

Bases de Données I (J. Wijsen)

9 janvier 2014

NOM + PRÉNOM :

Orientation + Année :

Cet examen contient 10 questions. Durée : 2 heures et 30 minutes.

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. Soit $SPJRU^{\neq}$ l'algèbre qui étend l'algèbre SPJRU avec les sélections de forme $\sigma_{A \neq B}(R)$ et $\sigma_{A \neq c}(R)$, où A, B sont des attributs et c est une constante. La sémantique est évidente :

$$\begin{aligned}\sigma_{A \neq B}(R) &= \{t \in R \mid t(A) \neq t(B)\}; \\ \sigma_{A \neq c}(R) &= \{t \in R \mid t(A) \neq c\}.\end{aligned}$$

Noter que les algèbres $SPJRU^{\neq}$ et SPJRU ne contiennent pas la différence $R - S$ entre deux relations R et S .

Démontrez que l'algèbre SPJRUD (avec la différence) est plus puissante que l'algèbre $SPJRU^{\neq}$.

.../7

Question 3 La Ligue de diamant de l'IAAF est une compétition d'athlétisme organisée annuellement par l'Association internationale des fédérations d'athlétisme (IAAF). Elle est composée de 14 meetings dans 14 villes différentes. Toutes les épreuves sont individuelles. Les résultats sont stockées dans la table suivante. On apprend, par exemple, qu'en 2012, Kevin Borlée a gagné les 400m hommes à Bruxelles, alors qu'il était troisième à Lausanne. En 2011, Kevin était deuxième dans cette discipline à Bruxelles derrière son frère.

A chaque meeting, un athlète ne peut participer qu'à une seule discipline. Noter qu'en 2012, Usain Bolt a couru les 100m à Bruxelles et les 200m à Lausanne.

Un athlète est déterminé par son nom et prénom. Un athlète peut changer de pays d'une année à l'autre, mais il ne peut pas changer de pays pendant une même année civile. Un athlète ne change jamais de genre.

<i>Ville</i>	<i>Date</i>	<i>Année</i>	<i>Nom</i>	<i>Prénom</i>	<i>Genre</i>	<i>Pays</i>	<i>Discipline</i>	<i>Résultat</i>	<i>Classement</i>
Bruxelles	7 sep.	2012	Soumaré	Myriam	F	FRA	200m	22.63	1
Bruxelles	7 sep.	2012	Blake	Yohann	M	JAM	200m	19.54	1
Bruxelles	7 sep.	2012	Bolt	Usain	M	JAM	100m	9.86	1
Bruxelles	7 sep.	2012	Carter	Nesta	M	JAM	100m	9.96	2
Bruxelles	7 sep.	2012	Borlée	Kevin	M	BEL	400m	44.75	1
Bruxelles	7 sep.	2012	Borlée	Jonathan	M	BEL	400m	45.02	2
Lausanne	23 août	2012	Bolt	Usain	M	JAM	200m	19.58	1
Lausanne	23 août	2012	Borlée	Kevin	M	BEL	400m	45.27	3
Bruxelles	16 sep.	2011	Borlée	Jonathan	M	BEL	400m	44.78	1
Bruxelles	16 sep.	2011	Borlée	Kevin	M	BEL	400m	44.97	2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Donnez les dépendances fonctionnelles que l'on peut raisonnablement imposer.

Question 4 La table *BESTS* stocke la meilleure performance (*PB* = personal best) d'un(e) athlète dans une discipline. Chaque athlète est déterminé par son nom et prénom. La table *HOMMES* stocke les athlètes masculins ; la table *FEMMES* stocke les athlètes féminines. La colonne *Nat* donne la nationalité (dans cette question, on suppose que la nationalité d'un(e) athlète ne change pas). La table *RECORDS* stocke le record du monde dans chaque discipline. Pour chaque ligne dans la table *RECORDS*, on peut retrouver le nom du teneur (cq. de la teneuse) du record du monde en utilisant les tables *BESTS*, *HOMMES* et *FEMMES*. **Noter que le chiffre 18.29 intervient dans trois records** : dans le triple saut, le record du monde est de 18.29m pour les deux sexes, et 18.29s est le record du monde dans les 200m hommes.

