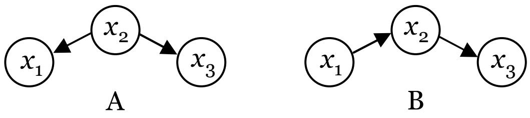
Name: ....

Student ID#: ....

Statistical Pattern Recognition (CE-725)
Department of Computer Engineering
Mini Exam #2 (Graphical Models) - Spring 2012
100 points – 50 minutes

1. (10 points) Consider the following Bayesian nets, which all their variables are binary. Having $P_A(x_1=1,x_2=1,x_3=1) = 0.1$, find $P_B(x_1=1,x_2=1,x_3=1)$.

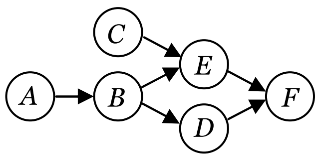


Sol:

$P_A(x_1,x_2,x_3)=P(x_2)P(x_1|x_2)P(x_3|x_2)$. Using Bayesian rule we have, $P(x_1 \mid x_2) = \dfrac{P(x_2 \mid x_1)P(x_1)}{P(x_2)}$. Then, $P_A(x_1,x_2,x_3)=P(x_1)P(x_2|x_1)P(x_3|x_2)=P_B(x_1,x_2,x_3)$. Therefore, $P_B(x_1=1,x_2=1,x_3=1)=0.1$.

2. (30 points) This problem will concern the following Bayesian net. All variables in this network are Boolean, and**,**

- Nodes with a single parent take the value of their parent with probability 3/4 otherwise they take the other value.
- Nodes with two parents take the value of the first parent with probability 1/2 otherwise they take the value of the second parent.
- $P(A=1)=p$, $P(C=1)=q$.



a. If some node X has a single parent Y, and $P(Y=1) = a$, what is a simple expression for $P(X=1)$? Assume that there is no child of X to worry about.

Sol: $P(X=1)=P(X=1|Y=0)P(Y=0)+ P(X=1|Y=1)P(Y=1)=1/4\ (1-a)+3/4a=1/4\ +1/2a$.

b. If some node X has a two independent parents Y, Z, and $P(Y=1)=a$, $P(Z=1)=b$, what is a simple expression for $P(X=1)$?A ssume that there is no child of X to worry about.

Sol:

$$
\begin{aligned}
P(X=1) &= P(X=1|Y=0,Z=0)P(Y=0)P(Z=0)+P(X=1|Y=1,Z=0)P(Y=1)P(Z=0) \\
&\quad +P(X=1|Y=0,Z=1)P(Y=0)P(Z=1)+P(X=1|Y=1,Z=1)P(Y=1)P(Z=1) \\
&= 0+{}^{1}\!/_{2}\,a(1-b)+{}^{1}\!/_{2}(1-a)b+ab={}^{1}\!/_{2}(a+b)
\end{aligned}
$$

c. What is $P(F=1)$ in the above graph?

Sol:

$$
\begin{gathered}
P(B=1)=1/4+1/2\,p \\
P(D=1)=1/4+1/2\,P(B=1)=3/8+1/4\,p \\
P(E=1)=1/2\,P(B=1)+1/2\,P(C=1)=1/8+1/4\,p+1/2\,q \\
P(F=1)=1/2\,P(D=1)+1/2\,P(E=1)=3/16+1/8\,p+1/16+1/8\,p+1/4\,q=1/4+1/4\,p+1/4\,q
\end{gathered}
$$

3. (20 points) Consider an HMM with states $\{w_0, w_1, w_2\}$, observations $\{v_0, v_1, v_2\}$, transition probabilities a_{ij} and symbol probabilities b_{jk}, where,

$$a_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0.2 & 0.3 & 0.5 \\ 0.4 & 0.5 & 0.1 \end{pmatrix} \qquad b_{jk} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.7 & 0.3 \\ 0 & 0.4 & 0.6 \end{pmatrix}$$

a. Suppose the initial hidden state at $t=0$ is w_1. Starting from $t=1$, what is the probability of generating the particular sequence $\{v_2, v_1, v_0\}$ by the given HMM?

Sol:

The probability of observing the given sequence is 0.03678. See the following table for the details.

	$t_1=v_2$	$t_2=v_1$	$t_3=v_0$
w_0	0	0	(.1239*.2+.03*.4)*1=.03678
w_1	.3*.3=.09	(.09*.3+ .3*.5)*.7=.1239	0
w_2	.5*.6=.3	(.09*.5+.3*.1)*.4=.03	0

b. Given the above sequence ($\{v_2, v_1, v_0\}$), what is the most probable sequence of hidden states?

