

Le modèle relationnel de base de données
Normalisation et Formes normales

1 – Normalisation d'un schéma

Processus de décomposition de relations sans perte d'informations, en se basant sur la notion de dépendance fonctionnelle.

Objectifs :

- Éviter de stocker des données redondantes,
- Simplifier la mise à jour des données,
- Éviter les problèmes d'incohérences.

Exemple: soit la relation R suivante :

R	A	B	C	D
	a1	b1	c1	d1
	a1	b2	c2	d2
	a1	b2	c2	d3
	a2	b1	c3	d4
	a2	b2	c2	d5
	a3	b1	c1	d6

1 – Normalisation d'un schéma

R contient les DF suivantes : $C \rightarrow B$, $A, B \rightarrow C$, $D \rightarrow A, B, C$. On peut en déduire que l'attribut D est une clé candidate de la relation.

- Grâce à la DF $C \rightarrow B$, il est inutile de stocker trois fois le couple $\langle b2, c2 \rangle$, et deux fois le couple $\langle b1, c1 \rangle$.
- La relation R peut être transformée en deux relations : $R1(\underline{C}, B)$ et $R2(\underline{D}, A, C)$
- Sans perte d'informations : la jointure de R1 avec R2 est équivalente à la relation R.

R1	B	C
	b1	c1
	b2	c2
	b1	c3

R2	A	C	D
	a1	c1	d1
	a1	c2	d2
	a1	c2	d3
	a2	c3	d4
	a2	c2	d5
	a3	c1	d6

2 - Formes normales

Définition des trois premières formes normales

- On parle de degré de normalisation d'une relation, chacun correspond à un niveau d'élimination de redondance (ou forme normale : 1NF, 2NF et 3NF)
- Les trois premières formes normales ont pour objectif de permettre la décomposition de relations sans perte d'informations, à partir de la notion de dépendance fonctionnelle.

1ère Forme Normale :

Une relation est en première forme normale si tous les attributs sont atomiques.

Exemple : **PERSONNE**(NO INSEE, NOM, AGE, ADRESSE) n'est pas en 1NF si ADRESSE est composé de VILLE, RUE et NUMERO.

Personne (NO INSEE, Nom, AGE) et **Adresse-Personne** (NO INSEE, VILLE, RUE, NUMERO) est en 1ère forme normale

2 - Formes normales

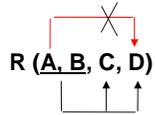
Définition des trois premières formes normales

2ième Forme Normale :

Une relation est en deuxième forme normale si :

- Elle est en première forme normale et
- Il n'y a pas de dépendance fonctionnelle entre un sous-ensemble de la clé et un sous-ensemble des attributs non clés.
- Autrement dit, toutes les DF des clés vers les attributs non-clés sont élémentaires.

Il faut donc éviter la configuration suivante :

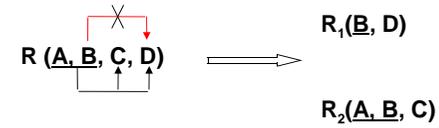


2 - Formes normales

Définition des trois premières formes normales

Décomposition en 2NF d'une relation :

Soit $r : R(U)$ avec $X \rightarrow Y$ et $U = XYZ$, alors $r = \Pi_{XY}(r) \bowtie \Pi_{XZ}(r)$ est une relation en deuxième forme normale.



Paraître (NoLivre, Nom Editeur, Titre, Année_Parution)
n'est pas en 2ième forme normale car NoLivre → Titre

Livre (Nolivre, Titre)

Paraître (Nolivre, Nom Editeur, Année_Parution)
sont des relations en 2ième forme normale

2 - Formes normales

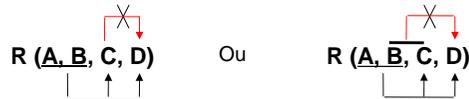
Définition des trois premières formes normales

3ième Forme Normale :

Une relation est en troisième forme normale si :

- Elle est en deuxième forme normale et
- Tout attribut n'appartenant pas à la clé ne dépend pas d'un attribut non clé.

→ Autrement dit, il faut éviter les configurations suivantes :

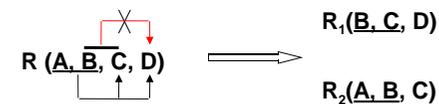
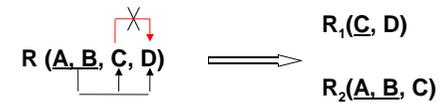


2 - Formes normales

Définition des trois premières formes normales

Décomposition en 3NF :

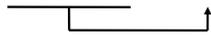
La aussi, la décomposition peut être obtenue comme suit :



2 - Formes normales

Exemple :

Livre (NoLivre, Titre, No_Collection, Nom_Collection)



n 'est pas en 3ième forme normale car Non_Collection est entièrement déterminée par la rubrique No_Collection

Livre (Nolivre, Titre, No_Collection)

Collection (No_Collection, Nom_Collection)

sont des relations en 3ième forme normale

2 - Formes normales

Propriétés d'une décomposition en troisième forme normale

- Les dépendances fonctionnelles sont des règles indépendantes du temps que doivent vérifier les valeurs des attributs. Il est nécessaire qu'une décomposition préserve ces règles.

