NTNU

Kunnskap for en bedre verden

Department of Computer Science

TDT4171 Metoder i kunstig intelligens

Exercise 1 - Uncertainty, Bayesian networks

Jostein Hjortland Tysse

Table of Contents

1 Siblings

Table 1 shows the probabilities for the number of siblings a person has.

Siblings	0	1	2	3	4	$5 \leq$
Probability	0.15	0.49	0.27	0.06	0.02	0.01

Table 1: Probabilities

a) What is the probability that a child has at most 2 siblings?

Let x be the number of silblings. We want to find $P(x \leq 2)$. We have a discrete probability distribution, so we can simply use that $P(x \leq 2) = P(x = 0) + P(x = 1) + P(x = 2)$.

$$P(x \leq 2) = P(x = 0) + P(x = 1) + P(x = 2) = 0.15 + 0.49 + 0.27 = 0.91$$

b) What is the probability that a child has more than 2 siblings given that he has at least 1 sibling?

We want to find $P(x > 2 \mid x \geq 1)$.

$$P(x > 2|x \geq 1) = \frac{P(x > 2) \wedge P(x \geq 1)}{P(x \geq 1)} = \frac{P(x > 2)}{P(x \geq 1)} = \frac{1 - P(x \leq 2)}{1 - P(x = 0)} = \frac{1 - 0.91}{1 - 0.15} = \frac{0.09}{0.85} \approx 0.10588$$

c) Three friends who are not siblings are gathered. What is the probability that they combined have three siblings?

We can see in Table 2 that there are 10 different ways that the friends can have 3 siblings combined.

x_1	x_2	x_3
3	0	0
0	3	0
0	0	3
2	1	0
2	0	1
1	2	0
1	0	2
0	2	1
0	1	2
1	1	1

Table 2: Combinations

We need to calculate the sum of these 10 probabilities. The three friends are not siblings, so we know that the variables are independent. Let's start with the scenario where one of the friends has 3 siblings and the other 0.

$$P(x_1 = 3 \wedge x_2 = 0 \wedge x_3 = 0) = 0.06 \cdot 0.15^2$$

There are 3 ways this can happen, so we get $3 \cdot 0.06 \cdot 0.15^2$ probability for this scenario.

Next, consider the scenario where one friend has 2 siblings, one has 1, and the last has 0.

$$P(x_1 = 2 \land x_2 = 1 \land x_3 = 0) = 0.27 \cdot 0.49 \cdot 0.15$$

There are 6 ways $(3 \cdot 2)$ this can happen, so we get $6 \cdot 0.27 \cdot 0.49 \cdot 0.15$ probability for this scenario.

The last scenario where the total is 3 siblings, is when everyone has 1 sibling.

$$P(x_1 = 1 \land x_2 = 1 \land x_3 = 1) = 0.49^3$$

We can now calculate the total probability.

$$P(\text{3 Siblings Combined}) = 3 \cdot 0.06 \cdot 0.15^2 + 6 \cdot 0.27 \cdot 0.49 \cdot 0.15 + 0.49^3 = 0.240769$$

d) Emma and Jacob are not siblings, but combined they have a total of 3 siblings. What is the probability that Emma has no siblings?

Let $x =$ number of siblings for Emma, and y the number of siblings for Jacob. Emma and Jacob are not siblings, so x and y are independent.

We want to find $P(x = 0 \mid x + y = 3)$.

$$P(x = 0 \mid x + y = 3) = \frac{P(x = 0 \land x + y = 3)}{P(x + y = 3)} = \frac{P(x = 0) \cdot P(y = 3)}{P(x + y = 3)}$$

There are 4 ways that Emma and Jacob can have 3 siblings combined. $(x, y) \to (0, 3), (3, 0), (1, 2), (2, 1)$

We then have $P(x + y = 3) = 0.15 \cdot 0.06 + 0.06 \cdot 0.15 + 0.49 \cdot 0.27 + 0.27 \cdot 0.49 = 0.2826$

We can now easily calculate the total probability.

$$P(x = 0 \mid x + y = 3) = \frac{P(x = 0) \cdot P(y = 3)}{P(x + y = 3)} = \frac{0.15 \cdot 0.06}{0.2826} = \frac{5}{157} \approx 0.03185$$

2 Bayesian Networks 1

Given the Bayesian network structure below, decide whether the statements are true or false. Justify each answer with an explanation.

