TDT4136 - Artificial Intelligence Methods
Assignment 1: uncertainty, bayesian networks

Elias Søvik Gunnarsson, MTIØT

January 22, 2021

Problem 1

The probability that a person has 0, 1, 2, 3, 4, or 5 or more siblings is 0.15, 0.49, 0.27, 0.06, 0.02, 0.01, respectively.

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

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

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

$$P(X > 2|X \geq 1) = \frac{P(X > 2 \wedge X \geq 1)}{P(X \geq 1)} = \frac{P(X > 2)}{P(X \geq 1)} = \frac{1 - P(X \leq 2)}{1 - P(X = 0)} = \frac{1 - 0.91}{1 - 0.15} = \frac{0.09}{0.85} = 0.11$$

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

$$\begin{aligned}P(X+Y+Z = 3) = P(X = 1 \wedge Y = 1 \wedge Z = 1) + 3P(X = 3 \wedge Y = 0 \wedge Z = 0) + 6P(X = 0 \wedge Y \\ = 1 \wedge Z = 2) = P(X = 1) \times P(Y = 1) \times P(Z = 1) + 3 \times P(X = 3) \times P(Y = 0) \times P(Z = 0) \\ + 6 \times P(X = 0) \times P(Y = 1) \times P(Z = 2) = 0.15^3 + 3 \times 0.06 \times 0.15^2 + 6 \times 0.15 \times 0.49 \times 0.27 \\ = 0.12\end{aligned}$$

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?*

$$\begin{aligned}P(X = 0|X+Y = 3) = \frac{P(X = 0 \wedge X + Y = 3)}{P(X + Y = 3)} = \frac{P(X = 0 \wedge X + Y = 3)}{2 \times P(X = 3 \wedge Y = 0) + 2 \times P(X = 2 \wedge Y = 1)} \\ = \frac{P(X = 0) \times P(Y = 3)}{2 \times P(X = 3) \times P(Y = 0) + 2 \times P(X = 2) \times P(Y = 1)} = \frac{0.15 \times 0.06}{2 \times 0.06 \times 0.15 + 2 \times 0.27 \times 0.49} \\ = 0.032\end{aligned}$$

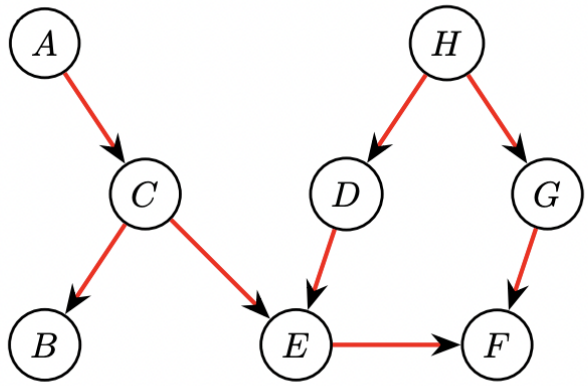

Figure 1: Bayesian network for problem 2.

Problem 2

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

Variables A and H have zero parents each which gives $1 \times 1 = 2$ numbers. B, C, D and G have one parent each which gives $4 \times 2 = 8$ numbers. E and F have two parent each which gives $2 \times 4 = 8$ numbers. In total the Bayesian network can be represented with $2 + 8 + 8 = 18$ numbers.

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

We have that $P(G|A) = P(G)$ and $P(A|G) = P(A)$, thus the random variables G and A are independent.

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

Path $E \longleftarrow D \longleftarrow H$ is blocked by conditioning on D. Path $E \longrightarrow F \longleftarrow G \longrightarrow H$ is blocked by conditioning on G. E and H are d-seperated by D and G, consequently, E and H are conditional independent given D and G.

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

By introducing the conditioning on F, the second path is not blocked by a collider around F since it is being conditioned and not d-separated. Therefore one can not guarantee that E and H are conditional independent given C, D and G.

