

420-C42

Langages d'exploitation des bases de données

Partie 14

Formes normales

- Dans le contexte des bases de données relationnelles, la normalisation est un processus visant à obtenir une solution quasi optimale d'un point de vue théorique.
- La normalisation est réalisée par un processus rigoureux basé sur un ensemble de règles appelées *formes normales*.
- Le processus vise à :
 - éliminer toutes les redondances
 - découper adéquatement les structures ayant des dépendances
 - garantir l'intégrité des données
- On dit qu'une BDR est normalisée si elle respecte les formes normales.

Formes normales

Sommaire

- Formellement, il existe 9 formes normales :
 - 1^{re} forme normale : 1NF
 - 2^e forme normale : 2NF
 - 3^e forme normale : 3NF
 - forme normale de clé élémentaire : EKNF
 - forme normale de Boyce Codd : BCNF
 - 4^e forme normale : 4NF
 - 5^e forme normale : 5NF
 - forme normale de domaine clé : DKNF
 - 6^e forme normale : 6NF
- Nous verrons seulement les 3 premières.
- Les 4 dernières formes sont moins mises en pratique puisqu'elles sont plus abstraites, plus difficiles à identifier et elles apportent généralement un découpage sévère qui est abrogé lors du processus de dénormalisation.

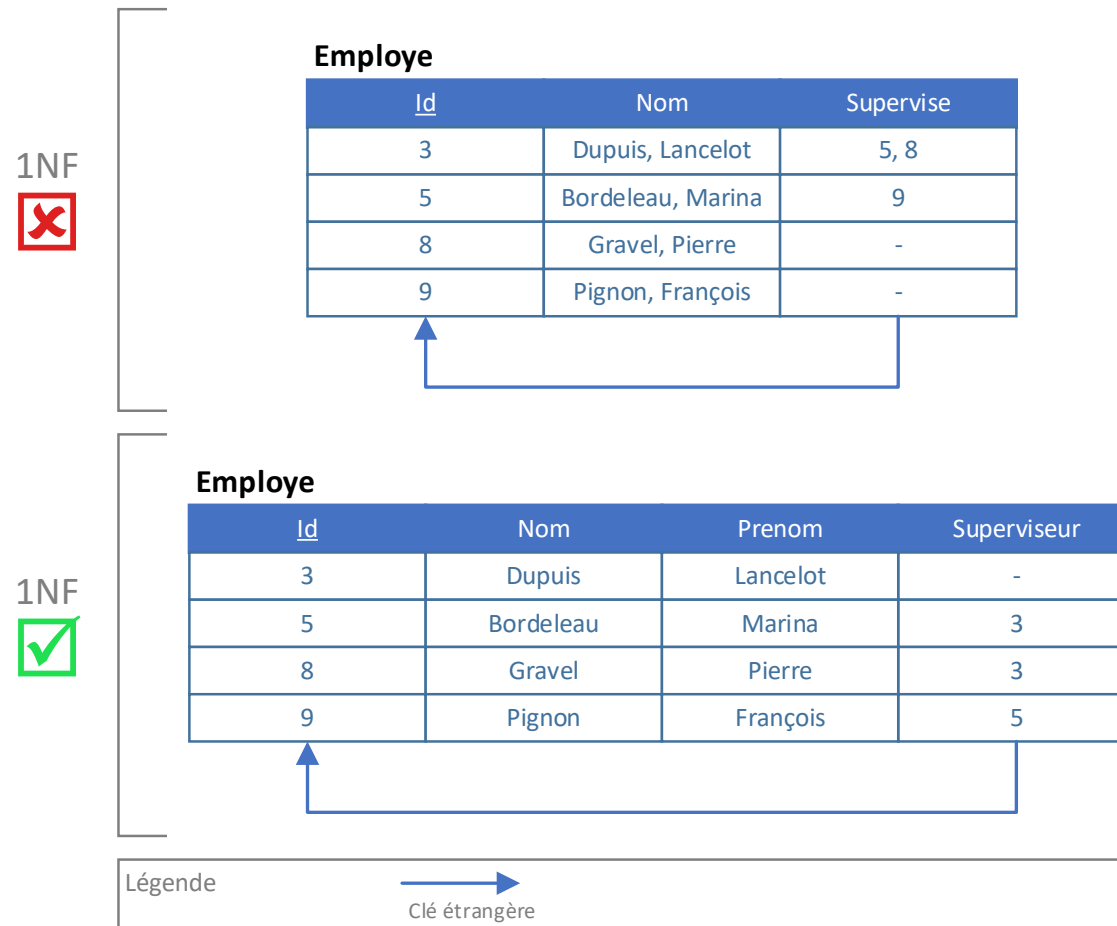
Première forme normale

1NF

- Une table respecte la 1^{re} forme normale si et seulement si :
 - tous les attributs sont atomiques.
- Impacts :
 - On remarque que la prémisse initiale du modèle relationnel est la priorité de la normalisation.
 - Cette caractéristique est l'un des piliers pour qu'une BDR soit sans perte d'intégrité.

