

Examen du module MGDE

Durée : 3 heures

Exercice 1

On considère le réseau bayésien suivant :

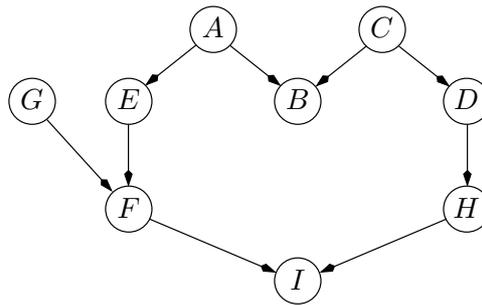


FIG. 1 – Un réseau bayésien avec cycle.

- Q 1.1** Donnez la décomposition de la loi jointe $P(A, B, C, D, E, F, G, H, I)$.
- Q 1.2** Dessinez l'arbre de jonction qu'obtiendrait l'algorithme de Shafer-Shenoy en utilisant la séquence d'élimination $A, D, B, C, I, G, E, F, H$. Vous préciserez à côté de chacune des cliques les matrices (probabilités) stockées par ces dernières.
- Q 1.3** Les variables F et H sont-elles indépendantes ? Justifiez votre réponse.
- Q 1.4** Lors de la phase d'initialisation (calcul de toutes les probabilités a priori), peut-on exécuter l'algorithme de Pearl sans effectuer de coupe-cycle ? Justifiez votre réponse. Le cas échéant, indiquez les calculs que réaliserait cet algorithme sans coupe, sinon indiquez quelle coupe vous feriez puis indiquez les calculs de Pearl avec cette coupe.
- Q 1.5** Les variables F et H sont-elles indépendantes conditionnellement à une information e_B sur B ? Justifiez votre réponse.
- Q 1.6** Décrivez les calculs que ferait Pearl pour obtenir $P(I|e_B)$. Si vous faites une ou plusieurs coupes, indiquez la(les)quelle(s).

Exercice 2

Le but de cet exercice est de créer un diagramme d'influence permettant d'aider le service du personnel d'une entreprise.

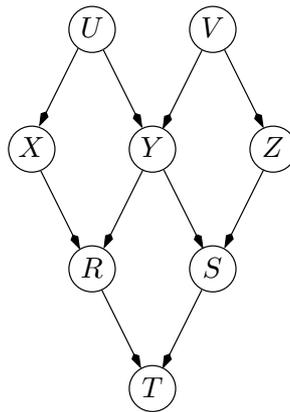
Le service reçoit des CV de la part de candidats. Le niveau réel de ces derniers est séparé en 4 catégories : excellent, bon, moyen, mauvais. Ce niveau n'est pas connu du service du personnel mais il a une influence non négligeable sur la qualité des CV reçus. Lors des lectures des CV, le service répartit ces derniers en 3 catégories A, B, C. Selon l'évaluation obtenue, on décide ou non de procéder à un entretien du candidat. Cet entretien a un coût de 50 €. À l'issue de chaque entretien, l'interviewer écrit un rapport qui contient ses impressions de concernant le candidat (bon/moyen/mauvais). Ce rapport est ensuite transmis au chef du personnel qui décide ou non de recruter le candidat. Cette décision a le coût suivant pour l'entreprise :

décision	recrutement				pas de recrutement			
niveau réel	excellent	bon	moyen	mauvais	excellent	bon	moyen	mauvais
coût	-1000 €	-700 €	-200 €	700 €	500 €	200 €	0 €	0 €

Q 2.1 Dessinez le diagramme d'influence représentant le problème de décision ci-dessus.

Exercice 3

La loi jointe $p(u, v, x, y, z, r, s, t)$ des 8 v.a. (U, V, X, Y, Z, R, S, T) admet le graphe d'indépendance suivant :



Q 3.1 Comment $p(u, v, x, y, z, r, s, t)$ se décompose-t-elle ?

Q 3.2 Qu'est-ce que le critère de d-séparation permet d'affirmer concernant les propriétés suivantes :

$$\begin{array}{ccccccc}
 X \perp\!\!\!\perp Z? & X \perp\!\!\!\perp Z \mid Y? & X \perp\!\!\!\perp Z \mid R? & X \perp\!\!\!\perp Z \mid (R, S)? \\
 X \perp\!\!\!\perp Z \mid (R, V)? & X \perp\!\!\!\perp Z \mid (R, V, S)? & X \perp\!\!\!\perp Z \mid (R, Y, V)? &
 \end{array}$$

(lorsque la réponse est non, on indiquera une chaîne active)

Q 3.3 L'ordre d'énumération sur les variables utilisé a été $UVXYZRST$; quelles sont les propriétés d'indépendances conditionnelles qui ont servi à construire le graphe ?

Q 3.4 À partir de ces dernières propriétés et des propriétés générales de la relation $\cdot \perp\!\!\!\perp \cdot \mid \cdot$, montrez successivement que :

$$Z \perp\!\!\!\perp X \mid (V, Y), \quad R \perp\!\!\!\perp Z \mid (X, V, Y) \quad \text{et} \quad X \perp\!\!\!\perp Z \mid (R, Y, V)$$

Q 3.5 Quel est le graphe moral dérivé du graphe d'indépendance donné ? Que peut-on dire du cycle $XUVZSR$? Quel est le graphe moral triangulé ? Quelles sont ses cliques ?

Q 3.6 Énumérez les sommets puis les cliques en appliquant l'algorithme de recherche de la cardinalité maximum et construisez un arbre de jonction.

[On notera $C_i, i = 1, \dots, 5$ les cliques et $S_l, l = 1, \dots, 4$, les séparateurs de cliques].

Q 3.7 Utilisant la décomposition de la question 1, proposez un système de potentiels de cliques et de séparateurs de cliques ; montrez, selon le cas, que ce système est cohérent (c'est-à-dire en équilibre selon Jensen) ou qu'il ne l'est pas.

Exercice 4

Soit la décomposition additive généralisée suivante :

$$u(A, B, C, D, E, F, G) = u_1(A, B, C) + u_2(B, D, F) + u_3(C, E) + u_4(B, C, D) + u_5(F, G).$$

Q 4.1 Dessinez un réseau GAI correspondant.

Q 4.2 On veut savoir quel est le septuplet maximisant la fonction d'utilité, autrement dit on cherche les valeurs de A , B , etc, maximisant :

$$\max_{A,B,C,D,E,F,G} u(A, B, C, D, E, F, G).$$

Bien entendu, le max sur tous les attributs est égal au max sur un premier attribut du max sur un deuxième attribut du max sur un troisième attribut, etc, de la fonction d'utilité puisque celle-ci est définie sur le produit cartésien des attributs. Autrement dit, obtenir le max sur tous les attributs revient à appliquer une séquence d'élimination (en utilisant l'opérateur max) sur l'utilité jointe $u(A, B, C, D, E, F, G)$. On voit donc qu'il existe un parallèle avec le calcul des probabilités dans un réseau bayésien/arbre de jonction.

En utilisant ce parallèle, indiquez une séquence d'élimination efficace pour calculer le max sur tous les attributs et précisez les calculs que vous feriez à chaque élimination.

Q 4.3 Déduisez-en un algorithme permettant de calculer une instantiation des attributs maximisant une utilité décomposable selon un arbre GAI.