Université de Montpellier - Master informatique	Janvier 2021
Théorie des bases de connaissances (HMIN 312)	

Correction Partie 1 de l'examen

Durée totale : 2 heures
Répartition conseillée : 40 à 45 mn pour la partie 1

PARTIE 1 : REGLES EXISTENTIELLES (notation / 20)

Exercice 1 (Chase) - 13 pts

Soit la base de connaissances $\mathcal{K}_1 = (F_1, \mathcal{R}_1 = \{R_1, R_2\})$ avec :
$F_1 = \{p(a,b), r(a)\}$
$R_1 = p(x,y) \rightarrow \exists z\ p(y,z)$
$R_2 = r(x) \wedge p(x,y) \rightarrow p(y,y)$

Question 1 Définir la saturation de F_1 par $\mathcal{R}_1$ en utilisant l'*oblivious chase*.

Cette saturation est infinie :

$$\{p(a,b), r(a), p(b,z_0), p(b,b), p(b,z_1)\} \cup \{p(z_i, z_{i+2}) | i \geq 0\}$$

Question 2 Définir la saturation de F_1 par $\mathcal{R}_1$ en utilisant le *restricted chase*. Si l'ordre d'application des règles change le résultat, considérez les différents cas possibles.

Selon qu'on applique d'abord R_2 ou R_1, la saturation est finie ou infinie :
- R_2 d'abord : on obtient $\{p(a,b), r(a), p(b,b)\}$ par R_2, ce qui rend l'application de R_1 redondante.
- R_1 d'abord : on obtient "une chaine infinie" (au lieu de 2 avec l'oblivious chase) :
$\{p(a,b), r(a), p(b,z_0), p(b,b)\} \cup \{p(z_i, z_{i+1}) | i \geq 0\}$

Question 3 Définir la saturation de F_1 par $\mathcal{R}_1$ en utilisant le *core chase* (on supposera que la base de faits courante est mise sous la forme d'un core à chaque application de règle).

On obtient $\{p(a,b), r(a), p(b,b)\}$. L'ordre d'application des règles est indifférent : même si on commence par R_1, l'atome ajouté $p(b,z_0)$ est supprimé par le calcul du core après l'application de R_2.

Question 4 La base de connaissances $\mathcal{K}_1$ possède-t-elle un modèle universel fini ? Justifier votre réponse.

Oui, celui calculé par le core chase, cf. Q3.

Question 5 L'ensemble de règles $\mathcal{R}_1$ est-il un "finite expansion set", c'est-à-dire assure-t-il que pour toute base de faits F, $(F, \mathcal{R}_1)$ a un modèle universel fini?

Non. Par exemple, si l'on prend $F = \{p(a,b)\}$, seule R_1 est applicable et le core chase est infini (donc la KB n'a pas de modèle fini).

Question 6 Utilisez la base de connaissances $\mathcal{K}_2$ ci-dessous pour montrer que le core chase est *strictement* plus puissant que le restricted chase en termes de terminaison.

$\mathcal{K}_2 = (F_2, \mathcal{R}_2 = \{R_1, R_3\})$ avec :
$F_2 = \{p(a,b), r(b)\}$
$R_1 = p(x,y) \rightarrow \exists z\ p(y,z)$
$R_3 = r(x) \wedge p(x,y) \rightarrow p(x,x)$

Sur cet exemple, le core chase s'arrête en produisant $\{p(a,b), r(b), p(b,b)\}$ mais pas le restricted chase, quel que soit l'ordre des règles : à l'étape 1, R_3 n'est pas applicable, par contre R_1 l'est, ce qui produit $p(b,z_0)$; cet atome est redondant par rapport à $p(b,b)$ qui sera ensuite produit par R_3, mais "c'est trop tard".).

Exercice 2 (Règles existentielles linéaires) - 7 pts

Soit l'ensemble de règles $\mathcal{R} = \{R_1, R_2\})$ avec :
$R_1 = p(x,y) \rightarrow \exists z\ p(y,z) \wedge q(z)$
$R_2 = s(x,y) \rightarrow \exists z\ p(y,z) \wedge t(z)$.

Question 1 Soit la requête conjonctive booléenne $q() = \exists u \exists v\ p(u,v) \wedge t(u) \wedge t(v)$. Donner l'ensemble des réécritures de q avec $\mathcal{R}$. Expliquer pourquoi il n'y a pas d'autres réécritures.

$\mathcal{Q} = \{q_0, q_1\}$ avec $q_0 = q$ et $q_1 = s(x_0, u) \wedge t(u)\}$.
q ne s'unifie pas avec la tête de R_1 (voir pourquoi), q_1 est produite par réécriture de q avec R_2, et ne peut être à son tour réécrite.

Question 2 Une règle existentielle est dite *linéaire* si son corps est composé d'un seul atome (comme R_1 et R_2). Montrer que l'ensemble des réécritures (non-isomorphes) d'une requête conjonctive booléenne avec un ensemble de règles linéaires est toujours fini. On rappelle que deux ensembles d'atomes sont isomorphes si on passe de l'un à l'autre par un renommage bijectif de leurs variables (q_1 et q_2 sont isomorphes s'il existe une bijection b de l'ensemble des variables de q_1 dans l'ensemble des variables de q_2 telle que $b(q_1) = q_2$).

Si les règles sont linéaires, chaque étape de réécriture d'une requête q_i produit une requête dont la taille (en nombre d'atomes) est inférieure ou égale à celle de q_i. Comme toute réécriture s'obtient par un nombre fini d'étapes de réécritures à partir de la requête initiale (soit q), on en conclut que (a) : toute réécriture de q a une taille inférieure ou égale à celle de q. :

Or, (b) : le nombre d'ensembles d'atomes non isomorphes de taille inférieure ou égale à $|q|$ (la taille de q) que l'on peut construire à partir d'un nombre fini de prédicats et de constantes (et un nombre quelconque de variables) est lui-même fini.

De (a) et (b) on conclut que l'ensemble des réécritures d'une CQ booléenne avec un ensemble (fini) de règles linéaires est fini.