CS 188 Spring 2010	Introduction to Artificial Intelligence	Written 5 Bayes net 2

Due: Thursday 4/1 in 283 Soda Drop Box by 11:59pm (no slip days)

Policy: Can be solved in groups (acknowledge collaborators) but must be written up individually

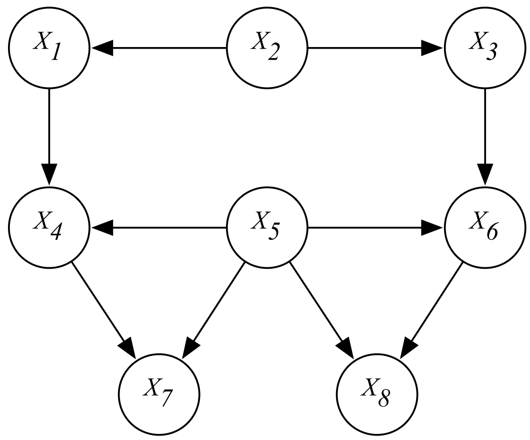
First name	
Last name	
SID	
Login	
Collaborators	

For staff use only:

Q1.	Bayes Nets and Independence	/10
Q2.	Variable Elimination	/10
	Total	/20

Q1. [10 pts] Bayes Nets and Independence

Consider the Bayes net with graph specified below, where all the random variables have domain $\{0, 1\}$.



(a) [2 pts] Find all the conditional independence relations implied by the graph alone of the form $X_i \perp\!\!\!\perp X_j | X_1, X_3, X_7$ (if any).

(b) [2 pts] Assume the following about the conditional probability tables: All of the prior distributions (i.e. $\mathbb{P}(X_2 = \cdot)$ and $\mathbb{P}(X_5 = \cdot)$) are uniform (probability 1/2 for each of the two outcomes) and all the conditional probabilities, except for $X_4|X_1, X_5$ and $X_6|X_3, X_5$, are also uniform. The conditional probabilities of $X_4|X_1, X_5$ and $X_6|X_3, X_5$ are defined in the table below:

x_5	x_1	$\mathbb{P}(X_4 = 0 \| X_5 = x_5, X_1 = x_1)$
0	0	1/3
0	1	1/4
1	0	1/4
1	1	1/3

x_5	x_3	$\mathbb{P}(X_6 = 0 \| X_5 = x_5, X_3 = x_3)$
0	0	1/3
0	1	1/4
1	0	1/3
1	1	1/4

Draw the Bayes net graph, which is a sub-graph of the original graph, which has the minimal number of edges, yet can still represent the probability distribution under consideration.

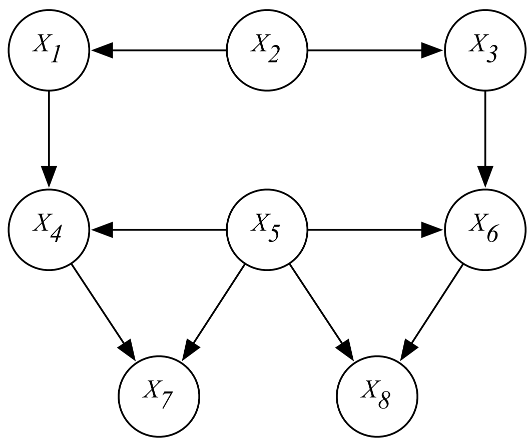
(c) [2 pts]

Assume the conditional probability tables from part (b). Find all the conditional independence relations of the form $X_i \perp\!\!\!\perp X_j | X_1, X_3, X_7$ (if any). *Hint: if you did part (b) correctly, you can determine them from the Bayes net structure you obtained in part (b).*

(d) [2 pts] Compute

$$p_i = \mathbb{P}(X_5 = i | X_1 = 0, X_2 = 0, X_3 = 0, X_4 = 0, X_6 = 0, X_7 = 0, X_8 = 0).$$

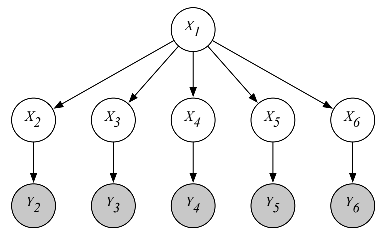
(e) [2 pts] We now reconsider the original Bayes net structure:



Write down a minimal Bayes net on the same set of vertices such that there are directed edges ($X_i \rightarrow X_j$) only if $i > j$, and such that the new Bayes net structure is capable of representing any distribution that can be represented with the original Bayes net structure. What is the number of random variables participating in the largest conditional probability table in your new Bayes net? In this question, you should not use the tables provided in (b), only the structure.

Q2. [10 pts] Variable Elimination

Consider the following Bayes net,



and suppose we want to compute $\mathbb{P}(X_6|Y_2 = y_2, \ldots, Y_6 = y_6)$, 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

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

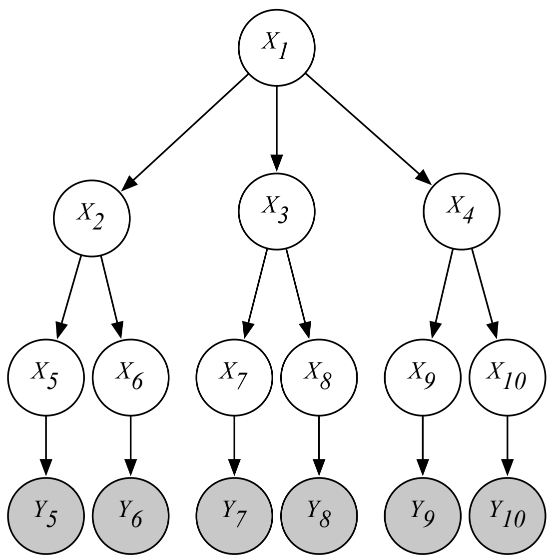
so normalizing over X_6 gives $\mathbb{P}(X_6|Y_2 = y_2, \ldots, Y_6 = y_6)$.

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).

(a) [4 pts] 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?

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

(c) [3 pts] 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.