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 :

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

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

Par P.3, $X \perp\!\!\!\perp (Y,W) \mid Z \quad \Longrightarrow \quad X \perp\!\!\!\perp Y \mid (Z,W)$;
Par P.1, $X \perp\!\!\!\perp Y \mid (Z,W) \quad \Longrightarrow \quad Y \perp\!\!\!\perp X \mid (Z,W)$;
Par P.4, $\left\{ \begin{array}{c} Y \perp\!\!\!\perp W \mid Z \\ Y \perp\!\!\!\perp X \mid (Z,W) \end{array} \right. \Longrightarrow Y \perp\!\!\!\perp (X,W) \mid Z$;
Par P.1, $Y \perp\!\!\!\perp (X,W) \mid Z \quad \Longrightarrow \quad (X,W) \perp\!\!\!\perp Y \mid Z.$

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 $(a, b \in (0,1))$:

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

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

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).

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

Comme ses éléments sont les produits de ses marges qui sont $[p(d)] = \begin{bmatrix} a \\ \bar{a} \end{bmatrix}$ et $[p(i,)] = \begin{bmatrix} b & \bar{b} \end{bmatrix}$, on a : $I \perp\!\!\!\perp D$.
Les éléments des tableaux $[p(d,i/O)]$ et $[p(d,i/N)]$ sont respectivement proportionnels à ceux de $[p(d,i,O)]$ et $[p(d,i,N)]$; ceux-ci n'étant pas les produits de leurs marges, ceux-là ne le sont pas non plus ; donc $NON\,[I \perp\!\!\!\perp D \mid C]$.
Pour l'ordre d'énumération (D, I, C), on obtient donc le graphe d'indépendance :

$$\boxed{I} \longrightarrow \boxed{C} \longleftarrow \boxed{D}$$

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

Probabilité pour que la rupture de la pièce soit due à ce qu'elle était défectueuse :
$P(D = O/C = O) = \dfrac{P(D = O, C = O)}{P(C = O)} = \dfrac{a}{a + \bar{a}.b} = \dfrac{a}{1 - \bar{a}.\bar{b}}$.

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.

Probabilité pour que la rupture de la pièce soit due à ce qu'elle était défectueuse :
$P(D = O/C = O, I = O) = \dfrac{P(D = O, C = O, I = O)}{P(C = O, I = O)} = \dfrac{ab}{a.b + \bar{a}.b} = a.$

Puisque $a < \dfrac{a}{1 - \bar{a}.\bar{b}}$, savoir que l'entretien a été insuffisant diminue la probabilité que la panne puisse être attribuée à un défaut de la pièce.

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).

$$100 \times [p(q,e)] = 100 \times [p(q,e,h) + p(q,e,b)] = \begin{array}{c|c} Q \setminus E & \begin{array}{cc} h & b \end{array} \\ \begin{array}{c} h \\ b \end{array} & \begin{bmatrix} 40 & 10 \\ 10 & 40 \end{bmatrix} \end{array} ;$$

les lignes (et colonnes) n'étant pas proportionnelles, $NON\,[E \perp\!\!\!\perp Q]$.

$$100 \times [p(h,e,t)] = \begin{array}{c|c} T \setminus E & \begin{array}{cc} h & b \end{array} \\ \begin{array}{c} h \\ b \end{array} & \begin{bmatrix} 36 & 1 \\ 4 & 9 \end{bmatrix} \end{array}$$

n'a pas ses lignes (et colonnes) pas proportionnelles, donc $[p(e,t/h)]$ non plus : $NON\,[T \perp\!\!\!\perp E \mid Q]$.

De même, $100 \times [p(q,h,t)] = \begin{array}{c|cc} Q \setminus T & h & b \\ \hline h & 36 & 4 \\ b & 1 & 9 \end{array}$ n'a pas ses lignes (et colonnes) pas proportionnelles, donc $[p(q,t/h)]$ non plus : $NON\,[T \perp\!\!\!\perp Q \mid E]$.

Le graphe du RB associé aux trois variables pour l'ordre d'énumération (Q,E,T) est donc complet.

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 ?

Probabilité qu'une personne ait un rendement élevé dans son travail

- sachant qu'elle a un Q.I. élevé :

$$P(T=h/Q=h) = \frac{P(T=h,Q=h)}{P(Q=h)} = \frac{37}{50} = 0,74.$$

- sachant qu'elle a un niveau d'études élevé

$$P(T=h/E=h) = \frac{P(T=h,E=h)}{P(E=h)} = \frac{37}{50} = 0,74.$$

- sachant qu'elle a à la fois un Q.I. et un niveau d'études élevés

$$P(T=h/Q=h,E=h) = \frac{P(T=h,Q=h,E=h)}{P(Q=h,E=h)} = \frac{36}{40} = 0,90.$$

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 $p(.\,||\,\hat{x})$, 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 $[p(q,e,t\,||\,\hat{q})] = [p(q,e,t\ /\ \hat{q})]$.

