Examen du module MGDE

Durée : 3 heures

Exercice 1

Montrer que la relation d'indépendance conditionnelle $\bullet \perp\!\!\!\perp \bullet \mid \bullet$, (qui, par définition, vérifie les propriétés P.1 à P.5 du cours) vérifie nécessairement la propriété suivante :

$$\textit{Mixage :}\ [X \perp\!\!\!\perp (Y,W) \mid Z \ \ \textit{et}\ \ Y \perp\!\!\!\perp W \mid Z] \Longrightarrow (X,W) \perp\!\!\!\perp Y \mid Z.$$

[suggestion : montrer d'abord que $Y \perp\!\!\!\perp X \mid (Z,W)$]

Exercice 2

Une pièce de moteur automobile (bielle) cassera (variable C) certainement prématurément (avant 100.000 km) si elle est défectueuse (variable D) ou s'il y a eu insuffisance d'entretien (variable I). Les 3 variables aléatoires (D,I,C) ont toutes pour valeurs $\{O(ui), N(on)\}$ et leur loi jointe est donnée par les tableaux suivants $(a,b \in (0,1))$:

$$[p(d,i,O)] = \begin{array}{c|c} D \setminus I & \begin{array}{cc} O & N \end{array} \\ \begin{array}{c} O \\ N \end{array} & \begin{bmatrix} a.b & a.\bar{b} \\ \bar{a}.b & 0 \end{bmatrix} \end{array}$$

$$[p(d,i,N)] = \begin{array}{c|c} D \setminus I & \begin{array}{cc} O & N \end{array} \\ \begin{array}{c} O \\ N \end{array} & \begin{bmatrix} 0 & 0 \\ 0 & \bar{a}.\bar{b} \end{bmatrix} \end{array}$$

Q 2.1 Est-ce que : $I \perp\!\!\!\perp D$? $I \perp\!\!\!\perp D \mid C$? Tracer le graphe du RB associé aux trois variables pour l'ordre d'énumération (D,I,C).

Q 2.2 Quelle est la probabilité pour que la rupture de la pièce soit due à ce qu'elle était défectueuse ?

Q 2.3 On apprend que l'entretien a été insuffisant. Quelle est maintenant la probabilité pour que la rupture de la pièce soit due à ce qu'elle était défectueuse ? Comparer cette probabilité avec la précédente.

Exercice 3

Les données statistiques indiquent que le quotient intellectuel (variable Q), le niveau d'études (variable E) et le rendement dans le travail (variable T) dans la population française peuvent être décrits par un modèle probabiliste où les 3 variables aléatoires (Q,E,T) ont toutes pour valeurs $\{h(aut), b(as)\}$ et ont une loi jointe $p(q,e,t)$ donnée par les tableaux suivants :

$$100 \times [p(q,e,h)] = \begin{array}{c|c} Q \setminus E & \begin{array}{cc} h & b \end{array} \\ \begin{array}{c} h \\ b \end{array} & \begin{bmatrix} 36 & 1 \\ 1 & 4 \end{bmatrix} \end{array}$$

$$100 \times [p(q,e,b)] = \begin{array}{c|c} Q \setminus E & \begin{array}{cc} h & b \end{array} \\ \begin{array}{c} h \\ b \end{array} & \begin{bmatrix} 4 & 9 \\ 9 & 36 \end{bmatrix} \end{array}$$

Q 3.1 Est-ce que : $E \perp\!\!\!\perp Q$? $T \perp\!\!\!\perp E \mid Q$? $T \perp\!\!\!\perp Q \mid E$? Tracer le graphe du RB associé aux trois variables pour l'ordre d'énumération (Q, E, T).

Q 3.2 Quelles sont dans cette population les probabilités qu'une personne ait un rendement élevé dans son travail :
- sachant qu'elle a un Q.I. élevé ?
- sachant qu'elle a un niveau d'études élevé ?
- sachant qu'elle a à la fois un Q.I. et un niveau d'études élevés ?

