

# Bases de Données I (J. Wijsen)

10 janvier 2019

NOM + PRÉNOM :

Orientation + Année :

Cet examen contient 12 questions. Durée : 3 heures.

**Question 1** La *division*  $\div$  est une opération algébrique binaire définie comme suit. Soient  $R$  et  $S$  deux noms de relation tels que  $sorte(S) \subseteq sorte(R)$ . On définit  $sorte(R \div S) = sorte(R) \setminus sorte(S)$ . Le résultat de  $R \div S$  contient chaque tuple  $t$  sur  $sorte(R \div S)$  tel que pour chaque tuple  $s$  de  $S$ , le tuple  $t \cup s$  est dans  $R$ .

Par exemple,

$$R \begin{array}{c|cccc} A & B & C & D \\ \hline a & b & 1 & 11 \\ a & b & 2 & 22 \\ a & b & 3 & 33 \\ c & d & 1 & 11 \\ c & d & 3 & 33 \\ e & f & 1 & 11 \\ e & f & 2 & 22 \end{array} \div S \begin{array}{c|cc} C & D \\ \hline 1 & 11 \\ 2 & 22 \end{array} = \begin{array}{c|cc} A & B \\ \hline a & b \\ e & f \end{array}$$

Donnez une expression SPJRUD qui est équivalente à  $R \div S$  ou argumentez pourquoi une telle expression n'existe pas. Noter que l'on ne peut pas supposer que  $sorte(R) = ABCD$ , ni que  $sorte(S) = CD$  (ce n'est vrai que dans l'exemple).

---

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

**Question 2** Argumentez pourquoi la division  $\div$  ne peut pas être exprimée en algèbre SPJRU (i.e., l'algèbre relationnelle sans la différence).

---

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

**Question 3** Pour chacune des trois requêtes suivantes, cochez la case qui précède une expression correcte. Si vous cochez la deuxième case (et seulement dans ce cas), donnez, dans le cadre qui suit, une base de données  $\mathcal{I}$  et un domaine **dom** tel que  $\mathbf{dom} \supseteq \mathit{adom}(\mathcal{I})$  et l'interprétation de la requête par rapport à  $\mathit{adom}(\mathcal{I})$  est différente de l'interprétation par rapport à **dom**. Par exemple,

—  $\{x, y \mid \neg R(x, y) \wedge S(y)\}$

Cette requête est *domain independent*.

Cette requête n'est pas *domain independent*.

$R$	$A$	$B$	$S$	$B$	— Résultat par rapport à $\mathit{adom}(\mathcal{I}) : \{(b, b)\}$
	$a$	$b$		$b$	— Résultat par rapport à $\{a, b, c\} : \{(b, b), (c, b)\}$

—  $\{x, y \mid \neg R(x, y) \wedge S(x) \wedge S(y)\}$

Cette requête est *domain independent*.

Cette requête n'est pas *domain independent*.

.../6

1.  $\{x \mid \forall y(R(x, y) \rightarrow \exists zR(z, x))\}$

Cette requête est *domain independent*.

Cette requête n'est pas *domain independent*.

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

Cette requête est *domain independent*.

Cette requête n'est pas *domain independent*.

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

Cette requête est *domain independent*.

Cette requête n'est pas *domain independent*.

#### Question 4

**Conjecture 1** Soit  $\Sigma$  un ensemble de dépendances multivaluées sur un ensemble  $\mathcal{A}$  d'attributs. Pour toute dépendance fonctionnelle  $X \rightarrow Y$  sur  $\mathcal{A}$ , si  $\Sigma \models X \rightarrow Y$ , alors  $Y \subseteq X$ .

Cochez une case :

- Cette conjecture est correcte.
- Cette conjecture est fausse.

Si vous avez coché la première case, donnez une preuve de la conjecture. Si vous avez coché la deuxième case, donnez un contre exemple.

---

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

**Question 5** On utilise une table pour encoder les prix de bières en bouteille disponibles auprès de différents distributeurs (Carrefour, Colruyt, Delhaize...). Chaque bière est produite par une brasserie (Brasserie de la Senne, Brasserie d'Achoffe, Chimay...). Chaque brasserie est située dans une seule ville (attribut *Ville*). Chaque bière porte un nom unique (attribut *Bière*); le degré alcoolique est exprimé en pourcentage volumique (attribut *Alcool*). Les bières sont disponibles en bouteilles de 25, 33, 37.5, 75 ou 100 centilitres. Certaines bières sont disponibles en plusieurs bouteilles; par exemple, la Chimay Bleue est disponible en bouteilles de 33 et 75 centilitres. Toutes ces bières et bouteilles ne sont pas forcément disponibles auprès de chaque distributeur. Cependant, pour une bière donnée, la gamme de bouteilles disponible auprès d'un même distributeur ne varie pas au cours d'une même année civile (sauf si le distributeur décide de suspendre toute vente de cette bière). Par exemple, si Delhaize vend la Chimay Bleue en 33 centilitres en janvier 2016 et en 75 centilitres en février 2016, alors Delhaize doit aussi vendre la Chimay Bleue en 75 centilitres en janvier 2016 et en 33 centilitres en février 2016. La première ligne encode qu'une bouteille de 33 centilitres de Chimay Bleue était vendue par Delhaize au prix de 2.10 EUR en mai 2016. Les prix ne varient pas au cours d'un même mois.

