Bases de Données I, Mons, 9 janvier 2025

NOM + PRÉNOM :

Orientation + Année :

Cet examen comprend 9 questions et doit être terminé dans une durée précise de 3 heures. Les questions sont censées être claires. Aucune clarification supplémentaire ne sera fournie pendant l'examen. Si une question vous semble ambiguë ou incomplète, veuillez formuler vos hypothèses et répondez en fonction de celles-ci.

Question 1 La table ci-dessous sert à enregistrer des concerts en plein air en Belgique. L'endroit d'un concert est déterminé par les attributs *Commune* et *Place*. Les attributs *Date* et *Heure* indiquent le jour et l'heure de début de chaque concert. Le jour de la semaine (lundi, mardi,...) est également enregistré. L'attribut *Tickets* répertorie les sites web où les billets peuvent être achetés, et l'attribut *Prix* spécifie le coût d'un billet.

Le montage et le démontage d'une scène pour un concert nécessitent un temps considérable. Pour cette raison, il est impossible qu'un même artiste donne des concerts à deux endroits différents le même jour. De plus, un même lieu ne peut pas accueillir deux artistes différents le même jour. Cependant, un artiste peut donner deux concerts au même endroit le même jour, par exemple, un premier concert l'après-midi, suivi d'un deuxième concert le soir. Le prix d'entrée peut varier selon l'heure.

Les différents sites web sont tenus de facturer le même prix pour un même concert.

Les communes sont responsables d'assurer la sécurité des spectateurs, ce qui nécessite un effort policier important. Pour cette raison, une commune ne peut pas accueillir des concerts à deux endroits différents le même jour.

Commune	*Jour*	*Date*	*Place*	*Artiste*	*Tickets*	*Prix*	*Heure*
Mons	samedi	7 juin 2025	Place du Parc	Stromae	www.fnacspectacles.com	100 EUR	14:00
Mons	samedi	7 juin 2025	Place du Parc	Stromae	www.ticketac.com	100 EUR	14:00
Mons	samedi	7 juin 2025	Place du Parc	Stromae	www.fnacspectacles.com	110 EUR	20:00
Ixelles	samedi	7 juin 2025	Place Flagey	Adele	www.ticketac.com	250 EUR	21:00
Ixelles	samedi	7 juin 2025	Place Flagey	Adele	www.billetreduc.com	250 EUR	21:00
Ixelles	dimanche	8 juin 2025	Place Flagey	Adele	www.billetreduc.com	220 EUR	14:00
Ixelles	dimanche	15 juin 2025	Place Flagey	Stromae	www.billetreduc.com	150 EUR	14:00
Liège	dimanche	15 juin 2025	Place du Parc	Adele	www.billetreduc.com	260 EUR	21:00

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

.../8

Date → Jour
Commune, Date → Artiste, Place
Commune, Date, Heure → Prix
Artiste, Date → Commune

Consequences logiques :
Artiste, Date, Heure → Prix
Artiste, Date → Place

Question 2 Les tables ci-dessous servent à enregistrer les emprunts dans une bibliothèque. Cette bibliothèque peut posséder zéro, un, ou plusieurs exemplaires d'un même livre. Chaque exemplaire a un numéro unique (*ENr*). Chaque livre est identifié par son ISBN (International Standard Book Number), un identifiant unique permettant une identification standardisée au niveau international. Notez que deux éditions différentes d'une même œuvre peuvent avoir des ISBN différents. Chaque membre de la bibliothèque reçoit une (et une seule) carte avec un identifiant unique (*NuméroCarte*). Pour chaque exemplaire emprunté, la table *EMPRUNTS* enregistre l'identité de l'emprunteur, la date de l'emprunt, ainsi que la date où le livre a été retourné. Une valeur PER (Pas Encore Retourné) dans la colonne *DateRetour* signifie que l'exemplaire a été emprunté mais n'a pas encore été rendu. Un livre ne peut pas être emprunté et retourné le même jour. On peut faire les suppositions suivantes :

- il n'y a pas deux lecteurs différents ayant à la fois le même nom, prénom et date de naissance ;
- la bibliothèque possède au moins un exemplaire de chaque livre figurant dans la table *LIVRES*.

