

Bases de Données II, Mons, 7 juin 2023

NOM + PRÉNOM :

Orientation + Année :

Cet examen contient 9 questions. Durée : exactement 3 heures. Les questions sont censées être claires. Aucune clarification supplémentaire ne sera donnée pendant l'examen. Si une question vous semble ambiguë ou incomplète, écrivez vos hypothèses et répondez en fonction de celles-ci. **Il est permis de détacher la dernière page.**

Un fleuriste livre des bouquets à la maison. Chaque espèce de fleur (tulipe, rose...) a un prix exprimé en centimes d'euro. Par exemple, le prix d'une rose est de 150 centimes d'euro, indépendamment de sa couleur ; voir la ligne

```
<fleur fnom="rose" prix="150"/>
```

Un bouquet rassemble des fleurs de différentes espèces et couleurs. Par exemple, le bouquet *Exotique* est composé d'un seul tournesol et trois iris bleus. Voir les lignes :

```
<bouquet bnom="Exotique">
  <fleur fnom="tournesol" couleur="jaune" nombre="1"/>
  <fleur fnom="iris" couleur="bleu" nombre="3"/>
</bouquet>
```

La figure 1 montre le DTD et la figure 2 le document XML.

La balise `ventes-par-jour` est utilisée pour encoder les ventes journalières. Par exemple, à la date du 18 septembre 2022, le fleuriste a vendu trois bouquets *Valentin* et six bouquets *Printemps*. Voir les lignes :

```
<date mois="sep 2022" jour="18">
  <vente bnom="Valentin">3</vente>
  <vente bnom="Printemps">6</vente>
</date>
```

Question 1 Écrivez une requête en **XPath** qui renvoie le nom de chaque bouquet contenant au moins deux couleurs distinctes. **Il n'est pas permis d'utiliser des fonctions d'agrégation telles que `count`, `min` et `max`.** Évitez l'usage des axes *parent* et *ancestor*.

Pour le document XML de la figure 2, la réponse est comme suit :

```
bnom="Belge"
bnom="Exotique"
```

Noter que le bouquet *Printemps* ne contient que des fleurs jaunes et n'est donc pas dans la réponse.

.../5

Question 2 Écrivez une requête en **XPath** qui renvoie le nom de chaque bouquet qui n'a aucune fleur en commun avec le bouquet *Printemps*. **Il n'est pas permis d'utiliser des fonctions d'agrégation telles que count, min et max.** Évitez l'usage des axes *parent* et *ancestor*.

Pour le document XML de la figure 2, la réponse est comme suit :

```
bnom="Exotique"
```

Noter que le bouquet *Printemps* contient des roses et des tulipes. Le bouquet *Exotique* est le seul qui ne contient ni de roses ni de tulipes.

.../5

Question 3 Écrivez une requête en **XQuery** qui renvoie le mois avec le plus grand nombre de bouquets vendus.

Pour le document XML de la figure 2, la réponse est :

```
sep 2022
```

.../10

Question 4 Écrivez un programme en XSLT qui renvoie, pour chaque mois (sans doublons), le nombre total de bouquets vendus. **Il n'est pas permis d'utiliser `xsl:for-each` et `xsl:if`. Il n'est pas non plus permis d'utiliser la fonction `distinct-values`.** L'usage de la fonction `sum` est permis. Les résultats doivent être groupés comme suit :

```
<ventes-par-mois>
  <nombre-de-bouquets mois="sep 2021">4</nombre-de-bouquets>
  <nombre-de-bouquets mois="juin 2022">7</nombre-de-bouquets>
  <nombre-de-bouquets mois="sep 2022">17</nombre-de-bouquets>
</ventes-par-mois>
```

Noter que 17 est obtenu comme $3 + 6 + 2 + 6$. L'ordre des mois n'a pas d'importance.

.../10

Question 5 Simplifiez la requête (UCQ) suivante ou expliquez pourquoi aucune simplification n'est possible :

$$\begin{cases} \text{Answer}(x) \leftarrow R(x, y), R(y, z), R(z, x), R(x, u), R(u, v), R(v, x) \\ \text{Answer}(y) \leftarrow R(y, x), R(x, y) \end{cases}$$

Détaillez les calculs.

.../10

Question 6 Un *graphe dirigé simple* est un couple $G = (V, E)$ avec V un ensemble fini de *sommets* et E un ensemble d'*arcs* (u, v) avec $u, v \in V$, $u \neq v$. On traite seulement des graphes sans sommets isolés, i.e., chaque sommet a au moins un arc sortant ou entrant. Chaque arc du graphe est coloré en une et une seule couleur.

Dans une base de données, un prédicat EDB R d'arité 3 est utilisé pour stocker les arcs du graphe et leur couleur. Par exemple, $R(a, c, \text{red})$ signifie qu'il existe un arc dirigé de a à c avec la couleur rouge.