Pour l'intervention $Q=\hat{q}=h$, le graphe ne change pas et mettre en Q la loi certaine en h, équivaut à introduire dans le RB initial l'information $Q=h$; donc $[p(q,e,t\,||\,\hat{q})] = [p(q,e,t\ /\ \hat{q})]$.

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

$$p(q,e,t\,||\,\hat{e}) = \begin{cases} p(q).p(t/q,e) & \text{si } e=\hat{e}(=h) \\ 0 & \text{si } e\neq\hat{e}(=h) \end{cases}$$

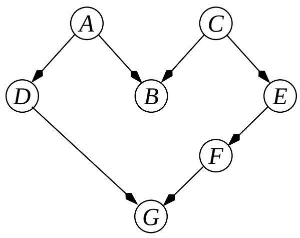
alors que :

$$p(q,e,t/\hat{e}) = \begin{cases} \dfrac{p(q).p(e/q)p(t/q,e)}{p(e)} & \text{si } e=\hat{e}(=h) \\ 0 & \text{si } e\neq\hat{e}(=h) \end{cases}$$

comme $NON\,[E \perp\!\!\!\perp Q]$, $p(\hat{e}/q) \neq p(\hat{e})$, pour $Q = h$ ou $Q = b$ (en fait pour les deux) et donc $[p(q, e, t \perp\!\!\!\perp \hat{e})] \neq [p(q, e, t/\hat{e})]$.
$[p(t)] = \left[\dfrac{42}{100}, \dfrac{58}{100}\right]$
$[p(t/\hat{e})] = \left[\dfrac{74}{100}, \dfrac{26}{100}\right]$
et $[p(t \perp\!\!\!\perp \hat{e})] = \left[\dfrac{50}{100}, \dfrac{50}{100}\right]$.

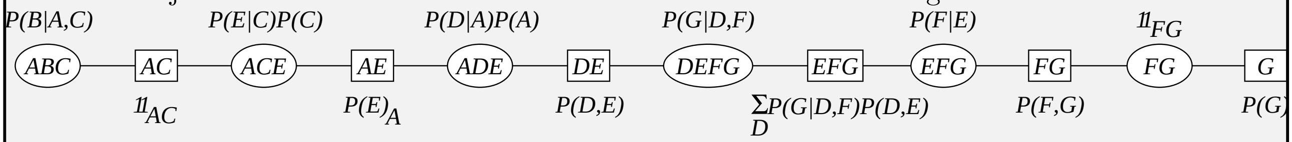
Exercice 4

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 $P(X)_Y$).

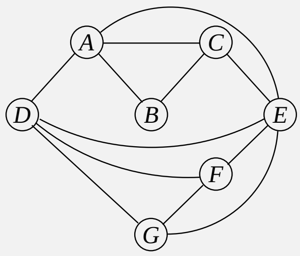
L'arbre de jonction suivant est obtenu en éliminant les variables de la gauche vers la droite :



Notons que dans le séparateur EFG, l'expression $\sum_D P(G|D, F)P(D, E)$ ne peut être réduite. En effet, le produit $P(G|D, F)P(D, E)$ est égal à $P(G|D, E, F)P(D, E)$ mais pas à $P(G|D, E, F)P(D, E|F)$ car E et F sont des variables dépendantes. En revanche, lorsque cette somme est multipliée par $P(F|E)$, on obtient $P(F|E) \sum_D P(G|D, F)P(D, E) = \sum_D P(G|D, F)P(D, E)P(F|E) = \sum_D P(G|D, E, F)P(D, E)P(F|D, E) = \sum_D P(G|D, E, F)P(D, E, F) = \sum_D P(G, D, E, F) = P(G, E, F)$, d'où le $P(F, G)$ contenu dans le séparateur FG.

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

Pour obtenir le graphe triangulé, il suffit de désorienter les arcs du réseau bayésien d'origine et de rajouter les arêtes qui n'existeraient pas déjà entre toutes les variables des cliques de l'arbre de jonction. On obtient donc le graphe triangulé suivant :

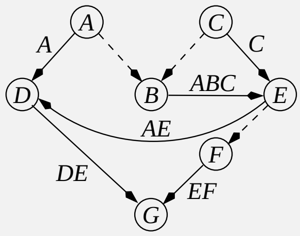


Q 4.3 Si chacune des variables est booléenne, quel est le nombre de multiplications et d'additions pour calculer $P(G)$ avec la séquence d'élimination ci-dessus ?

