CS 6730 Probabilistic Graphical Models *Due:* 15 Feb 2018

Home Work 2

Instructor: Harish Guruprasad Ramaswamy – BSB350
TAs: Jom Kuriakose and Manika Agarwal

1. Consider the following distribution over 3 random variables.

$$P(X_1 = x_1, X_2 = x_2, X_3 = x_3) = \frac{1}{Z}\psi(x_1, x_2)\psi(x_2, x_3)\psi(x_1, x_3)$$

where $\psi(a, b) = 1$ if $a \neq b$ and 0 if $a = b$. If the three random variables take values in $\{0, 1\}$, what is Z? If they take values in $\{0, 1, 2\}$, what is Z?

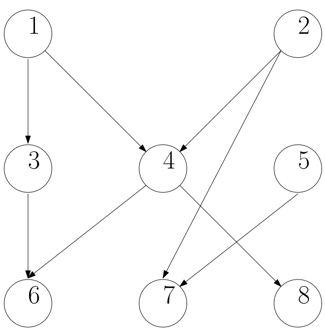
2. Consider the following factorisation of a distribution P over 4 random random variables that take values in $\{1, 2, \ldots, a\}$,

$$P(X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4) = \frac{1}{Z}\psi_1(x_1, x_2)\psi_2(x_2, x_3)\psi_3(x_3, x_4)\psi_4(x_1, x_4)\psi_5(x_2, x_4)\,.$$

Let $\mathcal{H}$ be the Markov network for the factorisation above. If Q is a distribution that factorised according to $\mathcal{H}$, what can be said about the form of the distribution Q? How many parameters would be required to represent such a distribution Q? Compare this with the number of parameters required to represent the distribution P. (You may assume a is large enough that $a - 1 \approx a$ and a^k is negligible when compared to a^{k+1}.)

3. Give the Markov network for the Naive Bayes model, with the class label Y taking values in the set $\{1, \ldots, n\}$ and feature values $X_1, \ldots, X_n$ taking values in $\{0, 1\}$. Give two distinct settings of the factors in the Markov network, so that $P(X_i = 1|Y = j) = 0.9$ if $i = j$ and 0.1 otherwise. Finally, give the Markov network for just the set of feature random variables $X_1, X_2, \ldots, X_n$, and also the factors that make up such a network.

4. Consider the Bayesian network given below:



Consider a distribution over 8 random variables $X_1, \ldots, X_8$ given by the Bayesian network $\mathcal{G}$ to the left.

   (a) Consider a Markov network $\mathcal{H}'$ such that it is equal to $\mathcal{G}$ with the direction of the edges removed. Show that there exists an implied conditional independence in $\mathcal{H}'$ that is not implied by $\mathcal{G}$, i.e. give an element in $\mathcal{I}(\mathcal{H}') \setminus \mathcal{I}(\mathcal{G})$.

   (b) Give the Markov network $\mathcal{H}$ with the fewest edges such that $\mathcal{I}(\mathcal{H}) \subseteq \mathcal{I}(\mathcal{G})$. i.e. such that any distribution that factorises according to the Markov network $\mathcal{H}$, also factorises according to $\mathcal{G}$.

   (c) In this Markov network, what is the largest (in terms of number of variables) conditional independence that can be implied, when X_1 and X_4 are given.

   (d) Are there any conditional independences in $\mathcal{I}(\mathcal{G})$, that is not in $\mathcal{I}(\mathcal{H})$? List them.

5. Let $\mathcal{H}$ be an undirected graph with n nodes. Let $T(\mathcal{H})$ be the set of all chordal graphs with n-nodes that contain all edges in $\mathcal{H}$. The tree width of $\mathcal{H}$ is defined as $\min\{\text{Max-Clique}(\mathcal{H}') - 1 : \mathcal{H}' \in T(\mathcal{H})\}$.

   (a) What is the tree-width of the $n \times n$ grid graph containing n^2 nodes? Give proof. (Hint: Answer scales linearly with n.)

(b) What is the tree width of the cycle graph with n nodes? Give proof. (Hint: Answer does not depend on n.)

6. Consider the factorised distribution given as follows:

$$P(X_1, X_2, X_3, X_4, X_5, X_6) = \frac{1}{Z}\psi_1(X_1, X_2)\psi_2(X_2, X_3)\psi_3(X_3, X_4)\psi_4(X_4, X_5)\psi_5(X_5, X_6)\psi_6(X_6, X_1)$$

   (a) Give the Markov network $\mathcal{H}$ for the above factorisation.
   (b) Let $\mathcal{H}'$ be a triangulated graph of $\mathcal{H}$, with minimal added edges. Give the form of a distribution that factors according to $\mathcal{H}'$.
   (c) Express the distribution P above using the form of factors in the triangulated graph $\mathcal{H}'$.
   (d) Are there any conditional independences in $\mathcal{H}$ that are not present in $\mathcal{H}'$? i.e. give $\mathcal{I}(\mathcal{H}) \setminus \mathcal{I}(\mathcal{H}')$.

7. Consider the following method of constructing an ordering from a connected, chordal graph, called the MC-ordering.

   (1) Initially all nodes in the graph are unmarked.
   (2) Mark any node on the graph, and add it to an empty stack S.
   (3) Repeat step until all nodes are marked.
       - Choose the unmarked node with the largest number of marked neighbours, mark it, and push it to stack S. (Break ties arbitrarily)
   (5) Pop the nodes from the stack, and form the ordering.

   Consider a 6 node graph $\mathcal{H}$, with edges given by $\{(1,2),(1,3),(1,4),(1,5),(1,6),(2,3),(3,4),(4,5),(5,6)\}$.

   (a) Show $\mathcal{H}$ is triangulated and give an MC ordering of the graph.
   (b) Consider a Markov network with such a graph, the 6 random variables $X_1, X_2, \ldots, X_6$ are such that they take values in $\{0, 1, \ldots, a-1\}$. How many total computation steps are required to compute the partition function using the MC elimination ordering?
   (c) For the Markov network above, how many computation steps would be required to compute the partition function of the order of variables to be eliminated are $1, 2, 3, 4, 5, 6$?

   Bonus: Show that if a Markov network is chordal, and variable elimination is done according to the MC ordering no new edges are induced.

8. Write a python code for computing the partition function of an unnormalised Markov network model, using the variable elimination technique, which takes an elimination ordering as input. A template code, and the format of the input, and a few test cases is given as a supplementary file.