

# Probabilistic Artificial Intelligence

## Solutions to Problem Set 3

Nov 2, 2018

### 1. Variable elimination

---

In this exercise you will use variable elimination to perform inference on a bayesian network. Consider the network in figure 1 and its corresponding conditional probability tables (CPTs).

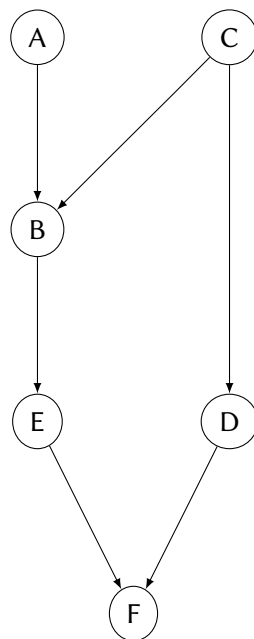


Figure 1: Bayesian network for problem 1.

$$P(A = t) = 0.3 \tag{1}$$

$$P(C = t) = 0.6 \tag{2}$$

Table 1: CPTs for problem 1.

<i>A</i>	<i>C</i>	$P(B = t)$	<i>C</i>	$P(D = t)$	<i>B</i>	$P(E = t)$	<i>D</i>	<i>E</i>	$P(F = t)$
<i>f</i>	<i>f</i>	0.2	<i>f</i>	0.9	<i>f</i>	0.2	<i>f</i>	<i>f</i>	0.95
<i>f</i>	<i>t</i>	0.8	<i>t</i>	0.75	<i>t</i>	0.4	<i>f</i>	<i>t</i>	1
<i>t</i>	<i>f</i>	0.3					<i>t</i>	<i>f</i>	0
<i>t</i>	<i>t</i>	0.5					<i>t</i>	<i>t</i>	0.25

Assuming a query on  $A$  with evidence for  $B$  and  $D$ , i.e.  $P(A | B, D)$ , use the variable elimination algorithm to answer the following queries. Make explicit the selected ordering for the variables and compute the probability tables of the intermediate factors.

1.  $P(A = t | B = t, D = f)$
2.  $P(A = f | B = f, D = f)$
3.  $P(A = t | B = t, D = t)$

Consider now the ordering,  $C, E, F, D, B, A$ , use again the variable elimination algorithm and write down the intermediate factors, this time without computing their probability tables. Is this ordering better or worse than the one you used before? Why?

### Solution

From the problem statement, we want to calculate  $P(\mathbf{Q} | \mathbf{E})$  and we know the following:

$$\mathbf{Q} = \{A\} \tag{3}$$

$$\mathbf{E} = \{B, D\} \tag{4}$$

$$\mathbf{X} \setminus \{\mathbf{Q}, \mathbf{E}\} = \{C, E, F\} \tag{5}$$

The variables to eliminate are given in equation 5. A good ordering must be chosen for them, a possibility is:  $[F, E, C]$ . First let's write down the initial factors derived from the CPTs:

Table 2: Initial factors

$A$	$g_1(A)$	$C$	$g_2(C)$	$C$	$D$	$g_3(C, D)$	$B$	$E$	$g_4(B, E)$
$f$	0.7	$f$	0.4	$f$	$f$	0.10	$f$	$f$	0.8
$t$	0.3	$t$	0.6	$f$	$t$	0.90	$f$	$t$	0.2
				$t$	$f$	0.25	$t$	$f$	0.6
				$t$	$t$	0.75	$t$	$t$	0.4

$A$	$B$	$C$	$g_5(A, B, C)$	$D$	$E$	$F$	$g_6(D, E, F)$
$f$	$f$	$f$	0.8	$f$	$f$	$f$	0.05
$f$	$f$	$t$	0.2	$f$	$f$	$t$	0.95
$f$	$t$	$f$	0.2	$f$	$t$	$f$	0.00
$f$	$t$	$t$	0.8	$f$	$t$	$t$	1.00
$t$	$f$	$f$	0.7	$t$	$f$	$f$	1.00
$t$	$f$	$t$	0.5	$t$	$f$	$t$	0.00
$t$	$t$	$f$	0.3	$t$	$t$	$f$	0.75
$t$	$t$	$t$	0.5	$t$	$t$	$t$	0.25

### 1.1.1 Eliminate $F$

$$g_7(D, E) = \sum_f \prod_{j \in \{6\}} g_j = \sum_f g_6(D, E, f) \quad (6)$$

The resulting factor is presented in table 3. Note that this marginalization trivializes the factor  $g_7(D, E)$  because the conditional query is independent of  $F$ .

