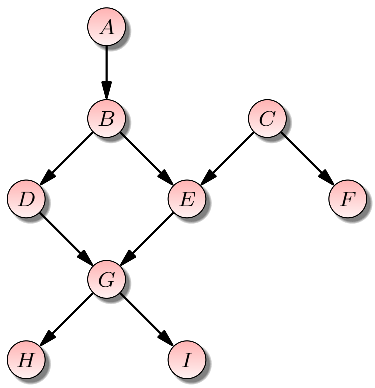
Examen du module de raisonnement dans l'incertain

Durée : 1 heure 30

Documents autorisés : 1 feuille A4 recto-verso

Exercice 1 (4 points) — *Séparations*

On considère le réseau bayésien suivant, de structure $\mathcal{G}$:



Qu'est-ce que le critère de *d*-séparation permet d'affirmer concernant les propriétés ci-dessous ? Vous justifierez votre réponse.

Q 1.1 $\langle A \perp_{\mathcal{G}} F \rangle$?

Q 1.2 $\langle A \perp_{\mathcal{G}} F | E \rangle$?

Q 1.3 $\langle A \perp_{\mathcal{G}} F | I \rangle$?

Q 1.4 $\langle D \perp_{\mathcal{G}} F | B, G \rangle$?

Exercice 2 (8 points) — *Inférence*

Q 2.1 En appliquant l'algorithme de Kjaerülff, quel arbre d'élimination obtient-on pour le réseau bayésien de l'exercice 1 ? Si plusieurs variables ont le même poids, on éliminera en premier la première dans l'ordre alphabétique.

Q 2.2 Déduisez-en un arbre de jonction. Indiquez à côté des cliques les distributions de probabilité conditionnelles que vous stockez dans celles-ci.

Q 2.3 On suppose que toutes les variables sont booléennes. Combien de multiplications et d'additions (sur des nombres réels) doit-on réaliser pour calculer, selon l'algorithme de Shafer-Shenoy, les messages dans les deux sens de tous les séparateurs ? On supposera, que, quand on doit combiner une clique avec plusieurs séparateurs, on combine d'abord les séparateurs puis on combine le résultat avec la clique.

Exercice 3 (8 points) — *Apprentissage*

Soit trois variables aléatoires booléennes A, B, C dont on a observé les occurrences ci-dessous :

A	B	C
a_1	b_1	c_1
a_2	b_2	c_1
a_2	b_1	c_2
a_2	b_2	c_2
a_2	b_2	c_1
a_1	b_1	c_1
a_2	b_2	c_1
a_2	b_1	c_1
a_1	b_2	c_2

On suppose ici que toute probabilité conditionnelle portant sur les variables A, B, C se déduit par normalisation des fréquences observées dans la base, c'est-à-dire qu'en multipliant ces fréquences par une constante de telle sorte qu'elles somment à 1, on obtient des probabilités. Par exemple, on observe 3 instances de a_1 et 6 instances de a_2. Par conséquent, $P(A = a_1) = 3 \times k$ et $P(A = a_2) = 6 \times k$, avec la constante $k = 1/9$ afin d'obtenir $P(A = a_1) + P(A = a_2) = 1$.

Q 3.1 En utilisant des tests d'indépenance conditionnels fondés sur les probabilités, appliquez l'algorithme PC pour apprendre le squelette du réseau bayésien ayant généré cette base.

Q 3.2 Appliquez les règles R1, R2, R3 de PC. Quel CPDAG obtient-on ?

Q 3.3 Déterminez un réseau bayésien compatible avec le CPDAG trouvé dans la question précédente. Estimez par maximum de vraisemblance les tables de probabilité conditionnelles des nœuds du réseau.