Cours : IFT-4102-7025

TP2 comptant pour 20%(4102) et 15%(7025)

À remettre le vendredi 27 mars 2015 à 12h30mn via email

Pr. B. Chaib-draa

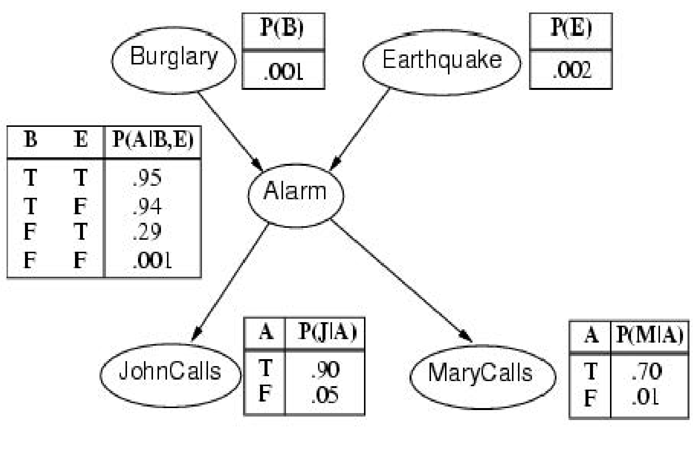
Question 1 : Soit à échanger le contenu de deux registres A et B. On dispose pour cela d'un opérateur $Assign(x, v, lv, ov)$ qui assigne la valeur v localisée en lv à x qui lui-même contenait auparavant ov.

a) En utilisant l'opérateur $Contains(x, ov)$ qui précise que le registre x contient la valeur ov, donner alors les préconditions et les effets de l'opérateur $Assign(x, v, lv, ov)$.

b) En supposant qu'il y a un registre additionnel qui contient au départ la valeur C et qu'on pourrait utiliser pour faire l'échange, donner alors le plan initial de départ.

c) Donner alors un plan pour échanger A et B en explicitant les différentes étapes, particulièrement en ce qui concerne l'ordre entre les différentes actions.

Question 2 : Soit le réseau Bayésien et les tables de probabilités conditionnelles tel que le montre la figure suivante :



Cette figure a été explicitée en cours, et on voudrait s'en servir pour exprimer quelques probabilités.

1. Exprimer la distribution jointe $P(B, E, A, J, M)$ en termes de probabilités conditionnelles données par le réseau.

2. On voudrait connaitre la probabilité de Burglary, sachant que John a appelé et Mary a appelé

   a. Quelles sont les variables cachées ? Que doit-on faire avec de telles variables ?

b. Calculez $P(B \mid J, M)$, i.e.,
calculez $P(Burglary = True \mid JohnCalls = True, MaryCalls = True)$

Question 3 : On voudrait représenter à l'aide d'un réseau Bayésien le système de carburant (Fuel) de véhicule à l'aide 3 variables : B pour batterie ; F pour Fuel et G pour jauge. Dans ce contexte, $B = 0$ signifie que la batterie est à plat; et $B = 1$ qu'elle est chargée. De même $F = 1$ signifie que le réservoir est plein et que $F = 0$ qu'il est vide. Finalement, $G = 1$ signifie plein de carburant et $G = 0$, pas de carburant.

Les probabilités sont comme suit :

$p(B=1) = 0.9$; $p(F=1) = 0.9$

$p(G=1 \mid B=1, F=1) = 0.8$

$p(G=1 \mid B=1, F=0) = 0.2$

$p(G=1 \mid B=0, F=1) = 0.2$

$p(G=1 \mid B=0, F=0) = 0.1$

a) Donner le réseau bayésien liant B, F, G
b) Supposons qu'on observe la variable G et qu'on découvre que $G = 0$; quelle est la probabilité pour que le réservoir soit vide ? Commentez cette probabilité par rapport à $p(F=0)$.
c) Supposons maintenant qu'on observe à la fois G et B et qu'il sont tous les deux à 0 ; dites quelle est la probabilité pour que le réservoir soit vide. Commentez cette probabilité par rapport à la probabilité trouvée en b).
d) Supposons qu'au lieu de lire directement G, c'est le conducteur (D pour Driver) qui se charge de nous dire si G reflète le plein ($D = 1$) ou il reflète le vide ($D = 0$). Donnez alors le nouveau réseau bayésien liant B, F, G, D
e) En réalité, notre conducteur est peu fiable puisque les probabilités donnent :

$$p(D=1 \mid G=1) = 0.9$$
$$p(D=0 \mid G=0) = 0.9$$

Supposons que le Driver nous disent que $D = 0$; évaluez alors la probabilité pour que le réservoir soit vide sachant uniquement cette observation.

f) En plus de $D = 0$ on a aussi $B = 0$; évaluez la probabilité pour que le réservoir soit vide et notez que cette probabilité est plus faible que la probabilité déterminée en e). Quelle est l'intuition derrière ça ?

Question 4 : Soit un HMM avec trois états, trois sorties et dont les modèles de transition $P(X_{t+1} \mid X_t)$ et de senseurs $P(E_t \mid X_t)$ sont comme suit :

X_t \ X_{t+1}	a	b	c
a	0.5	0.4	0.1
b	0.1	0.5	0.4
c	0.4	0.1	0.5

et

X_t \ E_t	p	q	r
a	0.7	0.1	0.2
b	0.2	0.7	0.1
c	0.1	0.2	0.7

On suppose que l'on a une distribution uniforme pour l'état initial X_0, calculez à la main la séquence d'états cachés la plus probable pour la séquence observée (p, p, r, r, q, r), et ce en utilisant l'algorithme de Viterbi.

Question 5 : Un patient souffrant d'un mal, rend visite au médecin. Le médecin suspecte trois maladies comme diagnostique pour ce malade : D_1, D_2, D_3 qui sont marginalement indépendantes l'une de l'autre. Il y a 4 symptômes que le Médecin voudrait tester pour trouver la cause la plus probable à l'origine du mal dont souffre le patient : S_1, S_2, S_3, S_4. Les symptômes sont conditionnellement dépendants aux 3 maladies comme suit : S_1 dépend seulement de D_1, S_2 dépend de D_1 et de D_2. S_3 dépend de D_1 et de D_3, tandis que S_4 dépend seulement de D_3. En supposant que les toutes les variables sont booléennes et qu'elles soient donc vraies ou fausses,

1. Donner le réseau Bayésien correspondant à ce problème de diagnostique
2. Exprimer la probabilité jointe $\mathbf{P}(D_1, D_2, D_3, S_1, S_2, S_3, S_4)$ en un produit de probabilités conditionnelles
3. Quel est le nombre de paramètres indépendants que requiert cette probabilité jointe ?
4. En supposant qu'il n'y a pas d'indépendance conditionnelle entre les variables, combien de paramètres indépendants sont alors requis (pour la même probabilité jointe) ?
5. Quelle est la couverture de Markov (Markov Blanket) de S_2 ?
6. Si on observe le quatrième symptôme ($S_4 = true$), pour quelle maladie a-t- on un gain d'information ?
7. Supposons qu'on observe le deuxième symptôme ($S_2 = true$), qu'est-ce que l'observation du quatrième symptôme ($S_4 = true$) nous dit maintenant ?