LIVRES

ISBN	*Titre*	*Auteur*	*AnnéePublication*	*Langue*
1111	Het verdriet van België	Hugo Claus	1983	néerlandais
2222	Het verdriet van België	Hugo Claus	2008	néerlandais
3333	Le chagrin des Belges	Hugo Claus	2014	français
4444	À la lumière de nos jours	Clarisse Sabard	2022	français
5555	Matilda	Roald Dahl	2022	anglais

LECTEURS

NuméroCarte	*Nom*	*Prénom*	*DateNaissance*
L1	Vanloo	Julien	10/02/1993
L2	Dupont	Jean	31/12/2001
L3	Witsel	Axel	12/01/1989

EXEMPLAIRES

ISBN	*ENr*	*DateAcquisition*
1111	E11	14/01/2020
1111	E12	14/01/2020
2222	E21	14/01/2020
2222	E22	14/01/2020
3333	E31	14/01/2020
4444	E41	15/02/2021
4444	E42	16/04/2022
5555	E51	09/08/2024

EMPRUNTS

Lecteur	*Livre*	*DateEmprunt*	*DateRetour*
L1	E11	02/12/2024	PER
L1	E42	02/12/2024	16/12/2024
L2	E31	16/12/2024	PER
L2	E42	17/12/2024	23/12/2024

Donnez pour chaque table la clé primaire, les clés étrangères et les contraintes de type UNIQUE, en utilisant la syntaxe vue au cours.

.../8

LIVRES PK(*ISBN*)
LECTEURS PK(*NuméroCarte*)
LECTEURS UNIQUE(*Nom, Prénom, DateNaissance*)
EMPRUNTS PK(*Livre, DateEmprunt*)
EXEMPLAIRES PK(*ENr*)
EXEMPLAIRES FK(*ISBN*) REFS *LIVRES*
EMPRUNTS FK(*Lecteur*) REFS *LECTEURS*
EMPRUNTS FK(*Livre*) REFS *EXEMPLAIRES*

Question 3 Pour la base de données de la question 2, écrivez une requête en algèbre relationnelle pour répondre à la question :

> Quels sont les numéros ISBN des livres en français ou en néerlandais dont aucun exemplaire n'a été emprunté ?

Pour la base de données de la question 2, le résultat est :

ISBN
2222

En effet, parmi les livres en néerlandais, le livre ISBN 2222 est le seul à n'avoir jamais été emprunté. De plus, tous les livres en français ont été empruntés au moins une fois. Il est important de noter que le livre 4444 n'apparaît pas dans le résultat, même si son exemplaire E41 n'a jamais été emprunté. La raison en est que l'exemplaire E42 a bien été emprunté.

…/10

$$L := \pi_{ISBN}(\sigma_{Langue=\text{néérlandais}}(LIVRES) \cup \sigma_{Langue=\text{français}}(LIVRES))$$
$$E := \pi_{ISBN}(EXEMPLAIRES \bowtie \rho_{Livre \to ENr}(EMPRUNTS))$$

La requête demandée est :

$$L - E.$$

Question 4 Disons qu'un lecteur est *polyglotte* s'il a emprunté des livres dans différentes langues. Pour la base de données de la question 2, rédigez une requête en calcul relationnel qui renvoie le nom et le prénom de chaque lecteur qui n'est pas polyglotte.

Pour la base de données de la question 2, le résultat est comme suit :

Nom	*Prénom*
Dupont	Jean
Witsel	Axel

En effet, le seul lecteur polyglotte est Vanloo Julien, ayant emprunté des livres en néerlandais (exemplaire E11) et en français (exemplaire E42). Le résultat contient donc tous les autres lecteurs.

Formatez votre requête de manière à ce qu'elle soit facile à lire et à comprendre. Ajoutez, si besoin, une explication en français.

$\boxed{\ldots/10}$

Soit $\varphi(\ell, w)$ la formule suivante qui exprime que ℓ est l'identifiant d'un lecteur ayant emprunté un livre en langue w :

$$\varphi(\ell, w) := \exists e \exists x \exists z \exists i \exists q \exists t \exists r \exists u \, (\mathit{EMPRUNTS}(\ell, e, x, z) \wedge \mathit{EXEMPLAIRES}(i, e, q) \wedge \mathit{LIVRES}(i, t, r, u, w)) \, .$$

La requête demandée est :

$$\big\{n, p \mid \exists \ell \exists w' \left(\exists y \, (\mathit{LECTEURS}(\ell, n, p, y)) \, \wedge \forall w \left(\varphi(\ell, w) \rightarrow w = w'\right)\right)\big\} \, .$$

La formule exprime qu'il existe un lecteur avec l'identifiant ℓ, portant le nom n et le prénom p, ainsi qu'une langue w', tel que tous les exemplaires empruntés par ℓ sont des livres en w'.

Voici une autre solution.

$$\{n, p \mid \exists \ell \exists y \, (\mathit{LECTEURS}(\ell, n, p, y) \wedge \forall w_1 \forall w_2 \, ((\varphi(\ell, w_1) \wedge \varphi(\ell, w_2)) \rightarrow w_1 = w_2)) \, .$$

Ici $\varphi(\ell, w_1)$ est la formule obtenue à partir de $\varphi(\ell, w)$ en remplaçant w par w_1.