On définit un *chemin dirigé* de u_1 à u_n comme une séquence $\langle u_1, u_2, \dots, u_n \rangle$ de sommets telle que $n \geq 2$ et pour chaque $i \in \{1, 2, \dots, n-1\}$, (u_i, u_{i+1}) est un arc du graphe. La *longueur* de ce chemin est de $n-1$, i.e., la longueur est le nombre d'arcs parcourus. Noter que la condition $n \geq 2$ dans la définition précédente implique que la longueur d'un chemin n'est jamais zéro. Noter aussi qu'un sommet peut se répéter dans un chemin.

On dira qu'un chemin dirigé $\langle u_1, u_2, \dots, u_n \rangle$ est *tricolore* si $n \geq 4$ et pour chaque $i \in \{1, 2, \dots, n-3\}$, les trois arcs (u_i, u_{i+1}) , (u_{i+1}, u_{i+2}) et (u_{i+2}, u_{i+3}) sont de trois couleurs distinctes. En mots simples, la longueur d'un chemin tricolore est d'au moins 3 et chaque trois arcs qui se succèdent sur un chemin tricolore sont de trois couleurs distinctes.

Pour le graphe de la Figure 3, le chemin $\langle f, d, a, b, f, g, h, i, j \rangle$ est tricolore : les couleurs parcourues sont, de gauche à droite, $\langle \text{blue, red, green, blue, pink, red, green, blue} \rangle$; chaque trois couleurs qui se succèdent sur ce chemin sont mutuellement distinctes. Noter que les trois occurrences de **blue** sur ce chemin sont bien séparées l'une de l'autre par (au moins) deux autres couleurs.

Un chemin tricolore est *non-extensible* s'il est impossible de le rendre plus longue, en ajoutant un arc à gauche ou à droite, sans perdre la propriété d'être tricolore. De façon formelle, si $\langle u_1, u_2, \dots, u_n \rangle$ est un chemin tricolore non-extensible, alors pour tous les arcs $(t, u_1), (u_n, v)$ de E , les chemins $\langle t, u_1, u_2, \dots, u_n \rangle$ et $\langle u_1, u_2, \dots, u_n, v \rangle$ ne sont pas tricolores.

Le chemin tricolore $\langle f, d, a, b, f, g, h, i, j \rangle$ est non-extensible ; notamment, si l'on ajoutait b ou e à gauche de ce chemin, le chemin ne serait plus tricolore (parce que les arcs (b, f) et (e, f) ont la même couleur que (f, d)).

$\langle h, g, h, i, j \rangle$ est un autre chemin tricolore non-extensible ; ce chemin parcourt, de gauche à droite, les couleurs $\langle \text{blue, red, green, blue} \rangle$. On ne peut pas ajouter g à gauche de ce chemin, car $\langle g, h, g, h, i, j \rangle$ n'est pas tricolore.

$\langle b, e, f, g, h, g \rangle$ est un autre chemin tricolore non-extensible ; ce chemin parcourt, de gauche à droite, les couleurs $\langle \text{green, blue, pink, red, blue} \rangle$. On ne peut pas ajouter a à gauche de ce chemin, ni h à droite.

Écrivez un programme en Datalog avec \neq et négation stratifiée pour le prédicat IDB binaire T tel que $T(x, y)$ est vrai si x et y sont des sommets tels qu'il existe un chemin tricolore **non-extensible** de x à y .

Avant de donner votre programme, expliquez en français la stratégie suivie par votre programme. Avant chaque règle (ou groupe de règles), expliquez en français ce que cette règle (ou ces règles) vont calculer.

.../20

Question 7 On dira qu'un chemin dirigé $\langle u_1, u_2, \dots, u_n \rangle$ est *palindromique* si $n \geq 4$ est un nombre pair (soit $m = \frac{n}{2}$) et

- le premier arc (u_1, u_2) a la même couleur que le dernier arc (u_{n-1}, u_n) ;
- le deuxième arc (u_2, u_3) a la même couleur que l'avant-dernier arc (u_{n-2}, u_{n-1}) ;
- ...
- l'arc (u_{m-1}, u_m) a la même couleur que l'arc (u_{m+1}, u_{m+2}) ; et
- **deux arcs successifs n'ont pas la même couleur.**

Noter que la séquence des couleurs d'un chemin palindromique forme un palindrome : les couleurs de gauche à droite sont les mêmes que les couleurs de droite à gauche. Par exemple,

red, green, blue, red, blue, red, blue, red, blue, green, red.

Écrivez un programme en Datalog avec \neq et négation stratifiée pour le prédicat IDB unaire L tel que $L(x)$ est vrai si x est un sommet tel qu'il existe un chemin palindromique de x à x .

Pour le graphe de la Figure 3, le programme doit renvoyer $L(a)$. En effet, le chemin $\langle a, c, d, a \rangle$ est palindromique :

- l'arc (a, c) a la même couleur que (d, a) ;
- (a, c) et (c, d) ont une couleur différente; et
- (c, d) et (d, a) ont une couleur différente.

.../10

Question 8 Pour le schéma

$$\{ABCE, ABDF, ABEF, AC, CE, BD, DF\},$$

1. déterminez si, oui ou non, ce schéma est α -acyclique ;
2. si le schéma est α -acyclique, donnez un arbre de jointure pour ce schéma.