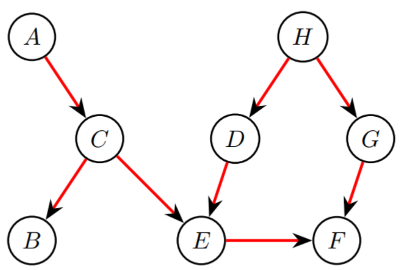

Figure 1: Bayesian Network

a) If every variable in the network has a Boolean state, then the Bayesian network can be represented with 18 numbers.

The amount of numbers needed for nodes in a Bayesian network with only Boolean states depends on the number of parents of each node. If a node has k parents, we need 2^k numbers. Let's investigate the network.

Node	**Parents (k)**	2^k
A	0	1
B	1	2
C	1	2
D	1	2
E	2	4
F	2	4
G	1	2
H	0	1

Table 3: Numbers to represent the network

We can sum the amount of numbers. $1+2+2+2+4+4+2+1=18$

Thus, the statement is **True**

To solve the following exercises, we are going to find out if the nodes are *D-separated*. As described in [1], a trail $X_1 \rightleftharpoons \cdots \rightleftharpoons X_n$ is active given Z if: "for all *v-structures* $X_{i-1} \rightarrow X_i \leftarrow X_{i+1}$, then X_i or one of its descendants are in Z, and no other node along the trail is in Z". If the nodes are *D-separated*, it implies conditionally independence.

b) $G \perp\!\!\!\perp A$

There are two trails from A to G, $A-C-E-F-G$ and $A-C-E-D-H-G$.

Let's start with $A-C-E-F-G$. We have a *v-structure*, $C \rightarrow E \leftarrow D$ and we don't have E or any descendants of E in the evidence set (which is empty). That means that the trail $A-C-E-F-G$ is not active.

When we look at the trail $A-C-E-D-H-G$, we find the *v-structure*, $E \rightarrow F \leftarrow G$. Since F is not in the evidence set, the trail $A-C-E-D-H-G$ is not active either.

Thus, there are no active trails between A and G, which means that they are independent, so the statement is **True**.

c) $E \perp\!\!\!\perp H \mid \{D, G\}$

There are two trails from E to H, $E-F-G-H$ and $E-D-H$.

Let's start with $E-F-G-H$. We have a *v-structure*, $E \rightarrow F \leftarrow G$. Since F is not in the evidence set, the trail $A-C-E-D-H-G$ is not active.

Now, we look at the trail $E-D-H$. This evidence trail is active iff D is not observed. Since D is in the evidence set, the trail is not active.

By D-separation, we find that the statement is **True**.

d) $E \perp\!\!\!\perp H \mid \{C, D, F\}$

We still have the same two trails from E to H, $E - F - G - H$ and $E - D - H$.

$E - D - H$ is still not active, since we are given D.

$E - F - G - H$ however, is actually an active trail. The *v-structure*, $E \rightarrow F \leftarrow G$ is active since F **is in the evidence set**. G is not observed, so we have no other node along the trail which is in the evidence set, thus the trial is active, and the nodes are *not* D-separated. We can state that the nodes are not necessarily independent, but we cannot know for sure, since the implication only goes one way. We need the probability tables to be able to verify for sure that they are not independent, but the tables are not given. We can conclude that the statement is **False**, since we cannot be sure that E and H are always conditionally independent.

3 Bayesian Networks 2

The Bayesian network below contains only binary states. The conditional probability for each state is listed. From the Bayesian network, calculate the following probabilities:

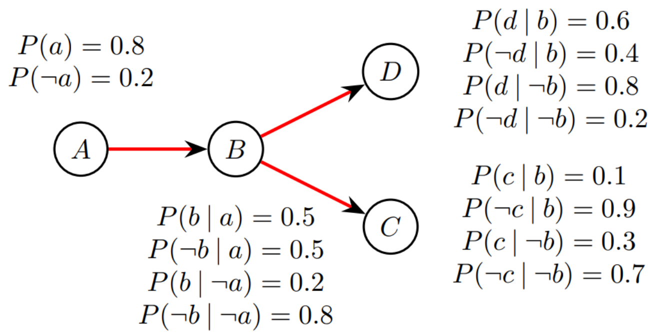

Figure 2: Bayesian Network with probabilities

a) $P(b)$

$$P(b) = P(b \mid a) \cdot P(a) + P(b \mid \neg a) \cdot P(\neg a) = 0.5 \cdot 0.8 + 0.2 \cdot 0.2 = \frac{11}{25} = 0.44$$

b) $P(d)$

We use the result from a) for $P(b)$. We also know that $P(\neg b) = 1 - P(b) = 0.56$

$$P(d) = P(d \mid b) \cdot P(b) + P(d \mid \neg b) \cdot P(\neg b) = 0.6 \cdot 0.44 + 0.8 \cdot 0.56 = \frac{89}{125} = 0.712$$

c) $P(c \mid \neg d)$

To find the values, we can first find the value of $P(b \mid \neg d)$ and then use that to calculate

$$P(c \mid \neg d) = P(c \mid b) \cdot P(b \mid \neg d) + P(c \mid \neg b) \cdot P(\neg b \mid \neg d)$$