Table 3: Intermediate factor  $g_7(D, E)$

$D$	$E$	$g_7(D, E)$
$f$	$f$	$0.05 + 0.95 = 1$
$f$	$t$	$0.00 + 1.00 = 1$
$t$	$f$	$1.00 + 0.00 = 1$
$t$	$t$	$0.75 + 0.25 = 1$

### 1.1.2 Eliminate $E$

$$g_8(B, D) = \sum_e g_4(B, e)g_7(D, e) = \sum_e f_1(B, D, e) \quad (7)$$

A product of the factors,  $f_1(B, D, E) = g_4(B, E)g_7(D, E)$ , is presented in table 4.

Table 4: Product factor  $f_1(B, D, E)$

$B$	$D$	$E$	$f_1(B, D, E)$
$f$	$f$	$f$	$1 \times 0.8 = 0.8$
$f$	$f$	$t$	$1 \times 0.2 = 0.2$
$f$	$t$	$f$	$1 \times 0.8 = 0.8$
$f$	$t$	$t$	$1 \times 0.2 = 0.2$
$t$	$f$	$f$	$1 \times 0.6 = 0.6$
$t$	$f$	$t$	$1 \times 0.4 = 0.4$
$t$	$t$	$f$	$1 \times 0.6 = 0.6$
$t$	$t$	$t$	$1 \times 0.4 = 0.4$

After marginalizing  $E$ , the resulting factor is in table 5.

Table 5: Intermediate factor  $g_8(B, D)$

$B$	$D$	$g_8(B, D)$
$f$	$f$	1
$f$	$t$	1
$t$	$f$	1
$t$	$t$	1

Therefore, the query is independent of  $E$  or  $F$  given the evidence variables.

### 1.1.3 Eliminate $C$

$$f_2(C, D) = g_2(C)g_3(C, D) \quad (8)$$

$$f_3(A, B, C, D) = f_2(C, D)g_5(A, B, C) \quad (9)$$

$$g_9(A, B, D) = \sum_c g_2(c)g_3(c, D)g_5(A, B, c) = \sum_c f_3(A, B, c, D) \quad (10)$$

Tables 6 and 7 show the product of the factors containing  $C$ .

Table 6: Product factor  $f_2(C, D)$

$C$	$D$	$f_2(C, D)$
$f$	$f$	$0.10 \times 0.4 = 0.04$
$f$	$t$	$0.90 \times 0.4 = 0.36$
$t$	$f$	$0.25 \times 0.6 = 0.15$
$t$	$t$	$0.75 \times 0.6 = 0.45$

Table 7: Product factor  $f_3(A, B, C, D)$

$A$	$B$	$C$	$D$	$f_3(A, B, C, D)$
$f$	$f$	$f$	$f$	$0.04 \times 0.8 = 0.032$
$f$	$f$	$f$	$t$	$0.36 \times 0.8 = 0.288$
$f$	$f$	$t$	$f$	$0.15 \times 0.2 = 0.030$
$f$	$f$	$t$	$t$	$0.45 \times 0.2 = 0.090$
$f$	$t$	$f$	$f$	$0.04 \times 0.2 = 0.008$
$f$	$t$	$f$	$t$	$0.36 \times 0.2 = 0.072$
$f$	$t$	$t$	$f$	$0.15 \times 0.8 = 0.120$
$f$	$t$	$t$	$t$	$0.45 \times 0.8 = 0.360$
$t$	$f$	$f$	$f$	$0.04 \times 0.7 = 0.028$
$t$	$f$	$f$	$t$	$0.36 \times 0.7 = 0.252$
$t$	$f$	$t$	$f$	$0.15 \times 0.5 = 0.075$
$t$	$f$	$t$	$t$	$0.45 \times 0.5 = 0.225$
$t$	$t$	$f$	$f$	$0.04 \times 0.3 = 0.012$
$t$	$t$	$f$	$t$	$0.36 \times 0.3 = 0.108$
$t$	$t$	$t$	$f$	$0.15 \times 0.5 = 0.075$
$t$	$t$	$t$	$t$	$0.45 \times 0.5 = 0.225$

Table 8 shows the resulting factor after marginalizing  $C$ .