Reprenons le joint tree obtenu en 1°). Aucune multiplication n'est à effectuer dans la clique de gauche. Dans la clique suivante, on commence par multiplier $P(E|C)$ par $P(C)$, soit 4 multiplications, puis on expanse la matrice obtenue de manière à ce qu'elle ait la dimension ACE, puis on en fait le produit avec $_{AC}$, soit 8 multiplications. De la même manière, dans la clique suivante, il y a aussi 12 multiplications. Dans les cliques restantes, il n'y a qu'un seul produit effectué (celui de la matrice stockée dans la clique avec celle du séparateur à sa gauche). En tout, il y a donc $12 + 12 + 8 + 8 + 4 = 44$ multiplications. Les additions correspondent aux projections des cliques vers les séparateurs qui sont à leur droite. Il y a donc $8 + 8 + 8 + 16 + 8 + 4 = 52$ additions.

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.

Pour effectuer les mêmes calculs que Shafer-Shenoy, Pearl aurait utilisé le graphe ci-dessous :



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

D et F sont bien d-séparés, en effet, dans le réseau bayésien d'origine, il y a deux chaînes reliant ces deux nœuds : $DABCEF$ et DGF. Or ces chaînes ne sont pas d-connectantes car les arcs adjacents à B et à G sont convergents.

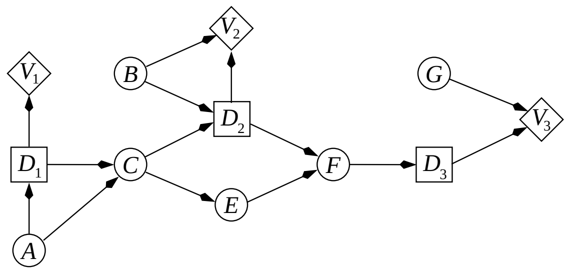
Q 4.6 Si l'on ne tient pas compte de la séquence d'élimination ci-dessus, Pearl a-t-il absolument besoin d'effectuer des coupes-cycles pour calculer $P(G)$? Justifiez votre réponse.

Pearl n'a pas besoin d'effectuer de coupe-cycles. En effet, ces derniers sont nécessaires uniquement dansle cas où les messages arrivant sur les puits des cycles sont dépendants. Or le message de D vers G et celui de F vers G sont forcément indépendants puisque D et F sont d-séparés. Autrement dit, les calculs effectués par un Pearl « amélioré » devraient être :

1. A envoie le message $P(A)$ à D et B. De même, C envoie le message $P(C)$ à B et E.
2. B calcule $P(B) = \sum_A \sum_C P(B|A,C)P(A)P(C)$.
3. D calcule $P(D) = \sum_A P(D|A)P(A)$ et envoie ce message à G.
4. E calcule $P(E) = \sum_C P(E|C)P(C)$ et envoie ce message à F.
5. F calcule $P(F) = \sum_E P(F|E)P(E)$ et envoie ce message à G.
6. G calcule $P(G) = \sum_D \sum_F P(G|D,F)P(D)P(F)$.

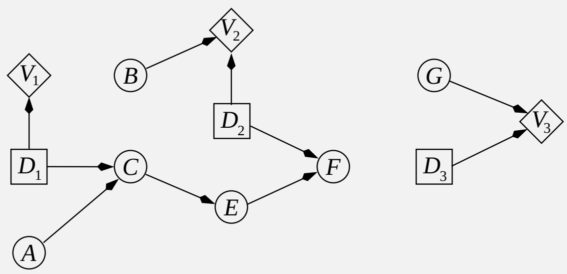
Exercice 5

Soit le diagramme d'influence suivant :



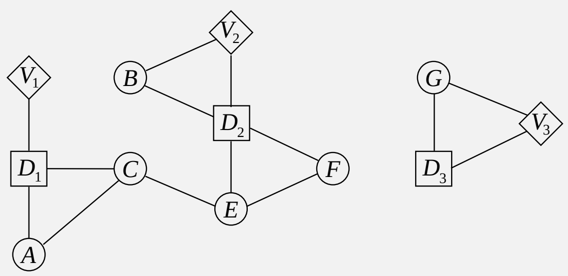
quest Tracez le réseau de valuation correspondant.

