Probabilistic Graphical Models

Solution to HWK#2 (part B)
Assigned Sunday, Feb. 9, 2020
Due: Thursday, March 5, 2020

Problem 7

Consider the distribution

$$p(y|x_1,\ldots,x_T)p(x_1)\prod_{t=2}^{T}p(x_t|x_{t-1})$$

where all variables are binary.

1. *Draw a junction tree for this distribution and explain the computational complexity of computing* $p(x_T)$*, as suggested by the junction tree algorithm.*

2. *By using an approach different from the plain JTA above, explain how* $p(x_T)$ *can be computed in time that scales linearly with* T*.*

Solution:

Part (1)

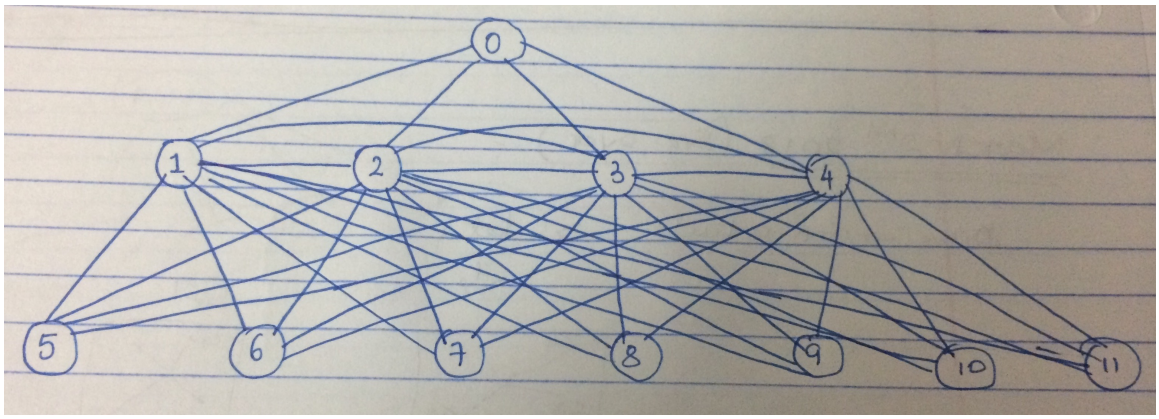
The JT consists of a single clique on all the variables (due to marrying all the $x_1, \ldots, x_T$). The complexity of inference from the JT would suggest that this has complexity of order $O(2^T)$.

Part (2)

If we sum first over y, the remaining distribution on $x_1, \ldots, x_T$ is simply a linear chain, for which inference of $p(x_T)$ is O (T) time

Problem 8. Inference in a Fully Observed Bayesian Network

Solution: Part (a)



The moralized graph is shown above. If the above picture is too messy, consider the original graph. Make all the edges undirected. Connect Node 1 to 2, 1 to 3, 1 to 4, 2 to 3, 2 to 4 and 3 to 4. This gives you the moralized graph.