Problem 3

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

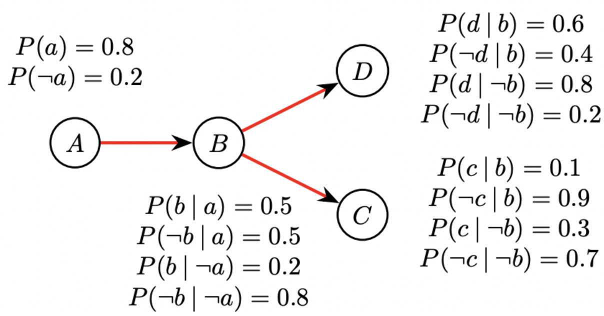

Figure 2: Bayesian network for problem 3.

a) $P(b) = P(b|a) \times P(a) + P(b|\neg a) \times P(\neg a) = 0.5 \times 0.8 + 0.2 \times 0.2 = 0.44$

b) It follows from a) that $P(\neg b) = 1 - 0.44 = 0.56$
$P(d) = P(d|b) \times P(b) + P(d|\neg b) \times P(\neg b) = 0.6 \times 0.44 + 0.8 \times 0.56 = 0.712$

c) Similarly from b) we have that $P(\neg d) = 1 - 0.712 = 0.288$. D and C are conditional independent given B which results in $P(C, D|B) = P(C|B) \times P(D|B)$.
$P(c|\neg d) = \frac{P(c,\neg d)}{P(\neg d)} = \frac{P(c,\neg d) \times P(b) + P(c,\neg d) \times P(\neg b)}{P(\neg d)} = \frac{P(c|b) \times P(\neg d|b) \times P(b) + P(c|\neg b) \times P(d|\neg b) \times P(\neg b)}{P(\neg d)} = \frac{0.1 \times 0.4 \times 0.44 + 0.3 \times 0.2 \times 0.56}{0.288} = 0.178$

d) $P(a|\neg c, d) = \alpha P(a, \neg c, d) = \alpha \sum_b P(a, b, \neg c, d) = \alpha \sum_b P(a) \times P(b|a) \times P(\neg c|b) \times P(d|b) = \alpha \times P(a) \times \sum_b P(b|a) \times P(\neg c|b) \times P(d|b) = \alpha(0.8(0.5 \times 0.9 \times 0.6 + 0.5 \times 0.7 \times 0.8)) = \alpha(0.44)$. Similar calculations for $P(\neg a|\neg c, d) = \alpha(0.1112)$. Hence $P(a|\neg c, d) = 0.798$

Problem 4

Code for Problem 4 is implemented in ***assignment_1.py***.

a) Kahn's algorithm was chosen for the topological ordering of the node. The implementation is based on the pseudocode described in Wikipedia's page about *Topological sorting* [1].

b) The following screenshot in Figure 3 presents the output from running the implementation of the inference by enumeration algorithm found in *Artificial Intelligence: A Modern Approach* on the query from Problem 3c [2].

c) A model of the Monty Hall problem is visualized in Figure 4 with the three random variables *ChosenByGuest*, *OpenedByHost* and *Prize*. Each of these variables has three states, namely door *0*, *1* and *2* for convenience. The results from the implementation of the inference by enumeration algorithm on the Monty Hall problem query is shown in Figure 5. From the results one can conclude that it is to ones advantage to switch the choice after the host opens the door.

```
Probability distribution, P(A)
+----------+----------+
|   A(0)   |  0.8000  |
+----------+----------+
|   A(1)   |  0.2000  |
+----------+----------+

Probability distribution, P(B | A)
+----------+----------+----------+
|    A     |   A(0)   |   A(1)   |
+----------+----------+----------+
|   B(0)   |  0.5000  |  0.2000  |
+----------+----------+----------+
|   B(1)   |  0.5000  |  0.8000  |
+----------+----------+----------+

Probability distribution, P(C | B)
+----------+----------+----------+
|    B     |   B(0)   |   B(1)   |
+----------+----------+----------+
|   C(0)   |  0.1000  |  0.3000  |
+----------+----------+----------+
|   C(1)   |  0.9000  |  0.7000  |
+----------+----------+----------+

Probability distribution, P(D | B)
+----------+----------+----------+
|    B     |   B(0)   |   B(1)   |
+----------+----------+----------+
|   D(0)   |  0.6000  |  0.8000  |
+----------+----------+----------+
|   D(1)   |  0.4000  |  0.2000  |
+----------+----------+----------+

Probability distribution, P(C | !D)
+----------+----------+
|   C(0)   |  0.1778  |
+----------+----------+
|   C(1)   |  0.8222  |
+----------+----------+

(tdt4171) eliassovikgunnarsson@MacBook-Pro assignment_1 %
```

Figure 3: Results from Problem 3c.

