

Bases de Données I (J. Wijsen), 11 janvier 2016

NOM + PRÉNOM :

Orientation + Année :

Cet examen contient 11 questions. Durée : 3 heures.

Question 1 Pour chacune des cinq requêtes suivantes, cochez la case qui précède une expression correcte. Calcul du score : la bonne case cochée= +1, la mauvaise case cochée= -1.5, pas de case cochée= 0. Si le score total est négatif, il sera converti en 0.

.../5

1. $\{x \mid \forall z \exists y (R(x, y) \wedge R(y, z))\}$
Cette requête est *domain independent* n'est pas *domain independent*.
2. $\{x \mid \neg S(x) \wedge \neg \forall y (S(y) \rightarrow S(y))\}$
Cette requête est *domain independent* n'est pas *domain independent*.
3. $\{x, y \mid \exists z (T(x, y, z) \vee R(x, y))\}$
Cette requête est *domain independent* n'est pas *domain independent*.
4. $\{x \mid \exists y \exists z (R(x, y) \vee R(z, x))\}$
Cette requête est *domain independent* n'est pas *domain independent*.
5. $\{x \mid \forall z (\neg R(x, z) \vee R(z, x))\}$
Cette requête est *domain independent* n'est pas *domain independent*.

Question 2 Est-ce que l'affirmation suivante est vraie ou fausse ?

Pour un schéma-de-relation $\{A, B, C, D\}$, on a $\{A \rightarrow B\} \models A \rightarrow C$.

Démontrez ou donnez un contre-exemple.

.../5

Cette affirmation est vraie fausse (cochez une case).

Question 3 On enregistre dans une table les prix actuels des fruits dans différents magasins. Chaque magasin appartient à une chaîne de magasins (Aldi, Colruyt, Lidl. . .). Pour chaque variété de fruits, les magasins peuvent adapter le prix selon l'origine : Belgique (B) ou Importation (I). Les fruits d'origine belge portent l'étiquette "Produit de Belgique" et sont en général plus chers (mais de meilleure qualité d'après le premier ministre). Par exemple, les Colruyt d'Anderlecht vendent la Jonagold d'origine belge au prix de 2.50 EUR/kg, et la Jonagold importée à 2.30 EUR/kg. Certaines variétés, comme la Pink Lady (une variété de pomme), ne sont pas cultivées en Belgique, et seront donc toujours importées.

La gamme de fruits offerte peut varier d'un magasin à l'autre. Par exemple, le Colruyt d'Uccle ne vend que la Jonagold d'origine belge.

Évidemment, les prix des fruits peuvent varier d'une chaîne à l'autre. Par exemple, Aldi est en général moins cher que Colruyt. En plus, une même chaîne peut différencier les prix selon la commune. Cependant, par une législation belge du marché des fruits, aucune chaîne ne peut faire varier ses prix à l'intérieur d'une même région. Pour rappel, la Belgique est divisée en trois Régions : la Région flamande (FLA), la Région de Bruxelles-Capitale (BRU), la Région wallonne (WAL). Par exemple, les prix chez Colruyt seront toujours les mêmes à Anderlecht et à Uccle (deux communes en Région de Bruxelles-Capitale), mais les prix chez Colruyt peuvent être différents entre Anderlecht et Mons (commune en Région wallonne).

Les prix sont aussi convertis en ancien franc belge (pour mémoire, 1 EUR = 40.3399 BEF).

<i>Origine</i>	<i>Fruit</i>	<i>Magasin</i>	<i>Variété</i>	<i>Ville</i>	<i>Région</i>	<i>EUR</i>	<i>BEF</i>
B	Pomme	Colruyt	Jonagold	Anderlecht	BRU	2.50	100.85
I	Pomme	Colruyt	Jonagold	Anderlecht	BRU	2.30	92.78
B	Pomme	Colruyt	Jonagold	Uccle	BRU	2.50	100.85
B	Pomme	Colruyt	Jonagold	Mons	WAL	3.50	141.19
B	Pomme	Aldi	Jonagold	Mons	WAL	3.00	121.02
I	Pomme	Aldi	Pink Lady	Mons	WAL	3.50	141.19
B	Poire	Aldi	Conférence	Mons	WAL	1.50	60.51
B	Poire	Lidl	Conférence	Mons	WAL	1.49	60.11

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

.../5

Question 4 Soit Σ un ensemble de dépendances fonctionnelles sur un ensemble \mathcal{A} d'attributs. On dit qu'une relation R sur \mathcal{A} est une *relation d'Armstrong pour le schéma-DF* (\mathcal{A}, Σ) si les deux conditions suivantes sont vérifiées :

1. R satisfait toute dépendance fonctionnelle qui est une conséquence logique de Σ ; et
2. R ne satisfait aucune dépendance fonctionnelle qui n'est pas une conséquence logique de Σ .

En symboles, R est une relation d'Armstrong pour (\mathcal{A}, Σ) si pour chaque dépendance fonctionnelle σ sur \mathcal{A} ,

$$R \models \sigma \iff \Sigma \models \sigma.$$

Soit $\mathcal{A}_0 = \{A, B, C\}$ et $\Sigma_0 = \{A \rightarrow B, AB \rightarrow C, C \rightarrow B\}$. Construisez une relation d'Armstrong pour $(\mathcal{A}_0, \Sigma_0)$, et argumentez pourquoi votre relation est en effet une relation d'Armstrong.

