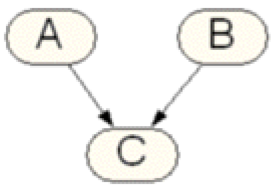
Maxime SCHMITT 1719088

INF8225 - Rapport TP1

Question 1 :

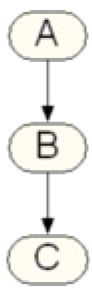
Explaining away : Phénomène traduisant la situation où, lorsqu'on fixe l'état d'un nœud, une variation de probabilité d'un nœud parent de celui-ci entrainera une variation inverse de probabilité pour les autres nœuds parents du nœud fixé. Cela se retrouve pour le cas d'un réseau bayésien tel que celui ci-contre.



Avec ces notations, si on prend l'exemple suivant :
- A : Probabilité d'être malade
- B : Probabilité d'un problème familial (accident d'un proche, etc.)
- C : Probabilité d'être absent en cours

Si on fixe C à « vrai » (l'élève est absent) alors, s'il s'avère que B est « vrai » (il y a eu un problème dans la famille de l'élève), la probabilité que A soit « vrai » (l'élève est malade) va diminuer puisque l'absence est déjà expliquée par le problème familial.
Le code pour cet exemple est dans le script nommé *explainingAway.m*.

Serial blocking : Phénomène traduisant la situation où, lorsqu'on fixe l'état d'un nœud au milieu d'une chaîne, les variations de probabilité des nœuds le précédant dans la chaîne n'ont plus d'effet sur celle des nœuds lui succédant. Avec les notations du réseau bayésien ci-contre :

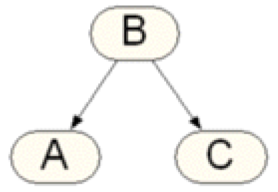


Si on fixe B on peut alors dire que B bloque la propagation des variations de A vers C. A et C sont alors conditionnellement indépendants à B fixé :
P(A,B|C) = P(A,B,C)/P(B) = P(A)P(B|A)P(C|B)/P(B) = P(A)[P(A|B)P(B)/P(A)]P(C|B)/P(B)
= P(A|B)P(C|B)

Avec ces mêmes notations, on prend l'exemple suivant :
- A : Probabilité qu'un étudiant étudie correctement le cours
- B : Probabilité qu'un étudiant réussisse l'examen final du cours
- C : Probabilité qu'un étudiant valide le cours

Si on fixe B à « vrai » (l'étudiant a réussi l'examen final) alors le fait de savoir que A est « vrai » (l'étudiant a étudié correctement le cours) n'aura pas d'influence sur la probabilité que l'étudiant valide le cours ou non. (Du moins pas au travers de B, on pourrait imaginer que A influe encore sur C par une autre variable D qui représenterait la probabilité de réussir les laboratoires qui serait influencée par l'étude ou non du cours).
Le code pour cet exemple est dans le script nommé *serialBlocking.m*.

<u>Divergent blocking :</u> Même phénomène que pour le serial blocking mais dans le cas où le nœud fixé n'est pas partie d'une chaîne mais plutôt d'une structure « en arbre » comme illustré dans le réseau bayésien illustré sur la figure ci-contre.



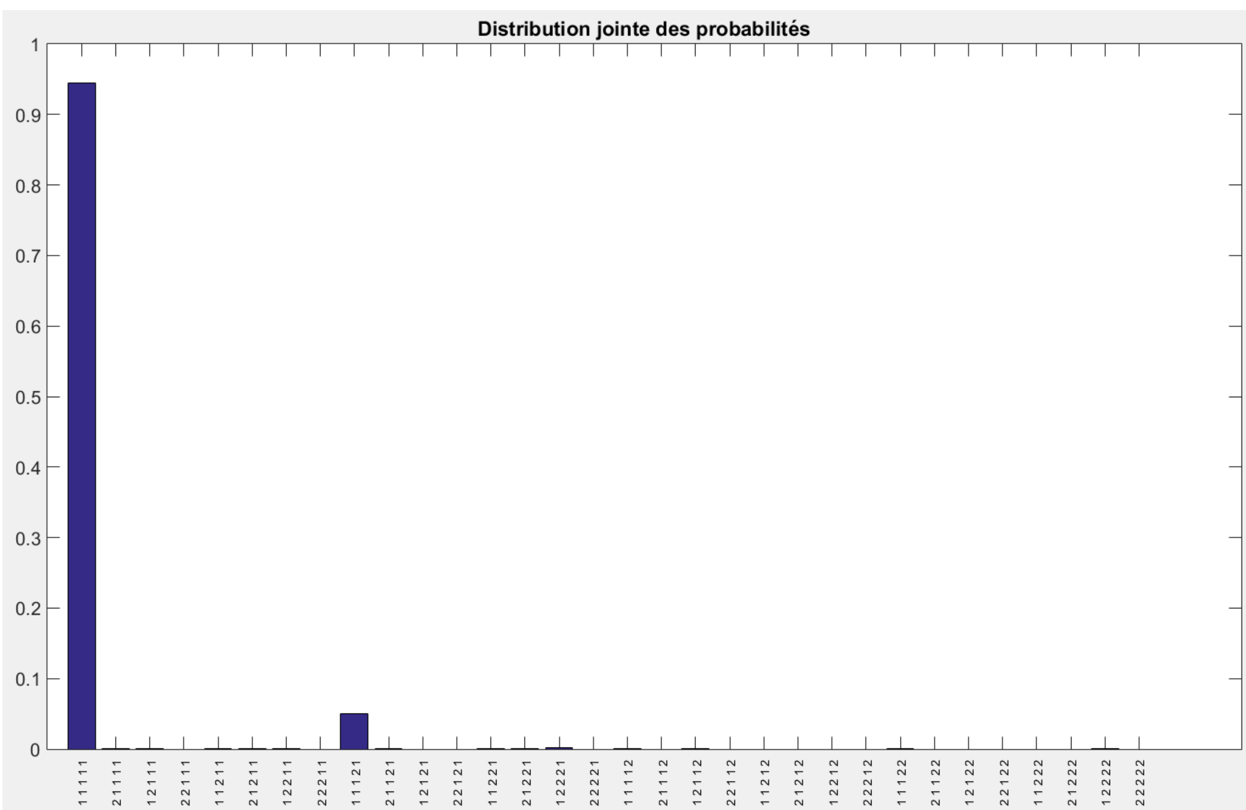
Si on fixe B on peut alors dire que B bloque la propagation des variations de A vers C. A et C sont alors conditionnellement indépendants à B fixé :
P(A,C|B) = P(A,B,C)/P(B) = P(B)P(A|B)P(C|B)/P(B) **= P(A|B)P(C|B)**

