Mohammadsadegh Abolhasani - u1416052
HW3 CS6190 - Spring 23

Analytical problems

1. First, we will see if a and b are marginally dependent or not.

$$p(a=0) = 0.192 + 0.144 + 0.048 + 0.216 = 0.6,\ p(a=1) = 0.192 + 0.064 + 0.048 + 0.096 = 0.4$$

$$p(b=0) = 0.192 + 0.144 + 0.192 + 0.064 = 0.592,\ p(b=1) = 0.048 + 0.216 + 0.048 + 0.096 = 0.408$$

$$p(a=0, b=0) = 0.192 + 0.144 = 0.336,\ p(a=0) * p(b=0) = 0.6 * 0.592 = 0.3552$$

$$p(a=1, b=0) = 0.192 + 0.064 = 0.256,\ p(a=1) * p(b=0) = 0.4 * 0.592 = 0.2368$$

$$p(a=0, b=1) = 0.048 + 0.216 = 0.264,\ p(a=0) * p(b=1) = 0.6 * 0.408 = 0.2448$$

$$p(a=1, b=1) = 0.048 + 0.096 = 0.144,\ p(a=1) * p(b=1) = 0.4 * 0.408 = 0.1632$$

The inequality that we see tells us that a and b are not independent. Now, we will compute the conditional probabilities w.r.t. c. We know that $p(a, b|c) = p(a|c)p(b|c)$, therefore we will first calculate $p(a|c)$ and $p(b|c)$, and then we will finally calculate $p(a, b|c)$:

$$p(c=0) = 0.192 + 0.048 + 0.192 + 0.048 = 0.48,\ p(c=1) = 0.144 + 0.216 + 0.064 + 0.096 = 0.52$$

$$p(a=0, c=0) = 0.192 + 0.048 = 0.240,\ p(a=0, c=1) = 0.144 + 0.216 = 0.360$$

$$p(a=1, c=0) = 0.192 + 0.048 = 0.240,\ p(a=1, c=1) = 0.064 + 0.096 = 0.160$$

$$p(b=0, c=0) = 0.192 + 0.192 = 0.384,\ p(b=0, c=1) = 0.144 + 0.064 = 0.0.268$$

$$p(b=1, c=0) = 0.048 + 0.048 = 0.096,\ p(b=1, c=1) = 0.216 + 0.096 = 0.312$$

We have for example:

$$p(a=0|c=0) = \frac{p(a=0, c=0)}{p(c=0)} = \frac{0.240}{0.480} = 0.5$$

We will calculate all of the combinations and set up a table to show $p(a|c)$ & $p(b|c)$ and $p(a, b|c)$:

	$c=0$	$c=1$
$a=0$	$p(a\|c) = 0.5$	$p(a\|c) = 0.69$
$a=1$	$p(a\|c) = 0.5$	$p(a\|c) = 0.31$
$b=0$	$p(b\|c) = 0.8$	$p(b\|c) = 0.4$
$b=1$	$p(b\|c) = 0.2$	$p(b\|c) = 0.6$

	$c=0$	$c=1$
$a=0, b=0$	$p(a, b\|c) = 0.4$	$p(a, b\|c) = 0.277$
$a=1, b=0$	$p(a, b\|c) = 0.4$	$p(a, b\|c) = 0.123$
$b=1, a=0$	$p(b, b\|c) = 0.1$	$p(b, b\|c) = 0.415$
$b=1, a=1$	$p(b, b\|c) = 0.1$	$p(b, b\|c) = 0.185$

Now, we check the independency. Take $a=0, b=1, c=0$ for example:

$$p(a=0, b=1, c=0) = 0.1,\ p(a=0|c=0) * p(b=1|c=0) = 0.5 * 0.2 = 0.1$$

And this will hold true foe all the combinations of a, b, c, and therefore, we see that they become independent.