Using Bayes' Rule and the values calculated in a) and b), we get the following

$$P(b \mid \neg d) = \frac{P(\neg d \mid b) \cdot P(b)}{P(\neg d)} = \frac{0.4 \cdot 0.44}{1 - \frac{89}{125}} = \frac{11}{18} \approx 0.6111$$

It then follows that $P(\neg b \mid \neg d) = 1 - P(b \mid \neg d) = \frac{7}{18} \approx 0.3889$, and we can now calculate $P(c \mid \neg d)$

$$P(c \mid \neg d) = P(c \mid b) \cdot P(b \mid \neg d) + P(c \mid \neg b) \cdot P(\neg b \mid \neg d) = 0.1 \cdot \frac{11}{18} + 0.3 \cdot \frac{7}{18} = \frac{8}{45} \approx 0.1778$$

d) $P(a \mid \neg c, d)$

To find the probability, we can use *Inference by enumeration* as described in [2]. Figure 3 shows the evaluation process for calculating the probabilities. Let's calculate the values:

$$P(a \mid \neg c, d) = \alpha \cdot 0.8 \cdot (0.5 \cdot 0.9 \cdot 0.6 + 0.5 \cdot 0.7 \cdot 0.7) = \alpha \cdot \frac{11}{25}$$

$$P(\neg a \mid \neg c, d) = \alpha \cdot 0.2 \cdot (0.2 \cdot 0.9 \cdot 0.6 + 0.8 \cdot 0.7 \cdot 0.7) = \alpha \cdot \frac{139}{1250}$$

It follows that

$$P(A \mid \neg c, d) = \alpha \langle \frac{11}{25}, \frac{139}{1250} \rangle$$

$$\alpha = \frac{1}{\frac{11}{25} + \frac{139}{1250}} = \frac{1250}{689}$$

$$\Rightarrow P(A \mid \neg c, d) = \langle \frac{550}{689}, \frac{139}{689} \rangle$$

That is, we have found the solution $P(a \mid \neg c, d) = \frac{550}{689} \approx 0.7983$

