

## Bases de Données I (J. Wijsen)

13 janvier 2015

NOM + PRÉNOM :

Orientation + Année :

Cet examen contient 10 questions. Durée : 3 heures.

**Question 1** Considérez la requête

$$\{y \mid \exists x \exists z ((R(x, y) \vee R(y, z)) \wedge \neg R(x, y))\}.$$

Cochez la case qui précède une expression correcte :

- Cette requête est *domain independent*.
- Cette requête n'est pas *domain independent*.

Expliquez en détail.

---

.../5
-------

**Question 2** Soit  $SPJRU$  l'algèbre que l'on obtient en enlevant la Différence de l'algèbre  $SPJRUD$ . Démontrez que l'algèbre  $SPJRUD$  (avec la différence) est plus puissante que l'algèbre  $SPJRU$ .

---

.../5
-------

**Question 3** On enregistre dans une table les prix journaliers de livres en ligne auprès de plusieurs vendeurs qui sont actifs sur le territoire européen. Chaque livre est identifié de façon unique par son *International Standard Book Number* (ISBN). La première ligne enregistre que le livre avec ISBN 0-486-43947-X coûtait 20.00 EUR auprès d'Amazon à la date du 11 août 2014. Le prix d'un livre peut varier d'un jour au lendemain, mais ne change pas au cours d'une même journée. Évidemment, le prix d'un livre peut varier d'un vendeur à l'autre.

Le prix en Dollar (USD) est obtenu en multipliant le prix en Euro (EUR) par la parité EUR/USD en vigueur. Cette parité est fixée chaque jour par la Banque centrale européenne, mais peut varier d'un jour au lendemain. A partir de la première rangée de la table, on peut conclure qu'à la date du 11 août 2014, la parité EUR/USD était de  $26.80/20.00 = 1.34$  (i.e., 1.00 EUR valait 1.34 USD). Le 12 août 2014, la parité était de  $26.00/20.00 = 1.30$ .

Un livre peut avoir un, deux, ou plusieurs auteurs. Par exemple, le livre de 2010 intitulé "Databases" est co-écrit par R. Elmasri et L. C. Grove. Deux livres différents peuvent avoir le même titre, mais l'agence officielle d'attribution des ISBN veille à ce qu'aucune personne ne soit auteur ou co-auteur de deux livres portant des ISBN différents mais ayant le même titre et année de publication (*PubYear*). Donc, les livres publiés dans une même année par une même personne auront tous des titres distincts.

Pour un livre à plusieurs auteurs, les auteurs doivent toujours être listés dans le même ordre, qui n'est pas forcément l'ordre alphabétique. L'attribut *Order* encode cet ordre : le premier auteur porte le chiffre 1, le deuxième auteur porte le chiffre 2... Si un livre n'a qu'un seul auteur, cet auteur est premier auteur.

Supposons que chaque auteur est identifié de façon unique par son nom.

<i>ISBN</i>	<i>Title</i>	<i>Author</i>	<i>Order</i>	<i>PubYear</i>	<i>Vendor</i>	<i>Date</i>	<i>EUR</i>	<i>USD</i>
0-486-43947-X	Algebra	L. C. Grove	1	1983	Amazon	11 Aug 2014	20.00	26.80
0-486-43947-X	Algebra	L. C. Grove	1	1983	Proxis	11 Aug 2014	21.00	28.14
0-486-43947-X	Algebra	L. C. Grove	1	1983	Amazon	12 Aug 2014	20.00	26.00
0-486-43947-X	Algebra	L. C. Grove	1	1983	Proxis	12 Aug 2014	20.00	26.00
978-0136086208	Databases	R. Elmasri	1	2010	Amazon	12 Aug 2014	80.00	104.00
978-0136086208	Databases	L. C. Grove	2	2010	Amazon	12 Aug 2014	80.00	104.00
978-0596809157	Cookiebook	R. Elmasri	1	2012	Proxis	12 Aug 2014	80.00	104.00

Quelles sont les dépendances fonctionnelles que l'on peut raisonnablement imposer sur ces données ?

---

.. /10
--------

**Question 4** Les mêmes données de la question 3 sont maintenant stockées dans quatre tables  $V$ ,  $T$ ,  $A$  et  $E$ . La table  $V$  stocke les prix journaliers en EUR par vendeur. La table  $T$  stocke les titres des livres, ainsi que leurs années de publication. La table  $A$  stocke les auteurs des livres et indique l'ordre des auteurs pour des livres à plusieurs auteurs. Au lieu de stocker des prix en USD, la table  $E$  stocke le cours de change entre EUR et USD.

