

INFO-H-303 Bases de données
Séance d'exercices 10
Normalisation

15 décembre 2011

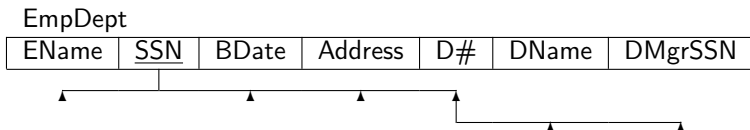
But de la normalisation

EmpDept

EName	<u>SSN</u>	BDate	Address	D#	DName	DMgrSSN
Smith	1234	21/07/39	...	1	Research	1234
Narayan	6668	18/01/43	...	1	Research	1234
English	4534	8/05/53	...	2	Account	4534
Wong	9788	30/11/49	...	3	Admin	9788
Zelaya	6677	23/08/60	...	3	Admin	9788

- ▶ Objectif : construire un schéma relationnel évitant la redondance
- ▶ La redondance implique des anomalies lors de
 - ▶ l'insertion (nouvel employé, nouveau département)
 - ▶ la suppression (du dernier employé d'un département)
 - ▶ la modification (changement de manager)

Dépendances fonctionnelles (DF)



DF1 : $SSN \rightarrow \{EName, BDate, Address, D#\}$

DF2 : $D\# \rightarrow \{DName, DMgrSSN\}$

Soit $R(A_1, \dots, A_n)$ avec $X, Y \subseteq \{A_1, \dots, A_n\}$

Il y a une **dépendance fonctionnelle** $X \rightarrow Y$ (X détermine Y)
si pour chaque paire de tuple t_1, t_2 de R,
si $t_1[X] = t_2[X]$ alors $t_1[Y] = t_2[Y]$.

Dépendances fonctionnelles : exemple



DF1 : $SSN \rightarrow \{EName, BDate, Address, D\# \}$

DF2 : $D\# \rightarrow \{DName, DMgrSSN \}$

EmpDept

EName	<u>SSN</u>	BDate	Address	D#	DName	DMgrSSN
Smith	1234	21/07/39	...	1	Research	1234
Narayan	6668	18/01/43	...	1	Research	1234
English	4534	8/05/53	...	2	Account	4534
Wong	9788	30/11/49	...	3	Admin	9788
Zelaya	6677	23/08/60	...	3	Admin	9788

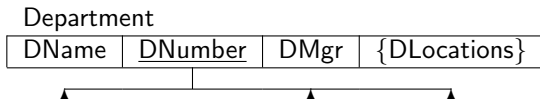
Première forme normale

Une relation R est en **première forme normale** si :

- ▶ R respecte la définition du modèle relationnel
- ▶ R ne possède pas d'attribut composés ou multivalués

Toutes les relations que l'on a vu jusqu'à présent respectent la première forme normale.

Première forme normale : exemple



Department

DName	<u>DNumber</u>	DMgr	{DLocations}
Research	5	333445555	{Bellaire, Sugarland, Houston}
Administration	4	987654321	{Stafford}
Headquarters	1	888665555	{Houston}

↓ 1NF Normalization

Department

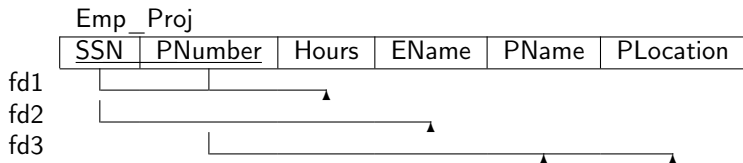
DName	<u>DNumber</u>	<u>DLocations</u>	DMgr
Research	5	Bellaire	333445555
Research	5	Sugarland	333445555
Research	5	Houston	333445555
Administration	4	Stafford	987654321
Headquarters	1	Houston	888665555

Deuxième forme normale

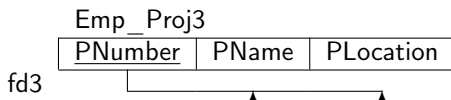
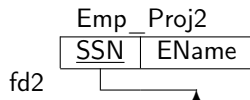
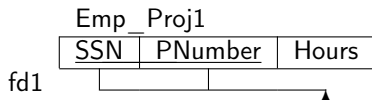
Une relation R est en **deuxième forme normale** si :

- ▶ R est en première forme normale
- ▶ il n'y a pas d'attribut ne faisant pas partie d'une clé (*non-prime attribute*) qui dépend d'une partie d'une des clés de R

Deuxième forme normale : exemple



↓ 2NF Normalization

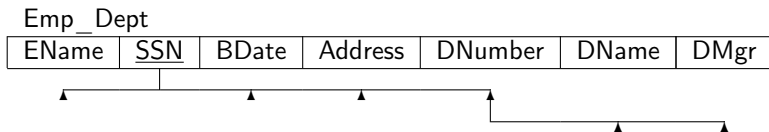


Troisième forme normale

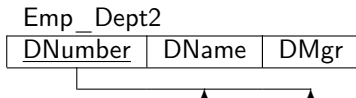
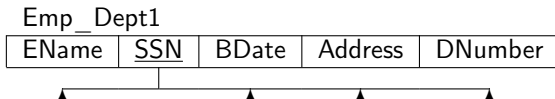
Une relation R est en **troisième forme normale** si :

- ▶ R est en deuxième forme normale
- ▶ il n'y a pas d'attribut ne faisant pas partie d'une clé qui dépend transitivement de cette clé

Troisième forme normale : exemple



↓ 3NF Normalization



Forme normale de Boyce-Codd (BCNF)

Une relation R est en **BCNF** si :

- ▶ R est en troisième forme normale
- ▶ la partie gauche de chaque DF est une clé candidate entière

La plupart des relations en troisième forme normale sont en BCNF.

Décomposition

<u>Professeur</u>	<u>Département</u>	<u>Faculté</u>
Esteban	CoDE	Polytech
Jean	DI	Sciences
Hugues	CoDE	Polytech
Olivier	LISA	Polytech

<u>Professeur</u>	<u>Faculté</u>
Esteban	Polytech
Jean	Sciences
Hugues	Polytech
Olivier	Polytech
<u>Département</u>	<u>Faculté</u>
CoDE	Polytech
DI	Sciences
LISA	Polytech

Perte d'information

<u>Professeur</u>	<u>Département</u>
Esteban	CoDE
Jean	DI
Hugues	CoDE
Olivier	LISA
<u>Professeur</u>	<u>Faculté</u>
Esteban	Polytech
Jean	Sciences
Hugues	Polytech
Olivier	Polytech

Perte de DF

<u>Professeur</u>	<u>Département</u>
Esteban	CoDE
Jean	DI
Hugues	CoDE
Olivier	LISA
<u>Département</u>	<u>Faculté</u>
CoDE	Polytech
DI	Sciences
CoDE	Polytech
LISA	Polytech

Bonne décomposition

Décomposition sans perte

Lors de la décomposition, il faut veiller à ne perdre ni information ni dépendance fonctionnelle.

Soit une relation R décomposée en deux relations R_1 et R_2 . Si l'ensemble des attributs communs de R_1 et R_2 est une clé d'une des deux relations, alors la décomposition est sans perte d'information.

Si dans une relation R on peut trouver trois ensembles d'attributs A, B et C tel qu'il existe une dépendance fonctionnelle $A \rightarrow B$, alors R peut être décomposée en deux relations $R_1(A, B)$ et $R_2(A, C)$ sans perte d'information.