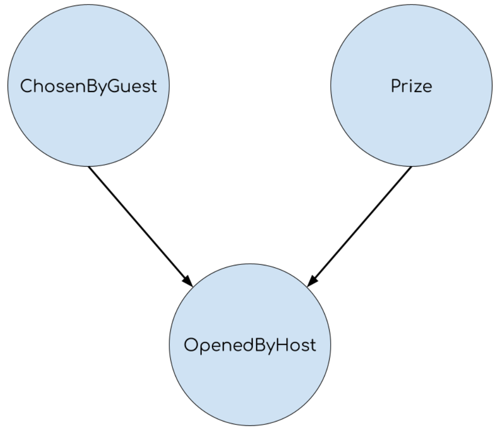

Figure 4: Visual model of the Monty Hall problem as a Bayesian Network.

```
Probability distribution, P(P)
+----------+----------+
|   P(0)   |  0.3333  |
+----------+----------+
|   P(1)   |  0.3333  |
+----------+----------+
|   P(2)   |  0.3333  |
+----------+----------+

Probability distribution, P(C)
+----------+----------+
|   C(0)   |  0.3333  |
+----------+----------+
|   C(1)   |  0.3333  |
+----------+----------+
|   C(2)   |  0.3333  |
+----------+----------+

Probability distribution, P(O | P,C)
+----------+----------+----------+----------+----------+----------+----------+----------+----------+----------+
|    P     |   P(0)   |   P(1)   |   P(2)   |   P(0)   |   P(1)   |   P(2)   |   P(0)   |   P(1)   |   P(2)   |
+----------+----------+----------+----------+----------+----------+----------+----------+----------+----------+
|    C     |   C(0)   |   C(0)   |   C(0)   |   C(1)   |   C(1)   |   C(1)   |   C(2)   |   C(2)   |   C(2)   |
+----------+----------+----------+----------+----------+----------+----------+----------+----------+----------+
|   O(0)   |  0.0000  |  0.0000  |  0.0000  |  0.0000  |  0.5000  |  1.0000  |  0.0000  |  1.0000  |  0.5000  |
+----------+----------+----------+----------+----------+----------+----------+----------+----------+----------+
|   O(1)   |  0.5000  |  0.0000  |  1.0000  |  0.0000  |  0.0000  |  0.0000  |  1.0000  |  0.0000  |  0.5000  |
+----------+----------+----------+----------+----------+----------+----------+----------+----------+----------+
|   O(2)   |  0.5000  |  1.0000  |  0.0000  |  1.0000  |  0.5000  |  0.0000  |  0.0000  |  0.0000  |  0.0000  |
+----------+----------+----------+----------+----------+----------+----------+----------+----------+----------+

Probability distribution, P(P | C=0,O=2)
+----------+----------+
|   P(0)   |  0.3333  |
+----------+----------+
|   P(1)   |  0.6667  |
+----------+----------+
|   P(2)   |  0.0000  |
+----------+----------+

(tdt4171) eliassovikgunnarsson@MacBook-Pro assignment_1 %
```

Figure 5: Implementation of inference by enumeration algorithm on query from Monty Hall problem.

Bibliography

[1] Wikipedia, "Topological sorting — Wikipedia, the free encyclopedia," http://en.wikipedia.org/w/index.php?title=Topological%20sorting&oldid=1001259852, 2021, [Online; accessed 22-Jebruary-2021].

[2] P. N. Stuart Russell, *Artificial Intelligence: A Modern Approach (4th Edition) (Pearson Series in Artifical Intelligence)*, 4th ed. Language: English, 2020.