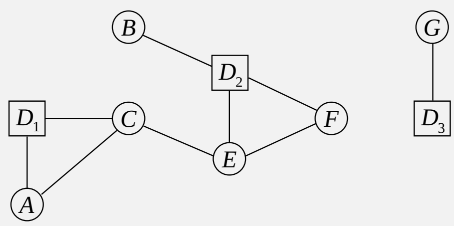
L'obtention du réseau de valuation consiste à éliminér les arcs entrant dans les nœuds de décision. D'où le graphe suivant :



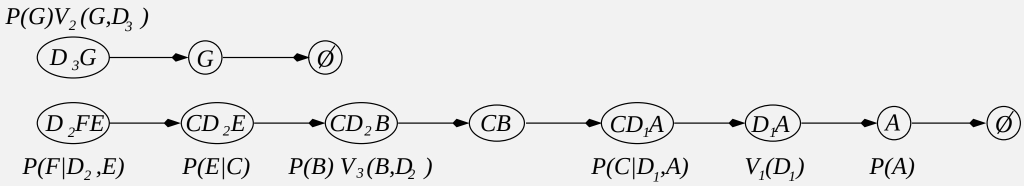
Q 5.1 En supposant que l'on ait la contrainte temporelle $A \prec D_1 \prec \{B, C\} \prec D_2 \prec \{E, F, G\} \prec D_3$, triangulez le réseau du 1°) et construisez le strong junction tree correspondant. Vous noterez à côté des cliques les tables que vous stockerez dans celles-ci.

On commence par moraliser le réseau (figure de gauche) puis on élimine les noeuds d'utilité (figure de droite) :



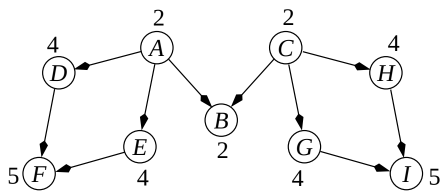


L'élimination de D_3 crée la clique D_3G, puis l'élimination de G crée la clique G. L'élimination de F crée une clique D_2FE. L'élimination de E crée la clique CD_2E, celle de D_2 la clique CD_2B, celle de B la clique BC. Enfin les éliminations suiccessives de C, D_1 et A créent successivement les cliques ACD_1, AD_1 et A. On obtient donc *in fine* l'arbre de jonction suivant :



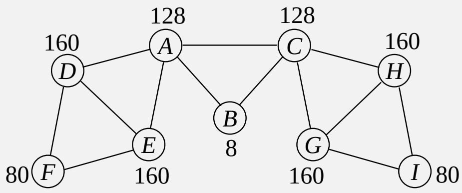
Exercice 6

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

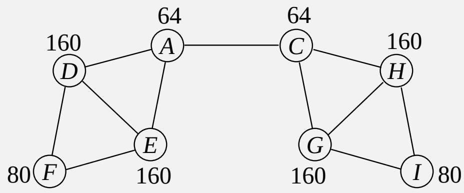


Q 6.1 Triangulez ce réseau en utilisant la méthode de Kjærulff (la méthode vue en cours) et déduisez-en un arbre de jonction (vous dessinerez le graphe à chacune des étapes de l'algorithme et noterez les poids de Kjærulff à côté de chaque nœud).

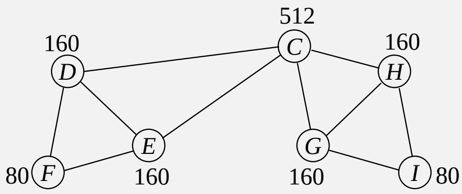
La méthode de Kjærulff conduit à l'élimination suivante :



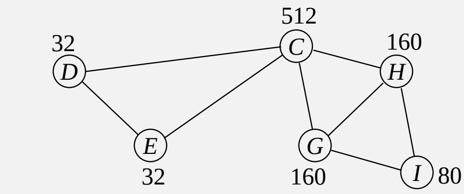
après moralisation



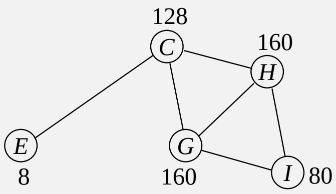
après suppression de B



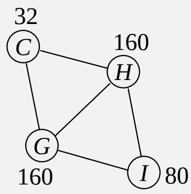
après suppression de A



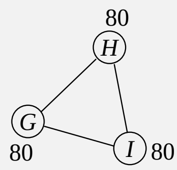
après suppression de F



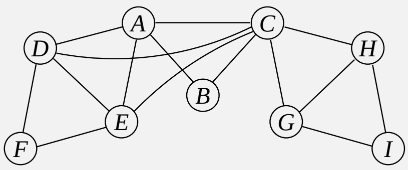
après suppression de D



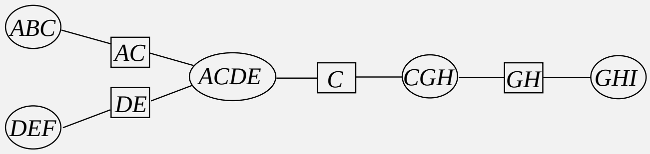
après suppression de E



après suppression de C



graphe triangulé final



un arbre de jonction possible

Q 6.2 Quelle est la triangulation optimale ?

On peut remarquer que le graphe moral est déjà triangulé. C'est donc la triangulation optimale.