<i>Bière</i>	<i>Alcool</i>	<i>Brasserie</i>	<i>Ville</i>	<i>Bouteille</i>	<i>Distributeur</i>	<i>Mois</i>	<i>Année</i>	<i>EUR</i>
Chimay Bleue	9	Chimay	Scourmont	33	Delhaize	mai	2016	2.10
Chimay Bleue	9	Chimay	Scourmont	75	Delhaize	mai	2016	5.10
Chimay Bleue	9	Chimay	Scourmont	33	Carrefour	mai	2016	2.00
Chimay Bleue	9	Chimay	Scourmont	33	Delhaize	juin	2016	2.15
Chimay Bleue	9	Chimay	Scourmont	75	Delhaize	juin	2016	5.15
Chimay Bleue	9	Chimay	Scourmont	33	Carrefour	juin	2016	2.05
Stouterik	4.5	Brass. de la Senne	Bruxelles	33	Carrefour	mai	2016	2.20
Stouterik	4.5	Brass. de la Senne	Bruxelles	33	Carrefour	juin	2016	2.20
Chimay Bleue	9	Chimay	Scourmont	33	Delhaize	mai	2017	2.99
Chimay Bleue	9	Chimay	Scourmont	75	Carrefour	mai	2017	5.99
Chimay Rouge	7	Chimay	Scourmont	75	Delhaize	mai	2017	6.00
Chimay Rouge	7	Chimay	Scourmont	33	Carrefour	mai	2017	2.00

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

.../5

**Question 6** La table *BIÈRES* liste toutes les bières. La table *STOCK* liste le nombre de bouteilles en stock chez différents distributeurs. La table *PRIX* liste les prix des bières qui sont disponibles chez Colruyt. L'attribut *MPA* (Meilleur Prix Ailleurs) stocke le plus bas prix que l'on trouve en dehors de Colruyt. Si le prix *MPA* est inférieur au prix *Colruyt*, Colruyt promet à ses clients de rembourser la différence ☺.

<i>BIÈRES</i>	<i>Bière</i>	<i>Brasserie</i>	<i>Alcool</i>
	Chimay Bleue	Chimay	9
	Chimay Rouge	Chimay	7
	Stouterik	Brass. de la Senne	4.5
	Brusseleir	Brass. de la Senne	8

<i>STOCK</i>	<i>Bière</i>	<i>Distributeur</i>	<i>Bouteille</i>	<i>Quantité</i>
	Chimay Bleue	Spar	33	8000
	Chimay Bleue	Spar	75	800
	Stouterik	Spar	33	200
	Chimay Bleue	Colruyt	33	7000
	Chimay Rouge	Colruyt	33	8000
	Chimay Rouge	Colruyt	75	1000

<i>PRIX</i>	<i>Bière</i>	<i>Bouteille</i>	<i>Colruyt</i>	<i>MPA</i>
	Chimay Bleue	33	2.10	2.15
	Chimay Rouge	33	2.10	2.20
	Chimay Rouge	75	6.00	6.15

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

---

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

**Question 7** Pour la base de données de la question 6, écrivez une requête en algèbre relationnelle pour répondre à la question :

Quelles sont les deux bières les plus fortes ? Plus précisément, dans le cas où plusieurs bières aient le même degré alcoolique, on demande d'afficher chaque bière telle que le nombre de bières strictement plus fortes est  $\leq 1$ .

Pour la base de données de la question 6, le résultat est :

<i>Bière</i>
Chimay Bleue
Brusseleir

Pour cette question, on peut utiliser la sélection  $\sigma_{A < B}(E)$  avec  $A, B \in \text{sorte}(E)$ . La sémantique est :

$$\llbracket \sigma_{A < B}(E) \rrbracket^{\mathcal{I}} = \{t \in \llbracket E \rrbracket^{\mathcal{I}} \mid t(A) < t(B)\}.$$

---

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

**Question 8** Pour la base de données de la question 6, écrivez une requête en calcul relationnel qui rend chaque paire  $\langle \text{Distributeur} : d, \text{Brasserie} : b \rangle$  tel que le distributeur  $d$  a **toutes** les bières de la brasserie  $b$  en stock (peu importe les tailles des bouteilles).

Pour la base de données de la question 6, le résultat est comme suit :

<i>Distributeur</i>	<i>Brasserie</i>
Colruyt	Chimay

Notez que Spar ne stocke pas de Chimay Rouge, et qu'aucun distributeur stocke toutes les bières de la Brasserie de la Senne.

---

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





**Question 11** Situez et expliquez le protocole **Redo/No-Undo**.

---

.../5

**Question 12** Considérez l'exécution suivante :

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

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

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