UMass CMPSCI 383 (AI) HW4: Chapters 13–14

Yifu Liu

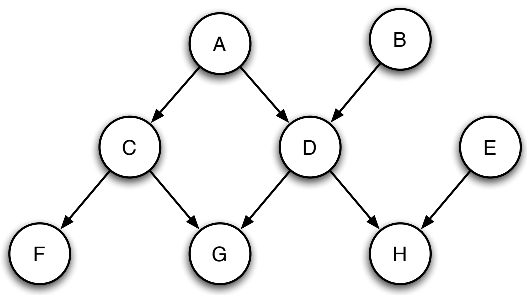
Assigned: Oct 26 2017; Due: Nov 6 2017 @ 11:55 PM EST

Abstract

Submit a (.zip) file to both Moodle and Gradescope containing your latex (.tex) file and rendered pdf. All written HW responses should be done in latex (use sharelatex.com or overleaf.com).

1 Reading Independence Relationships from BN (30 pts)

Consider the following Bayesian network:



1. Suppose that all the variables are Boolean. How many individual probabilities are needed to specify an arbitrary joint probability distribution over eight variables (without reference to the network) (5 pts)?
   2^8
   Each variable needs 2 values, which are True and False. If we have an additional variable, so we have to double the probabilities. So, that need 2^8 probabilities in this case.

2. How many individual probabilities are needed to specify the joint probability distribution if we assume the conditional independence relations encoded in the Bayesian network above? Explain how the joint distribution can be recovered from this more compact representation (5 pts).
   D, H and G, each of them has 4 probabilities; C and F each of them has 2 probabilities; A, B, and E, each of them has 1 probability.

3. Which of the following probabilistic relations are implied by the structure of the above Bayesian network? There may be multiple. Justify your answers (5 pts each).
   (a) $P(E|G) = P(E)$
   True, it's implied by the structure of the Bayesian network, because they are implying d-separated rule.
   (b) $P(C|D) = P(C)$
   False, there C and D are the parents of G, so they are not conditionally independent.
   (c) $P(D|A, B, C, E, F, G, H) = P(D|A, B, G, H)$
   False, because H and G are respectively dependent on E and C. P(H|D) inside P (D | A,B,G,H) is different from P(H|D, E) inside P (D | A,B,C,E,F,G,H).
   Same reason for P(G|D) inside P(D|A,B,G,H) would be different from P(G|C,D) inside P (D | A,B,C,E,F,G,H).

   (d) $P(C, E|D) = P(C|D)P(E|D)$
   True. Since C and E are d-seperated. There is a undirected path CADHE has a diverging node A and CGDHE has a D. Because they are d-separated, so they are conditionally independent.

2 Computing Posteriors (25 pts)

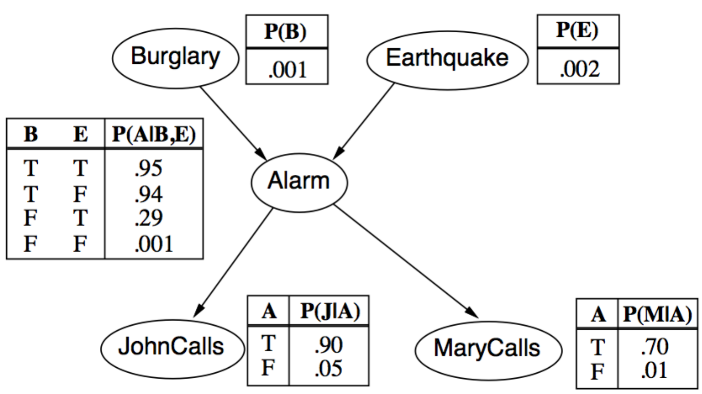
There is a test that detects whether you are suffering from the flu, but it is only 90% accurate in patients who actually have the flu. Also, the test comes back negative in only 95% of the patients who do not have the flu. From the overall records, it is known that only 8% of the population at large will get the flu.