Sol:

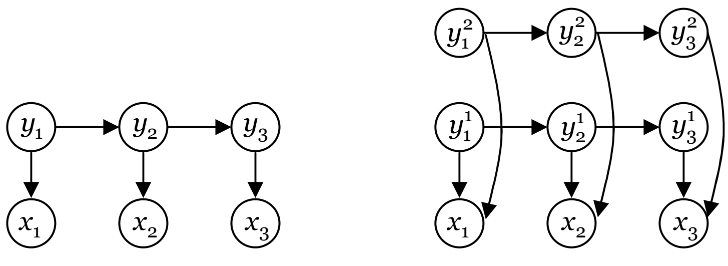
The most probable sequence of hidden states is $\{w_1, w_2, w_1, w_0\}$. See the following Viterbi table for more details.

	$t_1=v_2$	$t_2=v_1$	$t_3=v_0$
w_0	0	0	max{.105*.2,.018*.4}*1=.021
w_1	.3*.3=.09	max{.09*.3, .3*.5}*.7=.105	0
w_2	.5*.6=.3	max{.09*.5,.3*.1}*.4=.018	0

4. (40 points) Consider the hidden Markov model in Fig. (a) that consists of three observations $\{x_1, x_2, x_3\}$ and three hidden states $\{y_1, y_2, y_3\}$.

a. Write down the joint probability distribution $P(x_1,x_2,x_3,y_1,y_2,y_3)$.

Sol: $\pi(y_1)P(x_1 | y_1)P(y_2 | y_1)P(x_2 | y_2)P(y_3 | y_2)P(x_3 | y_3)$



a) HMM b) FHMM

b. In an HMM, the observations are generated from a sequence of hidden states. However, in some applications, like speech, biology and NLP, the observations are known to be generated from the interaction of multiple independent state sequences. This gives rise to the Factorial HMM (FHMM) model depicted in Fig. (b). Fig. (b) shows two independent state sequences: $\{y_1^1, y_2^1, y_3^1\}$ and $\{y_1^2, y_2^2, y_3^2\}$. At each time step, say t, the observation, x_t is generated based on the values of the hidden states in both sequences: y_t^1 and y_t^2. Write down the joint probability $P(x_1,x_2,x_3,y_1^1,y_2^1,y_3^1,y_1^2,y_2^2,y_3^2)$.

Sol:

$$\left(\pi\left(y_1^1\right)\pi\left(y_1^2\right)P\left(x_1|y_1^1,y_1^2\right)\right)\times\left(P\left(y_2^1|y_1^1\right)P\left(y_2^2|y_1^2\right)P\left(x_2|y_2^1,y_2^2\right)\right)\times\left(P(y_3^1 | y_2^1)P(y_3^2 | y_2^2)P\left(x_3|y_3^1,y_3^2\right)\right)$$

c. The FHMM is an example of a Bayesian network. Use the conditional independence property in

HMM or Bayesian networks to give a simpler expression of the conditional probability $P(y_3^1, y_1^2 \mid y_2^1, y_2^2)$.

Sol:

$$y_3^1 \perp y_1^2 \mid y_2^1, y_2^2 \Rightarrow P\left(y_3^1, y_1^2 \middle| y_2^1, y_2^2\right) = P(y_3^1 \mid y_2^1, y_2^2) P(y_1^2 \mid y_2^1, y_2^2) = P(y_3^1 \mid y_2^1) P(y_1^2 \mid y_2^2)$$

$$, P(y_1^2 \mid y_2^2) = \frac{P(y_2^2 \mid y_1^2) P(y_1^2)}{\sum_{y_1^2} P(y_2^2 \mid y_1^2) P(y_1^2}$$

d. Can we write $P\left(y_2^1, y_2^2 \middle| x_1, x_2, x_3\right)$ as $P\left(y_2^1 \middle| x_1, x_2, x_3\right) P\left(y_2^2 \middle| x_1, x_2, x_3\right)$? Explain your answer.

Sol:

No, because knowing either of the y_2^i lets us explain away the other one (You can also use d-separation rule, considering that x_1, x_2, and x_3 are h2h states, and they are given).

e. Can we write $P(y_2^1, y_2^2)$ as $P(y_2^1) P(y_2^2)$? Explain your answer.

Sol:

Yes, because we have no observations (You can also use d-separation rule, considering that x_1, x_2, and x_3 are h2h states, and they are unknown).