Avec les mêmes notations, on propose l'exemple suivant :
- B : Probabilité d'avoir mal à la tête
- A : Probabilité de se reposer
- C : Probabilité de prendre un médicament

Si on fixe B à « vrai » (la personne à mal à la tête) alors le fait de savoir que A est « vrai » (elle décide de se reposer) n'influence pas la probabilité qu'elle prenne également un médicament.
Le code pour cet exemple est dans le script nommé *divergentBlocking.m*.

Question 2 :

a) Voir fichier *mkDetecterCambriolageDgm.m*.
b) Voir fichier *mkDetecterCambriolageDgm.m* pour le code. Ci-dessous l'histogramme de probabilité jointe du modèle :



c) Résultats des marginales conditionnelles demandées (voir fichier *mkDetecterCambriolageDgm.m* pour le détail du code) :

P(Cambriolage = V| MarieAppelle = V, JeanAppelle = F) = 0.005656
P(Cambriolage = V| MarieAppelle = F, JeanAppelle = V) = 0.000749
P(Cambriolage = V| MarieAppelle = V, JeanAppelle = V) = 0.002770
P(Cambriolage = V| MarieAppelle = F, JeanAppelle = F) = 0.000741
P(Cambriolage = V| MarieAppelle = V) = 0.005648
P(Cambriolage = V| JeanAppelle = V) = 0.000992

d) Probabilité marginales inconditionnelles des variables du modèle :

P(Cambriolage = V) = 0.001000
P(Tremblement = V) = 0.002000
P(Alarme = V) = 0.003166
P(MarieAppelle = V) = 0.052691
P(JeanAppelle = V) = 0.001152

e) Équations demandées :

$$\boldsymbol{P(J)} = \sum_{C,T,A,M} P(C,T,A,M,J) = \sum_{\boldsymbol{C,T,A,M}} \boldsymbol{P(C)P(T)P(A|C,T)P(M|A)P(J|T,A)}$$

$$P(C|J=V) = \frac{P(C,J=V)}{P(J=V)} = \frac{\sum_{T,A,M} P(C,T,A,M,J=V)}{\sum_{C,T,A,M} P(C)P(T)P(A|C,T)P(M|A)P(J|T,A)}$$
$$= \frac{\sum_{T,A,M} P(C)P(T)P(A|C,T)P(M|A)P(J=V|T,A)}{\sum_{C,T,A,M} P(C)P(T)P(A|C,T)P(M|A)P(J|T,A)}$$

Soit le résultat suivant :

$$\boldsymbol{P(C|J=V)} = \frac{\sum_{\boldsymbol{T,A,M}} \boldsymbol{P(C)P(T)P(A|C,T)P(M|A)P(J=V|T,A)}}{\sum_{\boldsymbol{C,T,A,M}} \boldsymbol{P(C)P(T)P(A|C,T)P(M|A)P(J|T,A)}}$$

Question 3 :

Algorithme EM :

On a les statistiques suivantes :

A	M	J	#
*	0	0	N_1
*	0	1	N_2
*	1	0	N_3
*	1	1	N_4
			N

1) **Phase préliminaire :** On fixe aléatoirement les valeurs initiales des paramètres du modèle

P(A=1) = p_A
P(A=0) = 1 – p_A
P(M=1|A=1) = $p_{M,A=1}$
P(M=0|A=1) = 1 – $p_{M,A=1}$

