Lif10 – Fondements des bases de données
TD5 – Relation d'Armstrong, couvertures et dépendances d'inclusion

Licence informatique – Automne 2014–2015

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

Algorithme 1: calcul d'une relation d'Armstrong

```
Data : R un schéma de relation, F un ensemble de DF sur R.
Result : Une relation d'Armstrong r pour F.
for A ∈ R do
    t[A] := 0
r := {t}
i := 1
for X ∈ Cl(F) \ R do
    for A ∈ R do
        if A ∈ X then
            t[A] := 0
        else
            t[A] := i
    r := r ∪ {t}
    i := i + 1
return r
```

Exercice 1 : relation d'Armstrong (†)

Soit un schéma de relation R et un ensemble F de DF sur ce schéma. Une *relation d'Armstrong* pour F est une relation r sur R vérifiant *exactement* F^+. Soit le schéma $R = ABCDE$ et l'ensemble de DF $F = \{A \to BC, D \to E, C \to D\}$ pour la suite.

1. (†) Calculer l'ensemble des fermés de F, défini par $Cl(F) = \{X^+ \mid X \subseteq R\}$. Pour cela, utilisez une exploration systématique des sous-ensembles de R en commençant par ceux de taille 1, puis 2 ... Penser à ne pas visiter les sur-ensembles des clés pour limiter le nombre de calculs.
2. (†) Utiliser l'algorithme 3 pour calculer une relation d'Armstrong r.
3. Supposons que l'on supprime le tuple correspondant au fermé BE de la relation r obtenue à la question précédente. Il y a-t-il des DFs nouvellement vérifiées qui ne l'étaient pas ?
4. Même question que précédemment avec le fermé DE.
5. Quand $F = \emptyset$, combien il y a-t-il de tuples dans la relation d'Armstrong ? Même question quand F est l'ensemble de toutes les DFs possibles sur R ?

Algorithme 2: calcul d'une couverture minimum

Data : F un ensemble de DF
Result : G une couverture minimum de F

$G := \emptyset$
for $X \to Y \in F$ **do**
    $G := G \cup \{X \to X^+\}$;
for $X \to X^+ \in G$ **do**
    if $G - \{X \to X^+\} \vdash X^+$ **then**
        $G := G - \{X \to X^+\}$;
return G;

Algorithme 3: réduction des attributs d'un ensemble minimum de DFs

Data : Un ensemble *minimum* de DF F sur R.
Result : F avec un nombre minimal d'attributs

$Min := F$
`// Réduction des parties gauches`
for $X \to Y \in Min$ **do**
    $W := X$
    for $A \in X$ **do**
        if $Min \models (W \setminus \{A\}) \to X$ **then** $W := W \setminus \{A\}$
    $Min := (Min \setminus \{X \to Y\}) \cup \{W \to Y\}$
`// Réduction des parties droites`
for $X \to Y \in Min$ **do**
    $W := Y$
    for $A \in Y$ **do**
        $G := (Min \setminus \{X \to Y\}) \cup \{X \to (W \setminus \{A\})\}$
        if $G \models X \to Y$ **then** $W := W \setminus \{A\}$
    $Min := (Min - \{X \to Y\}) \cup \{X \to W\}$
return Min

Exercice 2 : couverture minimale (†)

1. (†) Soient $F_1 = \{A \to C, AC \to D, E \to AD, E \to F\}$ et $F_2 = \{A \to CD, E \to AF\}$. Déterminer si chaque DF de F_2 peut être inférée par F_1 et réciproquement.
2. En utilisant l'algorithme 2 déterminer une couverture minimale de F_1. Comment comparer la couverture ainsi obtenue à F_2 ?
3. En utilisant l'algorithme 2 déterminer si $F = \{A \to BC, D \to E, C \to D\}$ est minimal.

Exercice 3 : Couvertures non redondantes

Soit l'ensemble F de DF suivant le schéma $R = ABCDEFG$:

$D \to A$	$D \to C$	$D \to E$
$D \to F$	$CE \to G$	$AG \to F$
$ADG \to B$	$BG \to ADE$	$BF \to DG$

1. Démontrer que F n'est ni optimum, ni minimum, ni non-redondante.
2. Calculer une couverture canonique de F (couv. minimum + réduction, algorithme 3).

Exercice 4 : dépendances d'inclusion (†)

Soit $I = \{R[ABC] \subseteq S[EFG], S[F] \subseteq T[J], S[FE] \subseteq T[LK]\}$ l'ensemble de dépendances d'inclusion sur le schéma de bases de données $\mathbf{D} = \{R, S, T\}$ avec $R = ABCD, S = EFGH$ et $T = IJKL$.

1. (†) Prouver $I \models R[AB] \subseteq T[KL]$ à l'aide du système de Casanova donné en figure 1, en précisant σ quand la règle de permutation et projection est utilisée.

2. Enumérer les différentes DI que l'on peut produire par application de la règle de projection et permutation sur $R[ABC] \subseteq S[EFG]$. Les dénombrer dans le cas général.
3. (†) Prouver la propriété suivante d'intéraction entre DFs et DIs

$$\{R[XY] \subseteq S[TU], S : T \to U\} \models R : X \to Y$$

4. Même question pour la propriété

$$\{R[XY] \subseteq S[TU], R[XZ] \subseteq S[TV], S : T \to U\} \models R[XYZ] \subseteq S[TUV]$$

$$\frac{}{R[X] \subseteq R[X]}\ \gamma_R \text{ (réflexivité)} \qquad \frac{R[X] \subseteq S[Y] \qquad S[Y] \subseteq T[Z]}{R[X] \subseteq T[Z]}\ \gamma_T \text{ (transitivité)}$$

$$\frac{R[A_1 \dots A_n] \subseteq S[B_1 \dots B_n]}{R[A_{\sigma(1)} \dots A_{\sigma(k)}] \subseteq S[B_{\sigma(1)} \dots B_{\sigma(k)}]}\ \gamma_P \text{ (permutation \& projection)}$$

avec σ une permutation d'un sous-ensemble de $\{1 \dots n\}$

FIGURE 1 – Axiomatisation de Casanova pour les DIs