Examen du module MGDE

Durée : 3 heures

Exercice 1 (4 points) — *Le dilemne du pauvre amoureux transi*

Théophile est amoureux de Juliette mais pauvre. Il veut lui déclarer sa flamme alors que ses finances lui permettent peu : il doit choisir entre l'emmener au cinéma et lui offrir des fleurs.
Juliette est belle mais simple : si Théophile passe le test et lui offre ce qu'elle attend d'un prince charmant, de manière certaine, elle lui accordera son amour.
Une étude statistique poussée a montré que 50% des femmes belles mais simples veulent un bouquet de fleur comme gage d'amour, 30% préfèrent le cinéma et enfin 20% considèrent que ces 2 cadeaux se valent.

Q 1.1 (1 point) Décrire les 3 variables et les relations causales permettant de représenter le dilemne de Théophile. Ne pas oublier les paramètres du modèle.
Remarques : Il n'y a pas de variables cachées. Il manque les paramètres de la loi a priori pour le choix de Théophile.

Q 1.2 (1 point) Théophile a expliqué la situation à son meilleur ami Boris et lui a dit avoir décidé de l'amener au cinéma. Quelque temps plus tard, Boris voit Théophile et Juliette se promener main dans la main et conclut que l'histoire s'est bien terminée. Que peut-il en déduire sur ce que préférait Juliette comme gage d'amour ?

Q 1.3 (1 point) Boris ne peut pas s'empêcher de se demander ce qui ce serait passé si Théophile avait fait l'autre choix. Exprimer cette interrogation sous la forme d'une contrefactuelle et résolvez la.

Q 1.4 (1 point) Est-ce que la décision d'offrir un bouquet était la bonne, finalement ?

Exercice 2 (5 points)

Soit X_i, $i \in \{1, 2, 3, 4, 5\}$, des variables ayant toutes pour domaine $\{1, 2, 3, 4\}$. On considère l'équation :

$$\min_{x_1,x_2,x_3} \left[x_1 + 0,2x_2 - \log((x_2 + x_3) \times x_1) + \max_{x_4,x_5} \left[x_1x_4 + x_4x_3 + x_4^2x_5 \right] \right]. \tag{1}$$

Q 2.1 (3 points) En observant que $\log ab = \log a + \log b$, écrivez la séquence de calculs permettant de résoudre l'équation (1) avec la séquence d'élimination x_5, x_4, x_2, x_1, x_3. Vous effectuerez ces calculs et donnerez la solution du l'équation (1). Vous pourrez vous aider du tableau ci-dessous qui fournit les logarithmes de tous les entiers de 1 à 32 (par exemple, $\log 25$ se trouve à l'intersection de la colonne 1 et de la ligne 24 et vaut donc 3,218876).

	1	2	3	4	5	6	7	8
0	0.000000	0.693147	1.098612	1.386294	1.609438	1.791759	1.945910	2.079442
8	2.197225	2.302585	2.397895	2.484907	2.564949	2.639057	2.708050	2.772589
16	2.833213	2.890372	2.944439	2.995732	3.044522	3.091042	3.135494	3.178054
24	3.218876	3.258097	3.295837	3.332205	3.367296	3.401197	3.433987	3.465736

Q 2.2 (1 point) Tracez l'arbre d'*élimination* correspondant aux calculs ci-dessus. Uniquement pour la phase de collecte, vous indiquerez à côté des séparateurs les opérateurs utilisés pour passer de la clique voisine au séparateur et du séparateur à l'autre clique. Par exemple, sur le graphe de la figure 1, on passe de la clique XYZ au séparateur YZ en calculant un min sur X, et on ajoute la table du séparateur à celle de la clique YZT.

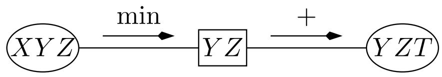


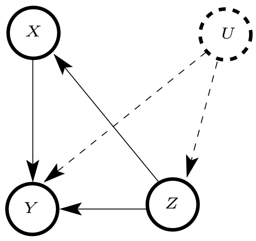
FIG. 1 – Opérateurs utilisés pendant la phase de collecte.

Q 2.3 (1 point) Tracez l'arbre de jonction correspondant.

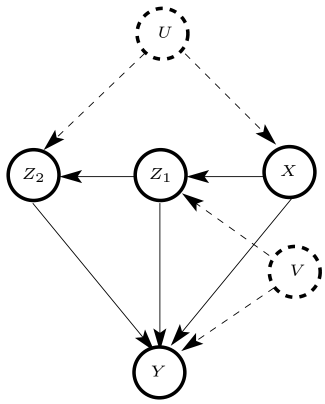
Exercice 3 (6 Points) — *Identification*

Dans les graphes suivants, peut-on calculer $P(y \mid do(x))$? Si oui, donnez son expression.

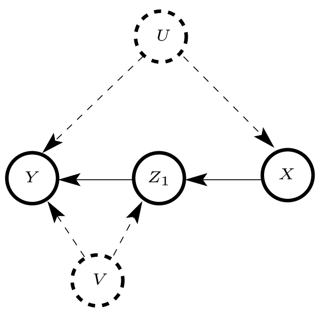
Q 3.1 (2 points)



Q 3.2 (2 points)



Q 3.3 (2 points)



Exercice 4 (10 points)

Dans cet exercice, on considère des graphes non orientés $G = (\mathcal{V}, \mathcal{E})$, où $\mathcal{V} = \{X_1, \ldots, X_n\}$ et $\mathcal{E} \subseteq \{(X_i, X_j) : i, j \in \{1, \ldots, n\}, i \neq j\}$. $\sigma : \{1, \ldots, n\} \mapsto \{1, \ldots, n\}$ représente une permutation.