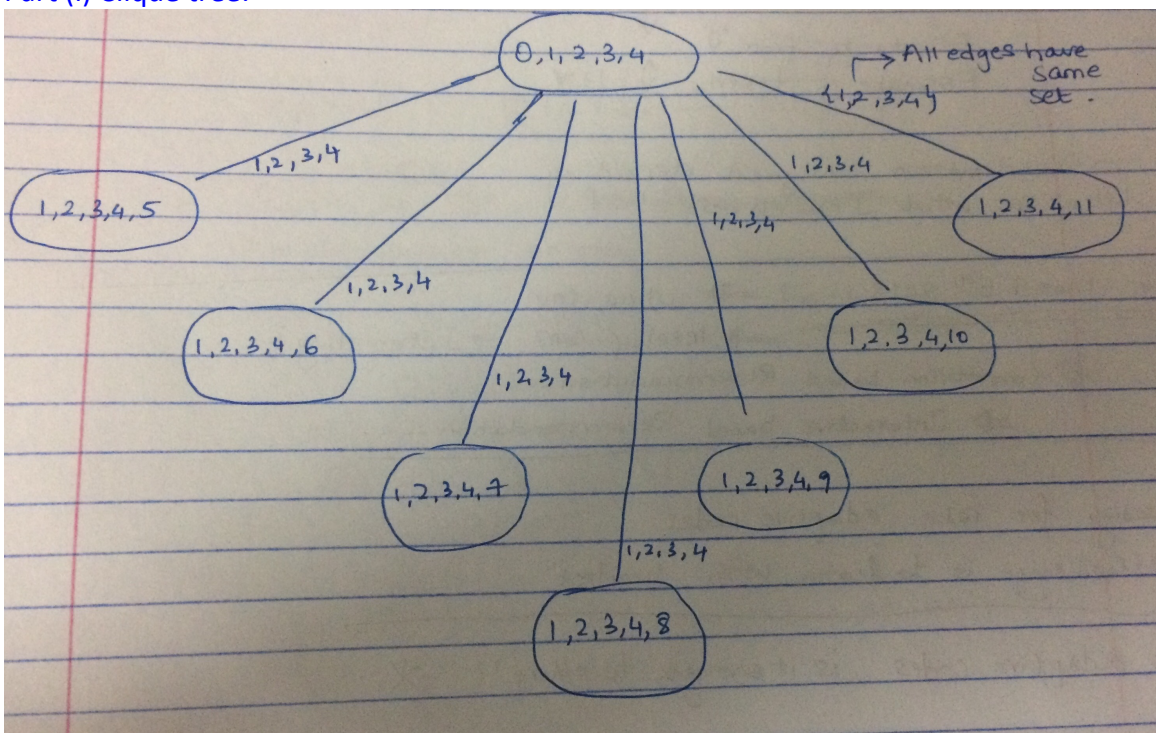
Node elimination order	Factor produced
11	$m_{11}(1,2,3,4) = p(x'_{11}\|x_1,x_2,x_3,x_4) = \sum_{x_{11}} p(x_{11}\|x_1,x_2,x_3,x_4)\delta(x_{11}-x'_{11})$
10	$m_{10}(1,2,3,4) = 1$
9	$m_9(1,2,3,4) = 1$
8	$m_8(1,2,3,4) = p(x'_8\|x_1,x_2,x_3,x_4)$
7	$m_7(1,2,3,4) = 1$
6	$m_6(1,2,3,4) = 1$
5	$m_5(1,2,3,4) = 1$
0	$m_0(1,2,3,4) = \sum_{x_0} p(x_1\|x_0)p(x_2\|x_0)p(x_3\|x_0)p(x_4\|x_0)$
4	$m_4(1,2,3) = \sum_{x_4} m_0(1,2,3,4)\, m_{11}(1,2,3,4) m_8(1,2,3,4)$
3	$m_3(1,2) = \sum_{x_4} m_4(1,2,3)$
2	$m_2(1) = \sum_{x_2} m_4(1,2)$

In the above table $m_{11}(1,2,3,4)$ is shorthand for $m_{x_{11}}(x_1, x_2, x_3, x_4)$ and x'_{11} and x'_8 indicate true values for nodes 8 and 11. Once the final $m_2(1)$ normalize to get the require query distribution.

Time complexity: : O(2^5) because one out of 5 variables from factors.

Part (b)

