INF8225 – Intelligence Artificielle: Techniques Probabilistes et d'Apprentissage

TP1

Éric Morissette

1631103

2 Février 2016

École Polytechnique de Montréal

Supposons ce graphe pour la question 1.

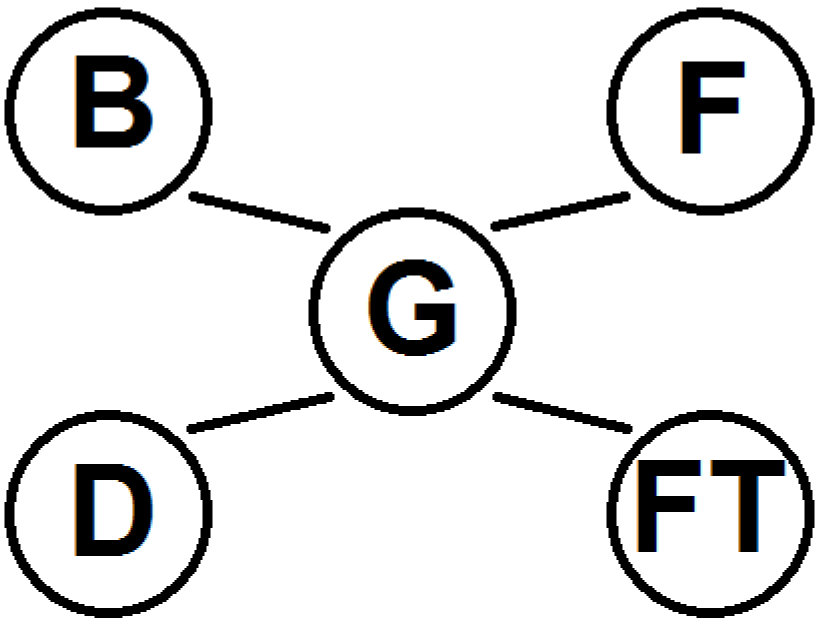
B = Battery

F = Fuel

G = Gauge

D = Distance

FT = Fill Tank



Question 1:

(1) Explaining Away
Pour voir le principe de "Explaining Away", il faut se concentrer sur les noeuds B, F et G. Pour ce faire, il faut voir les probabilités suivantes:
   a. p(F|B=0)=0.100000
   b. p(F|B=1)=0.100000
   c. p(F|B=1,G=1)=0.027027

Avec a) et b), on voit que la valeur de B n'influence pas F, ils sont donc indépendants.
Par contre, avec c), on voit que si B et G sont vrai, alors la probabilité de F baisse, car G est "expliqué" par B.

Le concept d'un "Explaining Away" est issu du fait que si G est vrai, alors il est probable que B et/ou F soit vrai, pour notre cas, on voit que si l'on ne connait pas la valeur de G mais que l'on connait la valeur de B, on peut rien tirer pour F. Par contre, si on sait que B et G sont vrais, alors il est plus probable que F soit faut, car la valeur de G ne peut être expliquée que par B ou F, et nous savons que B est vrai. Alors F est probablement faux.

(2) Serial Blocking

Pour comprendre le principe de “Serial Blocking”, on doit se concentrer sur les noeuds B, G et D. Et considérer les probabilités suivantes:

a. p(D|B=0)=0.902500
b. p(D|B=1)=0.765000
c. p(D|B=0, G=0)=0.950000
d. p(D|B=1, G=0)=0.950000

Avec a) et b), on voit que la valeur de B a une influence sur D. Par contre, si l'on fixe la valeur de G à faux et qu'on modifie la valeur de B, on perd l'influence de B sur D. Il s'agit d'un “Serial Blocking”, ce qui veut dire que si le chemin entre deux noeuds passe par un point avec une valeur fixée, on peut ignorer tous les noeuds au dessus de ce dernier, à condition qu'il n'y ait qu'un seul chemin entre les noeuds.

Le concept d'un “Serial Blocking” vient du fait que si on connait la valeur d'une variable à quelque part dans la chaîne entre deux noeuds, alors tout ce qui est au dessus du noeud fixe n'a plus d'importance sur les noeuds en dessous du noeud fixe.

(3) Divergent Blocking

Pour analyser un “Divergent Blocking”, il faut regarder les noeuds G, D et FT ainsi que les probabilités suivantes:

a. p(FT|D=0)=0.446292
b. p(FT|D=1)=0.571525
c. p(FT|D=0,G=0)=0.200000
d. p(FT|D=1,G=0)=0.200000

Avec a) et b), on voit très bien que D a une influence sur la probabilité de FT, ce qui est logique car si l'on y pense, le noeud G influence D et FT et si D est confirmé ou infirmé, cela influence les probabilités de FT.

Avec c) et d), on voit que si la valeur de G est fixée, faire varier la valeur de D n'aura pas d'influence sur la probabilité de FT car l'influence est bloquée par le noeud fixe G.