$V$	<i>ISBN</i>	<i>Vendor</i>	<i>Date</i>	<i>EUR</i>
	0-486-43947-X	Amazon	11 Aug 2014	20.00
	0-486-43947-X	Proxis	11 Aug 2014	21.00
	0-486-43947-X	Amazon	12 Aug 2014	20.00
	0-486-43947-X	Proxis	12 Aug 2014	20.00
	978-0136086208	Amazon	12 Aug 2014	80.00
	978-0596809157	Proxis	12 Aug 2014	80.00

$E$	<i>Date</i>	<i>EURvsUSD</i>
	11 Aug 2014	1.34
	12 Aug 2014	1.30

$T$	<i>ISBN</i>	<i>Title</i>	<i>PubYear</i>
	0-486-43947-X	Algebra	1983
	978-0136086208	Databases	2010
	978-0596809157	Cookiebook	2012

$A$	<i>ISBN</i>	<i>Author</i>	<i>Order</i>
	0-486-43947-X	L. C. Grove	1
	978-0136086208	R. Elmasri	1
	978-0136086208	L. C. Grove	2
	978-0596809157	R. Elmasri	1

Donnez toutes les clés primaires et étrangères en utilisant la syntaxe du cours.

---

.../5
-------

**Question 5** Pour la base de données de la question 4, écrivez une requête en algèbre SPJRUD pour répondre à la question suivante :

Quel est le prix le plus élevé qui a jamais été demandé pour un livre écrit ou co-écrit par L. C. Grove ? Donnez aussi le titre du livre en question, le vendeur qui a demandé ce prix et la date à laquelle ce prix a été demandé.

À la date du 12 août 2014, Amazon a vendu le livre “Databases” de L. C. Grove au prix de 80 EUR. Aucun autre livre de L. C. Grove n’a jamais coûté plus cher. La réponse est donc :

<i>Title</i>	<i>Vendor</i>	<i>Date</i>	<i>EUR</i>
Databases	Amazon	12 Aug 2014	80.00

Pour cette question, on peut utiliser la sélection de type  $\sigma_{A < B}(R)$ , où  $A$  et  $B$  sont des colonnes de la relation  $R$  qui contiennent des prix. La sémantique est comme attendue :

$$\sigma_{A < B}(R) = \{t \in R \mid t(A) < t(B)\}.$$

Expliquez le fonctionnement de votre requête aussi en français.

---

.../10
--------

**Question 6** Pour la base de données de la question 4, écrivez une requête en calcul relationnel qui donne le nom de chaque auteur qui a été premier auteur de **tous** les livres dont il est auteur ou co-auteur. Pour la base de données de la question 4, la réponse contient R. Elmasri. Notez que L. C. Grove ne fait pas partie de la réponse, parce qu'il a aussi été deuxième auteur.

---

.../5
-------

**Question 7** Soient

$$U = ABCDEFG$$

$$\Sigma = \{ABC \rightarrow DEF, DE \rightarrow AB, EF \rightarrow BC, DF \rightarrow ACG, G \rightarrow C\}$$

Cochez chaque case (éventuellement plusieurs) qui précède une expression correcte :

- Le schéma  $(U, \Sigma)$  est en BCNF.
- Le schéma  $(U, \Sigma)$  n'est pas en BCNF.
- Le schéma  $(U, \Sigma)$  est en 3NF.
- Le schéma  $(U, \Sigma)$  n'est pas en 3NF.

Détaillez les calculs qui mènent à cette conclusion.

---

.../10
--------

**Question 8** La table *Aime* enregistre qui aime quelle bière. On dit que la personne  $x$  est plus sélective que la personne  $y$  si  $y$  aime toutes les bières que  $x$  aime et, en plus,  $y$  aime au moins une bière que  $x$  n'aime pas. Par exemple, Maria est plus sélective que Pierre, car Pierre aime toutes les bières que Maria aime (notamment Leffe et Westmalle) et, en plus, Pierre aime une bière que Maria n'aime pas (notamment Orval).

<i>Aime</i>	<i>Personne</i>	<i>Bière</i>
	Pierre	Orval
	Pierre	Leffe
	Pierre	Westmalle
	Jean	Orval
	Jean	Leffe
	Jean	Westmalle
	Maria	Leffe
	Maria	Westmalle

Écrivez une requête SQL qui s'appuie sur [NOT] EXISTS pour rendre chaque personne  $x$  qui est plus sélective que Pierre. Explicitez la logique derrière votre requête SQL en montrant la requête en *tuple relational calculus*. Pour l'exemple, le résultat est comme suit.

<i>Personne</i>
Maria

---

.../10
--------

**Question 9** Expliquez pourquoi, en pratique, les SGBD utilisent Stict 2PL au lieu de 2PL.

---

.../5

**Question 10** Considérez l'exécution suivante :

$$R_1(A)R_3(B)W_2(A)R_1(B)R_3(A)R_1(C)W_1(C)W_3(C)$$

Est-ce que cette exécution est possible en 2PL ? Complétez l'exécution avec des demandes de verrous ou argumentez pourquoi cette exécution n'est pas possible en 2PL.

.../5
-------

