

## Formes normales

1

## Formes normales

- Première forme normale
- Deuxième forme normale
- Troisième forme normale
- ...

2

## 1ère forme normale

- Définition
  - Une relation est en première forme normale si tous ses attributs sont atomiques (inhérent au modèle relationnel)
- Un attribut atomique n'est pas :
  - multivalué (liste de valeurs)
  - composé (structuré en sous-attributs)

3

## Deuxième forme normale

- Définition
  - Une relation est en deuxième forme normale ssi :
    - elle est en première forme normale
    - tout attribut non clé dépend de la totalité de toutes les clés
- Exemple
  - C(nc, dateexp, qtéexp, nb) pas en 2FN car nc, dateexp clé et  $nc \rightarrow nb$

4

## Troisième forme normale

- Objectif : élimination des redondances dues aux dépendances fonctionnelles déduites par transitivité
- Définition
  - Une relation est en troisième forme normale ssi :
    - elle est en deuxième forme normale
    - il n'existe aucune DF entre attributs non clé

5

## Comment calculer une forme normale

Soit  $R(A_1, \dots, A_n)$  et  $DF_R$  ens. de DF associé

- calculer  $DF_R^+$
- déterminer la (les) clé (s) de R
- partitionner les attributs en attributs clés (ils appartiennent à au moins une clé) et attributs non clés
- appliquer les définitions de forme normale (depuis la 1ère)

6

## Propriétés

- Toute relation R admet au moins une décomposition en 3FN qui préserve l'information et les DF
- Deux approches pour la calculer :
  - décomposition : on s'arrête dès que les relations dérivées sont en 3FN (mais pas de garantie sur les DF)
  - synthèse à partir de la couverture minimale

7

## Algorithme de synthèse en troisième forme normale

- Pré : connaissance du contenu de la relation universelle + DF
- Principe de l'algorithme :
  - A partir du graphe G des DF, calculer une couverture minimale C
  - Editer l'ensemble des attributs isolés dans une même relation (tous sont clés)
  - Recherche le plus grand ensemble X d'attributs qui détermine d'autres attributs
  - Editer la relation  $R(X, A_1, \dots, A_n)$
  - Supprimer les DF  $X \rightarrow A_1, \dots, X \rightarrow A_n$  du graphe de couverture minimale C
  - Supprimer les attributs isolés de C
  - Reprendre l'opération à partir de l'étape 3 jusqu'à ce que C soit vide

8

## Insuffisance de la 3FN

### Relation VINS(CRU, PAYS, REGION)

Chenas, France, Beaujolais

Juliéna, France, Beaujolais

Chablis, France, Bourgogne

Chablis, USA, Californie

avec les DF suivantes :

région → pays; cru, pays → région

VINS est en 3FN, pourtant il y a des redondances (ici on peut régler cela avec la 3FN Boyce Codd Kent)

9

## Conclusion sur la normalisation

- Permet d'affiner une conception de schéma
- Peut se coupler avec une démarche de conception "à la Merise"
- Inconvénient majeur : on suppose qu'on possède une couverture minimale des DF (si on a oublié une DF tout est faux)
- Autres formes normales avec d'autres types de dépendance
- Peut être remis en cause au niveau physique (dénormalisation)

10