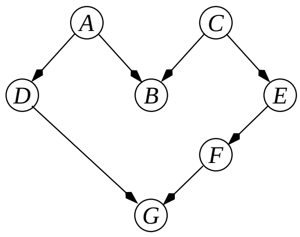
Q 3.3 On admet que le RB obtenu à la question Q 3.1 est un graphe causal, c'est-à-dire qu'une *intervention* sur l'une des variables, X, (imposer une valeur $\hat{x}$ à X) modifie le RB comme suit :
- le graphe associé s'obtient à partir du graphe du RB précédent en supprimant les arcs entrant en X ;
- la loi jointe, notée $p(. \parallel \hat{x})$, décomposable selon ce nouveau graphe, est celle obtenue en laissant inchangées les lois conditionnelles associées aux variables autres que X et en associant à X la loi certaine en $\hat{x}$.

Q 3.3.1 Pour l'intervention $Q = \hat{q} = h$, montrer que $[p(q, e, t \parallel \hat{q})] = [p(q, e, t \ / \ \hat{q})]$.

Q 3.3.2 Pour l'intervention $E = \hat{e} = h$, montrer que $[p(q, e, t \parallel \hat{e})] \neq [p(q, e, t/\hat{e})]$ puis comparer $[p(t)]$, $[p(t/\hat{e})]$ et $[p(t \parallel \hat{e})]$.

Exercice 4

Considérons le réseau bayésien suivant :



Q 4.1 Tracez l'arbre de jonction obtenu par Shafer-Shenoy si l'on utilise la séquence d'élimination B, C, A, D, E, F. Vous indiquerez à côté des cliques les probabilités que vous stockerez dans celles-ci, et à côté des séparateurs les résultats des calculs que vous aurez effectués dans les cliques (si un résultat vaut $P(X)$ et que ce dernier est stocké dans un séparateur (X, Y), vous le noterez $P(X)_Y$).

Q 4.2 Quel est le graphe triangulé correspondant à cette séquence d'élimination ?

Q 4.3 Si chacune des variables est booléenne, quel est le nombre de multiplications et d'additions pour calculer $P(G)$ avec la séquence d'élimination ci-dessus ?

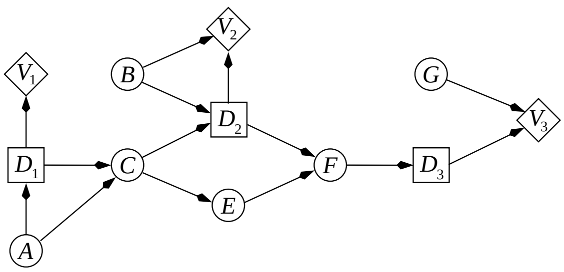
Q 4.4 Tracez le graphe qu'aurait utilisé Pearl pour effectuer les mêmes calculs que Shafer-Shenoy et indiquez à côté des arcs la taille des messages envoyés par Pearl.

Q 4.5 D et F sont-ils d-séparés ? Justifiez votre réponse.

Q 4.6 Si l'on ne tient pas compte de la séquence d'élimination ci-dessus, Pearl a-t-il absolument besoin d'effectuer des coupes-cycles pour calculer $P(G)$? Justifiez votre réponse.

Exercice 5

Soit le diagramme d'influence suivant :

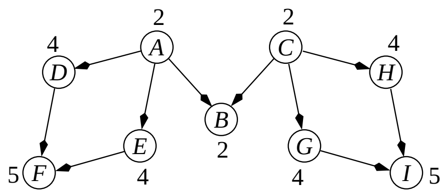


quest Tracez le réseau de valuation correspondant.

Q 5.1 En supposant que l'on ait la contrainte temporelle $A \prec D_1 \prec \{B, C\} \prec D_2 \prec \{E, F, G\} \prec D_3$, triangulez le réseau du 1°) et construisez le strong junction tree correspondant. Vous noterez à côté des cliques les tables que vous stockerez dans celles-ci.

Exercice 6

Soit le réseau bayésien suivant, où le nombre de modalités de chaque variable aléatoire est indiqué à côté du nœud correspondant :



Q 6.1 Triangulez ce réseau en utilisant la méthode de Kjærulff (la méthode vue en cours) et déduisez-en un arbre de jonction (vous dessinerez le graphe à chacune des étapes de l'algorithme et noterez les poids de Kjærulff à côté de chaque nœud).

Q 6.2 Quelle est la triangulation optimale ?