P(M=1|A=0) = $p_{M,A=0}$
P(M=0|A=0) = 1 - $p_{M,A=0}$
P(J=1|A=1) = $p_{J,A=1}$
P(J=0|A=1) = 1 - $p_{J,A=1}$
P(J=1|A=0) = $p_{J,A=0}$
P(J=0|A=0) = 1 - $p_{J,A=0}$

2) **Phase E (Espérance)**

a. On calcule les postérieurs

P(A=0|M=0,J=0)	$\frac{(1-p_A)(1-p_{M,A=0})(1-p_{J,A=0})}{(1-p_A)(1-p_{M,A=0})(1-p_{J,A=0})+p_A.(1-p_{M,A=1})(1-p_{J,A=1})}$
P(A=0|M=0,J=1)	$\frac{(1-p_A)(1-p_{M,A=0}).p_{J,A=0}}{(1-p_A)(1-p_{M,A=0}).p_{J,A=0}+p_A.(1-p_{M,A=1}).p_{J,A=1}}$
P(A=0|M=1,J=0)	$\frac{(1-p_A).p_{M,A=0}.(1-p_{J,A=0})}{(1-p_A)(1-p_{M,A=0})(1-p_{J,A=0})+p_A.(1-p_{M,A=1})(1-p_{J,A=1})}$
P(A=0|M=1,J=1)	$\frac{(1-p_A).p_{M,A=0}.p_{J,A=0}}{(1-p_A).p_{M,A=0}.p_{J,A=0}+p_A.p_{M,A=1}.p_{J,A=1}}$

Exemple de développement pour P(A=0|M=0,J=0) :

$$\boldsymbol{P(A=0|M=0,J=0)} = \frac{P(A=0,M=0,J=0)}{P(M=0,J=0)} = \frac{P(A=0)P(M=0|A=0)P(J=0|A=0)}{\sum_A P(A,M=0,J=0)}$$
$$= \frac{P(A=0)P(M=0|A=0)P(J=0|A=0)}{\sum_A P(A)P(M=0|A)P(J=0|A)}$$
$$= \frac{\boldsymbol{(1-p_A)(1-p_{M,A=0})(1-p_{J,A=0})}}{\boldsymbol{(1-p_A)(1-p_{M,A=0})(1-p_{J,A=0})+p_A.(1-p_{M,A=1})(1-p_{J,A=1})}}$$

b. On calcule les statistiques équivalentes

A	M	J	#
0	0	0	N_1* P(A=0|M=0,J=0)
0	0	1	N_2* P(A=0|M=0,J=1)
0	1	0	N_3* P(A=0|M=1,J=0)
0	1	1	N_4* P(A=0|M=1,J=1)
1	0	0	N_1* (1 - P(A=0|M=0,J=0))
1	0	1	N_2* (1 - P(A=0|M=0,J=1))
1	1	0	N_3* (1 - P(A=0|M=1,J=0))
1	1	1	N_4* (1 - P(A=0|M=1,J=1))
			N

3) **Phase M (Maximisation)** :

a. On met à jour les probabilités conditionnelles et inconditionnelles du réseau

On montre le calcule pour P(A=0) comme exemple :

$$\boldsymbol{P(A=0)} = \sum_{M,J} P(A=0,M,J)$$
$$= \boldsymbol{N_1} * \mathbf{P(A=0|M=0,J=0)} + \boldsymbol{N_2} * \mathbf{P(A=0|M=0,J=1)} + \boldsymbol{N_3} * \mathbf{P(A=0|M=1,J=0)} + \boldsymbol{N_4} * \mathbf{P(A=0|M=1,J=1)}$$

b. <u>On mesure la vraisemblance (log vraisemblance)</u>

A	M	J	N	P	N*ln(P)
*	0	0	N_1	P_1	N_1*ln(P_1)
*	0	1	N_2	P_2	N_2*ln(P_2)
*	1	0	N_3	P_3	N_3*ln(P_3)
*	1	1	N_4	P_4	N_4*ln(P_4)
					LV

On trouve les P_i en compilant le réseau de Bayes. On montre les calculs pour P_1 à la première iteration comme exemple :

$$\boldsymbol{P_1} = P(M=0, J=0) = \sum_A P(A, M=0, J=0) = \sum_A P(A)P(M=0|A)P(J=0)|A)$$
$$= (\mathbf{1} - \boldsymbol{p_A})(\mathbf{1} - \boldsymbol{p_{M,A=0}})(\mathbf{1} - \boldsymbol{p_{J,A=0}}) + \boldsymbol{p_A} \cdot (\mathbf{1} - \boldsymbol{p_{M,A=1}})(\mathbf{1} - \boldsymbol{p_{J,A=1}})$$

Et LV est la somme des *N*ln(P)*.

c. <u>On compare le dernier LV calculé avec le précédent</u>

   i. Si la différence est inférieure à un seuil, on garde le modèle actuel
   ii. Si la différence est supérieure à ce seuil, on reprend à l'étape E
   (Lors de la première itération, on n'a pas de valeurs à laquelle comparer LV, on reprend donc à la phase E par défaut.)