Base de données et structures de données

CM2-1 : Dépendances fonctionnelles et normalisation

Mickaël Martin Nevot

V1.0.3

Cette œuvre de Mickaël Martin Nevot est mise à disposition sous licence Creative Commons Attribution - Utilisation non commerciale - Partage dans les mêmes conditions.

BY-NC-SA

Base de données et structures de données

I. Prés.
II. BD et SGBD
III. DF et FN
IV. Merise
V. LDD
VI. LMD
VII. LCT
VIII. Droits
IX. LDSP
X. SQL avancé

 BY-NC-SA

Dépendances fonctionnelles

- **Dans une relation, il y a dépendance fonctionnelle (DF) entre deux ensembles d'attributs quand, à chaque valeur du premier, est associé au maximum une valeur du second**
- Exemple :
  - $Étudiant \nrightarrow Note$:
    - Aegon Targaryen a deux notes différentes : 13 et 09
  - $Étudiant \rightarrow Matière$:
    - BD n'apparait que pour Aegon Targaryen, C++ pour Walter White et PHP pour Magnus Maximus

Étudiant	Matière	Note
Aegon Targaryen	BD	13
Aegon Targaryen	BD	09
Walter White	C++	15
Magnus Maximus	PHP	18

 BY-NC-SA

Dépendances fonctionnelles

- Vérification de dépendances fonctionnelles par comparaison de cardinalités (comptage)
- Exemple :
  - $Étudiant \nrightarrow Note$:

$$Card(Étudiant) \neq Card(Note) \Rightarrow 3 \neq 4$$

  - $Étudiant \rightarrow Matière$:

$$Card(Étudiant) = Card(Matière) \Rightarrow 3 = 3$$

Étudiant	Matière	Note
Aegon Targaryen	BD	13
Aegon Targaryen	BD	09
Walter White	C++	15
Magnus Maximus	PHP	18

 BY-NC-SA

Théorie des DF

- **Dépendance fonctionnelle totale (DFT)** : ← Ou DF minimale
  - L'ensemble des seuls attributs nécessaires pour déterminer un ensemble d'attributs
- Axiomes d'Armstrong simplifiés :
  - **Réflexivité** : $X \rightarrow X$
  - **Augmentation** : $X \rightarrow Y \;\Rightarrow\; XW \rightarrow Y$
  - **Transitivité** : $X \rightarrow Y,\; Y \rightarrow Z \;\Rightarrow\; X \rightarrow Z$
  - **Pseudo-transitivité** : $X \rightarrow Y,\; YZ \rightarrow W \;\Rightarrow\; XZ \rightarrow W$
  - **Union** : $X \rightarrow Y,\; X \rightarrow Z \;\Rightarrow\; X \rightarrow YZ$
  - **Décomposition** : $X \rightarrow YZ \;\Rightarrow\; X \rightarrow Y,\; X \rightarrow Z$

X, Y, W et Z sont des ensembles d'attributs

(cc) BY-NC-SA

DF et clefs

- **Clef candidate minimale** :
  - Un sous-ensemble d'attributs minimal ayant une dépendance fonctionnelle avec un sur-ensemble d'attributs
- **Clef primaire** :
  - Choix d'une clef parmi les clefs candidates minimales :
    - Priorité aux plus petits nombres d'attributs
    - Priorité des types numériques sur les types chaînes de caractères

Rappel : il peut exister plusieurs clefs candidates minimales dans une relation

 BY-NC-SA

Recherche d'une clef

Clef candidate minimale, et donc clef primaire

- Depuis la DF donnée par l'axiome de réflexivité ($X \rightarrow X$) :
  - Suppression un par un des attributs en se basant :
    - Sur l'ensemble des DF
    - Les axiomes d'Armstrong

- Exemple :
  - Réflexivité : $A, B, C, D \rightarrow A, B, C, D$
  - DF : $A, B \rightarrow C, D$; $B \rightarrow C, D$; $C \rightarrow D$
  - A ? Aucune DF ne détermine A : à garder
  - B ? *Idem supra* : à garder
  - C ? $B \rightarrow C$ (par décomposition) : à supprimer
  - D ? $B \rightarrow D$ (par décomposition) : à supprimer

(A, B) est donc clef candidate minimale

(cc) BY-NC-SA

Cas pratique

BY-NC-SA

Formes normales

- **Les formes normales sont des règles permettant d'éviter les redondances en base de données**
- Défauts des redondances :
  - Perte de place
  - Risques d'incohérences
  - Difficultés de mise à jour
- **Cinq formes normales** :
  - De plus en plus complexes
  - De moins en moins essentielles

Seules les **trois premières formes normales** sont traitées dans le cadre de ce cours

(cc) BY-NC-SA

Première forme normale

- 1NF
- Attribut **atomique** :
  - **Scalaire** (élémentaire)
  - **Non répétitive** (mono-valué : pas de liste de valeurs)
- Valeur **constante** (date de naissance plutôt que âge)

La notion d'élémentaire est relative : une adresse postale complète (a priori composée) peut être considérée comme élémentaire si elle est toujours manipulée comme telle (pas de tri par ville par exemple)

Il n'est pas dérangeant de séparer les attributs groupées, mais pas l'inverse !

