
TP7: Normalisation avec DBD

Durée : 2 séances

Dans ce TP, vous travaillez avec l'outil de normalisation **Database Design (DBD) Toolkit**, qui fait part du support pédagogique du livre *Fundamentals of Database Systems* de Ramez Elmasri et Shamkant B. Navathe. Cet outil en SWI-Prolog délivre une série d'algorithmes pour la normalisation : clôture d'un ensemble d'attributs pour un schéma avec ensemble de dépendances fonctionnelles *xplus*, identification des clés candidates, clôture d'un ensemble de DFs selon les axiomes d'Armstrong (*fplus*, *finfplus*), équivalence d'un ensemble de DFs (*equiv*), tester si une décomposition n'entraîne pas de tuples parasites - dans ce cas, elle s'appelle une *lossless join decomposition* (prédicat : *ljd*), et des algos pour tester et générer la 3NF et la BCNF (*is3NF*, *threenf*, *isBCNF*, *bcnf*).

Téléchargez DBD de Moodle. L'archive contient

- *ch5/dbd.pl* Le code source de DBD.
- **ch5/ch5.pdf** Un tutoriel pour l'utilisation de DBD avec des exemples d'utilisation de tous les prédicats. Seule adaptation nécessaire au M5 : vous lancez Prolog avec *swipl* au lieu de *pl*. En fin du tutoriel, une série d'exercices. L'exo 6 du tutoriel correspond à l'exo 1 de votre TD8, le 4 au 4.
- Le fichier *ex.pl* avec un modèle de ce qui est attendu, par exercice : un fichier *exoX.pl* (X : numéro de l'exo), avec le codage du schéma, des FDs ainsi que par question, un prédicat *answer*. Puis vous ajoutez en commentaire la capture du résultat de votre requête, et finalement d'éventuelles remarques/conclusions/observations de votre part.

Exercice 1 :

Faites les 14 exercices du tutoriel de DBD. Répondez également aux questions suivantes.

Question 1.1 : Comment peut-on quantifier la taille de la clôture d'un ensemble de DF ?

Question 1.2 : Vous pouvez facilement observer que DBD vous propose plus d'une décomposition en 3NF (il suffit de taper un point-virgule après l'affichage de la première réponse). Comment se distinguent-elles ? Comment faire un choix idéal parmi les propositions ? Comment reconnaître si l'une est meilleure que l'autre ?

Question 1.3 : Une décomposition n'est pas utile, si elle génère des *tuples parasites* lors de la jointure. L'algorithme 11.4 du livre d'Elmasri et Navathe, qui est implémenté dans les prédicats de 3NF de DBD, ne garantit pas la propriété qui, pour des raisons historiques, s'appelle *lossless-join*. Implicitement, l'exo 8 du tutoriel vous demande de décider quelle décomposition est préférable.

Question 1.4 : Prenez également en compte, si une décomposition conserve les dépendances fonctionnelles, ou non.