Théorie des bases de données et de connaissances (HAI933I)
Contrôle n°2 : Sujet et Correction

Durée : 40 mn. Sans documents.

Exercice 1 (10 pts)

Question 1 (1 pt). Qu'est-ce qu'un *modèle universel* d'une base de connaissances $\mathcal{K}$?

Correction: Voir cours.

Question 2 (1 pt). Quel est l'intérêt de cette notion pour répondre à des requêtes conjonctives (disons : booléennes pour simplifier) sur une base de connaissances ?

Correction: Voir cours.

Question 3 (6 pts). Soit la base de connaissances $\mathcal{K} = (F, \mathcal{R})$ où $F = \{s(a), p(a,b), q(b,b)\}$ et $\mathcal{R} = \{R_1, R_2, R_3\}$ où :

$R_1 : s(x) \to \exists z\ q(x,z) \wedge r(x)$
$R_2 : r(x) \wedge p(x,y) \to q(x,y)$
$R_3 : q(x,y) \to \exists z\ q(y,z)$

On effectue le *restricted chase* sur $\mathcal{K}$ (rappel en annexe). On procède en largeur. Dire :
- quelles sont les règles appliquées aux 3 premières étapes de largeur et quels sont les nouveaux faits obtenus.
- quel est le résultat du chase.

Correction: On demande un algorithme qui procède en largeur.
Etape 1 : Seule R_1 est appliquée. R_3 n'est pas applicable à cause de $q(b,b)$: l'homomorphisme $\{x \mapsto a, y \mapsto b\}$ du corps de R_3 dans F s'étend à un homomorhisme de tous les atomes de R_3 dans F : $\{x \mapsto a, y \mapsto b, z \mapsto b\}$. On obtient $F_1 = F \cup \{q(a,z_0), r(a)\}$.
Etape 2 : R_2 et R_3 sont appliquées (peu importe dans quel ordre ici) et on obtient $F_2 = F_1 \cup \{q(a,b), q(z_0,z_1)\}$.
Etape 3 : R_3 est appliquée et on obtient $F_3 = F_2 \cup \{q(z_1,z_2)\}$. On voit que le chase va construire un “chemin infini” partant de a.
Le résultat du chase est :

$$F^\star = \{s(a), p(a,b), q(b,b), r(a), q(a,b), q(a,z_0)\} \cup \{q(z_i, z_{i+1}) | i \geq 0\}$$

Question 4 (2 pts). Montrer que $\mathcal{K}$ admet un modèle universel fini.

Correction: Le core chase s'arrête sur $\mathcal{K}$, qui admet donc un modèle universel fini. Le résultat du core chase est :

$$F^\star = \{s(a), p(a,b), q(b,b), r(a), q(a,b)\}$$

Intuitivement : le chemin infini partant de a construit par le restricted chase peut se replier sur $\{q(a,b), q(b,b)\}$.

Exercice 2 (10 pts)

On considère l'ensemble $\mathcal{R}$ composé de ces deux règles :

$R_1 : r(x) \to \exists z\ p(x,z) \wedge p(z,x)$
$R_2 : p(x,x) \to r(x)$

Question 1 (3 pts). Dessiner le graphe de dépendance des positions des prédicats associé à $\mathcal{R}$ (rappel en annexe). En vous appuyant sur ce graphe, dire si l'ensemble de règles est *weakly-acyclic* (wa) ou non.

Correction: Voir le cours pour le graphe obtenu. Ce graphe contient au moins un circuit dangereux, c'est-à-dire comportant un arc spécial, par exemple $[(r,1),(p,1),(r,1)]$. $\mathcal{R}$ n'est donc pas wa.

Question 2 (3 pts). Donner le graphe de dépendance des règles associés à $\mathcal{R}$ (rappel en annexe). Vous justifierez chaque arc ou absence d'arc. En vous appuyant sur ce graphe, dire si l'ensemble de règles est *acyclic GRD* (aGRD) ou non.

Correction: Il est facile de voir que R_1 dépend de R_2 car le corps de R_1 est isomorphe à la tête de R_2. Par contre R_2 ne dépend pas de R_1 car une application de R_1 ne peut pas contribuer à créer un nouvel homomorphisme du corps de R_2. Plus précisément, il n'existe par d'unificateur (au sens unificateur par pièce - *piece unifier*) entre le corps de R_2 et la tête de R_1 : en effet, si on cherche à unifier $p(x_2, x_2)$ [on renomme x de R_2 pour ne pas avoir de variable commune avec R_1] avec l'un des atomes de la tête de R_1, la variable existentielle z se retrouve unifiée avec x (et x_2), ce qui n'est pas possible car, par définition d'un unificateur par pièce, une variable existentielle de R_1 ne peut pas être unifiée avec un autre terme de R_1 (mais seulement avec des variables de R_2, qui de plus n'apparaissent pas dans les atomes non-unifiés de R_2 - ce qui n'est pas le problème ici). On observe aussi qu'aucune des règles ne dépend d'elle-même, ce qui est évident car pour chaque règle le corps et la tête n'ont pas de prédicat commun.
Le graphe de dépendances des règles de $\mathcal{R}$ a donc un seul arc : (R_2, R_1). On en conclut que $\mathcal{R}$ est aGRD.

Question 3 (1 pt). Que concluez-vous des questions précédentes concernant l'arrêt du chase avec $\mathcal{R}$? Justifier votre réponse.

Correction: La propriété aGRD est une condition suffisante pour l'arrêt de n'importe quelle variante de chase. Le chase s'arrête donc avec $\mathcal{R}$ (pour toute base de faits et toute variante de chase).

Question 4 (3 pts). Est-il possible qu'un ensemble de règles ne soit ni wa ni aGRD et que le chase s'arrête pourtant sur toute base de faits ? Justifier précisément votre réponse.

Correction: C'est possible, chacune des propriétés n'étant qu'une condition suffisante d'arrêt du chase. Pour justifier la réponse, donner un exemple d'ensemble de règles ni aGRD ni wa pour lequel toute variante de chase s'arrête sur toute base de fait.

Annexe (rappels)

Le **restricted chase** n'effectue l'application d'une règle $R : B \rightarrow H$ dans une base de faits F selon un homomorphisme h de B dans F que si h ne peut pas être étendu à un homomorphisme de $B \cup H$ dans F.

Le **graphe de dépendance des positions des prédicats** associé à un ensemble de règles $\mathcal{R}$ a :

- un sommet pour chaque position d'un prédicat apparaissant dans $\mathcal{R}$
- pour toute règle $R \in \mathcal{R}$ et toute variable frontière x en position (p, i) du corps de R, un arc de (p, i) vers chaque position de x dans la tête de R et un arc spécial de (p, i) vers chaque position d'une variable existentielle de R.

Le **graphe de dépendance des règles** associé à un ensemble de règles $\mathcal{R}$ a un sommet par règle de $\mathcal{R}$ et un arc de R_i vers R_j si R_j dépend de R_i (une application de R_i peut permettre un nouvel homomorphisme du corps de R_j).