Diminue la complexité de traitement, permet une recherche plus rapide et facilite les mises à jour régulières

(cc) BY-NC-SA

Normalisation en 1NF

- Exemple :

Cas 1 : classique

Clef primaire

Attributs non-clefs

Film	Réalisateur	Acteurs
Le Parrain	Francis Ford Coppola	Marlon Brando, Al Pacino, etc.
Bienvenue à Gattaca	Andrew Niccol	Ethan Hawke, Uma Thurman, Jude Law, etc.
Captain Fantastic	Matt Ross	Viggo Mortensen, etc.

Film	Réalisateur
Le Parrain	Francis Ford Coppola
Bienvenue à Gattaca	Andrew Niccol
Captain Fantastic	Matt Ross

Film	Acteur
Le Parrain	Marlon Brando
Le Parrain	Al Pacino
…	…
Captain Fantastic	Viggo Mortensen
Captain Fantastic	…

(cc) BY-NC-SA

Normalisation en 1NF

- Exemple : Cas 2 : nombre maximum fini de valeurs

BG	Téléphones
Regé-Jean Page	+1 424 288 2000, +1 725 228 5100
Chris Hemsworth	+61 2 9663 0100, +61 4 8882 7343
Mickaël Martin Nevot	+33 7 57 13 08 69, +33 4 42 93 90 90

BG	Téléphone 1	Téléphone 2
Regé-Jean Page	+1 424 288 2000	+1 725 228 5100
Chris Hemsworth	+61 2 9663 0100	+61 4 8882 7343
Mickaël Martin Nevot	+33 7 57 13 08 69	+33 4 42 93 90 90

(cc) BY-NC-SA

Deuxième forme normale

- 2NF
- Modèle en 1NF
- **Un attribut non-clef ne dépend pas que d'une partie de la clef primaire**
- DF interdite : $X \rightarrow A$

Une partie de la clef primaire

Un attribut n'appartenant pas à la clef primaire

DF

Nom	Ville	Région
Jul	Marseille	Sud PACA
Alonzo	Marseille	Sud PACA
SCH	Aubagne	Sud PACA
Seth Gueko	Saint-Ouen-l'Aumône	Île-de-France

Pas 2NF

 BY-NC-SA

Troisième forme normale

- 3NF
- Modèle en 2NF
- **Pas de dépendance transitive : un attribut non-clef ne dépend pas d'un ou plusieurs attributs ne participant pas à la clef primaire**
- DF interdite : $X \rightarrow A$

Un attribut n'appartenant pas à la clef primaire

N'est pas, ou ne contient pas, la clef primaire

DF

Nom	Ville	Région
Jul	Marseille	Sud PACA
SCH	Aubagne	Sud PACA
Seth Gueko	Saint-Ouen-l'Aumône	Île-de-France

Pas 3NF

(cc) BY-NC-SA

Normalisation en 3NF

Et 2NF

- Pour chaque relation :
  1. Trouver une DF interdite $X \rightarrow A$
  2. Maximiser (augmenter) la DF $X \rightarrow Y$
  3. Supprimer les attributs de Y de la relation
  4. Nouvelle relation : X, clef primaire, Y, attributs non-clefs

Ajouter tous les attributs déterminés par X

Ne pas oublier ensuite les clefs étrangères !

- Exemple :
  - Relation : | A | B | C | D |

Pas 3NF

  - DF : $A, B \rightarrow C, D$; $B \rightarrow C, D$; $C \rightarrow D$
  - $B \rightarrow C \Rightarrow B \rightarrow$ C, D | A | B |

3NF

| B | C | D |

Pas 3NF

  - $C \rightarrow D$

3NF

A	*B*
B	*C*
C	D

Clefs étrangères : en italique

(cc) BY-NC-SA

Cas pratique

(cc) BY-NC-SA

Aller plus loin

- Forme normale de Boyce-Codd (BCNF), 4NF, 5NF

Crédits

Auteur

Mickaël Martin-Nevot
mmartin.nevot@gmail.com

- Laurent Carmignac

Carte de visite électronique

Relecteurs

Cours en ligne sur : www.mickael-martin-nevot.com