

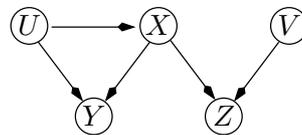
# rattrapage du module MGDE

Durée : 2 heures

*Seuls documents autorisés :  
Aucun document*

## Exercice 1 (5 points)

La loi jointe  $p(u, x, v, y, z)$  des 5 v.a.  $(U, X, V, Y, Z)$  admet le graphe d'indépendance  $G$  de la figure ci-dessous :



**Q 1.1** L'ordre d'énumération sur les variables utilisé a été  $UXVYZ$  ;

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

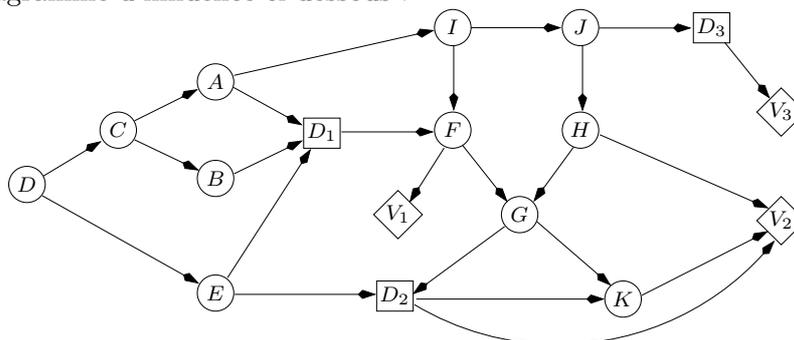
Quelles sont les relations d'indépendance et d'indépendance conditionnelle qui ont servi à construire le graphe ?

**Q 1.2** Qu'est-ce que le critère de d-séparation permet d'affirmer concernant les propriétés suivantes :

- $V \perp\!\!\!\perp (X, Y)$  ?     $U \perp\!\!\!\perp (V, X)$  ?     $U \perp\!\!\!\perp V \mid (Y, Z)$  ?
- $U \perp\!\!\!\perp V \mid (X, Y)$  ?     $U \perp\!\!\!\perp V \mid (X, Y, Z)$  ?

## Exercice 2 (5 points)

On considère le diagramme d'influence ci-dessous :

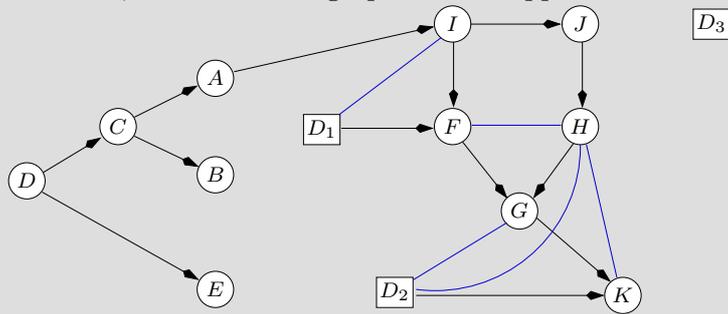


**Q 2.1** Déterminez l'ordre partiel temporel  $\prec$  de ce diagramme d'influence, c'est-à-dire que vous indiquerez les ensembles  $D_k$  et  $I_k$  comme vu en cours.

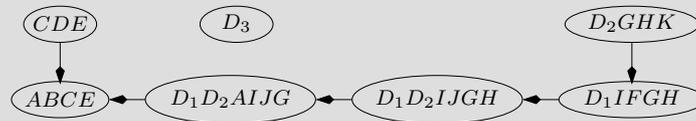
$\{A, B, E\} \prec D_1 \prec \{G\} \prec D_2 \prec \{J\} \prec D_3 \prec \{C, D, I, F, H, K\}$

**Q 2.2** Déterminez un strong junction tree pour ce diagramme d'influence. Vous indiquerez la séquence d'élimination que vous aurez utilisée.

On commence par supprimer les arcs entrant dans les nœuds de décision (de manière à obtenir un réseau de valuation). Ensuite, on moralise le graphe et on supprime les nœuds d'utilité :

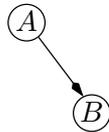


Ensuite, on peut supprimer les variables selon une séquence compatible avec la question 1. Par exemple, la séquence d'élimination  $D, C, K, F, H, I, D_3, J, D_2, G, D_1, A, B, E$  produit le strong junction tree suivant :



**Exercice 3 (5 points)**

Soit le réseau bayésien et la base de données (de taille 8) :



A	1	0	0	1	?	1	0	0
B	0	0	?	1	1	?	1	0

où les '?' représentent des valeurs manquantes.