Q 4.1 (3 points) Soit l'algorithme d'élimination :

Fonction Elim (G, σ)

$\mathcal{E}' \longleftarrow \mathcal{E}$
pour i variant de 1 à n **faire**
  pour X_j, X_k adjacents à $X_{\sigma(i)}$ dans G **faire**
    si $(X_j, X_k) \notin \mathcal{E}$ **alors** $\mathcal{E} \longleftarrow \mathcal{E} \cup \{(X_j, X_k)\}$; $\mathcal{E}' \longleftarrow \mathcal{E}' \cup \{(X_j, X_k)\}$ **finsi**
  fait
  supprimer toutes les arêtes adjacentes à $X_{\sigma(i)}$ dans $\mathcal{E}$
fait
renvoyer le graphe $G' = (\mathcal{V}, \mathcal{E}')$

On appelle cycle de G' un ensemble de nœuds $\{X_{i_1}, \ldots, X_{i_p}\}$ tel que : i) pour tout $k \in \{1, \ldots, p-1\}$, $(X_{i_k}, X_{i_{k+1}}) \in \mathcal{E}'$; ii) $(X_{i_p}, X_{i_1}) \in \mathcal{E}'$; iii) il n'existe pas $j, k \in \{1, \ldots, p\}$, $j \neq k$, tels que $X_{i_j} = X_{i_k}$. La longueur d'un cycle est le nombre d'arêtes qui le constitue, autrement dit, la longueur du cycle $\{X_{i_1}, \ldots, X_{i_p}\}$ est $p + 1$. Montrez que le graphe G' renvoyé par la fonction **Elim** est triangulé, i.e., dans tout cycle de longueur 4 ou plus, il existe deux nœuds X, Y non voisins dans le cycle tels que $(X, Y) \in \mathcal{E}'$.

Q 4.2 (2 points) On appelle graphe markovien associé à un arbre de jonction $\mathcal{J}$ le graphe $G = (\mathcal{V}, \mathcal{E})$ tel que $\mathcal{V}$ est l'ensemble des variables contenues dans les cliques de $\mathcal{J}$ et $\mathcal{E}$ est l'ensemble des arêtes (X_i, X_j) telles qu'il existe une clique dans $\mathcal{J}$ contenant X_i et X_j. Montrez que tout graphe markovien associé à un arbre de jonction est triangulé.

Q 4.3 Considérons l'algorithme d'élimination orienté suivant :

Fonction DirectElim $(G = (\mathcal{V}, \mathcal{E}), \sigma)$

$\mathcal{A} \longleftarrow \emptyset$
pour i variant de 1 à n **faire**
  pour tous X_k adjacents à $X_{\sigma(i)}$ dans G **faire** $\mathcal{A} \longleftarrow \mathcal{A} \cup \left\{\overrightarrow{X_{\sigma(i)} X_k}\right\}$ **fait**
  pour tous X_j, X_k adjacents à $X_{\sigma(i)}$ dans G **faire**
    si $(X_j, X_k) \notin \mathcal{E}$ **alors** $\mathcal{E} \longleftarrow \mathcal{E} \cup \{(X_j, X_k)\}$; **finsi**
  fait
  supprimer toutes les arêtes adajacentes à $X_{\sigma(i)}$ dans $\mathcal{E}$
fait
renvoyer le graphe orienté $G' = (\mathcal{V}, \mathcal{A})$

Le graphe $G' = (\mathcal{V}, \mathcal{A})$ renvoyé par **DirectElim** est un graphe orienté tel que, si l'on enlève les orientations des arcs de $\mathcal{A}$, le graphe résultant est triangulé.

Q 4.3.1 (1 point) Dessinez le graphe G' obtenu à l'issue de l'application de **DirectElim** sur le graphe de la figure 2 et la séquence d'élimination $\sigma = \{A, C, B, E, D, F, G\}$:

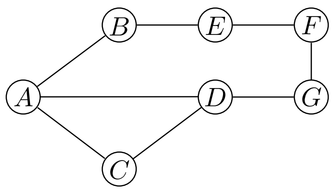


FIG. 2 – Un graphe non orienté à trianguler.

Q 4.3.2 (2 points) Montrer que, pour tout arc $\overrightarrow{XY} \in \mathcal{A}$, X est éliminé avant Y dans l'algorithme **DirectElim**.

Q 4.3.3 (2 points) Dans le graphe G' de la question 4.3.1, on veut supprimer l'arc $\overrightarrow{FG}$ de $\mathcal{A}$ de telle sorte que le graphe ainsi obtenu reste triangulé. Pour cela, à tout arc $\overrightarrow{YZ}$ de $\mathcal{A}$, on associe un nombre entier $x_{YZ} \in \{0, 1\}$ ayant pour signification $x_{YZ} = 1 \Longleftrightarrow$ l'arc $\overrightarrow{YZ}$ est supprimé. On impose donc que $x_{FG} = 1$. Soit le programme linéaire (Σ) défini par :

$$\left\{\begin{array}{l} x_{YZ} \leq x_{XY} + x_{XZ} \ \forall X, Y, Z \text{ tels que } \overrightarrow{XY}, \overrightarrow{XZ}, \overrightarrow{YZ} \in \mathcal{A} \\ x_{FG} = 1 \\ x_{YZ} \in \{0,1\} \ \forall \overrightarrow{YZ} \in \mathcal{A} \end{array}\right. \qquad (\Sigma)$$

Montrer que, pour toute solution du programme (Σ), le graphe $G'' = (\mathcal{V}, \mathcal{A}'')$ constitué uniquement des arcs pour lesquels les x_{YZ} valent 0 est un graphe triangulé. Suggestion : démonstration par récurrence sur les nœuds éliminés, selon leur ordre d'élimination.