Methods in Artificial
Intelligence
TDT4171
Spring 2021

Jørgen Rosager
Exercise 1

Norwegian University of Science and
Technology
Department for Computer Science

1 **a)**

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

b)

$$P(Siblings > 2 | Siblings \geq 1) = \frac{P(Siblings > 2 \wedge Siblings \geq 1)}{P(Siblings \geq 1)}$$
$$= \frac{P(Siblings > 2)}{P(Siblings \geq 1)}$$
$$= \frac{1 - P(Siblings \leq 2)}{1 - P(Siblings = 0)}$$
$$= \frac{0.09}{0.85} = \frac{9}{85} \approx 0.1059$$

c) Let S_1, S_2 and S_3 denote the siblings for each of the friends.

There are 3 permutations for who can have 3 siblings and the other 0 siblings, only 1 permutation where each have 1 sibling and $3 \cdot 2 = 6$ permutations for who can have 2 siblings, 1 sibling and 0 sibling (we first have 3 choices for the friend with 2 siblings, then 2 choices for the one with 1 sibling). This gives us:

$$P(S_1 + S_2 + S_3 = 3) = 3P(S_1 = 3)P(S_2 = 0)P(S_3 = 0)$$
$$+ P(S_1 = 1)P(S_2 = 1)P(S_3 = 1)$$
$$+ 6P(S_1 = 2)P(S_2 = 1)P(S_3 = 0)$$
$$= 3 \cdot 0.06 \cdot 0.15^2 + 0.49^3 + 6 \cdot 0.27 \cdot 0.49 \cdot 0.15$$
$$= 0.240769$$

d) Let S_E and S_J denote siblings for Emma and Jacob respectfully. We have:

$$P(S_E = 0 | S_E + S_J = 3) = \frac{P(S_E = 0 \wedge S_E + S_J = 3)}{P(S_E + S_J = 3)}$$
$$P(S_E = 0 \wedge S_E + S_J = 3) = P(S_E = 0)P(S_J = 3)$$
$$P(S_E + S_J = 3) = 2P(S_E = 0)P(S_J = 3) + 2P(S_E = 1)P(S_J = 2)$$
$$\implies P(S_E = 0 | S_E + S_J = 3) = \frac{P(S_E = 0)P(S_J = 3)}{2P(S_E = 0)P(S_J = 3) + 2P(S_E = 1)P(S_J = 2)}$$
$$= \frac{0.15 \cdot 0.06}{2 \cdot 0.15 \cdot 0.06 + 2 \cdot 0.49 \cdot 0.27}$$
$$= \frac{5}{157} \approx 0.031847$$

2   **a)** Every node can be represented by 2^k numbers where k is the number of parent nodes when each variable has a Boolean state. This gives us the following:

Variable	Numbers needed
A	1
B	2
C	2
D	2
E	4
F	4
G	2
H	1
Sum	18

As we can see the sum is 18 numbers needed and the statement is thus **true**.

For solving the next exercises I will use d-separation as described in this document from MIT. The following is a short version of the procedure paraphrased from the document:

1. Draw ancestral graph: Draw a graph with all variables mentioned in the statement with all of their ancestors.

2. Moralize the ancestral graph by "marrying" the parents: Draw a undirected edge between each pair of parents with a common child.

3. Disorient the graph: Replace each directed edge with an undirected edge.

4. Delete the givens and their edges: Remove any given variables (the evidence) and the edges connected to them.

Since d-separation is a sufficient condition for conditional independence, we can use this to determine if the statements are false or true with the given knowledge about the Bayesian network.

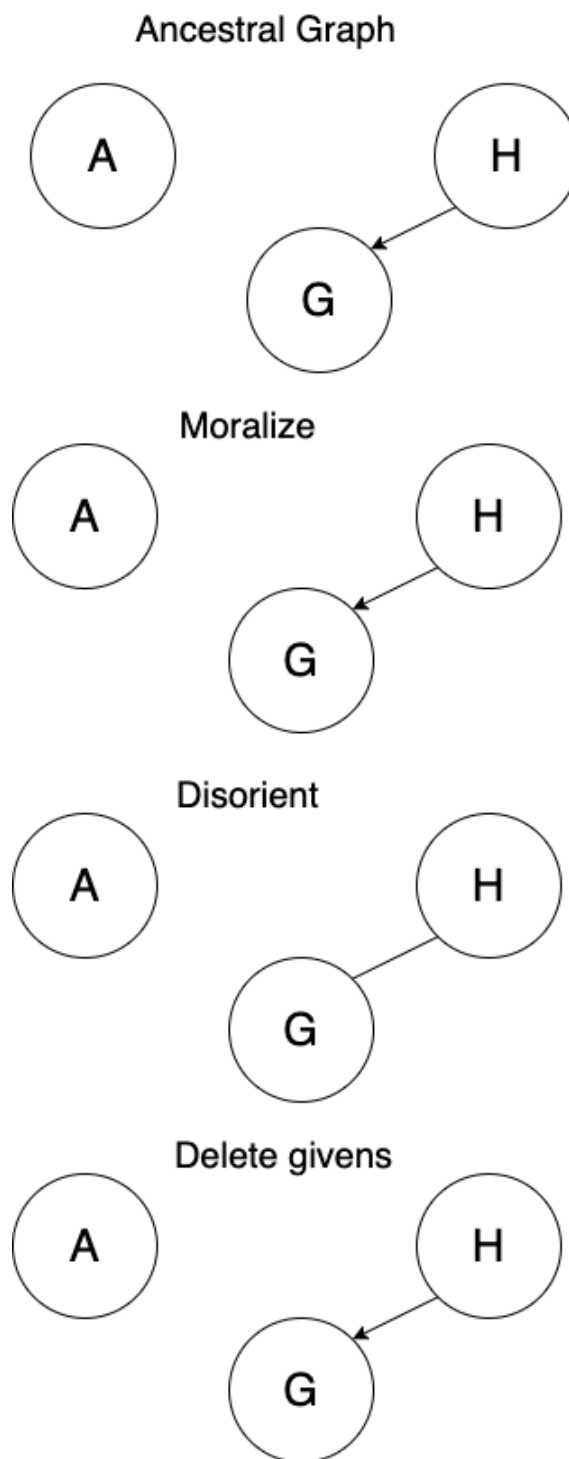**b)** **True**, we can see that G and A are d-separated in 1.