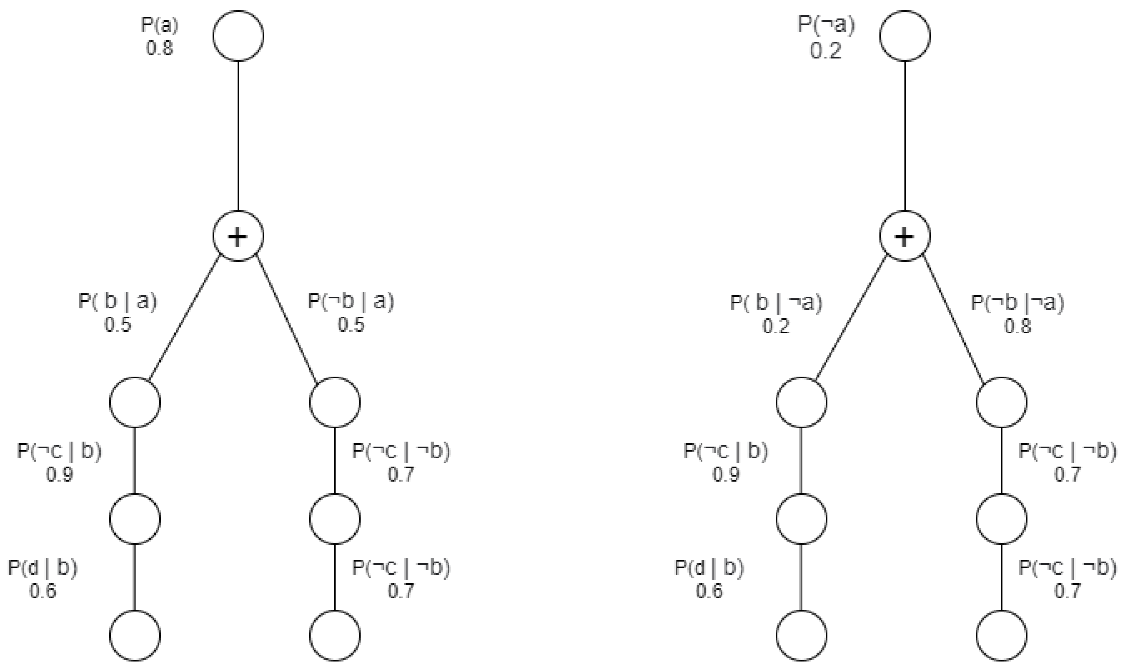

Figure 3: Inferation by Enumeration, Evaluation Process

4 Monty hall

Model the Monty Hall problem as a Bayesian Network using the following states $ChosenByGuest$, $OpenedByHost$, and $Prize$. Use your implementation of inference by enumeration, and the evidence described in the problem statement to answer the question; is it to your advantage to switch your choice? Answering this question entails calculating

$$P(Prize \mid ChosenByGuest = 1, OpenedByHost = 3)$$

The Monty Hall problem is modeled as a Bayesian network in Figure 4.

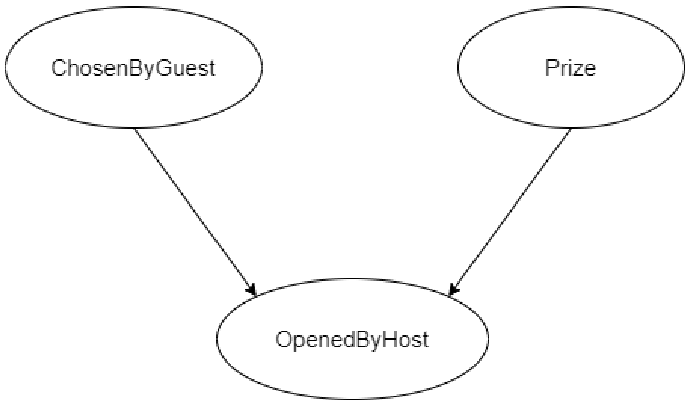

Figure 4: Monty Hall problem as a Bayesian Network

$P(ChosenByGuest = 1)$	$= \frac{1}{3}$
$P(ChosenByGuest = 2)$	$= \frac{1}{3}$
$P(ChosenByGuest = 3)$	$= \frac{1}{3}$

Table 4: Probabiliy Table for $ChosenByGuest$

$P(Prize = 1)$	$= \frac{1}{3}$
$P(Prize = 2)$	$= \frac{1}{3}$
$P(Prize = 3)$	$= \frac{1}{3}$

Table 5: Probabiliy Table for $Prize$

$ChosenByGuest$	1	2	3	1	2	3	1	2	3
$Prize$	1	1	1	2	2	2	3	3	3
$P(OpenedByHost = 1)$	0	0	0	0	0.5	1.0	0	1.0	0.5
$P(OpenedByHost = 2)$	0.5	0	1.0	0	0	0	1.0	0	0.5
$P(OpenedByHost = 3)$	0.5	1.0	0	1.0	0.5	0	0	0	0

Table 6: Conditional Probability Table for $P(OpenedByHost|Prize, ChosenByGuest)$

Results The posterior probabilities calculated by the code, is $\langle 0.3333, 0, 6667, 0, 0000\rangle$. The console output is shown in Figure 5. We then have the following:

$$P(Prize \mid ChosenByGuest = 1, OpenedByHost = 3) = \langle 0.3333, 0, 6667, 0, 0000\rangle$$

Meaning

$$P(Prize = 1 \mid ChosenByGuest = 1, OpenedByHost = 3) = 0.3333$$
$$P(Prize = 2 \mid ChosenByGuest = 1, OpenedByHost = 3) = 0.6667$$
$$P(Prize = 3 \mid ChosenByGuest = 1, OpenedByHost = 3) = 0.0000$$

```
Probability distribution, P(P | G=0, H=2)
+----------+----------+
|   P(0)   |  0.3333  |
+----------+----------+
|   P(1)   |  0.6667  |
+----------+----------+
|   P(2)   |  0.0000  |
+----------+----------+
```

Figure 5: Results from python script.

Conclusion The calculation shows that it always to your advantage to switch your choice of door, as the probabilities for the price being behind the first chosen and the remaining door is $\frac{1}{3}$ and $\frac{2}{3}$, respectively, after the host opens one door.

PS: Note that in the code, we use 0-index, so $P(Prize = 1)$ would be $P(0)$ in the code.

References

[1] Daphne Koller and Nir Friedman. *Probabilistic Graphical Models, Principles and Techniques.* 1. Massachusetts Institute of Technology, 2009.

[2] Stuart Russel and Peter Norvig. *Artificial Intelligence, A Modern Approach.* 3. Pearson Education, inc., 2010.