

LIFBDW2 – BASES DE DONNÉES AVANCÉES

TD2 – Introduction aux dépendances

Licence informatique – Automne 2019–2020

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

Exercice 1 : modélisation avec les dépendances (†)

Soit R le schéma de bases de données suivant :

- $Films = \{IDFilm, Titre, Annee, IDStudio\}$;
- $Reprises = \{IDReprise, IDOriginal, Similarite\}$;
- $Studios = \{IDStudio, Nom, Adresse\}$.

1. Trouver la dépendance fonctionnelle ou d'inclusion permettant de restreindre les extensions possibles de cette base, pour chacune des assertions suivantes :
 1. chaque film a un identifiant unique à partir duquel on connaît tous ses attributs ;
 2. la même année, deux films ne peuvent pas avoir le même titre ;
 3. chaque studio doit avoir effectivement participé à la réalisation d'un film ;
 4. un film peut être repris plusieurs fois, un film peut reprendre plusieurs films et pour chaque film repris par un autre, il y a un unique taux de similarité.
2. En supposant les contraintes précédentes satisfaites, peut-on également demander que plusieurs studios puissent participer à la réalisation d'un même film ? Justifier et proposer une solution au problème.

Exercice 2 : préservation des dépendances par les requêtes

Soit r une instance de R qui satisfait à la dépendance $R : X \rightarrow Y$ (soit $r \models X \rightarrow Y$) et s une instance quelconque. Pour chaque expression ci-dessous, indiquer en le justifiant si elle est vraie.

1. $\sigma_C(r) \models X \rightarrow Y$
2. $r \cup s \models X \rightarrow Y$
3. $r \setminus s \models X \rightarrow Y$
4. $\pi_W(r) \models X \rightarrow Y$
5. $r \times s \models X \rightarrow Y$
6. $r \bowtie s \models X \rightarrow Y$

Exercice 3 : axiomatisation des dépendances fonctionnelles

On rappelle les règles d'inférences suivantes pour les dépendances fonctionnelles.

$$\frac{Y \subseteq X}{X \rightarrow Y} \sigma_R \text{ (réflexivité)}$$

$$\frac{X \rightarrow Y \quad X \rightarrow Z}{X \rightarrow YZ} \sigma_C \text{ (composition)}$$

$$\frac{X \rightarrow Y}{WX \rightarrow WY} \sigma_A \text{ (augmentation)}$$

$$\frac{X \rightarrow YZ}{X \rightarrow Y} \sigma_D \text{ (décomposition)}$$

$$\frac{X \rightarrow Y \quad Y \rightarrow Z}{X \rightarrow Z} \sigma_T \text{ (transitivité)}$$

$$\frac{X \rightarrow Y \quad WY \rightarrow Z}{WX \rightarrow Z} \sigma_P \text{ (pseudo-transitivité)}$$

1. Donner une preuve que $\{AB \rightarrow C, A \rightarrow D, CD \rightarrow EF\} \models AB \rightarrow F$ en utilisant le système $\{\sigma_R, \sigma_A, \sigma_T\}$
2. La règle suivante est-elle correcte ?

$$\frac{XW \rightarrow Y \quad XY \rightarrow Z}{X \rightarrow (Z \setminus W)}$$

3. Montrer que toute preuve de $F \models X \rightarrow Y$ utilisant la règle σ_P peut être transformée en une preuve n'utilisant que σ_A et σ_T .
4. Montrer que toute preuve de $F \models X \rightarrow Y$ utilisant les règles σ_R, σ_A et σ_T peut être transformée en une preuve n'utilisant que σ_R et σ_P .
5. En déduire que le système $\{\sigma_R, \sigma_P\}$ est correct et complet pour l'inférence des DFs.

Exercice 4 : adéquation du système d'Armstrong

1. Démontrer que les règles du système d'Armstrong (réflexivité, transitivité et augmentation) sont justes en exploitant la définition de la satisfaction d'une dépendances.

Exercice 5 : vérification des dépendances en SQL

1. Prouver que $r \models X \rightarrow Y$ si et seulement si $|\pi_X(r)| = |\pi_{XY}(r)|$, en déduire une méthode qui permet de tester la satisfaction d'une dépendance fonctionnelle avec SQL. Commenter son efficacité par rapport à une méthode faisant intervenir seulement la sémantique des dépendances.
2. Prouver que $r, s \models R[X] \subseteq S[Y]$ si et seulement si $|\pi_X(r) \setminus \pi_Y(s)| = 0$, en déduire une requête SQL qui permet de tester la satisfaction d'une dépendance d'inclusion.