Première forme normale

1NF



Deuxième forme normale

2NF

- Une table respecte la 2^e forme normale si et seulement si :
 - elle respecte la 1^{re} forme normale,
 - tous les attributs non-clés sont complètement dépendants de la clé primaire.
autrement dit, les attributs qui ne sont pas membre de la clé primaire doivent être identifiables à partir de l'entièreté de la clé primaire et non pas seulement à partir d'une partie de la clé primaire
- Impacts :
 - Élimine plusieurs redondances.
 - Met en évidence certaines structures sous-jacentes.
 - On remarque que les tables n'ayant pas de clé primaire composée ne sont pas sujette à cette analyse. Ce qui est très souvent le cas.

Deuxième forme normale

2NF

2NF



EtudiantCours

<u>IdEtudiant</u>	<u>IdCours</u>	NomEtudiant	NomCours	Note
3	12	Dupuis	Physique II	90%
3	14	Dupuis	Philosophie I	88%
7	4	Gravel	Mathématique I	76%
9	12	Bordeleau	Physique II	90%

2NF



Etudiant

<u>Id</u>	Nom
3	Dupuis
3	Dupuis
7	Gravel
9	Bordeleau

Cours

<u>Id</u>	Nom
12	Physique II
14	Philosophie I
4	Mathématique I
12	Physique II

EtudiantCours

<u>Etudiant</u>	<u>Cours</u>	Note
3	12	90%
3	14	88%
7	4	76%
9	12	90%

Légende



Clé étrangère



Notion de détermination

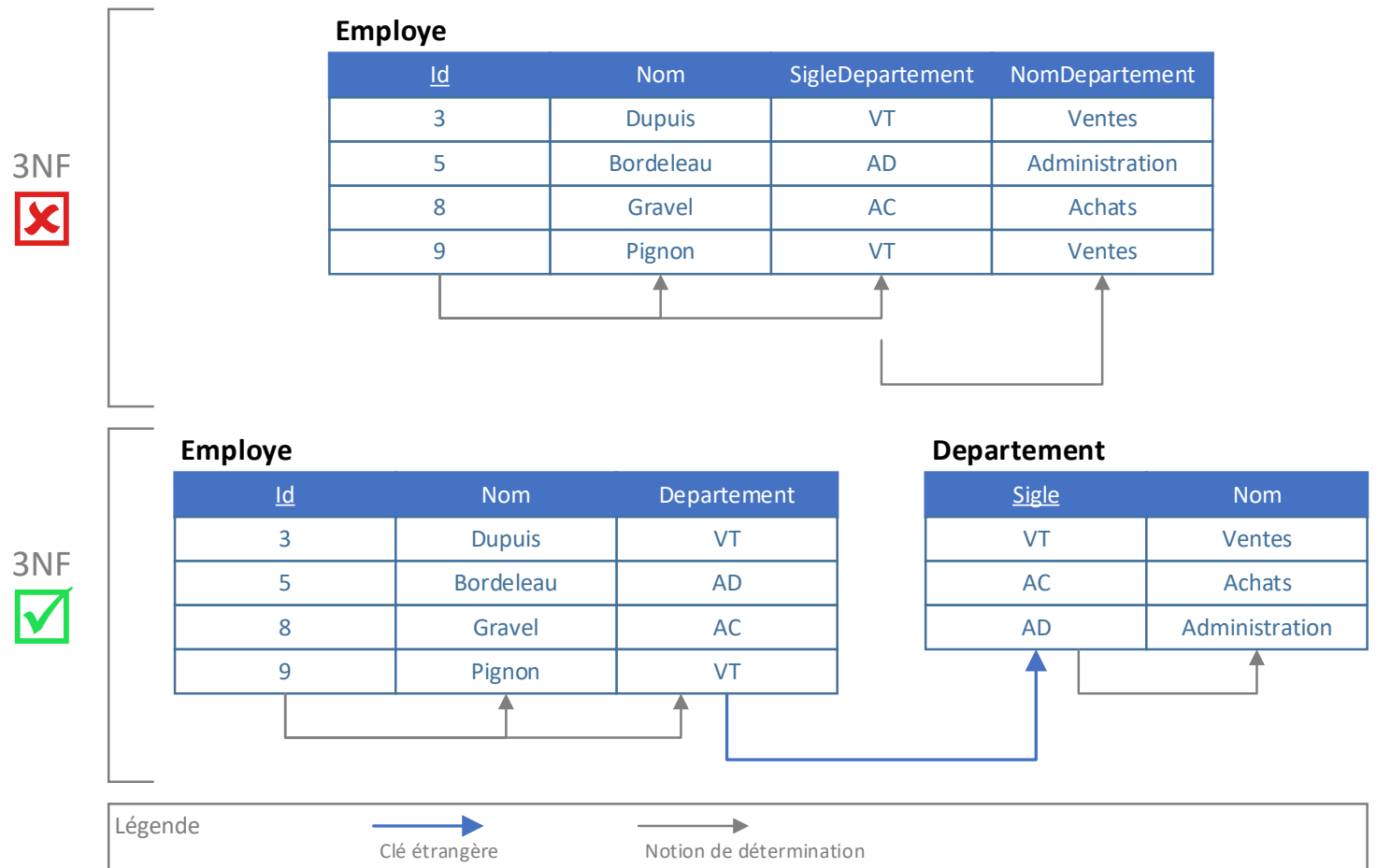
Troisième forme normale

3NF

- Une table respecte la 3^e forme normale si et seulement si :
 - elle respecte la 2^e forme normale,
 - tous les attributs non-clés sont dépendants non transitivement de la clé primaire.
autrement dit, les attributs qui ne sont pas membre de la clé primaire ne dépendent pas d'un attribut qui soit distinct de la clé primaire
- Impacts :
 - Élimine plusieurs redondances.
 - Met en évidence certaines structures sous-jacentes.

Troisième forme normale

3NF



Formes normales

Inconvénients

- Le processus de normalisation génère systématiquement un découpage de plus en plus sévère des données. Ce découpage amène une meilleure garantie de l'intégrité des données.
- Toutefois, cette situation amène un effet indésirable qui est directement lié à la performance des opérations due au nombre croissant de jonctions nécessaires.