.../5

Question 5 Pierre souhaite comparer des pommes et des poires (toutes cultivées en Belgique), en termes de magasin, variété, prix, goût et durée de conservation. À cette fin, le samedi, il achète toutes les variétés de pommes et de poires disponibles dans un nombre de magasins à Mons (table *ACHATS*) ; le prix est exprimé en EUR/kg. Noter que toutes les variétés ne sont pas forcément disponibles dans tous les magasins (dans l'exemple, pas de poires chez Colruyt, pas de Jonagold chez Lidl, pas de Doyenné chez Aldi). À plusieurs jours dans la semaine qui suit (lundi et vendredi dans l'exemple), il demande à trois amis (dénoté par *A, B, C*) de goûter les fruits et donner leurs avis. L'avis est un chiffre entre 0 (dégueulasse) et 5 (délicieux). Les avis peuvent varier d'un jour à l'autre parce que les fruits peuvent encore mûrir (voire pourrir). Les résultats de ce test de goût sont stockés dans la table *AVIS*. La première ligne, par exemple, indique que le lundi, les trois personnes trouvent que la Jonagold de Colruyt est délicieuse (score maximal de 5) ; le vendredi, les personnes *A* et *B* rapportent une légère baisse de qualité (de 5 à 4) pour ce fruit. La table *DÉGUSTATEURS* révèle l'identité des trois amis. La table *POMMES* stocke les variétés de pommes ; la table *POIRES* stocke les variétés de poires.

<i>ACHATS</i>				<i>AVIS</i>						
<i>Magasin</i>	<i>Variété</i>	<i>Prix</i>		<i>Jour</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>Magasin</i>	<i>Variété</i>	
Colruyt	Jonagold	3.50		lundi	5	5	5	Colruyt	Jonagold	
Aldi	Jonagold	3.00		lundi	4	3	3	Aldi	Jonagold	
Aldi	Conférence	1.50		lundi	3	4	4	Aldi	Conférence	
Lidl	Conférence	1.49		lundi	5	5	4	Lidl	Conférence	
Lidl	Doyenné	1.00		lundi	1	1	0	Lidl	Doyenné	
				vendredi	4	4	5	Colruyt	Jonagold	
				vendredi	3	2	2	Aldi	Jonagold	
				vendredi	3	4	4	Aldi	Conférence	
				vendredi	4	5	3	Lidl	Conférence	
				vendredi	0	1	0	Lidl	Doyenné	

<i>POMMES</i>		<i>POIRES</i>		<i>DÉGUSTATEURS</i>			
<i>Variété</i>		<i>Variété</i>		<i>Identifiant</i>	<i>Nom</i>	<i>Âge</i>	<i>Genre</i>
Jonagold		Conférence		A	Brigitte	24	F
		Doyenné		B	Fabian	23	M
				C	Jean	21	M

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

.../5

Question 6 Pour la base de données de la question 5, supposons que l'on veut savoir, pour chaque variété de fruits, quel est le magasin qui offre la meilleure qualité (peu importe le prix). Ce n'est pas évident, car les trois dégustateurs peuvent être en désaccord, et la qualité d'un fruit peut se dégrader (ou s'améliorer) d'un jour à l'autre. On décide donc d'afficher seulement les résultats qui font l'objet d'une unanimité.

Par exemple, pour la Conférence, il n'y a pas d'unanimité : le lundi, les trois dégustateurs donnent bien leur meilleur score à Lidl, mais le vendredi, le dégustateur C préfère la Conférence d'Aldi à celle de Lidl.

Concrètement, on décide de créer une table à deux colonnes, *Variété* et *Magasin*, qui contient une rangée $\langle \text{Variété} : v, \text{Magasin} : m \rangle$ si pour la variété v , à chaque jour de dégustation, tous les dégustateurs donnent leur meilleur score au magasin m .

Noter que les jours de dégustation et leur nombre ne sont pas connus à priori.

Pour la base de données de la question 5, le résultat est comme suit :

<i>Variété</i>	<i>Magasin</i>
Jonagold	Colruyt
Doyenné	Lidl

Noter que la deuxième ligne de cette table est correcte, car aucun magasin ne dépasse Lidl en ce qui concerne la qualité de la Doyenné.

Écrivez une requête en calcul relationnel qui construit la table demandée. On peut faire appel aux prédicats $<$ et \leq pour comparer les scores.

.../5

Question 9 Soient

$$\mathcal{A} = ABCDEFG$$

$$\Sigma = \{DE \rightarrow AFG, FG \rightarrow BC, CE \rightarrow AD, ADF \rightarrow EB, BCE \rightarrow DG\}$$

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

- Le schéma-DF (\mathcal{A}, Σ) est en BCNF.
- Le schéma-DF (\mathcal{A}, Σ) n'est pas en BCNF.
- Le schéma-DF (\mathcal{A}, Σ) est en 3NF.
- Le schéma-DF (\mathcal{A}, Σ) n'est pas en 3NF.

Détaillez les calculs qui mènent à cette conclusion.

.../7

Question 10 Situez et expliquez le protocole WOUND-WAIT.

.../5

Question 11 Considérez l'exécution suivante :

$$R_1(A)R_2(A)R_3(A)W_1(A)W_2(B)W_3(C)R_1(B)R_1(C)$$

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

$R_1(A)$		
	$R_2(A)$	
		$R_3(A)$
$W_1(A)$		
	$W_2(B)$	
		$W_3(C)$
$R_1(B)$		
$R_1(C)$		