.../10

Question 9 Pour le schéma $ABCDE$, déterminez si $\vDash [AB, AD, BE, CD]$ est une conséquence logique de l'ensemble $\{A \rightarrow BD, AB \rightarrow C, AD \rightarrow E, D \rightarrow C, DE \rightarrow A\}$. Détaillez les calculs.

.../10

```

<?xml version="1.0" encoding="utf-8"?>
<!DOCTYPE fleuriste [
<!ELEMENT fleuriste (fleurs, bouquets, ventes-par-jour)>
<!ELEMENT fleurs (fleur)*>
<!ELEMENT bouquets (bouquet)*>
<!ELEMENT bouquet (fleur)*>
<!ELEMENT fleur (#PCDATA)>
<!ELEMENT ventes-par-jour (date)*>
<!ELEMENT date (vente)*>
<!ELEMENT vente (#PCDATA)>
<!ATTLIST fleur fnom CDATA #REQUIRED>
<!ATTLIST fleur prix CDATA #IMPLIED>
<!ATTLIST fleur couleur CDATA #IMPLIED>
<!ATTLIST fleur nombre CDATA #IMPLIED>
<!ATTLIST bouquet bnom CDATA #REQUIRED>
<!ATTLIST date mois CDATA #REQUIRED>
<!ATTLIST date jour CDATA #REQUIRED>
<!ATTLIST vente bnom CDATA #REQUIRED>
]>

```

FIGURE 1 – DTD

```

<?xml version="1.0" encoding="utf-8"?>
<fleuriste>
  <fleurs>
    <!-- Les prix sont en centimes d'euro -->
    <fleur fnom="tulipe" prix="100"/>
    <fleur fnom="rose" prix="150"/>
    <fleur fnom="iris" prix="250"/>
    <fleur fnom="tournesol" prix="300"/>
  </fleurs>
  <bouquets>
    <bouquet bnom="Valentin">
      <fleur fnom="rose" couleur="rouge" nombre="10"/>
    </bouquet>
    <bouquet bnom="Belge">
      <fleur fnom="tulipe" couleur="noir" nombre="3"/>
      <fleur fnom="tulipe" couleur="jaune" nombre="4"/>
      <fleur fnom="tulipe" couleur="rouge" nombre="6"/>
    </bouquet>
    <bouquet bnom="Exotique">
      <fleur fnom="tournesol" couleur="jaune" nombre="1"/>
      <fleur fnom="iris" couleur="bleu" nombre="3"/>
    </bouquet>
    <bouquet bnom="Printemps">
      <fleur fnom="rose" couleur="jaune" nombre="10"/>
      <fleur fnom="tulipe" couleur="jaune" nombre="4"/>
    </bouquet>
  </bouquets>
  <ventes-par-jour>
    <date mois="sep 2021" jour="14">
      <vente bnom="Exotique">4</vente>
    </date>
    <date mois="juin 2022" jour="30">
      <vente bnom="Valentin">4</vente>
      <vente bnom="Printemps">1</vente>
      <vente bnom="Belge">2</vente>
    </date>
    <date mois="sep 2022" jour="18">
      <vente bnom="Valentin">3</vente>
      <vente bnom="Printemps">6</vente>
    </date>
    <date mois="sep 2022" jour="19">
      <vente bnom="Valentin">2</vente>
      <vente bnom="Belge">6</vente>
    </date>
  </ventes-par-jour>
</fleuriste>

```

FIGURE 2 – Document XML

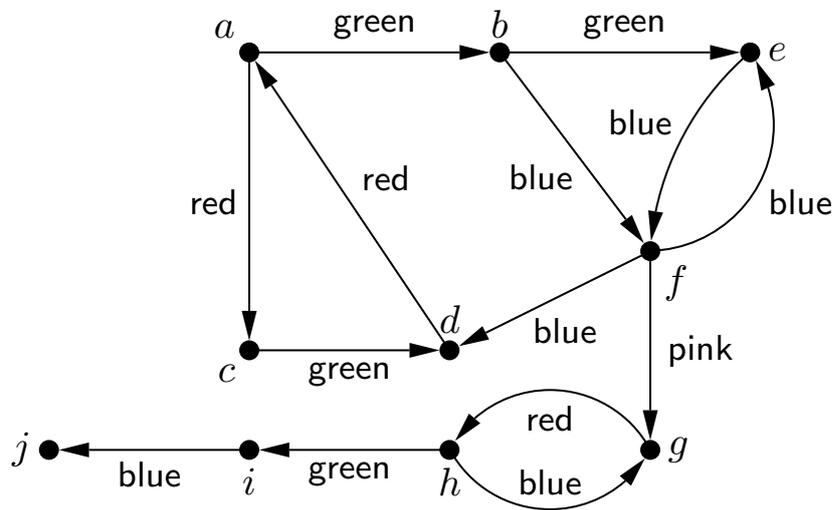


FIGURE 3 – Graphe dirigé avec des arcs colorés.