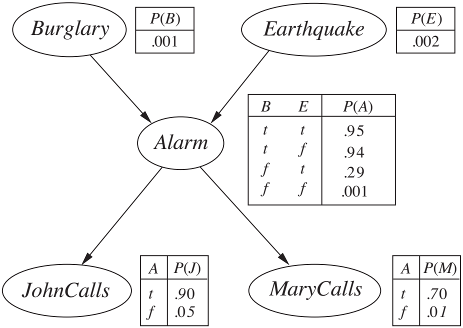
CSC384
Introduction to Artificial Intelligence: Uncertainty

November 13, 2014

Bayesian Networks
Inference

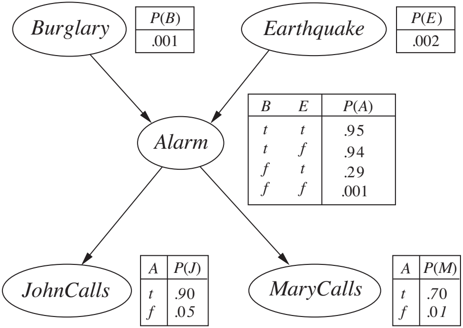


We have: $P(B, E, A, M, J) = P(B) * P(E) * P(A|B, E) * P(M|A) * P(J|A)$

Want to ask a queries in forms such as: $P(B = t|M = f, J = t) = ?$

Bayesian Networks

Inference



P(B)
.001

P(E)
.002

B	E	P(A)
t	t	.95
t	f	.94
f	t	.29
f	f	.001

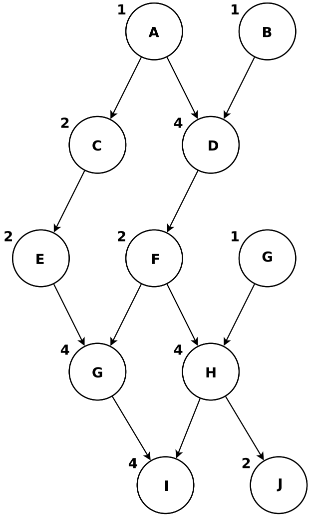
A	P(J)
t	.90
f	.05

A	P(M)
t	.70
f	.01

We have: $P(B, E, A, M, J) = P(B) * P(E) * P(A|B, E) * P(M|A) * P(J|A)$

Want to ask a queries in forms such as: $P(B = t|M = f, J = t) = ?$

$$
\begin{aligned}
P(A,B,C,D,E,F,G,H,I,J,K) = \\
P(A)* \\
P(B)* \\
P(C|A)* \\
P(D|A,B)* \\
P(E|C)* \\
P(F|D)* \\
P(G)* \\
P(H|E,F)* \\
P(I|F,G)* \\
P(J|H,I)* \\
P(K|I)
\end{aligned}
$$



$$P(A,B,C,D,E,F,G,H,I,J,K) = P(A)P(B)P(C|A)P(D|A,B)P(E|C)$$
$$P(F|D)P(G)P(H|E,F)P(I|F,G)P(J|H,I)P(K|I)$$

We may know that that $H = true, I = False$. This forms an evidence set. If we want to know $P(D|h,\neg i)$ we can calculate this value by summing out the relevant probabilities.

(Notation: let lower case letters such as a represent $A = t$ and $\neg a$ represent $A = f$)

$$\sum_{A,B,C,E,F,G,J,K} P(A,B,C,d,E,F,G,h,\neg i,J,K) = P(d,h,\neg i)$$

Bayesian Networks

Inference

$$\sum_{A,B,C,E,F,G,J,K} P(A, B, C, d, E, F, G, h, \neg i, J, K) = P(d, h, \neg i) \quad (1)$$

$$\sum_{A,B,C,E,F,G,J,K} P(A, B, C, -d, E, F, G, h, \neg i, J, K) = P(-d, h, \neg i) \quad (2)$$

$$P(d, h, \neg i) + P(\neg d, h \neg i) = P(h, \neg i) \quad (3)$$

$$P(d|h, \neg i) = P(d, h, \neg i)/P(h, \neg i) \quad (4)$$

$$P(\neg d|h, \neg i) = P(\neg d, h, \neg i)/P(h, \neg i) \quad (5)$$

Bayesian Networks

Inference

We can use properties of Bayes nets to decompose the earlier summation

$$\sum_{A,B,C,E,F,G,J,K} P(A,B,C,d,E,F,G,h,\neg i,J,K) = P(d,h,\neg i) =$$
$$P(A)P(B)P(C|A)P(d|A,B)P(E|C)P(F|d)P(G)P(h|E,F)$$
$$P(\neg i|F,G)P(J|h,\neg i)P(K|\neg i)$$

Rearrange the summations so that we only sum over relevant variables:

$$\sum_{A},\sum_{B},\sum_{C},\sum_{E},\sum_{F},\sum_{G},\sum_{J},\sum_{K} P(A)P(B)P(C|A)P(d|A,B)P(E|C)$$
$$P(F|d)P(G)P(h|E,F)P(\neg i|F,G)P(J|h,\neg i)P(K|\neg i)$$