Ancestral Graph

Moralize

Disorient

Delete givens

Figure 1: Using d-separation to determine independence for $G \perp A$.

c) True, we can see that E and H are d-separated in 2.

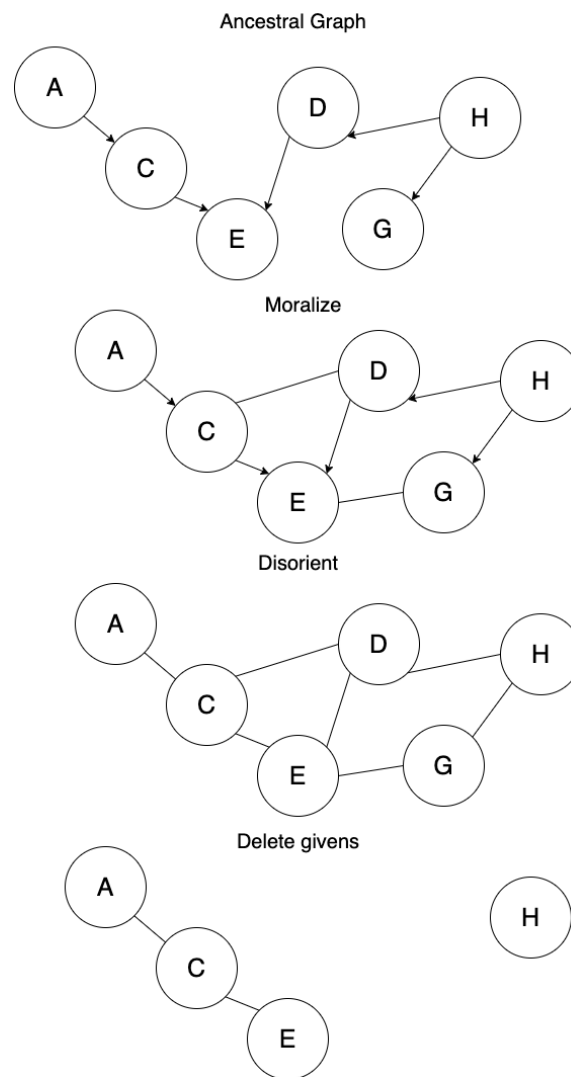Ancestral Graph

Moralize

Disorient

Delete givens

Figure 2: Using d-separation to determine independence for $E \perp H | \{D, G\}$.