### 1.1.4 Answering queries

$$P(a | b, d) \propto g_1(a)g_8(b, d)g_9(a, b, d) \quad (11)$$

$$g_1(A = f)g_8(B = t, D = f)g_9(A = f, B = t, D = f) = 0.7 \times 1 \times 0.128 = 0.0896 \quad (12)$$

$$g_1(A = t)g_8(B = t, D = f)g_9(A = t, B = t, D = f) = 0.3 \times 1 \times 0.087 = 0.0261 \quad (13)$$

Table 8: Intermediate factor  $g_9(A, B, D)$

$A$	$B$	$D$	$g_9(A, B, D)$
$f$	$f$	$f$	$0.032 + 0.030 = 0.062$
$f$	$f$	$t$	$0.288 + 0.090 = 0.378$
$f$	$t$	$f$	$0.008 + 0.120 = 0.128$
$f$	$t$	$t$	$0.072 + 0.360 = 0.432$
$t$	$f$	$f$	$0.028 + 0.075 = 0.103$
$t$	$f$	$t$	$0.252 + 0.225 = 0.477$
$t$	$t$	$f$	$0.012 + 0.075 = 0.087$
$t$	$t$	$t$	$0.108 + 0.225 = 0.333$

Re-normalizing,

$$P(A \mid B = t, D = f) = \langle 0.7744, \mathbf{0.2256} \rangle \quad (14)$$

$$g_1(A = f)g_8(B = f, D = f)g_9(A = f, B = f, D = f) = 0.7 \times 1 \times 0.062 = 0.0434 \quad (15)$$

$$g_1(A = t)g_8(B = f, D = f)g_9(A = t, B = f, D = f) = 0.3 \times 1 \times 0.103 = 0.0309 \quad (16)$$

Re-normalizing,

$$P(A \mid B = f, D = f) = \langle \mathbf{0.5841}, 0.4159 \rangle \quad (17)$$

$$g_1(A = f)g_8(B = t, D = t)g_9(A = f, B = t, D = t) = 0.7 \times 1 \times 0.432 = 0.3024 \quad (18)$$

$$g_1(A = t)g_8(B = t, D = t)g_9(A = t, B = t, D = t) = 0.3 \times 1 \times 0.333 = 0.0999 \quad (19)$$

Re-normalizing,

$$P(A \mid B = t, D = t) = \langle 0.7517, \mathbf{0.2483} \rangle \quad (20)$$

### 1.1.5 Another ordering

Let's consider now the ordering  $C, E, F$  for variable elimination. The intermediate factors are:

$$g_7(A, B, D) = \sum_c g_2(c)g_3(c, D)g_5(A, B, c) \quad (21)$$

$$g_8(B, D, F) = \sum_e g_4(B, e)g_6(D, e, F) \quad (22)$$

$$g_9(B, D) = \sum_f g_8(B, D, f) \quad (23)$$

**Is this ordering better?** Note that with the former ordering, only one of the factors has 3 variables, i.e.  $g_9(A, B, D)$  in equation 10. On the other hand, with the latter ordering we would have to compute the table for  $g_8(B, D, F)$  in equation 22 only to marginalize  $F$  in the next step. Therefore, the former ordering is better as it requires less operations to answer the given queries.

## 2. Belief propagation

---

In this exercise, you will implement the belief propagation algorithm for performing inference in Bayesian networks. As you have seen in the class lectures, the algorithm is based on converting the Bayesian network to a factor graph and then passing messages between variable and factor nodes of that graph until convergence.

You are provided some skeleton Python code in the .zip file accompanying this document. Take the following steps for this exercise.

1. Install the Python dependencies listed in README.txt, if your system does not already satisfy them. After that, you should be able to run demo.py and produce some plots, albeit wrong ones for now.
2. Implement the missing code in bprop.py marked with TODO. In particular, you have to fill in parts of the two functions that are responsible for sending messages from variable to factor nodes and vice versa, as well as parts of the function that returns the resulting marginal distribution of a variable node after message passing has terminated.
3. Now, set up the full-fledged earthquake network, whose structure was introduced in Problem Set 2 and is shown again in Figure 2. Here is the story behind this network:

While Fred is commuting to work, he receives a phone call from his neighbor saying that the burglar alarm in Fred's house is ringing. Upon hearing this, Fred immediately turns around to get back and check his home. A few minutes on his way back, however, he hears on the radio that there was an earthquake near his home earlier that day. Relieved by the news, he turns around again and continues his way to work.