2. The markov blanket of a node in a bayesian network contains much important information about the node including its parents, children, and other nodes that are parents of its children. Given a node x in a markov blanket and all other nodes in it, we know that these nodes are independent from other nodes in the graph.
   For the D-Separation algorithm, we will consider head-to-head and tail-to-tail relationships and consider every path this way between nodes to see if they are not blocked (which shows dependence) or if they are blocked (which shows independence). Here, we will do the exact same thing when roaming to the parents and children of a node.
   Now we consider a markov blanket of the node x. We need to show that conditional distribution for a node x in a directed graph, conditioned on all of the nodes in its Markov blanket, is independent of the remaining variables in the graph.
   Now, assuming different paths that might happen where this Independence is threatened, we have the following cases:
   Outside of the markov blanket, first we turn to the grandparents of node x. they might be dependent, but their children (who are the parents of x) are in x's markov blanket, Therefore, they are independent and thus, the grandparents will be independent, too.
   The next possible path is the relationship between node x and its grandchildren. Since all the children of node x are in its markov blanket, they are independent and so will happen to all their children, which means the grandchildren of node x are independent from it.
   The last step concerns other grandparents of node x's children and grandchildren of it's children. in this case we know that all parents and children of node x are in its markove blanket, so them, their children, and their parents are together in their own markov blanket. This means that for each of them, the path is blocked and so, they are independent.
   For any other node, either they have a path that lies in node$x$'s markov blanket, which according to the above cases is taken care of and they are independent, or they are not connected in a way which automatically means they are independent.
   Based on all above, we can see that node x is independent to all other nodes conditioned to the nodes in its markov blanket.

3. Here, we have:

$$p(a,b,c,d) = p(a)p(b)p(c|a,b)p(d|c)$$

   We need to show that $a \perp\!\!\!\perp b|\emptyset$, and for this, we need to show that a and b are independent. We have:

$$p(a,b) = \int p(a)p(b)p(c|a,b) \int p(d|c)$$

   By looking at the graph we have:

$$\int p(d|c) = 1, \int p(c|a,b) = 1$$

   So we have:

$$p(a,b) = p(a)p(b)$$

   Which tells us a and b are independent of each other. Now we have to prove that $a \not\perp\!\!\!\perp b|d$. We know:

$$p(a,b|d) = \frac{p(a,b,d)}{p(d)}$$

   and for $p(a,b,d)$ we have:

$$p(a,b,d) = \int_c p(a,b,c,d) = \int_c p(a)p(b)p(c|a,b)p(d|c)$$

   From the last part, we know:

$$\int p(c|a,b) = 1$$

Therefore we have:

$$p(a,b,d) = \int_c p(a,b,c,d) = \int_c p(a)p(b)p(d|c)$$

And at last:

$$p(a,b|d) = \frac{\int_c p(a)p(b)p(d|c)}{p(d)}$$

Therefore:

$$p(a,b|d) \neq p(a)p(b)$$

This proves that in general, $a \not\perp\!\!\!\perp b|d$.

4. Here, we have to convert the given figure 1 into an undirected graphical model. As we talked about this in class, we should do a moralization process, which includes adding a link between the parents. I have done this using Draw.io.

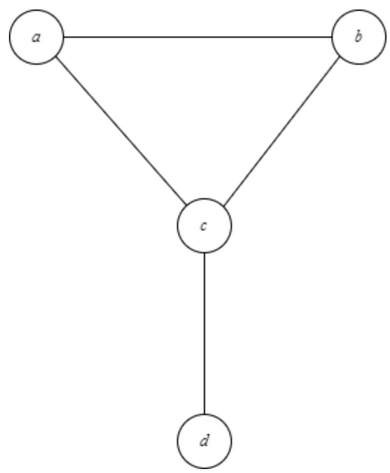


Figure 1: Indirect graph of figure 1 (problem 4)

In this graph we have:

$$p(a,b,c,d) = \psi(a,b)\psi(a,c)\psi(b,c)\psi(c,d)$$
$$\psi(a,b,c) = \psi(a,b)\psi(a,c)\psi(b,c)$$
$$\psi(a,b,c) = p(a)p(b)p(c|a,b)$$

And the same way we have:

$$\psi(c,d) = p(d|c)$$

Therefore:

$$p(a,b,c,d) = p(a)p(b)p(c|a,b)p(d|c)$$

5. Here, we will utilize the sum-product algorithm on the given factor graph. Assuming x_1 as the root node, we will once go from the leaf to the root and once the other way around. First, let us go from the leaf to the node:

$$\mu_{f_d} \rightarrow \mathbf{x}_5(x_5) = f_d(x_5)$$
$$\mu_{x_5} \rightarrow f_c(x_5) = f_d(x_5)$$
$$\mu_{f_c} \rightarrow x_4 = \sum_{x_5} f_c(x_4,x_5)\mu_{f_d} \rightarrow x_5(x_4)$$
$$\mu_{x_4} \rightarrow f_b(x_4) = \mu_{f_c} \rightarrow x_4(x_4)$$

$$\mu_{f_b} \to x_2(x_2) = \sum_{x_4} f_b(x_2, x_4)\mu_{x_4} \to f_b(x_2)$$

$$\mu_{x_3} \to f_a(x_3) = 1$$

$$\mu_{f_a} \to x_1(x_1) = \sum_{x_2} f_a(x_1, x_2, x_3)\mu_{x_2} \to f_a(x_2)\mu_{x_3} \to f_a(x_3) = \sum_{x_2}\sum_{x_3} f_a(x_1, x_2, x_3)\mu_{x_2} \to f_a(x_2)$$

