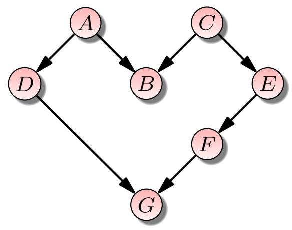
Master SID

Raisonnement dans l'incertain

TD n°6 : arbres de jonction

Exercice 1 – Le cœur du problème

Considérons le réseau bayésien suivant :



Q 1.1 Tracez l'arbre d'élimination 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 1.2 Tracez l'arbre de jonction correspondant à l'abre d'élimination que vous avez trouvé dans la question précédente.

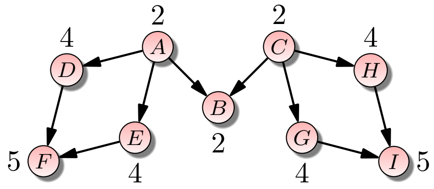
Q 1.3 Quel est le graphe triangulé correspondant à cette séquence d'élimination ?

Q 1.4 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 1.5 D et F sont-ils d-séparés ? Justifiez votre réponse.

Exercice 2 – Scie métrique

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

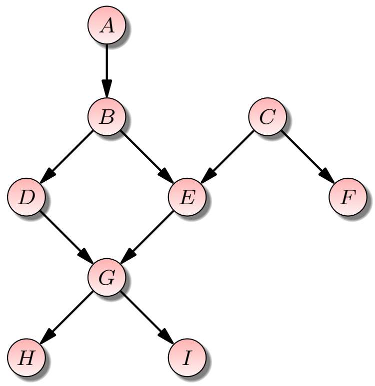


Q 2.1 Triangulez ce réseau en utilisant la méthode de Kjærulff (la méthode vue en cours) et déduisez-en un arbre de jonction (vous dessinerez le graphe à chacune des étapes de l'algorithme et noterez les poids de Kjærulff à côté de chaque nœud).

Q 2.2 Quelle est la triangulation optimale ?

Exercice 3 – arbre de jonction

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

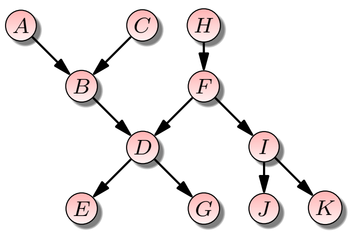


Q 3.1 En appliquant l'algorithme de Kjaerülff, quel arbre d'élimination obtient-on pour ce réseau bayésien ? Si plusieurs variables ont le même poids, on éliminera en premier la première dans l'ordre alphabétique.

Q 3.2 Déduisez-en un arbre de jonction. Indiquez à côté des cliques les distributions de probabilité conditionnelles que vous stockez dans celles-ci.

Exercice 4 – Démonstrations

Soit le réseau bayésien ci-dessous :



Q 4.1 Moralisez ce réseau.

Q 4.2 Triangulez le graphe moral en utilisant la séquence d'élimination suivante : $A, K, I, J, G, H, C, D, B, F, E$. Vous indiquerez pour chaque nœud éliminé la clique qui en résulte. Enfin, vous tracerez l'arbre d'élimination.

Q 4.3 Dessinez un arbre de jonction correspondant à cette séquence d'élimination et indiquez à côté des cliques les probabilités conditionnelles que vous stockerez dans ces cliques.

Q 4.4 Indiquez les contenus des messages transitant dans les deux sens des arêtes sur chaque séparateur pour le calcul des probabilités *a priori* par l'algorithme de Shafer-Shenoy.

Q 4.5 Montrer que, quel que soit l'arbre de jonction et quelles que soient deux cliques voisines C_i et C_j d'intersection S_{ij}, les variables de $C_i \backslash S_{ij}$ sont indépendantes de $C_j \backslash S_{ij}$ conditionnellement à S_{ij}.

S_{ij} sépare l'arbre de jonction en deux sous-arbres T_1 et T_2. Montrez que les variables de $T_1 \backslash S_{ij}$ sont indépendantes de $T_2 \backslash S_{ij}$ conditionnellement à S_{ij}.

Q 4.6 En utilisant la *d*-séparation, montrez que la propagation d'une information e_A concernant A dans l'arbre de jonction obtenu dans la question Q 4.4 ne nécessite le calcul que de deux nouveaux messages.

Exercice 5 – Démonstrations bis

Dans cet exercice, on considère des graphes non orientés $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, où $\mathcal{V} = \{X_1, \dots, X_n\}$ et $\mathcal{E} \subseteq \{(X_i, X_j) : i, j \in \{1, \dots, n\}, i \neq j\}$. $\sigma : \{1, \dots, n\} \mapsto \{1, \dots, n\}$ représente une permutation.