**Q 3.1** Quels sont les paramètres du réseau bayésien ?

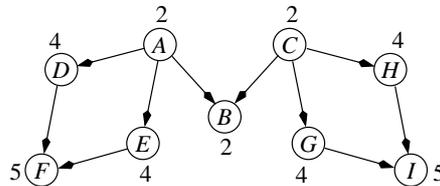
**Q 3.2** Calculer les paramètres de ce réseau bayésien pour la base de données obtenues en retirant les lignes où il existe au moins une valeur manquante.

**Q 3.3** Calculer, avec ces paramètres, la distribution de  $P(A \mid B = 1)$ . Était-il possible de calculer cette distribution directement dans la base ?

**Q 3.4** Appliquer l'algorithme EM pour apprendre les paramètres de ce réseau avec la base entière (faire au plus 3 itérations de l'algorithme si la convergence n'est pas atteinte.).

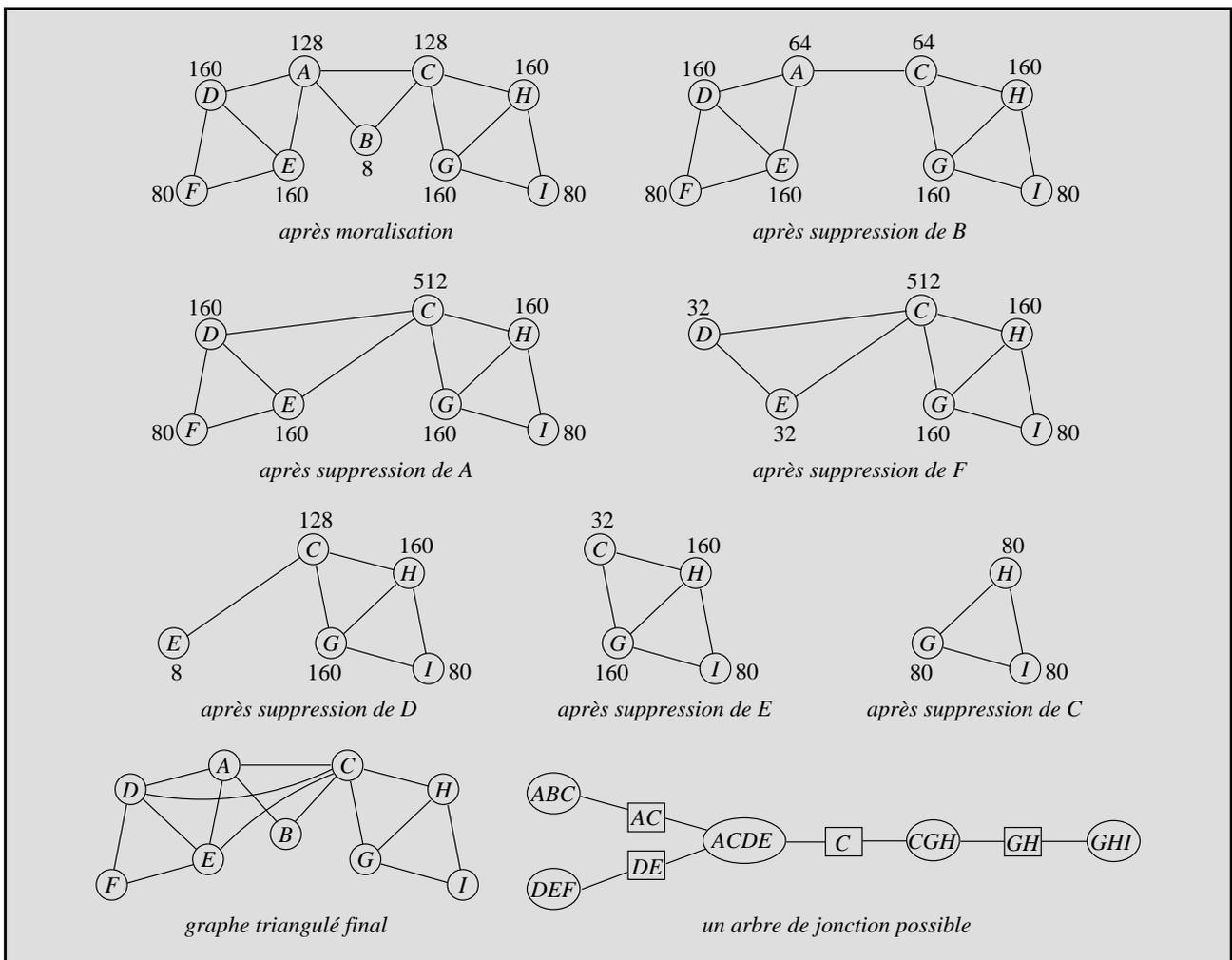
**Exercice 4 (5 points)**

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



**Q 4.1** 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 :



Q 4.2 Quelle est la triangulation optimale ?

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