Théorie des bases de connaissances (HMIN312)
Correction du contrôle n°1 (14/10/2020)

Durée : 1h15. Sans documents.
<u>**Toutes**</u> **vos réponses doivent être justifiées.**

Exercice 1

Soient deux ensembles d'atomes, où a est la seule constante :
$Q_1 = \{p(x_1, y_1), q(y_1, a)\}$
$Q_2 = \{p(x_2, y_2), q(y_2, z_2), p(x_2, u_2), q(u_2, z_2)\}$

Question 1. Existe-t-il un homomorphisme de Q_1 dans Q_2 ? De Q_2 dans Q_1 ?

Il n'y a pas d'homomorphisme de Q_1 dans Q_2 à cause de la constante a de Q_1 qu'on ne retrouve pas dans Q_2.

Il y a un homomorphisme $h = Q_2 \to Q_1$: $x_2 \mapsto x_1$, $y_2, u_2 \mapsto y_1$, $z_2 \mapsto a$.

Question 2. Si Q_1 et Q_2 représentent des requêtes conjonctives booléennes : a-t-on $Q_1 \sqsubseteq Q_2$? $Q_2 \sqsubseteq Q_1$?

On sait que : étant données deux BCQs Q et Q', $Q \sqsubseteq Q'$ ssi il existe un homomorphisme de Q' dans Q. On a donc $Q_1 \sqsubseteq Q_2$ mais pas $Q_2 \sqsubseteq Q_1$.

Rappel : Etant données deux BCQs Q et Q', on note $Q \sqsubseteq Q'$ si toute base de faits qui répond oui à Q répond aussi oui à Q'.

Exercice 2

Soit la base de connaissances $\mathcal{K} = (F, \mathcal{R})$ où F est une base de faits sans variables et $\mathcal{R} = \{R_1, R_2\}$ est un ensemble de règles positives :

$F = \{p(a, b), p(b, c)\}$
$R_1 : p(x, y) \to q(y)$
$R_2 : q(x) \wedge p(x, y) \to r(y)$

Question 1. Pour chacune des trois interprétations $\mathcal{I}_1$, $\mathcal{I}_2$ et $\mathcal{I}_3$ ci-dessous, dire quelle affirmation est vraie (et justifier votre réponse bien sûr) :

1. Ce n'est pas un modèle de $\mathcal{K}$
2. C'est un modèle de $\mathcal{K}$ mais pas un modèle universel de $\mathcal{K}$
3. C'est un modèle universel de $\mathcal{K}$.

$\mathcal{I}_1 = (D, .^{I_1})$ avec $D = \{a, b, c\}$, $p^{I_1} = \{(a, b), (b, c)\}$, $q^{I_1} = \{b, c\}$, $r^{I_1} = \{b, c\}$
$\mathcal{I}_2 = (D, .^{I_2})$ avec $D = \{a, b, c\}$, $p^{I_2} = \{(a, b), (b, c)\}$, $q^{I_2} = \{b, c\}$, $r^{I_2} = \{c\}$
$\mathcal{I}_3 = (D, .^{I_3})$ avec $D = \{a, b, c\}$, $p^{I_3} = \{(a, b), (b, c)\}$, $q^{I_3} = \{b\}$, $r^{I_3} = \{c\}$

$\mathcal{I}_2$ est un modèle universel de $\mathcal{K}$: en effet, c'est le modèle isomorphe à la saturation de F par $\{R_1, R_2\}$. $\mathcal{I}_1$ est un modèle de $\mathcal{K}$: il satisfait F et il satisfait chacune des règles

(R_1 car pour chaque couple $(d,d') \in p^{I_1}$, avec d et d' des éléments de D, on a $d \in q^{I_1}$; et R_2 car pour chaque couple $(d,d') \in p^{I_1}$ avec $d \in q^{I_1}$ on a $d' \in r^{I_1}$). Mais ce n'est pas un modèle universel de $\mathcal{K}$ car il ne s'envoie pas par homomorphisme dans tous les modèles de $\mathcal{K}$, par exemple il ne s'envoie pas dans $\mathcal{I}_2$. $\mathcal{I}_3$ n'est pas un modèle de $\mathcal{K}$, en effet il ne satisfait pas R_1 : on a $(b,c) \in p^{I_3}$ mais $c \notin q^{I_3}$.

Question 2. On rappelle qu'un modèle universel M de $\mathcal{K}$ assure la propriété suivante (P) : pour toute requête conjonctive booléenne Q, si M est un modèle de Q alors $\mathcal{K} \models Q$. Montrez que cette propriété n'est pas vraie si on considère un modèle de $\mathcal{K}$ qui n'est pas universel. Pour cela, vous prendrez la base de connaissances $\mathcal{K}$ de cet exercice, vous choisirez un modèle M de $\mathcal{K}$ non universel (éventuellement pris dans la question précédente si cela convient) et vous construirez un contre-exemple à (P) pour $\mathcal{K}$ et M.