Question 5 On ajoute un nouvel opérateur binaire $\oplus$ à l'algèbre SPJRUD. De façon formelle, $\oplus$ est défini comme suit :

Syntaxe : Si E_1 et E_2 sont des expressions algébriques telles que $sorte(E_1) = sorte(E_2)$, alors $E_1 \oplus E_2$ est une expression algébrique telle que $sorte(E_1 \oplus E_2) = sorte(E_1)$.

Sémantique : Soit $\mathcal{I}$ une instance de base de données.

$$[\![E_1 \oplus E_2]\!]^{\mathcal{I}} := \begin{cases} [\![E_1]\!]^{\mathcal{I}} \cup [\![E_2]\!]^{\mathcal{I}} & \text{si } [\![E_1]\!]^{\mathcal{I}} \subseteq [\![E_2]\!]^{\mathcal{I}} \text{ ou } [\![E_2]\!]^{\mathcal{I}} \subseteq [\![E_1]\!]^{\mathcal{I}}\,; \\ [\![E_1]\!]^{\mathcal{I}} \cap [\![E_2]\!]^{\mathcal{I}} & \text{sinon.} \end{cases}$$

En termes simples, $\oplus$ renvoie l'union des deux relations d'entrée si l'une des deux est incluse dans l'autre ; sinon, $\oplus$ renvoie l'intersection des deux relations d'entrée.

Par exemple, considérons les relations suivantes :

$$r_1 \begin{array}{|cc} A & B \\ \hline a & 1 \\ b & 2 \end{array} \qquad r_2 \begin{array}{|cc} A & B \\ \hline a & 1 \\ b & 2 \\ c & 3 \end{array} \qquad r_3 \begin{array}{|cc} A & B \\ \hline a & 1 \\ c & 3 \end{array} \qquad r_4 \begin{array}{|cc} A & B \\ \hline a & 1 \end{array}$$

Puisque $r_1 \subseteq r_2$, on obtient $r_1 \oplus r_2 = r_1 \cup r_2 = r_2$. Puisque $r_1 \nsubseteq r_3$ et $r_3 \nsubseteq r_1$, on obtient $r_1 \oplus r_3 = r_1 \cap r_3 = r_4$.

Cochez la case qui précède une expression correcte :

☐ L'opérateur $\oplus$ est monotone ;

☒ l'opérateur $\oplus$ est non-monotone.

Démontrez l'exactitude de la case cochée.

…/6

Soit r, s_1, et s_2 les relations suivantes :

$$r \begin{array}{|c} A \\ \hline b \end{array} \qquad s_1 \begin{array}{|c} A \\ \hline \end{array} \qquad s_2 \begin{array}{|c} A \\ \hline c \end{array}$$

Notez que s_1 est la relation vide et s_2 est non vide. Alors $s_1 \subseteq s_2$, mais $r \oplus s_1 \nsubseteq r \oplus s_2$. En effet, $r \oplus s_1 = r$ et $r \oplus s_2 = s_1$.

Question 6 L'algèbre SPJRUD$\oplus$ est définie comme l'algèbre résultante de l'ajout du nouvel opérateur $\oplus$, tel que défini dans la question 5, à l'algèbre SPJRUD. Cochez la case qui précède une expression correcte :

☒ Pour chaque expression E en SPJRUD$\oplus$, il existe une expression E' en SPJRUD telle que $E' \equiv E$. Le nouvel opérateur est donc redondant.

☐ L'algèbre SPJRUD$\oplus$ est plus expressive que SPJRUD.

Argumentez votre choix. Si vous avez cochez la première case, expliquez la construction de E' à partir de E.

$\boxed{\ldots/10}$

Supposons que R et S ont arité n. Soit $\vec{x} = x_1, x_2, \ldots, x_n$.

La formule φ suivante est vraie ssi $R \subseteq S$ ou $S \subseteq R$:

$$\varphi := \forall \vec{x}\, (R(\vec{x}) \rightarrow S(\vec{x})) \vee \forall \vec{x}\, (S(\vec{x}) \rightarrow R(\vec{x}))$$

$R \oplus S$ peut donc s'exprimer comme suit en calcul relationnel :

$$\{\vec{x} \mid (\varphi \wedge (R(\vec{x}) \vee S(\vec{x}))) \vee (\neg\varphi \wedge (R(\vec{x}) \wedge S(\vec{x})))\}.$$

Notez :
- si φ est fausse, cette requête équivaut à $R(\vec{x}) \wedge S(\vec{x})$; et
- si φ est vraie, cette requête équivaut à $R(\vec{x}) \vee S(\vec{x})$.