To build up the conditional probability tables (CPTs) for the network of Figure 2 you may make the following assumptions about the variables involved:

- All variables in the network are binary.

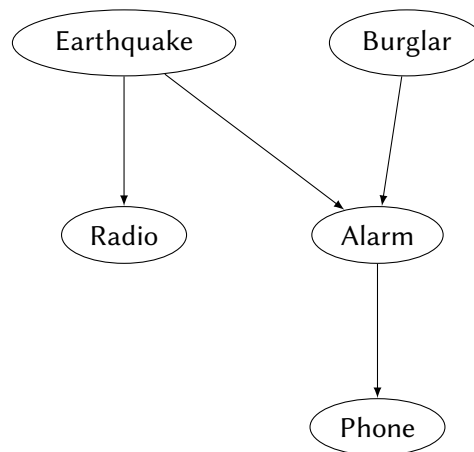


Figure 2: The earthquake network to be implemented.

- As can be seen from the network structure, burglaries and earthquakes are assumed to be independent. Furthermore, each of them is assumed to occur with probability 0.1%.
  - The alarm is triggered in the following ways: (1) When a burglar enters the house, the alarm will ring 99% of the time; (2) when an earthquake occurs, there will be a false alarm 1% of the time; (3) the alarm might go off due to other causes (wind, rain, etc.) 0.1% of the time. These three types of causes are assumed to be independent of each other.
  - The neighbor is assumed to call only when the alarm is ringing, but only does so 70% of the time when it is actually ringing.
  - The radio is assumed to never falsely report an earthquake, but it might fail to report an earthquake that actually happened 50% of the time. (This includes the times that Fred fails to listen to the announcement.)
4. After having set up the network and its CPTs, answer the following questions using your belief propagation implementation:
- (a) Before Fred gets the neighbor's call, what is the probability of a burglary having occurred? What is the probability of an earthquake having occurred?
  - (b) How do these probabilities change after Fred receives the neighbor's phonecall?
  - (c) How do these probabilities change after Fred listens to the news on the radio?

### Solution

The complete solution code can be found in the .zip file accompanying this file. Now we show how to derive the CPT from the problem statement and the expected results to check the obtained solution from your code.

### 2.1.1 Defining the CPTs

First, we show here the derivation of the conditional probability table (CPT) for the alarm variable  $A$ . Let us define the following quantities from the problem description:

- $f_b = 0.99$ : the probability that the alarm will ring because of a burglar,
- $f_e = 0.01$ : the probability that the alarm will ring because of an earthquake,
- $f_o = 0.001$ : the probability that the alarm will ring because of other causes.

Using the facts that these three causes are independent of each other, and that the alarm is triggered by either of them, we can compute the following probabilities:

$$P(A = f \mid B = f, E = f) = 1 - f_o = 0.999 \quad (24)$$

$$P(A = f \mid B = f, E = t) = (1 - f_e) \times (1 - f_o) = 0.98901 \quad (25)$$

$$P(A = f \mid B = t, E = f) = (1 - f_b) \times (1 - f_o) = 0.00999 \quad (26)$$

$$P(A = f \mid B = t, E = t) = (1 - f_b) \times (1 - f_e) \times (1 - f_o) = 0.0098901. \quad (27)$$

The CPTs for the alarm variable is presented in table 9 together with the other variables which can be obtained straightforwardly from the description.

Table 9: CPTs for problem 2.

$B$	$E$	$P(A = t)$	$A$	$P(P = t)$	$E$	$P(R = t)$
$f$	$f$	0.0010000	$f$	0.0	$f$	0.0
$f$	$t$	0.0109900	$t$	0.7	$t$	0.5
$t$	$f$	0.9900100				
$t$	$t$	0.9901099				

The probabilities of the independent variables are:

$$P(E = t) = 0.001 \quad (28)$$

$$P(B = t) = 0.001 \quad (29)$$

### 2.1.2 Queries

The correct results for the queries from point 5 are:

1.  $P(B = t) = 0.1\%$  and  $P(E = t) = 0.1\%$
2.  $P(B = t \mid P = t) = 49.5\%$  and  $P(E = t \mid P = t) = 0.6\%$
3.  $P(B = t \mid P = t, R = t) = 8.3\%$  and  $P(E = t \mid P = t, R = t) = 100\%$



### 3. Gibbs sampling

---

See .zip file on course website.