Le concept d'un “Divergent Blocking” vient du fait que si on connait la valeur de notre parent, alors la valeur des enfants de notre parent n'ont aucune influence sur nous.

Supposons ce graphe pour la question 2.

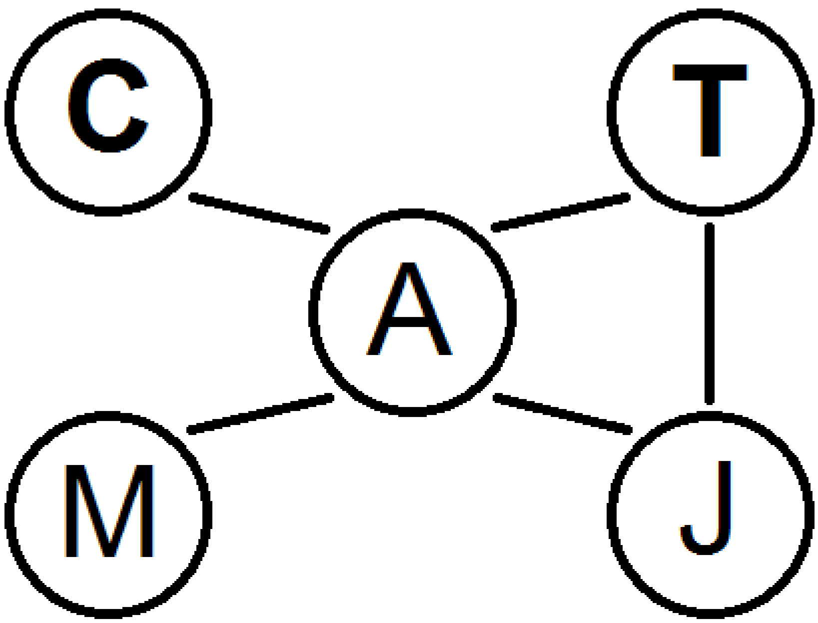
C = Cambriolage

T = Tremblement

A = Alarme

M = Marie Appellée

J = Jean Appellé



Question 2:

a) Fichier “mkDetecterCambriolageDgm.m”

b) [Histogramme]

c) Probabilités Marginales Conditionnelles
   1. p(C=1|M=1, J=0)=0.008337
   2. p(C=1|M=0, J=1)=0.044852
   3. p(C=1|M=1, J=1)=0.445950
   4. p(C=1|M=0, J=0)=0.000110
   5. p(C=1|M=1)=0.016284
   6. p(C=1|J=1)=0.235300

d) Probabilités Marginales Inconditionnelles
   1. C: p(C=1)=0.001000
   2. T: p(T=1)=0.002000

3. A: p(A=1)=0.002516
4. M: p(M=1)=0.052139
5. J: p(J=1)=0.001994

e)

1.

$$p(J = V) = \sum_{C,\, T,\, A,\, M} \{ p(C,\, T,\, A,\, J = V,\, M) \}$$

$$p(J = V) = \sum_{C,\, T,\, A,\, M} \{ \prod_{i=1}^{n} [p(N_i | Parents(N_i))] \}$$

$$p(J = V) = \sum_{C,\, T,\, A,\, M} \{ p(C) * p(T) * p(A \mid C,\, T) * p(M \mid A) * p(J \mid A,\, T) \}$$

2.

$$p(C = V \mid J = V) = \frac{p(C, J)}{p(J)}$$

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

$$p(C = V \mid J = V) = \frac{\sum_{T, A, M} \{ \prod_{i=1}^{n} [p(N_i | Parents(N_i))] \}}{p(J)}$$

$$p(C = V \mid J = V) = \frac{\sum_{T, A, M} \{ p(C) * p(T) * p(A \mid C, T) * p(M \mid A), p(J \mid A, T) \}}{p(J)}$$

Supposons le graphe de la question 2 mais avec seulement les nœuds A, M et J.

Question 3 :
Équations nécessaires à l'algorithme EM

$$(1) - p(M, J) = p(A) * p(M|A) * P(J|A)$$

$$(2) - p(A|M, J) = \frac{p(A, M, J)}{p(M, J)}$$

$$(3) - N_i = N_i * p(A|M, J)$$

$$(4) - \sum_{M, J} (N * log(p(M, J)))$$

Premièrement, pour utiliser l'algorithme EM, il faut prendre toutes les valeurs inconnues dans nos observations et leur assigner une valeur aléatoire arbitraire.

Deuxièmement, on prend le nombre d'échantillons d'observations N dans la formule (3) et ont les modifient selon les valeurs trouvées dans (1) et (2). Ce qu'on fait ici est l'ajustement des valeurs arbitraires mises initiallement.

Troisièmement, on peut calculer la vraisemblance des estimations avec (4).

Finalement, on répète le processus un nombre fini d'itérations, ou jusqu'à la convergence des valeurs.