Homework Assignment N°1

AML3

Thibault Douzon

Georgios Lioutas

November 30th, 2018

Contents

1 Exercise 1

1.1 Part a

From the topology of the graph, we deduce that:

$$P(S) = \sum_c P(D|C=c) \sum_d P(I|D=d) \sum_i P(S|I=i) \sum_g P(G=g|D=d, I=i)$$

Because $P(G|D=d, I=i)$ is a probabitlity distribution, $\sum_g P(G|D=d, I=i) = 1$. Our final formula is the following:

$$P(S) = \sum_c P(D|C=c) \sum_d P(I|D=d) \sum_i P(S|I=i)$$

$$P(S) = [s_0, s_1] = [0.516, 0.484]$$

1.2 Part b

From the graph we know that

$$P(G, I=i_0) = \sum_c P(D|C=c) \sum_d P(I|D=d) P(G|D=d, I=i0) \sum_s P(S|I=i_0)$$

Once again, $P(S|I=i_0)$ is a probability distribution, then

$$P(G, I=i_0) = \sum_c P(D|C=c) \sum_d P(I|D=d) P(G|D=d, I=i0)$$

$$P(G, I=i_0) = [p(g_0|i_0), p(g_1|i_0), p(g_2|i_0)] = [0.1384, 0.2272, 0.2024]$$

Because it is not a probability distribution, its sum is not 1. From Bayes rule we know that by dividing by $P(I=i_0)$ we would obtain $P(G|I=i_0)$ which is a distribution.

$$P(G|I=i_0) = \frac{P(G, I=i_0)}{P(I=i_0)} = \frac{P(G, I=i_0)}{\sum_g P(G, I=i_0)} = \frac{P(G, I=i_0)}{0.568}$$

$$P(G|I=i_0) = [0.244, 0.4, 0.356]$$

Which adds up to 1 accordingly.

1.3 Part c

Same process as before:

$$P(S, G=g_0) = \sum_c P(D|C=c) \sum_d P(I|D=d) \sum_i P(G=g_0|D=d, I=i) P(S|I=i)$$

$$P(S, G=g_0) = [p(s_0|g_0), p(s_1|g_0)] = [0.245, 0.148]$$

Because it is not a probability distribution, its sum is not 1. From Bayes rule we know that by dividing by $P(G=g_0)$ we would obtain $P(S|G=g_0)$ which is a distribution.

$$P(S|G=g_0) = \frac{P(S, G=g_0)}{P(G=g_0)} = \frac{P(S, G=g_0)}{\sum_s P(S, G=g_0)} = \frac{P(S, G=g_0)}{0.393}$$

$$P(S|G=g_0) = [0.624, 0.376]$$

Which adds up to 1 accordingly.

2 Exercise 2

In this exercise, the graph represents a Markov network. The function Φ is given by the different tables of values.

$$\Phi(C, D, I, G, S) = \Phi(C) \times \Phi(C, D) \times \Phi(D, I) \times \Phi(D, I, G) \times \Phi(I, S)$$

To get a specific value of Φ for a given parameter, we need to marginalize over all other possible parameters.
To compute probabilities we must then normalize Φ by the constant Z such as:

$$Z = \sum_X \Phi(X)$$

where X takes all possible joint probabilities.

$$Z = \sum_c \sum_d \sum_i \sum_g \sum_s \phi(C = c, D = d, I = i, G = g, S = s) = 69200$$

2.1 Part a

We first compute $\Phi(S)$:

$$\Phi(S) = [\Phi(s_0), \Phi(s_1)] = \sum_c \sum_d \sum_i \sum_g \phi(C = c, D = d, I = i, G = g, S)$$

$$\Phi(S) = [25840, 43360]$$

Now we can compute the probability $P(S)$:

$$P(S) = \Phi(S)/Z = [0.373, 0.627]$$

2.2 Part b

We now want to compute $\Phi(G, I = i_0)$. With the same formula as before:

$$\Phi(G, I = i_o) = \sum_c \sum_d \sum_s \Phi(C = c, D = d, I = i_o, G, S = s) = [13680, 19680, 15840]$$

The corresponding probability $P(G, I = i_0)$:

$$P(G, I = i_o) = \frac{1}{Z} \sum_c \sum_d \sum_s \Phi(C = c, D = d, I = i_o, G, S = s) = [0.198, 0.284, 0.229]$$

And now using Baye's rule we can finally obtain the probability distribution $P(G|I = i_0)$:

$$P(G|I = i_0) = \frac{P(G, I = i_0)}{P(I = i_0)} = \frac{P(G, I = i_0)}{\sum_g P(G, I = i_0)}$$

$$P(G|I = i_0) = [0.278, 0.4, 0.322]$$

2.3 Part c

We now want to compute $\Phi(S, G = g_0)$. With the same formula as before:

$$\Phi(S, G = g_o) = \sum_c \sum_d \sum_i \Phi(C = c, D = d, I = i, G = g_0, S) = [12640, 13420]$$

The corresponding probability $P(S, G = g_0)$:

$$P(S, G = g_o) = \frac{1}{Z} \sum_c \sum_d \sum_i \Phi(C = c, D = d, I = i, G = g_0, S) = [0.183, 0.194]$$

And now using Baye's rule we can finally obtain the probability distribution $P(S|G = g_0)$:

