Last name:________________ First name:________________ SID:__________ Class account login:__________

Collaborators: ____

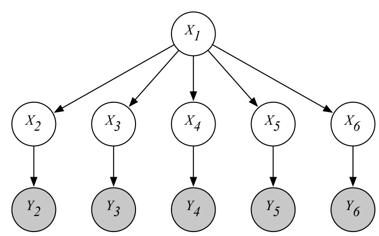
CS188 Spring 2011 Written 4: Bayes Nets and HMMs

Due: Monday, 4/11/11 at 5:29pm in 283 Soda Drop Box (no slip days).

Policy: Can be solved in groups (acknowledge collaborators) but must be written up individually. Recall to make a photo-copy of your solutions to allow you to resubmit for partial credit recovery. See course webpage for details.

1 [10 pts] Variable Elimination

Consider the following Bayes net,



and suppose we want to compute $P(X_6|Y_2 = y_2, \ldots, Y_6 = y_6)$, where all the random variables are binary. This exercise is about the variable elimination algorithm, one way of computing this quantity seen in class.

We start by an example, working on the elimination order X_2, X_3, X_1, X_4, X_5.

Start by inserting evidence , which gives the following initial factors:

$$p(X_1)p(X_2|X_1)p(X_3|X_1)p(X_4|X_1)p(X_5|X_1)p(X_6|X_1)p(y_2|X_2)p(y_3|X_3)p(y_4|X_4)p(y_5|X_5)p(y_6|X_6)$$

Eliminate X_2: $f_1(X_1, y_2) = \sum_{X_2} p(X_2|X_1)p(y_2|X_2)$, and get:

$$p(X_1)f_1(X_1, y_2)p(X_3|X_1)p(X_4|X_1)p(X_5|X_1)p(X_6|X_1)p(y_3|X_3)p(y_4|X_4)p(y_5|X_5)p(y_6|X_6)$$

Eliminate X_3: $f_2(X_1, y_3) = \sum_{X_3} p(X_3|X_1)p(y_3|X_3)$, and get:

$$p(X_1)f_1(X_1, y_2)f_2(X_1, y_3)p(X_4|X_1)p(X_5|X_1)p(X_6|X_1)p(y_4|X_4)p(y_5|X_5)p(y_6|X_6)$$

Eliminate X_1: $f_3(y_2, y_3, X_4, X_5, X_6) = \sum_{X_1} p(X_1)f_1(X_1, y_2)f_2(X_1, y_3)p(X_4|X_1)p(X_5|X_1)p(X_6|X_1)$, and get:

$$f_3(y_2, y_3, X_4, X_5, X_6)p(y_4|X_4)p(y_5|X_5)p(y_6|X_6)$$

Eliminate X_4: $f_4(y_2, y_3, y_4, X_5, X_6) = \sum_{X_4} f_3(y_2, y_3, X_4, X_5, X_6)p(y_4|X_4)$, and get:

$$f_4(y_2, y_3, y_4, X_5, X_6)p(y_5|X_5)p(y_6|X_6)$$

Eliminate X_5: $f_5(y_2, y_3, y_4, y_5, X_6) = \sum_{X_5} f_4(y_2, y_3, y_4, X_5, X_6)p(y_5|X_5)$, and get:

$$f_5(y_2, y_3, y_4, y_5, X_6)p(y_6|X_6)$$

Finally, we have

$$P(X_6, Y_2 = y_2, \ldots, Y_6 = y_6) = f_5(y_2, y_3, y_4, y_5, X_6)p(y_6|X_6),$$

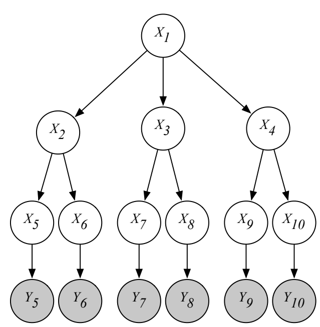
so normalizing over X_6 gives $P(X_6|Y_2 = y_2, \ldots, Y_6 = y_6)$.

Note: the dimensionality of the factors only depends on the number of *unobserved* variables involved in the factor. For example, $f_5(y_2, y_3, y_4, y_5, X_6)$ had dimensionality one (i.e. it has size $|\text{dom}(X_6)|$), and $f_4(y_2, y_3, y_4, X_5, X_6)$ has dimensionality two (i.e. size $|\text{dom}(X_5)| \cdot |\text{dom}(X_6)|$). This is what you should be looking at when determining the complexity of an elimination ordering (taking the max of the factor sizes across all the intermediate factors produced).

[4 pts] (a) Do the complete work for the ordering X_1, X_2, X_3, X_4, X_5. What is the size of the largest factor generated during variable elimination?

