow T$).

Exemple Scolarité : mise en BCNF

Application aux étudiants

On repart de l'ensemble de relations en 3NF :

- R_1(NoEtu*, NomEtu, PrenomEtu, DateNaissEtu, LoginEtu, MDPEtu)
- R_4(NoEtu*, CodeModule*, Annee*, NoteSession1, NoteSession2)
- R_5(NoHarpege*, NomEns, PrenomEns, Grade)
- R_6(CodeModule*, LibelleModule, VolCM, VolTD, VolTP, NoHarpege)

DF1 : NoEtu $\rightarrow$ NomEtu, PrenomEtu, DateNaissEtu, LoginEtu, MDPEtu
DF2 : LoginEtu $\rightarrow$ NoEtu, NomEtu, PrenomEtu, DateNaissEtu, MDPEtu
DF3 : NoHarpege $\rightarrow$ NomEns, PrenomEns, Grade
DF4 : CodeModule $\rightarrow$ LibelleModule, VolCM, VolTD, VolTP, NoHarpege
DF4b : CodeModule $\rightarrow$ LibelleModule, VolCM, VolTD, VolTP, NoHarpege, NomEns, PrenomEns, Grade (transitivité)
DF5 : NoEtu, CodeModule, Annee $\rightarrow$ NoteSession1, NoteSession2
Toutes les dépendances fonctionnelles élémentaires ont une super-clé dans leur déterminant $\rightsquigarrow$ schéma en BCNF

A vous de jouer : 3NF

Données 3NF

- M2 (NoMembre*, NomMembre, PrenomMembre)
- M3 (NoMembre*, NoEnfant*, NoContrat*)
- M5 (NoContrat*, CodeContrat)
- M6 (CodeContrat*, TypeContrat)
- DF1 : NoMem $\rightarrow$ NomMem, PrenomMem ; DF2 : NoCont $\rightarrow$ CodeCont
- DF3 : CodeCont $\rightarrow$ TypeCont DF4 : NoCont $\rightarrow$ TypeCont

Mise en BCNF

- Cette relation est-elle déjà en BCNF ?

Plan

Formes normales 4 et 5

Les premières formes normales (1NF, 2NF, 3NF, BCNF) : basées sur la notion de dépendance fonctionnelle.

Formes normales 4 et 5

Formes normales 4NF et 5NF : définies sur les dépendances plus générales :

- les dépendances multivaluées (4NF)
- les dépendances de jointure (5NF)

Plan

Dépendance multivaluée d'une relation $R(T)$

Définition pour α et β des attributs de $R(T)$

Pour chaque paire (t_1, t_2) de tuples de R égaux sur α, il existe (t_3, t_3) dans R tq.

- $[\alpha]t_1 = [\alpha]t_2 = [\alpha]t_3 = [\alpha]t_4$
- $[\beta]t_1 = [\beta]t_3$ et $[\beta]t_2 = [\beta]t_4$
- $[T \setminus (\alpha \cup \beta)]t_1 = [T \setminus (\alpha \cup \beta)]t_4$ et $[T \setminus (\alpha \cup \beta)]t_2 = [T \setminus (\alpha \cup \beta)]t_3$

On note $\alpha \twoheadrightarrow \beta$

En d'autres termes, pour (t_1, t_2) tuples de R égaux sur α

tuple	α	β	$T \setminus (\alpha \cup \beta)$
t_1	$a_1..a_n$	$b_1..b_m$	$d_1..d_k$
t_2	$a_1..a_n$	$c_1..c_m$	$e_1..e_k$
t_3	$a_1..a_n$	$b_1..b_m$	$e_1..e_k$
t_4	$a_1..a_n$	$c_1..c_m$	$d_1..d_k$

Les valeurs des attribut de β ne dépendent pas de celles de $T \setminus (\alpha \cup \beta)$

Dépendances multivaluées : exemple

Relation BAGUE

NoBague*	Taille*	Matiere*
1	41	or
1	42	or
1	40	or
2	40	or
2	42	argent
2	42	or
2	40	argent

Pas de dépendances fonctionnelles

$\rightsquigarrow$ 1NF, 2NF, 3NF,BCNF

Dépendances multivaluées : exemple

Relation BAGUE

NoBague*	Taille*	Matiere*
1	41	or
1	42	or
1	40	or
2	40	or
2	42	argent
2	42	or
2	40	argent

Pas de dépendances fonctionnelles
$\rightsquigarrow$ 1NF, 2NF, 3NF,BCNF

Redondances d'informations :

- Chaque bague existe dans certaines tailles et pour chaque taille dans les mêmes matières.
- Chaque bague existe dans certaines matière et pour chaque matière dans les mêmes tailles.