Décomposition préservant les dépendances fonctionnelles

Décomposition { R1, R2, ... Rn } d'une relation R telle que la fermeture transitive des DF de R est la même que celle de l'union des DF de { R1, R2, ... Rn }

2 – Formes normales

Propriétés d'une décomposition en troisième forme normale

Exemple : Soient les deux décompositions suivantes de la relation voiture

Voiture (NV, Marque, Type, Puissance, Couleur)

en

R1 (NV, Type, Couleur)

R2 (Type, Marque, Puissance)

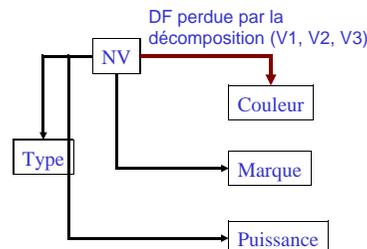
et

V1 (NV, Type)

V2 (Type, Puissance, Couleur)

V3 (Type, Marque)

La décomposition (R1, R2) préserve les dépendances fonctionnelles, alors que la décomposition (V1, V2, V3) ne les préserve pas



2 - Formes normales

Forme normale de Boyce-Codd

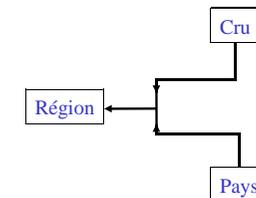
Exemple : Soit la relation VINS (Cru, Pays, Région) avec les dépendances fonctionnelles suivantes :

Région → Pays

(Cru, Pays) → Région

Cette relation est en troisième forme normale. Cependant, elle présente de nombreuses redondances

Vins	Cru	Pays	Région
	Chambertin	France	Bourgogne
	Juliéas	France	Beaujolais
	Morgon	France	Beaujolais



Extension et DF de la relation Vins

2 - Formes normales

Forme normale de Boyce-Codd

- Afin d'éliminer ces types de redondances, BOYCE et CODD ont introduit la forme normale qui porte leur nom

Forme normale de BOYCE-CODD :

Une relation est en BCNF si et seulement si les seules dépendances fonctionnelles élémentaires sont celles dans lesquelles une clé détermine un attribut

- Une décomposition en BCNF ne préserve pas en général les DF

Exemple : La relation VINS (Cru, Pays, Région) peut-être décomposée en deux relations :

Crus (Cru, Région)

Régions (Région, Pays)

La dépendance fonctionnelles (Cru, Pays) → Région est perdue

3 – Algorithme de normalisation

Algorithme de synthèse :

Entrée : Schémas ne contenant que des DF

Sortie : Schéma (R1, R2,..., Rn) avec Ri en 3NF, quel que soit i.

Étape 1 : Calculer une couverture minimale de F' notée MIN(F').

Exemple : Couverture minimale

A partir de l'ensemble de DF {A → B, B → A, A,B → C, C → B}, on obtient la couverture minimale suivante : {A → B, B → A, A → C, C → B}.

Étape 2 : Partitionner MIN(F') en groupes F'1, F'2,..., F'k tels que toutes les DF d'un même groupe aient la même partie gauche.

Exemple : condensation

En poursuivant, la condensation nous donne {A → B, C, B → A, C → B}.

3 – Algorithme de normalisation

Algorithme de synthèse :

Étape 3 : Pour chaque groupe F'i, i =1, ..., k, construire un schéma contenant les attributs de F'i et les DF de F'i.

Exemple : Schémas de relations : R1(A, B, C), R2(B, A) et R3(C, B).

Étape 4 : Regrouper progressivement et deux à deux les relations qui se réfèrent mutuellement via des clés étrangères potentielles,

Soit R₁(X₁, Y₁) et R₂(X₂, Y₂) avec X₁ ⊆ Y₂ et X₂ ⊆ Y₁ devient R₁₂(X₁, X₂, Y₁, Y₂).

Exemple : Dans l'exemple précédent, les deux schémas : R₁(A, B, C) et R₂(B, A) se réfèrent mutuellement. On les regroupe en R₁₂(A, B, C).

Cette nouvelle relation et R₃(C, B) se réfèrent mutuellement. Un second regroupement ne laisse plus que R₁₂₃(A, B, C), munie de trois clés candidates.