Prenons $\mathcal{I}_1$ qui est un modèle non universel de $\mathcal{K}$ et la requête $Q() = r(b)$. $\mathcal{I}_1$ est un modèle de Q et pourtant $\mathcal{K} \not\models Q$: par exemple le modèle $\mathcal{I}_2$ de $\mathcal{K}$ n'est pas un modèle de Q.

Exercice 3 (Règles positives)

Question 1. Soit une base de faits F sans variables. Soit $\mathcal{R}$ un ensemble de règles positives Datalog. La saturation de F par $\mathcal{R}$ est-elle nécessairement un *core* ?

Oui car la saturation de F par $\mathcal{R}$ n'a pas de variables.

Question 2. Même question avec une base de faits F qui peut comporter des variables mais qui est un *core*.

Non.
Exemple : $F = \{p(a,v), q(a,b)\}$, où v est la seule variable, et $\mathcal{R} = \{q(x,y) \rightarrow p(x,y)$.

Rappel : un ensemble d'atomes est un core s'il n'admet aucun homomorphisme dans l'un de ses sous-ensembles stricts.

Exercice 4 (Négation en OWA)

On considère des bases de faits et requêtes conjonctives munies de la négation. On se place en monde ouvert, autrement dit dans le cadre de la logique classique.
Soit $F = \{q(a), q(b), p(a,a), \neg p(b,b)\}$ où a et b sont des constantes.

Question 1. Soit $Q_1() = \exists x \exists y \exists z\ (p(x,y) \wedge q(z) \wedge \neg p(y,z))$. Déterminer $Q_1(F)$, l'ensemble des réponses à Q_1 sur F.

Tout modèle de F rend vrai $p(a,b)$ ou $\neg p(a,b)$. Dans le premier cas, on a la "bonne affectation" $\{x \mapsto a, y \mapsto b, z \mapsto b\}$, dans le second cas, on a la "bonne affectation" $\{x \mapsto a, y \mapsto a, z \mapsto b\}$. Donc tout modèle de F est un modèle de Q_1. Donc F répond oui à Q_1, autrement dit $Q_1(F) = \{()\}$.

Question 2. Même question avec $Q_2(x) = \exists y \exists z\ (p(x,y) \wedge q(z) \wedge \neg p(y,z))$.

$Q_2(F) = \{(a)\}$ car dans tous les modèles de F l'une des deux bonnes affectations précédentes envoie x sur a. Autrement dit, $F \models Q_2[x \mapsto a]$ où $Q_2[x \mapsto a]$ est la requête booléenne $\exists y \exists z\ (p(a,y) \wedge q(z) \wedge \neg p(y,z))$.

Question 3. Même question avec $Q_3(y) = \exists x \exists z\ (p(x,y) \wedge q(z) \wedge \neg p(y,z))$.

$Q_3(F) = \emptyset$ *car aucune des requêtes booléennes obtenues en remplaçant* y *par une constante n'est conséquence logique de* F.

Exercice 5 (Datalog avec négation)

On considère l'ensemble Π de règles Datalog avec négation ci-dessous, où `not` représente la négation du monde clos.

R_1 :	`C(X) :- A(X), not B(X).`
R_2 :	`D(X) :- A(X), not C(X).`
R_3 :	`B(X) :- A(X), D(X), not E(X).`

Question 1. Soit $D = \{A(a)\}$ où a est une constante. Montrer qu'il existe plusieurs dérivations satisfaisantes complètes à partir de D avec Π et donner les ensembles d'atomes obtenus.

La dérivation qui applique seulement R_1 *en produisant* $D_1 = \{A(a), C(a)\}$ *est satisfaisante et complète (on ne peut plus faire d'autres applications de règle). La dérivation qui applique* R_2 *puis* R_3 *en produisant* $D_2 = \{A(a), D(a), B(a)\}$ *est satisfaisante et complète.*

Question 2. Pouvez-vous en déduire que Π n'est pas stratifiable (et pourquoi) ?

Oui, car si Π *était stratifiable, toutes les dérivations satisfaisantes à partir d'une même base de données donneraient le même résultat.*

Question 3. Déterminer la stratifiabilité de Π en utilisant le graphe de dépendance des prédicats intensionnels.

Le graphe des prédicats intensionnels a pour sommets B, C et D, et pour arcs $(B, -, C)$, $(C, -, D)$ *et* $(D, +, B)$. *Il a donc un circuit avec un arc négatif, ce qui montre que* Π *n'est pas stratifiable.*

Rappels : Une dérivation à partir de D avec Π est une suite $D_0 \ldots D_n$ telle que $D_0 = D$ et pour $1 \leq i \leq n$, D_i est obtenue à partir de D_{i-1} en appliquant une règle de Π. Une telle dérivation est *complète* s'il n'y a aucune nouvelle application de règle possible sur D_n (si une règle de Π s'applique à D_n par un homomorphisme h, cette application a déjà été effectuée). Une telle dérivation est *satisfaisante* si toute règle appllquée par un homomorphisme h à une certaine étape reste applicable par le même homomorphisme h dans la suite de la dérivation (autrement dit, sur D_n).