Formes normales

Un article de Wikipédia, l’encyclopédie libre.

Une forme normale désigne un type de relation entre deux entités dans une base de données relationnelle, en informatique.

La normalisation des modèles de données permet de vérifier la robustesse de leur conception pour améliorer la modélisation (et donc obtenir une meilleure représentation) et faciliter la mémorisation des données (et donc éviter la redondance et les problèmes sous-jacents de mise à jour ou de cohérence). La normalisation s’applique à toutes les entités et aux relations porteuses de propriétés.

Il existe trois principaux types de formes normales :

- la *première forme normale* , où chaque attribut des entités contient une valeur atomique (non composée) ;

```
  exemple:
   produit / fournisseur
  ---------/--------------
     A     / dd,ff
```

Dans ce cas les valeurs du fournisseur sont composées (ne sont pas atomiques). Pour que cette relation soit en première forme normale, il faut décomposer les attributs de la colonne fournisseur comme suit :

```
solution:
   produit / fournisseur
  ---------/--------------
     A     / dd
     A     / ff
```

- la *deuxième forme normale* est une relation en première forme normale où chaque attribut qui n’appartient pas à la clé (l’ensemble des attributs permettant d’identifier de manière unique un tuple de l’entité) ne dépend pas uniquement d’une partie de la clé ;

```
exemple:
 produit / fournisseur  / adresse fournisseur
---------/-------------------------------------
   A     / dd           / 1, rue machin
   B     / dd           / 1, rue machin
   A     / ff           / 5, avenue bidule
```

Admettons que la clé de cette table soit une clé composite (produit - fournisseur). Dans le cas d'un changement d'adresse d'un fournisseur, il faudra faire preuve de beaucoup d'attention pour n'oublier aucun endroit où l'adresse est mentionnée. En effet, on constate que le champ adresse ne dépend que d'une partie de la clé: le champ fourniseur, ce qui induit la possibilité d'une redondance au sein de la table. Il convient donc de scinder la table en deux:

```
solution:
   produit / fournisseur                fournisseur / adresse fourn
  ---------/-------------              -------------/--------------
     A     / dd                            dd       / 1, rue machin
     B     / dd                            ff       / 5, avenue bid
     A     / ff
```

De cette manière, un changement d'adresse ne donne lieu qu'à une seule modification dans la table des fournisseurs.

- la *troisième forme normale* est une relation en deuxième forme normale où les attributs qui ne font pas partie de la clé ne dépendent pas d'attributs ne faisant pas non plus partie de la clé (les attributs sont donc complètement indépendants les uns des autres).

```
exemple:
 élève  / nationalité  / minerval
--------/-------------------------
 ww     / belge        / 1500
 xx     / belge        / 1500
 yy     / français     / 1500
 zz     / australien   / 1850
```

Le montant du minerval n'est pas dépendant de la clé de la table, à savoir l'identifiant de l'étudiant, mais est fonction de la nationalité de l'étudiant: s'il ne fait pas partie de la communauté européenne, il est plus élevé. De nouveau, il est préférable de scinder la table en deux:

```
 élève  / nationalité       nationalité  /  minerval
--------/------------      --------------------------
 ww     / belge             belge        /  1500
 xx     / belge             français     /  1500
 yy     / français          australien   /  1850
 zz     / australien
```

De cette manière, une modification de minerval pour un pays ne donnera lieu qu'à une seule modification.

Il existe également la BCNF (Boyce- Codd Normal Form), ainsi que la 4e et 5e forme normale.

Les différentes formes normales

1FN - première forme normale : toute propriété dépend de la clé
Une entité est en première forme normale si toutes ses propriétés sont élémentaires et s'il existe au moins une clé caractérisant chaque occurrence de l'entité. Toute propriété doit avoir un sens pour toute occurrence de la relation ou de l'entité (il doit y avoir dépendance fonctionnelle vis à vis de la clé). Cette normalisation peut amener à créer des sous-types spécialisés pour les propriétés non systématiquement pertinentes, pour les propriétés répétitives ou multi-valuées.
2FN - deuxième forme normale : toute propriété dépend de *toute* la clé
Une entité ou une relation en première forme normale est en deuxième forme normale si toute propriété dépend de la clé par une dépendance fonctionnelle élémentaire pleine : les propriétés doivent dépendre de l'ensemble des identifiants participant à la clé.
3FN - troisième forme normale : toute propriété dépend de *rien d'autre que* de la clé
Une entité ou une relation en deuxième forme normale est en troisième forme normale si toute propriété dépend de la clé par une dépendance fonctionnelle directe. : les propriétés ne doivent dépendre que de l'identifiant (pas de dépendance transitive). Cette normalisation peut amener à désimbriquer des entités cachées.
BCFN - forme normale de Boyce-Codd
Si une entité ou une relation en troisième forme normale a une clé concaténée, aucune des propriétés élémentaires de cette clé ne doit être en dépendance fonctionnelle d'une autre propriété. Cette normalisation conduit parfois à transformer une entité en relation ou à décomposer une relation en deux relations plus simples.
4FN - quatrième forme normale
Pour toute relation de dimension *n* en forme normale de Boyce-Codd, les relations de dimension *n-1* construites sur sa collection doivent avoir un sens. Il ne doit pas être possible de reconstituer les occurrences de la relation de dimension *n* par jointure de deux relations de dimension *n-1*. Cette normalisation conduit parfois à décomposer une relation complexe en deux relations plus simples.
5FN - cinquième forme normale
Pour toute relation de dimension *n* (avec *n* supérieur à 2) en quatrième forme normale, il ne doit pas être possible de retrouver l'ensemble de ses occurrences par jointure sur les occurrences des relations partielles prises deux à deux. Cette normalisation conduit parfois à décomposer une relation complexe en plusieurs relations plus simples.

- Avertissements