Now we will go through the path from the root to the leaf ($x_5 \to x_1$):

$$\mu_{x_1} \to f_a(x_1) = 1$$

$$\mu_{f_a} \to x_2(x_2) = \sum_{x_1}\sum_{x_3} f_a(x_1, x_2, x_3)\mu_{x_1} \to f_a(x_2)\mu_{x_3} \to f_a(x_2)$$

$$\mu_{f_a} \to x_3(x_2) = \sum_{x_1}\sum_{x_3} f_a(x_1, x_2, x_3)\mu_{x_1} \to f_a(x_2)\mu_{x_2} \to f_a(x_2)$$

$$\mu_{x_2} \to f_b(x_2) = \mu_{f_a} \to x_2 x_2$$

$$\mu_{f_b} \to x_4(x_4) = \sum_{x_2} f_b(x_4, x_2)\mu_{x_2} \to f_b(x_4)$$

$$\mu_{x_4} \to f_c(x_4) = \mu_{f_b} \to x_4(x_4)$$

$$\mu_{f_c} \to x_5(x_5) = \sum_{x_4} f_c(x_4, x_5)\mu_{x_4} \to f_c(x_5)$$

$$\mu_{x_5} \to f_d(x_5) = \mu_{f_c} \to x_5(x_5)$$

Now we can compute the marginal distribution of $p(x_4, x_5)$:

$$p(x_4, x_5) \propto \mu_{f_b} \to x_4(x_4)\mu_{f_d} \to x_5$$

$$p(x_4, x_5) = \frac{1}{Z}\sum_{x_2} f_b(x_4, x_2)\sum_{x_1}\sum_{x_3} f_a(x_1, x_2, x_3) f_d(x_5)$$

Where Z is a normalization constant

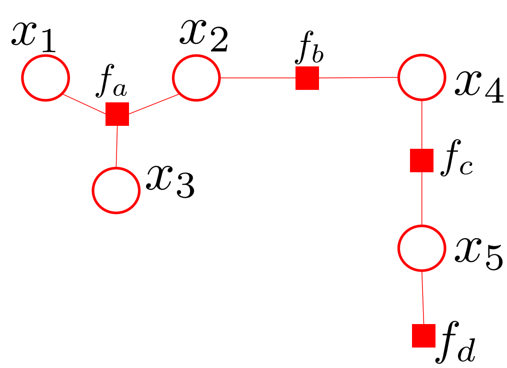


Figure 2: Factor graph.

6. A similar process to the previous question for the same factor graph. Here, we want to calculate the posterior distribution $p(x_4, x_5|x_2)$. Assuming x_1 as the root node, we have to go from the leaf to the root and once the other way around. As it is the exact same as the last question, I will skip this part. Now, we will assume x_2 is observed. Since x_4 and x_5 are neighbor nodes with one factor node

in between, and x_4 has one other neighbor that is x_2 which is now observed to have probability p, with one factor node in between. We have:

$$p(x_4, x_5|x_2 = p) = \frac{p(x_4, x_5, x_2 = p)}{p(x_2)}$$

$$p(x_4, x_5|x_2 = p) = \frac{\mu_{f_a} \to x_2(x_2)\mu_{f_d} \to x_5(x_5)}{\mu_{f_a} \to x_2(x_2)\mu_{f_b} \to x_2(x_2)}$$

$$p(x_4, x_5|x_2 = p) = \frac{\sum_{x_1}\sum_{x_3} f_a(x_1, x_3, x_2)mu_{x_1} \to f_a(x_2)\mu_{x_3} \to f_a(x_2)\mu_{f_d} \to x_5(x_5)}{\sum_{x_1}\sum_{x_3} f_a(x_1, x_3, x_2)mu_{x_1} \to f_a(x_2)\mu_{x_3} \to f_a(x_2)\sum_{x_2} f_b(x_4, x_2)\sum_{x_5} f_c(x_4, x_5)\mu_{f_b} \to (x_5)}$$

$$p(x_4, x_5|x_2 = p) = \frac{\mu_{f_d} \to x_5(x_5)}{\sum_{x_2} f_b(x_4, x_2)\sum_{x_5} f_c(x_4, x_5)\mu_{f_b} \to (x_5)}$$