Une requête équivalent plus courte est :

$$\{\vec{x} \mid (R(\vec{x}) \wedge S(\vec{x})) \vee (\varphi \wedge (R(\vec{x}) \vee S(\vec{x})))\}.$$

Cette formule est *domain independent* et donc, par le Théorème de Codd, équivalent à une expression en SPJRUD.

Question 7 Soit $\varphi = \exists x \exists y \forall z\,(R(x,y) \rightarrow \exists v\,(R(z,v) \vee R(v,z)))$, une formule en logique des prédicats. Soit $\mathcal{I}$ une instance de base de données qui contient la relation suivante :

$$\begin{array}{c|cc} R & A & B \\ \hline & 1 & 2 \\ & 2 & 1 \\ & 3 & 3 \end{array}.$$

Cochez la case qui précède une expression correcte :

☒ La formule φ est vraie dans $\mathcal{I}$.

☐ La formule φ est fausse dans $\mathcal{I}$.

☐ Puisque φ n'est pas *domain independent*, il n'est pas possible de savoir si φ est vraie ou fausse dans $\mathcal{I}$ sans connaître le domaine d'interprétation $\mathbf{dom}$.

Argumentez votre choix.

$\boxed{\ldots/8}$

En choisissant la valeur 1 pour x et y, la formule devient :

$$\forall z\,(R(1,1) \rightarrow \exists v\,(R(z,v) \vee R(v,z)))\,.$$

Cette formule exprime que pour toute constante $c \in \mathbf{dom}$, on aura :

$$R(1,1) \rightarrow \exists v\,(R(c,v) \vee R(v,c))\,.$$

La dernière formule est toujours vraie quel que soit c, puisque l'hypothèse $R(1,1)$ est fausse et qu'une implication avec une prémisse fausse est toujours vraie.

Une analyse plus approfondie nous apprend que φ est équivalent à :

$$\begin{aligned}
&\exists x \exists y \forall z\,(\neg R(x,y) \vee \exists v\,(R(z,v) \vee R(v,z)))\\
\equiv\,&(\exists x \exists y\,(\neg R(x,y))) \vee (\forall z \exists v\,(R(z,v) \vee R(v,z)))\\
\equiv\,&\neg\,(\forall x \forall y\,(R(x,y))) \vee (\forall z \exists v\,(R(z,v) \vee R(v,z)))\\
\equiv\,&(\forall x \forall y\,(R(x,y))) \rightarrow (\forall z \exists v\,(R(z,v) \vee R(v,z)))\\
\equiv\,&(\forall x \forall y\,(R(x,y))) \rightarrow (\forall x \exists v\,(R(x,v) \vee R(v,x)))
\end{aligned}$$

On peut vérifier que cette dernière formule, qui prend la forme d'une implication, est **vraie** pour tout domaine d'interprétation D, **quelle que soit la base de données**. Si la prémisse $(\forall x \forall y\,(R(x,y)))$ est fausse, l'implication est trivialement vraie. Regardons maintenant le cas où la prémisse $(\forall x \forall y\,(R(x,y)))$ est vraie. Cette prémisse exprime que la relation R est totale, c'est-à-dire qu'elle contient chaque élément du produit cartésien $D \times D$. Alors tout élément de D apparaît nécessairement comme premier ou second élément dans au moins une paire de R, donc la conclusion est vraie, et l'implication est donc toujours satisfaite.

Question 8 Soient

$$\begin{array}{rcl} \mathcal{A} & = & ABCDEF \\ \Sigma & = & \{BDE \rightarrow A, ABDE \rightarrow F, AEF \rightarrow C, F \rightarrow E\} \end{array}$$

Cochez la case qui précède une expression correcte :

- ☐ Le schéma $(\mathcal{A}, \Sigma)$ est en BCNF.
- ☐ Le schéma $(\mathcal{A}, \Sigma)$ est en 3NF mais n'est pas en BCNF.
- ☒ Le schéma $(\mathcal{A}, \Sigma)$ n'est pas en 3NF.

Détaillez les arguments qui mènent à cette conclusion.

$\boxed{\ldots/10}$

Puisque C n'apparaît pas à gauche d'une flèche, il est facile de vérifier que, pour tout $X \subseteq \mathcal{A}$, on a $X^{*,\Sigma} = (X \setminus \{C\})^{*,\Sigma}$. En conséquence, aucune clé ne peut contenir C.

Puisque $\{A, E, F\}^{*,\Sigma} = \{A, C, E, F\} \neq \mathcal{A}$ et que C n'est pas un attribut premier, la dépendance fonctionnelle $AEF \rightarrow C$ constitue une violation de la forme normale 3NF.

Question 9 Considérez l'exécution suivante :

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

Cochez la case qui précède une expression correcte :

- ☒ Cette exécution est possible en 2PL.
- ☐ cette exécution n'est pas possible en 2PL.

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

…/10

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