- Dans cet esprit, l'étape de dénormalisation permet d'apporter des ajustements qui pourront rendre les opérations dans un coût d'opération acceptable.

Formes normales

Dénormalisation

- La dénormalisation est un processus intentionnel visant à s'éloigner de la normalisation afin d'améliorer les performances globales du SGBD en réduisant la charge de traitement causée principalement par un grand nombre de jointures.
- Cette étape consiste à établir les compromis nécessaires qui tiennent compte des structures les plus sollicitées mais aussi des requêtes les plus récurrentes.
- Le processus de dénormalisation brise systématiquement le processus de normalisation et éloigne volontairement le modèle créé de son état optimum considérant l'intégrité des données.

Formes normales

Dénormalisation

- D'un point de vue formelle, cette tâche est plus complexe qu'il n'y paraît.
- Souvent, les concepteurs se contentent d'une analyse sommaire et intuitive et obtiennent un résultat sous-optimal mais satisfaisant.
- Une approche détaillée implique :
 - analyse détaillée des requis initiaux,
 - études opérationnelles réalisées par la mise en oeuvre d'un plan d'optimisation basé sur des simulations (complexes, exigeants et coûteux),
 - par contre, une telle approche permet d'établir les compromis optimaux.
- En général, ce sont les DBA chevronnés qui s'attaquent à cette problématique.

Normalisation

Processus

- Après avoir fait l'ébauche du modèle relationnel, le processus de normalisation s'applique ainsi :
 - 1NF : validation et correction
 - découpage des attributs en plusieurs attributs
 - ajout de tables et/ou de clés étrangères
 - 2NF : validation et correction
 - ajout de tables et/ou de clés étrangères
 - 3NF : validation et correction
 - ajout de tables et/ou de clés étrangères
 - selon le contexte et le niveau désiré :
 - on applique les autres formes normales
 - souvent jusqu'à la forme normale Boyce Codd (« BCNF »)
 - dénormalisation