7. The joint probability for all the nodes is:

$$p(x_1, x_2, x_3, x_4, x_5) = f_a(x_1, x_2, x_3)f_b(x_2, x_3)f_c(x_4, x_5)f_d(x_5)$$

For the purpose of numerical stability, we take the logarithm of this equation, and then attempt to apply the max-sum algorithm:

$$maxlnp(x_1, x_2, x_3, x_4, x_5) = max_{x_1,x_2,x_3}lnf_a(x_1, x_2, x_3)+max_{x_2,x_3}lnf_b(x_2, x_3)+max_{x_4,x_5}lnf_c(x_4, x_5)+max_{x_5}lnf_d(x_5)$$

We set the initialization amount to zero, now, we will go from the leaf to the root:

$$\mu_{f_d} \to x_5(x_5) = maxlnf_d$$

$$\mu_{x_5} \to f_c(x_5) = maxlnf_d$$

$$\mu_{f_d} \to x_4(x_4) = max_{x_5}lnf_c(x_4, x_5) + \mu_{x_5} \to f_d(x_5)$$

$$\mu_{x_4} \to f_b(x_4) = \mu_{f_c} \to x_4(x_4)$$

$$\mu_{f_b} \to x_2(x_2) = max_{x_4}lnf_b(x_4, x_2) + \mu_{x_4} \to f_b(x_4)$$

$$\mu_{x_3} \to f_a(x_3) = 0$$

$$\mu_{x_2} \to f_a(x_2) = \mu_{f_b} \to x_2(x_2)$$

$$\mu_{f_a} \to x_1(x_1) = max_{x_2,x_3}lnf_a(x_1, x_2, x_3) + \mu_{x_2} \to f_a(x_2) + \mu_{x_3} \to f_a(x_3)$$

Now we go from the root to the leaf:

$$\mu_{x_1} \to f_a(x_1) = 0$$

$$\mu_{x_3} \to f_a(x_1) = 0$$

$$\mu_{f_a} \to x_2(x_2) = max_{x_3,x_1}lnf_a(x_1, x_2, x_3) + \mu_{x_1} \to f_a(x_1) + \mu_{x_3} \to f_a(x_3)$$

$$\mu_{f_a} \to x_3(x_3) = max_{x_2,x_1}lnf_a(x_1, x_2, x_3) + \mu_{x_1} \to f_a(x_1) + \mu_{x_2} \to f_a(x_2)$$

$$\mu_{x_2} \to f_b(x_2) = \mu_{f_a} \to x_2(x_2)$$

$$\mu_{f_b} \to x_4(x_4) = max_{x_2}lnf_b(x_2, x_4) + \mu_{x_2} \to f_b(x_2)$$

$$\mu_{x_4} \to f_c(x_4) = \mu_{f_b} \to x_4(x_4)$$

$$\mu_{f_c} \to x_5(x_5) = max_{x_4}lnf_c(x_4, x_5) + \mu_{x_4} \to f_c(x_4)$$

$$\mu_{x_5} \to f_d(x_5) = \mu_{f_c} \to x_5(x_5)$$

According to what we have discussed in class, we need to find the x_{max} that maximizes the mentioned probability. For this, we need to keep track of the values each time that give us the max-sum values and terminate when $x^{max} = argmax_x[\sum \mu_{f_s} \to x(x)]$ and $p^{max} = max_x[\sum \mu_{f_s} \to x(x)]$ and we will have the maximised joint probability of all the nodes x_1, x_2, x_3, x_4, x_5.

8.

9. For this question, we will need to disregard the observed node and all its connections to the graph, and then look at the paths left to see if they show any connection or not. if we do not find any path, we will conclude their independence. Otherwise, they are connected and therefore, dependent. In figure 3, we want to know if node a is conditionally independent of node d observing the node e. here, Given node e is observed, there is still a path between a and d that passes through b, and hence, we cannot conclude that they are independent given e.
In figure 4, we want to know if node a is conditionally independent of node d observing the node b. Here, all paths from a to d through b will be blocked, ther will be no path between them, so we can conclude that node a and d are independent given node b.