Bayesian Networks

Inference

$$\sum_A, \sum_B, \sum_C, \sum_E, \sum_F, \sum_G, \sum_J, \sum_K P(A)P(B)P(C|A)P(d|A,B)P(E|C)$$
$$P(F|d)P(G)P(h|E,F)P(\neg i|F,G)P(J|h,\neg i)P(K|\neg i) =$$
$$\sum_A P(A), \sum_B P(B), \sum_C P(C|A)P(d|A,B), \sum_E P(E|C),$$
$$\sum_F P(F|d), \sum_G P(G)P(h|E,F)P(\neg i|F,G), \sum_J P(J|h,\neg i), \sum_K P(K|\neg i) =$$
$$\sum_A P(A), \sum_B P(B)P(d|A,B), \sum_C P(C|A), \sum_E P(E|C),$$
$$\sum_F P(F|d)P(h|E,F), \sum_G P(G)P(\neg i|F,G), \sum_J P(J|h,\neg i), \sum_K P(K|\neg i)$$

$$\sum_A P(A), \sum_B P(B)P(d|A,B), \sum_C P(C|A), \sum_E P(E|C),$$
$$\sum_F P(F|d)P(h|E,F), \sum_G P(G)P(\neg i|F,G), \sum_J P(J|h,\neg i), \sum_K P(K|\neg i)$$

Given the above note that:

$$\sum_K P(K|\neg i) = P(k|\neg i) + Pr(\neg k|\neg i) = c_1$$
$$\sum_J P(J|h,\neg i)c_1 = c_1 \sum_J P(J|h,\neg i) = c_1(P(j|h,\neg i) + P(\neg j|h,\neg i)) = c_1 c_2$$

Bayesian Networks

Inference

$$\sum_K P(K|\neg i) = P(k|\neg i) + Pr(\neg k|\neg i) = c_1$$

$$\sum_J P(J|h,\neg i)c_1 = c_1 \sum_J P(J|h,\neg i) = c_1(P(j|h,\neg i) + P(\neg j|h,\neg i)) = c_1 c_2$$

$$c_1 c_2 \sum_G P(G)P(\neg i|F,G) = c_1 c_2(P(g)P(\neg i|F,g) + P(\neg g)P(\neg i|F,\neg g)$$

$P(\neg i|F,g)$ depends on F and is not a single number!

Coming from the other direction:

Expanding in variable order will result in exponential increases in the size of the equation, but many repeated subterms

$$P(a)(\sum_B P(B)P(d|a,B), \sum_C P(C|a), \sum_E P(E|C),$$
$$\sum_F P(F|d)P(h|E,F), \sum_G P(G)P(\neg i|F,G), \sum_J P(J|h,\neg i), \sum_K P(K|\neg i))+$$
$$P(\neg a)\sum_B P(B)P(d|\neg a,B), \sum_C P(C|\neg a), \sum_E P(E|C),$$
$$\sum_F P(F|d)P(h|E,F), \sum_G P(G)P(\neg i|F,G), \sum_J P(J|h,\neg i), \sum_K P(K|\neg i) =$$
$$P(a)P(b)p(d|a,b)f_1 + P(\neg a)P(b)p(d|\neg a,b)f_2+$$
$$P(a)P(\neg b)p(d|a,\neg b)f_1 + P(\neg a)P(\neg b)p(d|\neg a,\neg b)f_2$$

Dynamic Programming

- Store values of computed subterms for reuse
- Can in some cases result in an exponential reduction in the amount of caluclation necessary!
- Common algorithm for doing this: Variable elimination

$$\sum_A P(A), \sum_B P(B)P(d|A,B), \sum_C P(C|A), \sum_E P(E|C),$$
$$\sum_F P(F|d)P(h|E,F), \sum_G P(G)P(\neg i|F,G), \sum_J P(J|h,\neg i), \sum_K P(K|\neg i)$$

Note that in this equation

$$\sum_K P(K|\neg i) = 1$$
$$\sum_J P(J|h,\neg i) = 1$$

And we can therefore remove these terms from the original equation (Why is this true?)

We approach variable elimination from the inside out as with the first approach we took.

$$\sum_K P(K|\neg i) = P(k|\neg i) + Pr(\neg k|\neg i) = 1$$

$$\sum_J P(J|h,\neg i)c_1 = c_1 \sum_J P(J|h,\neg i) = c_1(P(j|h,\neg i) + P(\neg j|h,\neg i)) = 1$$

$$\sum_G P(G)P(\neg i|F,G) = P(g)P(\neg i|F,g) + P(\neg g)P(\neg i|F,\neg g)$$

We continue by computing two different numbers for $P(\neg i|F,g)$, one for each value of F

$$t(f) = \sum_G P(G)P(\neg i|f,G)$$
$$t(\neg f) = \sum_G P(G)P(\neg i|\neg f,G)$$
$$t(f) = P(g)P(\neg i|f,g) + P(\neg g)P(\neg i|f,\neg g)$$
$$t(\neg f) = P(g)P(\neg i|\neg f,g) + P(\neg g)P(\neg i|\neg f,\neg g)$$

We can use this to sum out F

$$
\begin{aligned}
\sum_F P(F|d)P(h|E,F), \sum_G P(G)P(\neg i|F,G) &= \\
P(f|d)P(h|E,f), \sum_G P(G)P(\neg i|f,G) &+ \\
P(\neg f|d)P(h|E,\neg f), \sum_G P(G)P(\neg i|\neg f,G) & \\
&= \\
P(f|d)P(h|E,f)t(f) + P(\neg f|d)P(h|E,\neg f)t(\neg f) &= \\
\sum_F P(F|d)P(h|E,F)t(F) &=
\end{aligned}
$$

- At each stage we compute a table of values with entries for every instantiation of variables in the sum
- Table size is exponential in the number of variables
- The tables are called *Factors*
- $f(X, Y)$ denotes a factor with variable sets X and Y