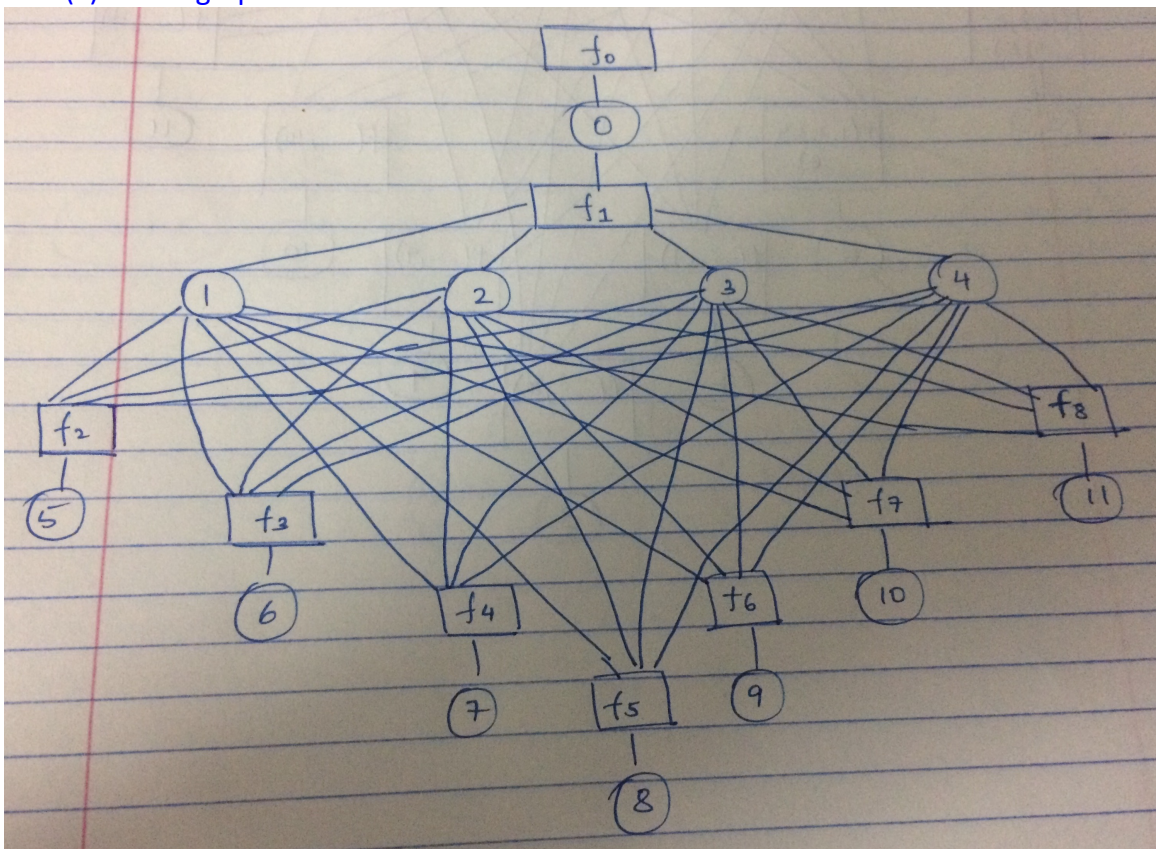
Part (i) Clique tree:



The message in the factor graphs correspond to the factors where all variables except the one corresponding to the variables present on the edge along which the message is being passed have been eliminated.

Time complexity: O(2^5) because one out of 5 variables from clique is eliminated during the message passing.

Part (ii) Factor graph:



From variable to factor nodes:

$\mu_{x_8 \to f_5} = 1$
$\mu_{x_2 \to f_1} = \prod_{i=2 \text{ to } 8} \mu_{f_i \to x_2}$
$\mu_{x_0 \to f_0} = \mu_{f_1 \to x_0}$

From factor to variable nodes:

$\mu_{f_2 \to x_1} = \sum_{x_2,x_3,x_4,x_5} f_2 \prod_{i=2 \text{ to } 5} \mu_{x_i \to f_2}$
$\mu_{f_1 \to x_0} = \sum_{x_1,x_2,x_3,x_4} f_1 \prod_{i=1 \text{ to } 4} \mu_{x_i \to f_1}$
$\mu_{f_0 \to x_0} = f_0$

Time complexity: Again O(2^5) because each factor node has 5 variables

Part (c)

See the solution code for estimating probability distribution.

Part (d)

The L-1 distance is 0.330792965088985. This question requires the est. prob. table generated in part c.

Part (e)

See the solution code for part e.

Query answers:

1. For query 1 the answer by VE, and true probability (sampling from true joint distribution, i.e., the "dataset.dat") are all approximately 0.516.

2. For query 2 the distribution over the symptoms is not mentioned here in the report.

3. For query 3 the answer by VE, and sampling from true probability are all approximately 0.097.

Time taken:

Query	**VE run time**	**Counting (sampling) run time**
Query-1	0.067s	26.409s
Query-2	0.114s	26.940s
Query-3	0.053s	30.112s