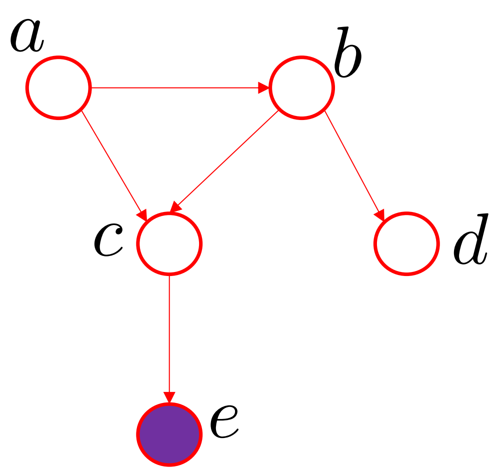


Figure 3: Model 1.

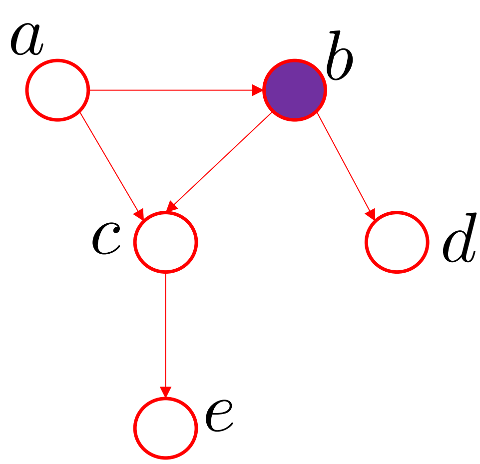


Figure 4: Model 2.

10.

(a) The only undirect graphical model on this graph comes from connecting nodes a and b. Here, since there is a path between a and b, even if c is observed and its connections are deleted, they will always be connected and therefore, not independent.

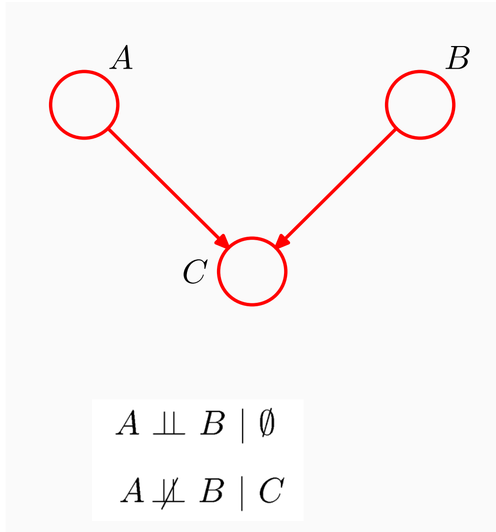
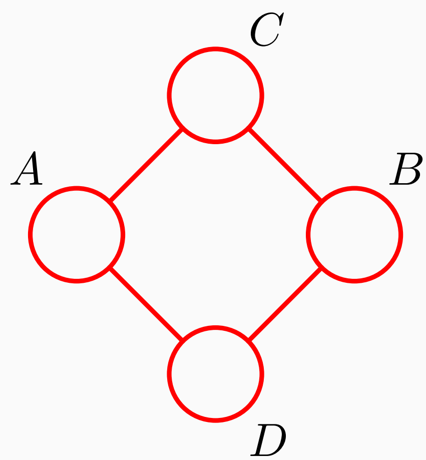


Figure 5: Directed.

(b) Here, we have an undirect graph and we have to get a directed graph. If we remove any path from this, a and b remain connected and it remains a path between them, so they are always dependent and there is no directed graphical model that makes them independent. If we want to add a directed path, it is either between a and b themselves that connects them so they are dependent, or between c and d which will be the same as we first said.
For $a \perp\!\!\!\perp b|c \cup d$ we can add a path between c and d to make this happen, but it will lead to a and b becoming independent so we do not have the same set of independence. The same goes for the third set $c \perp\!\!\!\perp d|a \cup b$. So, we cannot find a directed graph in this graphical model that has all three sets of independence together.



$$A \not\perp\!\!\!\perp B \mid \emptyset$$

$$A \perp\!\!\!\perp B \mid C \cup D$$

$$C \perp\!\!\!\perp D \mid A \cup B$$

Figure 6: Undirected.