

BD Relationnelles TD3

Normalisation, DFs et DMs

1. Un conseiller $C\# = C1$ de la compagnie d'assurances FIAM a trois clients caractérisés par ($CI\#, NOM$) avec trois valeurs ($c11, nom1$), ($c12, nom2$) et ($c13, nom3$). Proposez la clé et l'extension de la table:

C ($C\#, CI\#, NOM$).

Supposez que (i) que C n'a qu'un tuple en 0-NF, puis (ii) créez autant qu'il en faut en 1 NF. Ensuite, analysez d'une manière similaire la table:

E ($E\#, Email, Tel, Dipl$).

E décrit l'étudiant E1 possédant 3 adresses email, 3 téléphones et 3 diplômes.

2. Soit $PS = P \text{ JOIN } SP$, où P et SP sont les tables de la base $S-P$ du cours. PS montre donc la description de toute pièce et de toutes les fournitures de celles-ci.

- Montrez le schéma (l'intention) et les tuples (l'extension) de PS
- Montrez la clé et quelques dépendances fonctionnelles (DFs)
- Montrez que $(P\#, COLOR) \rightarrow (WEIGHT, COLOR)$ est une DF
- Montrez que $P\# \rightarrow PNAME, WEIGHT, COLOR$ est une DF
- Montrez que $P\# \rightarrow QTY$ n'est pas une DF
- Montrez que $P\# \rightarrow (PNAME, CITY, COLOR, WEIGHT)$ est une DF
- Proposez une décomposition sans perte par le Théorème de Heath.
- Est-ce que les tables résultantes sont en BCNF ?
- Montrez l'extension des tables résultantes de votre décomposition de PS dans la base $S-P$ et le résultat de leur jointure.

3. Considérez la base du cours sous l'intention (schéma) de la table universelle :

U ($S\#, SName, Status, City, P\#, PName, Weight, Color, PCity, Qty$)

Montrez la clé et l'extension 1NF de U , suivant les tuples de la base $S-P$ du cours. Montrez les redondances et les anomalies. En supposant les DFs du cours, si U n'est pas en BCNF, alors décomposez U sans pertes d'information. Montrez que cette décomposition préserve les DFs. Ensuite, montrez l'intention, la couverture *minimale*, c'est à dire par les DFs *élémentaires* (non-décomposables par les règles d'Armstrong), et l'extension de tables intermédiaires et finales. Continuez la normalisation jusqu'à ce que toute table soit en BCNF.

4. Supposez que dans U on a en plus la DF : **Weight \rightarrow Color**. Est-ce que le schéma de $S-P$ que vous avez produit dans (3) est toujours sans anomalies ? Sinon, proposez deux décompositions supplémentaires par le Th. de Heath. L'une en projections *indépendantes* et l'autre en celles *dépendantes*. Montrez que la décomposition en projections indépendantes préserve mieux la consistance de la BD.

5. En tant que conseiller de la FIAM ayant suivi ce cours de BDs à Dauphine, vous pensez mettre vos clients de dans la table : **C ($C\#, Nom, Tel, Adresse, A\#$)**. L'attribut $A\#$ identifie chaque assurance suscrite par le client. Les attributs (Nom, Tel) déterminent $C\#$ et vice versa. Les attributs (Nom, Tel, $A\#$) forment-ils la clé candidate ? Serait-elle minimale ? Est-ce que C serait en BCNF ? Sinon, proposez une décomposition de C sans perte d'info par le Th. de Heath. Montrez qu'il y a deux décompositions possibles. L'une est en pratique meilleure que l'autre. Pourquoi ?

6. Votre table C serait désormais **C ($C\#, Nom, Tel, Email, Adresse, A\#$)**, en supposant qu'un client peut avoir plusieurs numéros de tél et plusieurs adresses email. On suppose aussi qu'il y a une DF $Email \rightarrow C\#$. Identifiez des DFs et des DMs dans C . Si C n'est pas en 4 NF, identifiez les anomalies, puis décomposez C sans pertes d'info selon le Th. de Fagin généralisant celui de Heath pour les DMs.

7. Vous voulez ensuite créer la table suivante pour les assurances d'un client : **A ($A\#, N-Client, Montant, AV\#, Franchise, AM\#, V-Mobilier$)**. L'attribut $AV\#$ identifie l'assurance d'une voiture parmi toutes assurances de cette classe. Idem $AM\#$ identifie celle de maison. Une assurance identifiée par $A\#$ peut-être d'une voiture, d'une maison, les deux à la fois ou une autre non-spécifiée. Proposez une décomposition sans perte minimisant la présence de valeurs nulles, selon la généralisation du Th. De Heath aux jointures externes gauches (LEFT JOINT).

8. Vous avez décomposée la table **C ($C\#, Nom, CPostal, Ville$)** en deux tables **C ($C\#, Nom, CPostal$)** et **CV ($CPostal, Ville$)**, étant données les DFs $C\# \rightarrow (Nom, CPostal)$ et $CPostal \rightarrow Ville$. Pourquoi ? Après cet effort, on vous a dit que pas mal de vos futurs clients connaîtront la ville, mais pas le code postal. Est-ce que ceci pose

un problème pour le CS de votre base ? Si oui quelle approche vous pouvez appliquer pour le résoudre ?
Comment votre CS serait modifié alors et quel serait le désavantage apparu ?