

# Quelques Principes de Bon Database Design

Jef Wijsen

# La redondance

ENom	Sal	DNom	Budget
Ed	3000	MIS	5000K
An	3500	Marketing	2000K
Tim	3600	Marketing	?

- La valeur “?” est redondante : le budget du département *Marketing* est de 2000K...
- pourvu que chaque département n'ait qu'un seul budget.
- On écrira **DNom** → **Budget**.

# Autres dépendances fonctionnelles

ENom	Sal	DNom	Budget
Ed	3000	MIS	5000K
An	3500	Marketing	2000K
Tim	3600	Marketing	2000K

**ENom** → **Sal** : Un employé ne gagne qu'un seul salaire.

**ENom** → **DNom** : Un employé ne peut pas travailler pour deux départements différents.

**DNom** → **Budget** : Un département ne profite que d'un seul budget.

# Exemple plus compliqué

## HORAIRE

Prof	Cours	Local	Sem	Jour	Heure
J. Wijssen	Gestion de BDs	204	1	jeudi	15:15
J. Wijssen	Gestion de BDs	230	1	mercredi	10:15

Cours → Prof

Prof, Sem, Jour, Heure → Cours

Local, Sem, Jour, Heure → Cours

Cours, Sem, Jour, Heure → Local

# Dépendances fonctionnelles

- Une *dépendance fonctionnelle* (DF) est une contrainte de la forme

$$B_1 B_2 \dots B_n \rightarrow C$$

avec  $B_1, B_2, \dots, B_n, C$  des attributs.

- Une relation respecte cette DF si pour tous les tuples  $s, t$  de la relation, l'expression suivante est vraie :
  - si  $s(B_1) = t(B_1)$  et ... et  $s(B_n) = t(B_n)$
  - alors  $s(C) = t(C)$  .
- On dira que  $C$  est déterminé par  $B_1, B_2, \dots, B_n$ .

Notez : Si  $t = \langle a_1, a_2, \dots, a_m \rangle$  est un tuple sur le schéma

$R[A_1, A_2, \dots, A_m]$ , alors  $t(A_1) = a_1, t(A_2) = a_2, \dots$

# Observations

- Si une relation  $R$  respecte  $A \rightarrow B$  et  $B \rightarrow C$ ,  $R$  respecte nécessairement  $A \rightarrow C$ .
  - “Si  $A$  détermine  $B$ , et  $B$  détermine  $C$ ,  $A$  détermine aussi  $C$ .”
- Si une relation  $R$  respecte  $A \rightarrow B$ ,  $A \rightarrow C$  et  $BC \rightarrow D$ ,  $R$  respecte nécessairement  $A \rightarrow D$ .
  - Pour vous en convaincre, essayez de construire un contre-exemple, i.e. une relation qui respecte  $A \rightarrow B$ ,  $A \rightarrow C$  et  $BC \rightarrow D$ , mais qui ne respecte pas  $A \rightarrow D$ .
- ...

# Raisonner sur les DF (1/2)

- Soit  $\Sigma$  un ensemble de DF.  
Soit  $X$  un ensemble d'attributs.
- Quels attributs sont déterminés par  $X$  ?
- Pour répondre à cette question, il suffit d'appliquer la première règle une fois, puis la deuxième règle jusqu'au "point fixe" :
  1. Tout attribut de  $X$  est déterminé par  $X$ .
  2. *Transitivité* :  
Si la DF  $B_1 B_2 \dots B_n \rightarrow C$  est dans  $\Sigma$  et  
si chaque  $B_i$  est déjà déterminé par  $X$  ( $1 \leq i \leq n$ ),  
alors  $C$  est déterminé par  $X$ .

# Raisonner sur les DF (2/2)

## Exemple 1

Soit

$$\Sigma = \{ \quad A \rightarrow C, B \rightarrow D, BCD \rightarrow E, \\ AEF \rightarrow G, AG \rightarrow F, DE \rightarrow A \} .$$

Les attributs déterminés par  $AB$  sont :

- $A$  et  $B$  à cause de la première règle.
- $C$  à cause de la deuxième règle, suite à  $A \rightarrow C$ .
- $D$  à cause de la deuxième règle, suite à  $B \rightarrow D$ .
- $E$  à cause de la deuxième règle, suite à  $BCD \rightarrow E$ .

# Boyce-Codd Normal Form

- Un schéma de relation avec un ensemble  $\Sigma$  de DF est en *Boyce-Codd Normal Form* (BCNF) si la condition suivante est respectée :
  - si  $X \rightarrow C$  est une DF dans  $\Sigma$  avec  $C \notin X$ ,
  - alors  $X$  détermine tous les attributs du schéma.
- BCNF est un “certificat de qualité” : un schéma certifié BCNF ne contient pas de valeurs redondantes à cause des DF.

# Schéma qui n'est pas en BCNF

ENom	Sal	DNom	Budget
Ed	3000	MIS	5000K
An	3500	Marketing	2000K
Tim	3600	Marketing	?

Pas en BCNF :

- DNom détermine Budget, mais
- DNom ne détermine pas tous les attributs. En effet, ENom et Sal ne sont pas déterminés par DNom.

# Deux schémas qui sont en BCNF

ENom	Sal	DNom
Ed	3000	MIS
An	3500	Marketing
Tim	3600	Marketing

DNom	Budget
MIS	5000K
Marketing	2000K

# Exercice

*Monsieur Bricolage* est une entreprise qui loue des outils de bricolage (perceuses, bétonnières, . . .) au grand public sur base journalière.

NrOutil	Type	Prix	NrNat	Nom	Tel	Date
111	perceuse	34	60 10 07 123 12	Jean Leduc	56 65 34	7 dec 2002
111	perceuse	37	59 07 06 233 57	Anne Leblanc	56 65 34	8 dec 2002
232	bétonnière	233	60 10 07 123 12	Jean Leduc	26 23 11	9 jan 2002
333	perceuse	37	59 12 07 223 90	Anne Leblanc	99 33 34	8 dec 2002
999	marteau	3	59 12 07 223 90	Anne Leblanc	99 33 34	8 dec 2002

NrOutil dénote le numéro unique d'un outil. NrNat est le numéro national de l'emprunteur. Les autres attributs sont le type d'outil (Type), le prix journalier de location (Prix), le nom et le téléphone de l'emprunteur (Nom et Tel), la date de location (Date).

Il n'y a qu'un seul prix pour un type d'outil pour une journée donnée. Par exemple, le 8 décembre 2002, le prix d'une perceuse était fixé à 37 EUR. Bien sûr, le prix d'un type d'outil peut augmenter au cours du temps. Évidemment, il ne peut y avoir qu'un seul emprunteur par jour et par outil. Le type d'un outil ne change pas. Deux personnes peuvent avoir le même nom. Le nom et le numéro national d'une personne ne changent jamais. Le numéro de téléphone d'une personne peut évoluer. Néanmoins, si une personne loue deux outils au même jour, ces deux locations seront accompagnées du même numéro d'appel.