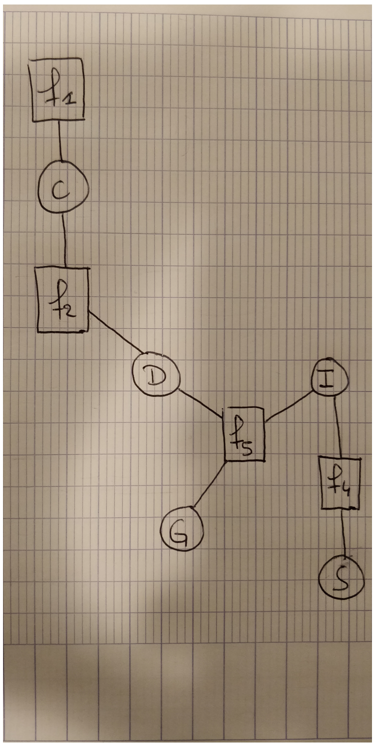
$$P(S|G = g_0) = \frac{P(S, G = g_0)}{P(G = g_0)} = \frac{P(S, G = g_0)}{\sum_s P(S, G = g_0)}$$

$$P(S|G = g_0) = [0.485, 0.515]$$

3 Exercise: Message Passing

3.1 Part a

To use sum-product algorithm, our graph needs to be a tree. Thus it is impossible to keep both $\Phi(D, I)$ and $\Phi(D, I, G)$. We need to combine them together into one single factor node.
Our factor graph is the following:



The factor f_5 is the combination of $\Phi(D, I)$ and $\Phi(D, I, G)$.
It is computed as follows: for every set of value for the variables, multiply together $\Phi(D, I)$ and $\Phi(D, I, G)$. We finally obtain the result table:

f_5	g_0	g_1	g_2
d_0, i_0	2	8	10
d_0, i_1	8	8	4
d_1, i_0	18	24	18
d_1, i_1	14	4	2

3.2 Part b

We now want to compute the probability $P(S)$ using message passing. Node S is now the root of the tree and we start apply message passing from the leaves.

1. $\mu_{f_1 \to C} = \Phi(C) = [3, 7]$
2. $\mu_{C \to f_2} = \Phi(C) = [3, 7]$
3. $\mu_{f_2 \to D} = \Phi(C, D) = [3 \times 2 + 7 \times 3, 3 \times 8 + 7 \times 7] = [27, 73]$
4. $\mu_{D \to f_5} = [27, 73]$
5. $\mu_{G \to f_5} = [1, 1]$
6. $\mu_{f_5 \to I} = \sum_d \sum_g \mu_{D \to f_5\, d} \times \mu_{G \to f_5\, g} \times f_5(d, g) = [4920, 2000]$
7. $\mu_{I \to f4} = [4920, 2000]$
8. $\mu_{f4 \to S} = \Phi(S) = [2 \times 4920 + 8 \times 2000, 8 \times 4920 + 2 \times 2000] = [25840, 43360]$

We can now normalize our last result to obtain a probability distribtution:

$$P(S) = [0.373, 0.627]$$

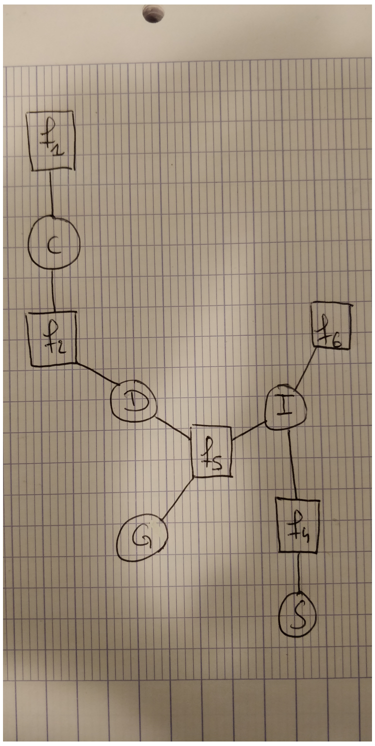
Which is the same result as in Exercise 2.

3.3 Part c

We now want to compute $P(G|I = i_0)$ also using message passing. Because we observe that $I = i_0$, we need to add a factor node to the graph signifying that the variable I is constrained to the value i_0.
We do so by adding a factor node f_6 only linked to I with the follow table of values:

f_6	i_0	i_1
	1	0

It expresses the fact i_0 prevails over i_1.
Our factor graph is now the following:



To apply the message passing, G is now the root of the tree. The messages starting at f_1 will be the same as in the last question until ther get to f_5.

1. $\mu_{f_1 \to C} = \Phi(C) = [3, 7]$
2. $\mu_{C \to f_2} = \Phi(C) = [3, 7]$
3. $\mu_{f_2 \to D} = \Phi(C, D) = [3 \times 2 + 7 \times 3, 3 \times 8 + 7 \times 7] = [27, 73]$
4. $\mu_{D \to f_5} = [27, 73]$
5. $\mu_{S \to f_4} = [1, 1]$
6. $\mu_{f_4 \to I} = [1 \times 2 + 1 \times 8, 1 \times 8 + 1 \times 2] = [10, 10]$
7. $\mu_{f_6 \to I} = [1, 0]$
8. $\mu_{I \to f_5} = [1 \times 10, 0 \times 10] = [10, 0]$
9. $\mu_{f_5 \to G} = \sum_d \sum_i \mu_{D \to f_5\, d} \times \mu_{I \to f_5\, g} \times f_5(d, i) = [13680, 19680, 15840]$

To obtain a probability distribution, we just have to normalize:

$$P(G|I = i_0) = [0.278, 0.4, 0.322]$$

Which is again the same as in the previous exercise.