1. You feel a bit feverish, and decide to get tested for swine flu. The test comes back positive. What is the posterior probability that you indeed have the flu? You must show your work for credit (10 pts).
   P(F|Postive) = (P(Positive|F)*P(F))/P(Positive)
   P(Positive) = P(Positive|F)*P(F) + P(Positive| F)*P(F)
   P(Positive) = 0.118 P(F|Postive) = 0.072/0.118 = 61%

2. The doctor decides to order a second test, which also comes back positive. What is the revised posterior probability that you have the flu? You must show your work for credit (10 pts).
   P = Positive
   P(F|P&P) = (P(P&P|F)P(F))/P(P&P)
   P(P&P) = P((P&P)|F)*P(F) + P(P&P|¬F)*P(¬F)
   P(F|P&P) = (0.9*0.9*0.08)/(0.9*0.9*0.08)+(0.05*0.05*0.92) = 96.57%.

3. Derive a formula for the posterior given the number of tests N. Assuming all your tests come back positive, how many independent positive tests do you need to obtain a posterior probability greater than 99%? You must show your work for credit (5 pts).
   $\frac{(.08)(.9)^N}{(.08)(.9)^N+(.92)(.05)^N} > .99$
   $(0.01) * (0.08)(0.9)^N > (0.99)(0.92)(0.05)^N$
   $(\frac{0.9}{0.05})^N > (\frac{0.92*0.99}{0.08*0.01})$
   By calculation, 3 times of positive test will result in the posterior probability greater than 99%.

3 Exact & Approximate Inference (45 pts)



1. Compute the conditional probability distribution, $P(E|m)$, using variable elimination. Identify the factors that would need to be computed during variable elimination and order your factors in a way that reduces the necessary the computation as much as possible (15 pts).
   $P(E|m) = \alpha \sum_B \sum_A \sum_J P(m|a)P(J|A)P(a|b,E)P(B)P(E)$
   $P(E|m) = \alpha P(E) \sum_B P(B) \sum_A P(m|a)P(a|b,E)$
   $f(E) = f(b)f(b,E) + f(\neg b)f(\neg b,E)$
   Function $f(E,e)$ for B and E respectively are (T,T),(T,F),(F,T),(F,F) => ((0.6655),(0.6586),(0.2101),(0.01069))
   Function $f(E)$, for E has True and False.
   $f(E = True) = 0.001 * 0.6655 + 0.999 * 0.2101 = 0.2105554$
   $f(E = False) = 0.001 * 0.6586 + 0.999 * 0.01069 = 0.01133791$
   Normalize:
   $P(E|m) = \alpha f(m|E)f_6(E) = 0.035881$ And (1 - 0.035881) = 0.964119

2. Assume we are performing direct sampling of our BN, and midway through the sampling, we need to sample A. Assume we sampled $\neg b$ and e so far. How do we sample a value for A?

(a) Write down the CPT that defines the distribution over A from which we will want to sample (5 pts).

A	$P(A \mid \neg b, e)$
T	0.29
F	0.71

If the generated numbers from [0,0.29), that will return True; But [0.29,1) that will return False.

(b) How can we sample from this distribution if all we have are random samples from a uniform distribution between 0 and 1: $U(0,1)$ (5 pts)? Map all sample results from 0 to 0.29 to True and map 0.29 to 1 to False.

3. Assume we are using Likelihood Weighting to compute the posterior distribution, $P(J, M \mid \neg e, a)$. Also assume that during the sampling process, we drew $\neg b$. What is the likelihood weight, w, that we will record in our table once we finish sampling the rest of the variables (10 pts)?

J	M	E	A	B	$weight$
T	T	T	T	F	0.999*1*1*0.9*0.7 = 0.62937
T	F	T	T	F	0.999*1*1*0.9*0.3 = 0.26973
F	T	T	T	F	0.999*1*1*0.1*0.7 = 0.6993
F	F	T	T	F	0.999*1*1*0.1*0.3 = 0.2997

4. Assume we are using Gibbs sampling to compute the posterior distribution, $P(J, M \mid \neg e, a)$, and we are starting with the event $[j, m, \neg b, e, a]$. Assume we randomly select A to be sampled. Write down the CPT that defines the distribution over A from which we will want to sample (10 pts). Use `\begin{table}...` to make a table (see 2.(a) for reference).
$P(b) = P(b) * P(j|a) * P(m|a) * P(a|b, \neg e) = 0.000592$
$P(\neg b) = P(\neg b) * P(j|a) * P(m|a) * P(a|\neg b, \neg e) = 0.000629$
$\alpha(0.000592 + 0.000629) = 1$

B	$P(B \mid MB(A))$
T	0.485
F	0.515