Q 5.1 Soit l'algorithme d'élimination :

Fonction Elim $(\mathcal{G}, \sigma)$

$\mathcal{E}' \longleftarrow \mathcal{E}$
pour i variant de 1 à n **faire**
    pour X_j, X_k adjacents à $X_{\sigma(i)}$ dans $\mathcal{G}$ **faire**
        si $(X_j, X_k) \notin \mathcal{E}$ **alors**
            $\mathcal{E} \longleftarrow \mathcal{E} \cup \{(X_j, X_k)\}$; $\mathcal{E}' \longleftarrow \mathcal{E}' \cup \{(X_j, X_k)\}$
        finsi
    fait
    supprimer toutes les arêtes adjacentes à $X_{\sigma(i)}$ dans $\mathcal{E}$
fait
renvoyer le graphe $\mathcal{G}' = (\mathcal{V}, \mathcal{E}')$

On appelle cycle de $\mathcal{G}'$ un ensemble de nœuds $\{X_{i_1}, \dots, X_{i_p}\}$ tel que :

1. pour tout $k \in \{1, \dots, p-1\}$, $(X_{i_k}, X_{i_{k+1}}) \in \mathcal{E}'$; $(X_{i_p}, X_{i_1}) \in \mathcal{E}'$;
2. il n'existe pas $j, k \in \{1, \dots, p\}$, $j \neq k$, tels que $X_{i_j} = X_{i_k}$.

La longueur d'un cycle est le nombre d'arêtes qui le constitue, autrement dit, la longueur du cycle $\{X_{i_1}, \dots, X_{i_p}\}$ est $p + 1$. Montrez que le graphe $\mathcal{G}'$ renvoyé par la fonction **Elim** est triangulé, c'est-à-dire que, dans tout cycle de longueur 4 ou plus, il existe deux nœuds X, Y non voisins dans le cycle tels que $(X, Y) \in \mathcal{E}'$.

Q 5.2 On appelle graphe markovien associé à un arbre de jonction $\mathcal{J}$ le graphe $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ tel que $\mathcal{V}$ est l'ensemble des variables contenues dans les cliques de $\mathcal{J}$ et $\mathcal{E}$ est l'ensemble des arêtes (X_i, X_j) telles qu'il existe une clique dans $\mathcal{J}$ contenant X_i et X_j. Montrez que tout graphe markovien associé à un arbre de jonction est triangulé.

Q 5.3 Considérons l'algorithme d'élimination orienté suivant :

Fonction DirectElim ($\mathcal{G} = (\mathcal{V}, \mathcal{E})$, σ)

$\mathcal{A} \longleftarrow \emptyset$
pour i variant de 1 à n **faire**
pour tous X_k adjacents à $X_{\sigma(i)}$ dans $\mathcal{G}$ **faire**
$\mathcal{A} \longleftarrow \mathcal{A} \cup \{X_{\sigma(i)} \rightarrow X_k\}$
fait
pour tous X_j, X_k adjacents à $X_{\sigma(i)}$ dans $\mathcal{G}$ **faire**
si $(X_j, X_k) \notin \mathcal{E}$ **alors**
$\mathcal{E} \longleftarrow \mathcal{E} \cup \{(X_j, X_k)\}$
finsi
fait
supprimer toutes les arêtes adajacentes à $X_{\sigma(i)}$ dans $\mathcal{E}$
fait
renvoyer le graphe orienté $\mathcal{G}' = (\mathcal{V}, \mathcal{A})$

Le graphe $\mathcal{G}' = (\mathcal{V}, \mathcal{A})$ renvoyé par **DirectElim** est un graphe orienté tel que, si l'on enlève les orientations des arcs de $\mathcal{A}$, le graphe résultant est triangulé.

Q 5.3.1 Dessinez le graphe $\mathcal{G}'$ obtenu à l'issue de l'application de **DirectElim** sur le graphe de la figure 1 et la séquence d'élimination $\sigma = \{A, C, B, E, D, F, G\}$:

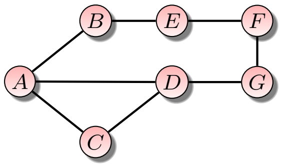


FIGURE 1 – Un graphe non orienté à trianguler.

Q 5.3.2 Montrer que, pour tout arc $X \rightarrow Y \in \mathcal{A}$, X est éliminé avant Y dans l'algorithme **DirectElim**.

Q 5.3.3 Dans le graphe $\mathcal{G}'$ de la question 5.3.1, on veut supprimer l'arc $F \rightarrow G$ de $\mathcal{A}$ de telle sorte que le graphe ainsi obtenu reste triangulé. Pour cela, à tout arc $Y \rightarrow Z$ de $\mathcal{A}$, on associe un nombre entier $x_{YZ} \in \{0, 1\}$ ayant pour signification $x_{YZ} = 1 \Longleftrightarrow$ l'arc $Y \rightarrow Z$ est supprimé. On impose donc que $x_{FG} = 1$. Soit le programme linéaire (Σ) défini par :

$$\begin{cases} x_{YZ} \leq x_{XY} + x_{XZ} \ \forall X, Y, Z \text{ tels que } X \rightarrow Y, X \rightarrow Z, Y \rightarrow Z \in \mathcal{A} \\ x_{FG} = 1 \\ x_{YZ} \in \{0, 1\} \ \forall Y \rightarrow Z \in \mathcal{A} \end{cases} \qquad (\Sigma)$$

Montrer que, pour toute solution du programme (Σ), le graphe $\mathcal{G}'' = (\mathcal{V}, \mathcal{A}'')$ constitué uniquement des arcs pour lesquels les x_{YZ} valent 0 est un graphe triangulé. Suggestion : démonstration par récurrence sur les nœuds éliminés, selon leur ordre d'élimination.