<i>BESTS</i>	<i>Nom</i>	<i>Prénom</i>	<i>Discipline</i>	<i>PB</i>
	Bolt	Usain	100m	9.58
	Bolt	Usain	200m	18.29
	Borlée	Kevin	400m	44.56
	Borlée	Jonathan	400m	44.43
	Borlée	Jonathan	200m	20.31
	Griffith Joyner	Florence	100m	10.49
	Gevaert	Kim	100m	11.04
	Gevaert	Kim	200m	22.20
	Johnson	Michael	400m	43.18
	Edwards	Jonathan	Triple saut	18.29
	Superwoman	Jessica	Triple saut	18.29

<i>HOMMES</i>	<i>Nom</i>	<i>Prénom</i>	<i>Nat</i>
	Bolt	Usain	JAM
	Borlée	Kevin	BEL
	Borlée	Jonathan	BEL
	Johnson	Michael	USA
	Edwards	Jonathan	GB

<i>RECORDS</i>	<i>Discipline</i>	<i>Genre</i>	<i>Résultat</i>
	100m	M	9.58
	100m	F	10.49
	200m	M	18.29
	400m	M	43.18
	Triple saut	M	18.29
	Triple saut	F	18.29

<i>FEMMES</i>	<i>Nom</i>	<i>Prénom</i>	<i>Nat</i>
	Griffith Joyner	Florence	USA
	Gevaert	Kim	BEL
	Superwoman	Jessica	JAP

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

.../7

Question 5 Pour la base de données de la question 4, écrivez une requête en algèbre SPJRUD qui donne pour chaque discipline la nationalité du teneur (de la teneuse) du record du monde, comme suit.

<i>Discipline</i>	<i>Genre</i>	<i>Nat</i>
100m	M	JAM
100m	F	USA
200m	M	JAM
400m	M	USA
Triple saut	M	GB
Triple saut	F	JAP

.../7

Question 6 Pour la base de données de la question 4, écrivez une requête en calcul relationnel qui donne le nom de chaque homme qui ne possède aucun record du monde, comme suit.

<i>Nom</i>	<i>Prénom</i>
Borlée	Kevin
Borlée	Jonathan

.../7

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

Détaillez les calculs qui mènent à cette conclusion.

.../9

Question 8 Les élèves d'une même classe sont évalués sur plusieurs branches. On dit qu'un élève x est dominé par un élève y si pour chaque branche, la note obtenue par x est strictement inférieure à la note obtenue par y dans cette même branche. Dans la table *Notes* suivante, les branches sont histoire, géographie et biologie. Bernard est dominé par Antoine. Christian n'est pas dominé par Antoine, car la note de Christian en géographie est supérieure à la note d'Antoine en géographie.

<i>Notes</i>	<i>Élève</i>	<i>Branche</i>	<i>Note</i>
	Antoine	histoire	6
	Antoine	géographie	6
	Antoine	biologie	8
	Bernard	histoire	5
	Bernard	géographie	5
	Bernard	biologie	7
	Christian	histoire	4
	Christian	géographie	9
	Christian	biologie	4
	Didier	histoire	3
	Didier	géographie	3
	Didier	biologie	3

Écrivez une requête SQL qui donne chaque paire (x, y) telle qu'élève x est dominé par y . Explicitez la logique derrière votre requête SQL en montrant la requête en *tuple relational calculus*. Pour l'exemple, le résultat est comme suit.

<i>Élève</i>	<i>EstDominéPar</i>
Didier	Antoine
Didier	Bernard
Didier	Christian
Bernard	Antoine

.../7

Question 9 Détaillez le protocole Wound-Wait et argumentez pourquoi ce protocole évite les verrous mortels.

.../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.

.../7

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