

LIFBDW2 – BASES DE DONNÉES AVANCÉES

TD6 – Formes normales, couvertures

Licence informatique – Automne 2020–2021

Les questions marquées du symbole (†) sont à préparer pour la séance

Exercice 1 : normalisation (†)

Donner la forme normale la plus avancée des schémas de relation suivants munis de l'ensemble Σ de contraintes. Si ce n'est pas fait, normaliser la relation en la décomposant.

1. $Employe(NomEmp, NomDept, NomChef)$
avec $\Sigma = \{NomEmp \rightarrow NomDept; NomDept \rightarrow NomChef\}$;
2. $Employe(NomEmp, NomEnfant, Salaire)$
avec $\Sigma = \{NomEmp \rightarrow Salaire; NomEmp \twoheadrightarrow NomEnfant\}$;
3. $Adresse(Rue, Ville, CodePostal)$
avec $\Sigma = \{Rue, Ville \rightarrow CodePostal; CodePostal \rightarrow Ville\}$;

Exercice 2 : définitions de la troisième forme normale

L'objectif de cette exercice est de montrer que les deux définitions **3FN-1** et **3FN-2** ci-après de la troisième forme normale sont équivalentes. On considère les paires $\langle R, F \rangle$ formées d'un schéma de relation R et d'un ensemble de F de dépendances fonctionnelles sur R . On donne les définitions suivantes :

Directe Une DF $X \rightarrow Y$ est *directe* ssi $\nexists Z. X \rightarrow Z \wedge Z \twoheadrightarrow X \wedge Z \rightarrow Y$.

Premier Un attribut $A \in R$ est *premier* s'il appartient à *au moins* une clé minimale de R .

2FN $\langle R, F \rangle$ est en **2FN** ssi il n'existe pas de DF non-triviale $X \rightarrow A \in F^+$ avec A non-premier et X sous-ensemble *propre*¹ d'une clé minimale.

3FN-1 $\langle R, F \rangle$ est en **3FN** ssi $\langle R, F \rangle$ est en **2FN** et toutes les DFs non-triviales $X \rightarrow A \in F^+$ avec A non-premier sont *directes*.

3FN-2 $\langle R, F \rangle$ est en **3FN** ssi pour toute DF non-triviale $X \rightarrow A \in F^+$, soit X est (super)clé, soit A est premier.

1. Montrer que si la condition **2FN** n'est pas respectée, alors la **3FN-2** ne l'est pas non plus.
2. Montrer que s'il existe une DF non-triviale non-directe $X \rightarrow A \in F^+$ avec A non-premier, alors la condition **3FN-2** n'est pas respectée.
3. Conclure des deux questions précédentes que la définition **3FN-2** implique la définition **3FN-1**.
4. Montrer que s'il existe une DF $X \rightarrow A$ avec A non-premier et X qui n'est pas (super)clé, alors la condition **3FN-1** n'est pas satisfaite.
5. Conclure sur l'équivalence des définitions **3FN-1** et **3FN-2**.

1. X est un sous-ensemble *propre* de Y noté $X \subsetneq Y$ ssi $X \subseteq Y \wedge X \neq Y$

Exercice 3 : modélisation de recettes

On souhaite créer une base de données de recettes de cuisine, décrites comme suit :

Une recette est identifiée par son numéro. Elle a un nom et un type particulier (soupe, entrée, dessert, ...). Elle utilise un ou plusieurs ingrédients (carottes, viande de boeuf, poivre, ...). Un ingrédient est identifié par son numéro et a un nom. Pour chaque ingrédient dans une recette, on précisera sa quantité.

1. Dresser l'inventaire des DF valides sur la relation universelle $R = \{NumR, NomR, TypeR, NumI, NomI, Qte\}$
2. Exhiber les problèmes de redondance à partir de cet exemple. La DF $NumR, NumI \rightarrow Qte$ engendre-t-elle un problème de redondance ?
3. Proposer de façon intuitive une décomposition sans pertes et sans redondance de R .
4. On veut maintenant ajouter à cette modélisation qu'à chaque recette correspond un ensemble d'ustensiles (attribut *Ustensile*). Cette information se traduit-elle par une nouvelle DF ?
5. Que penser de la décomposition $R_1 = (NumR, NomR, TypeR)$, $R_2 = (NumR, NumI, Qte)$, $R_3 = (NumI, NomI)$, $R_4 = (NumR, Ustensile)$.

Exercice 4 : modélisation d'une base de films

On souhaite modéliser des données sur des films. On analyse les données de la façon suivante.

Chaque film est identifié par un unique numéro d'exploitation. Un film est également déterminé par son titre et sa date de sortie. Un film peut être une reprise d'un ou plusieurs autres films avec lesquels il possède un taux de ressemblance. Chaque film a un réalisateur et un producteur, et implique un ou plusieurs studios et un ensemble d'acteurs. Pour chaque acteur, on souhaite stocker le nombre de minutes qu'il joue dans un film ainsi que son cachet pour le film.

Notons qu'un réalisateur peut-être producteur ou acteur dans un autre film, et inversement. Un réalisateur/producteur/acteur possède un identifiant unique, un nom, un prénom et une adresse. La connaissance du nom, du prénom et de l'adresse permet d'identifier uniquement une personne. Un studio est identifié par un numéro et possède un nom unique et une adresse.

1. Faire l'inventaire des attributs (relation universelle) et des contraintes (DF, DI) du problème.
2. Donner les clés (minimales) de la relation universelle
3. Proposer une décomposition sans pertes de la relation universelle en précisant les clés primaires et étrangères.