[3pts] (b) What are the most efficient orderings? What is the size of the largest factor these orderings generate? No need to show the complete work.

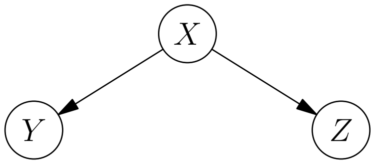
[3pts] (c) Now consider the following Bayes net:



What are the most efficient orderings for this tree? What is the size of the largest factor these orderings generate? No need to show the complete work. You can describe what the ordering should be for any query, or you can assume that we are solving for $P(X_5|Y_5, ..., Y_{10})$

2 [6 pts] Sampling

Consider the following Bayesian network.



X	$\Pr(X)$
0	0.5
1	0.5

Y	X	$\Pr(Y\|X)$
0	0	0.3
1	0	0.7
0	1	0.4
1	1	0.6

Z	X	$\Pr(Z\|X)$
0	0	0.1
1	0	0.9
0	1	0.5
1	1	0.5

Define the function $f(X,Y,Z)$ by

$$f(X,Y,Z) = (X+Y+Z)^2.$$

In the questions below, perform either rejection sampling or likelihood-weighted sampling by sampling the individual variables (as required by the sampling method) in the order (X,Y,Z). Use as many values as needed from the following sequence $\{a_i\}_{1\leq i\leq 15}$ of numbers generated independently and uniformly at random from $[0,1)$ as a source of randomness.

a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}
0.142	0.522	0.916	0.792	0.859	0.036	0.656	0.249	0.934	0.679	0.758	0.743	0.392	0.655	0.171

To sample a binary variable W with $\Pr(W=1)=p$ (and $\Pr(W=0)=1-p$) using a value a_i sampled uniformly at random from $[0,1)$, choose:

$$W = \begin{cases} 1 & \text{if } a_i < p; \\ 0 & \text{if } a_i \geq p. \end{cases}$$

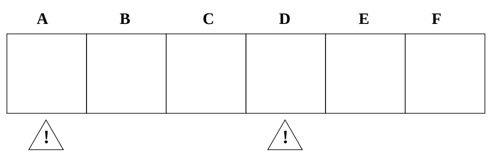
For example, if we first sample X we will set $X=1$ because $a_1 = 0.142 < Pr(X=1) = 0.5$.

[3 pts] (a) Estimate the expectation $\mathrm{E}[f(X,Y,Z)|Y=0,Z=1]$ using rejection sampling, continuing to collect samples until exactly two have been accepted. The sampling attempt should stop as soon as you discover that the sample will be rejected. Show your work.

[3pts] (b) Estimate the expectation $\mathrm{E}[f(X,Y,Z)|Y=0,Z=1]$ using two samples obtained using likelihood-weighting. Show your work. Reuse the sequence $\{a_i\}_{1\le i\le 15}$ (starting with a_1) as a source of randomness.

3 [6 pts] HMM: Search and Rescue

You are an interplanetary search and rescue expert who has just received an urgent message: a rover on Mercury has fallen and become trapped in Death Ravine, a deep, narrow gorge on the borders of enemy territory. You zoom over to Mercury to investigate the situation. Death Ravine is a narrow gorge 6 miles long, as shown below. There are volcanic vents at locations A and D, indicated by the triangular symbols at those locations.



The rover was heavily damaged in the fall, and as a result, most of its sensors are broken. The only ones still functioning are its thermometers, which register only two levels: *hot* and *cold*. The rover sends back evidence $E = hot$ when it is at a volcanic vent (A and D), and $E = cold$ otherwise. There is no chance of a mistaken reading. The rover fell into the gorge at position A on day 1, so $X_1 = A$. Let the rover's position on day t be $X_t \in \{A, B, C, D, E, F\}$. The rover is still executing its original programming, trying to move 1 mile east (i.e. right, towards F) every day. However, because of the damage, it only moves east with probability 0.5, and it stays in place with probability 0.5. Your job is to figure out where the rover is, so that you can dispatch your rescue-bot.

(a) (2 pt) Three days have passed since the rover fell into the ravine. The observations were ($E_1 = hot$, $E_2 = cold$, $E_3 = cold$). What is $P(X_3|hot_1, cold_2, cold_3)$, the probability distribution over the rover's position on day 3, given the observations?

You decide to attempt to rescue the rover on day 4. However, the transmission of E_4 seems to have been corrupted, and so it is not observed.

(b) (2 pt) What is the rover's position distribution for day 4 given the same evidence, $P(X_4|hot_1, cold_2, cold_3)$?

(c) (2 pt) All this computation is taxing your computers, so the next time this happens you decide to try approximate inference using particle filtering to track the rover. If your particles are initially in the top configuration shown below, what is the probability that they will be in the bottom configuration shown below after one day (after time elapses, but before evidence is observed)?