CSE 590ST: Statistical Methods in Computer Science

Homework 4

Due in class on June 2, 2004

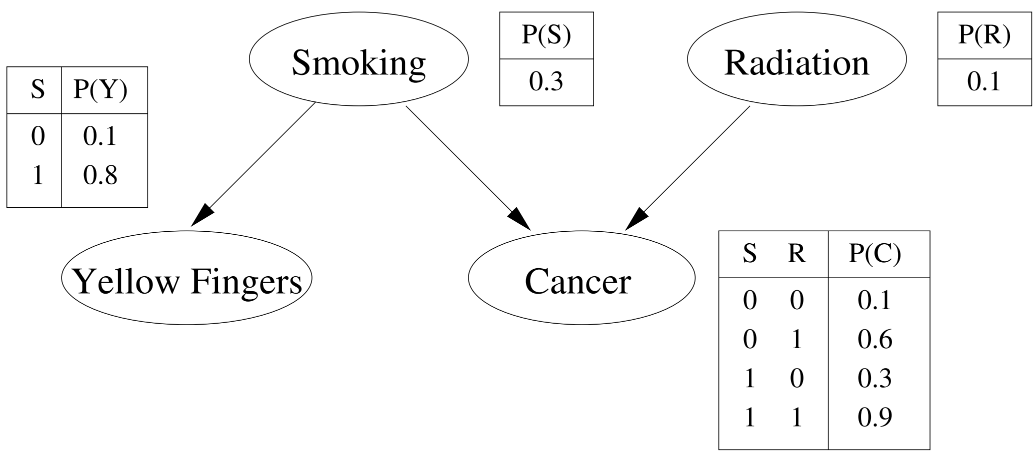
1. Suppose you use greedy search to learn the structure of a Bayesian network, with the search operators being all possible arc additions, deletions and reversals at each step. Suppose the network has n variables, and the maximum allowed number of parents per variable is a constant $k \ll n$. There is no missing data.

   (a) Show that the worst-case computational cost of scoring all the candidate networks at each step using a naive implementation is $O(n^3)$. (Ignore the issue of avoiding cycles.)

   (b) How can you reduce this to $O(n^2)$?

2. How would you generalize the EM algorithm to learn mixtures of Gaussians with unknown means, covariances and component priors?

3. Consider again the Bayesian network in Question 2 of Homework 1:



P(S)
0.3

P(R)
0.1

S	P(Y)
0	0.1
1	0.8

S	R	P(C)
0	0	0.1
0	1	0.6
1	0	0.3
1	1	0.9

   (a) Convert this Bayesian network to an equivalent Markov network, using one potential function per maximal clique.

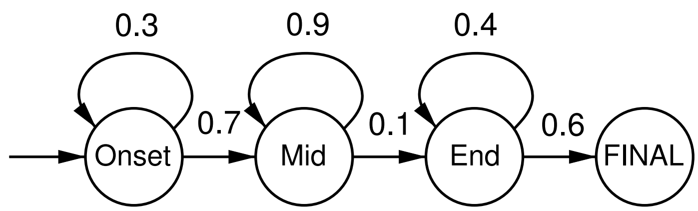
   (b) Choose a (maximal) clique. How does multiplying all values of its potential function by a constant c change the probability distribution represented by the Markov network? How does it change the value of the partition function?

   (c) Let G be the graph of this Markov network. What is the Markov blanket of Yellow Fingers in G?

   (d) Is Smoking independent of Radiation according to G?

4. Hidden Markov models are often used to recognize *phones*, the elementary sounds of speech. Each phone has three possible states: Onset, Mid and End, plus a final absorbing state. The space of possible acoustic observations is often partitioned into n regions $\{C_1, \ldots, C_n\}$. The figure below shows the HMM for the phone [m] (the first phone in, for example, the word "mother").

Phone HMM for [m]:



Output probabilities for the phone HMM:

Onset:	Mid:	End:
C1: 0.5	C3: 0.2	C4: 0.1
C2: 0.2	C4: 0.7	C6: 0.5
C3: 0.3	C5: 0.1	C7: 0.4

Calculate the most probable path through this HMM for the observation sequence $[C_1, C_2, C_3, C_4, C_4, C_6, C_7]$. Also give its probability.

5. In 1738, J. Bernoulli investigated the St. Petersburg paradox, which works as follows. You have the opportunity to play a game in which a fair coin is tossed repeatedly until it comes up heads. If the first head appears on the nth toss, you win 2^n dollars.

   (a) Show that the expected monetary value of this game is infinite.

   (b) How much would you, personally, pay to play the game?

   (c) Bernoulli resolved the apparent paradox by suggesting that the utility of money is measured on a logarithmic scale (i.e., $U(S_n) = a \log_2 n + b$, where S_n is the state of having \$$n$). What is the expected utility of the game under this assumption?

   (d) What is the maximum amount that it would be rational to pay to play the game, assuming that one's initial wealth is \$$k$?