Dépendances multivaluées : exemple

Relation BAGUE

NoBague*	Taille*	Matiere*
1	41	or
1	42	or
1	40	or
2	40	or
2	42	argent
2	42	or
2	40	argent

Pas de dépendances fonctionnelles
$\rightsquigarrow$ 1NF, 2NF, 3NF,BCNF

Dépendances multivaluées :
NoBague $\twoheadrightarrow$ Taille

Dépendances multivaluées : exemple

Relation BAGUE

NoBague*	Taille*	Matiere*
1	41	or
1	42	or
1	40	or
2	40	or
2	42	or
2	40	argent
2	42	argent

Pas de dépendances fonctionnelles
$\rightsquigarrow$ 1NF, 2NF, 3NF,BCNF

Décomposition sensée :

NoBague*	Taille*
1	41
1	42
1	40
2	40
2	42

NoBague*	Matiere*
1	or
2	or
2	argent

Dépendance multivaluée : propriétés

Définie pour α, β et $T \setminus (\alpha \cup \beta)$

- Si la clé de R ne contient qu'un attribut α : $\alpha \rightarrow \beta \rightsquigarrow \alpha \twoheadrightarrow \beta$

Pour une relation $R(T)$, X, Y, Z sous ensembles disjoints inclus dans T

- Complémentation : $X \twoheadrightarrow Y \Rightarrow X \twoheadrightarrow T - X - Y$
- Transitivité : $X \twoheadrightarrow Y$ et $Y \twoheadrightarrow Z \Rightarrow X \twoheadrightarrow Z - Y$

Exemple de complémentation

NoBague $\twoheadrightarrow$ Taille $\rightsquigarrow$ NoBague $\twoheadrightarrow$ Matiere

Quatrième forme normale

Définition de la dépendance multivaluée élémentaire

Une dépendance multivaluée $X \twoheadrightarrow Y$ est *élémentaire* si

- Y n'est pas vide et est disjoint de X
- Il n'existe pas de dépendance $X' \twoheadrightarrow Y'$ telle que $X' \subset X$ et $Y' \subset Y$

4^{eme} forme normale

Une relation est en quatrième forme normale (4NF) si et seulement si les seules dépendances multivaluées élémentaires sont celles dans lesquelles une clé détermine un attribut.

Décomposition pour se conformer à 4NF - Théorème de Fagin

$R(X*, Y*, Z*)$ avec $X \twoheadrightarrow Y$ rem : 3 attr. ds la clé	Passage en 4NF $\Longrightarrow$	$R_1 = [X*, Y*]R$ $R_2 = [X*, Z*]R$

Exemple Scolarité : mise en 4NF

Application aux étudiants

- R_1(NoEtu*, NomEtu, PrenomEtu, DateNaissEtu, LoginEtu, MDPEtu)
- R_4(NoEtu*, CodeModule*, Annee*, NoteSession1, NoteSession2)
- R_5(NoHarpege*, NomEns, PrenomEns, Grade)
- R_6(CodeModule*, LibelleModule, VolumeCM, VolumeTD, VolumeTP, NoHarpege)

Existe-t-il une dépendance multivaluée entre NoEtu, CodeModule et Annee ?
L'ensemble de relations est en 4NF.

A vous de jouer

Restriction à une relation

On s'intéresse aux valeurs stockées dans M4.

NoMembre	NoEnfant	NoContrat
1	4	011
1	3	011
2	4	124
2	3	124
1	4	123
1	3	123
1	4	012
1	3	012

- Cette relation contient-elle des dépendance multivaluées ?
- Comment rendre cette relation 4NF ?