d) False, we can see that E and H are **not** d-separated in 3. This means that we cannot **guarantee** independence. However, the specific probabilities involved could make them indepedent, but since we are given none probability tables we don't have enough information to determine this.
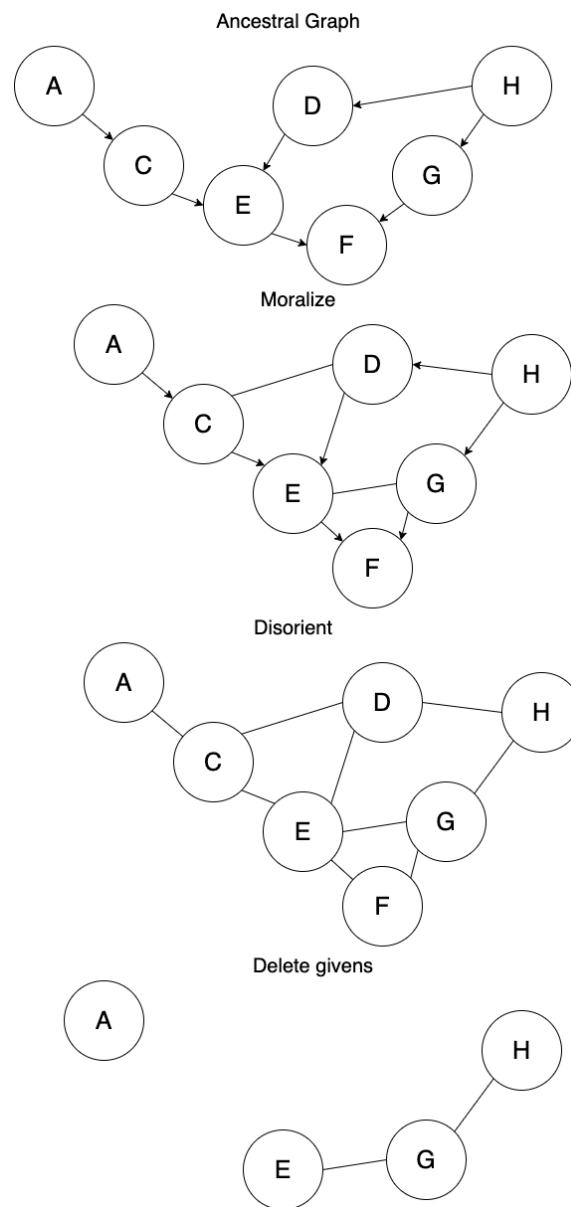
Ancestral Graph

Figure 3: Using d-separation to determine independence for $E \perp H | \{C, D, F\}$.

3 **a)**

$$P(b) = P(b|a)P(a) + P(b|\neg a)P(\neg a)$$
$$= 0.44$$
$$\implies P(\neg b) = 1 - P(b) = 0.56$$

b)

$$P(d) = P(d|b)P(b) + P(d|\neg b)P(\neg b)$$
$$= 0.712$$

c)

$$P(B|\neg d) = \alpha \langle P(\neg d|b)P(b), P(\neg d|\neg b)P(\neg b)\rangle$$
$$= \alpha \langle \frac{22}{125}, \frac{14}{125}\rangle = \langle \frac{11}{18}, \frac{7}{18}\rangle$$
$$\implies P(c|\neg d) = P(c|b)P(b|\neg d) + P(c|\neg b)P(\neg b|\neg d)$$
$$= \frac{11}{180} + \frac{7}{60} = \frac{8}{45} \approx 0.178$$

α is the normalizing factor.

d)

$$P(A|\neg c, d) = \alpha \langle P(a)\sum_b P(\neg b|a)P(\neg c|b)P(d|b), P(\neg a)\sum_b P(\neg b|\neg a)P(\neg c|b)P(d|b)\rangle$$
$$= \alpha \langle \frac{11}{25}, \frac{139}{1250}\rangle = \langle \frac{550}{689}, \frac{139}{689}\rangle \approx \langle 0.7983, 0.2017\rangle$$

α is the normalizing factor.

4   **c)** The Monty-Hall problem can be described by a bayesian network where *OpenedByHost* is conditional based on *Prize* and *ChosenByGuest*:
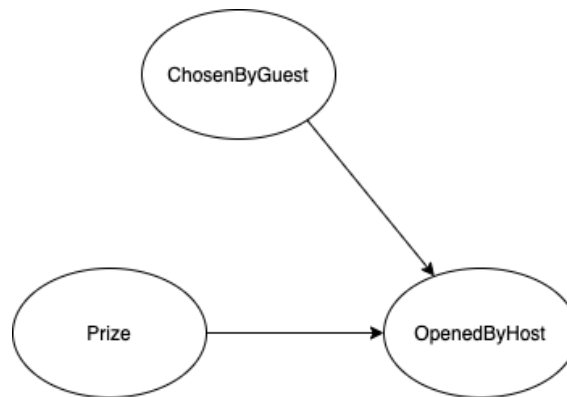

Figure 4: Bayesian network for the Monty Hall-problem.

The following logic was used when constructing the probability tables:

- The prize is chosen randomly between the 3 doors giving a probability of 1/3 to each.
- The guest is choosing randomly between the 3 doors also giving a probability of 1/3 to each.
- The door opened by the host is determined based on the choices of prize and guest, both of which the host knows. The host will never chose the door chosen by prize and guest so the probability of these doors are always 0. Otherwise we have two cases:
  1. If the prize and guest has the same door, the host has two options, giving a probability of 1/2 to both of them.
  2. If the prize and guest's door are not the same, the host only has one option left, giving a probability of 1 to that door.

The conditional probability tables are then given by:

	$P(Prize)$
$Door1$	0.333
$Door2$	0.333
$Door3$	0.333

Table 1: Probability distribution for *Prize*.

	$P(ChosenByGuest)$
$Door1$	0.333
$Door2$	0.333
$Door3$	0.333

Table 2: Probability distribution for *ChosenByGuest*.

Prize	0	1	2	0	1	2	0	1	2
ChosenByGuest	0	0	0	1	1	1	2	2	2
Door1	0	0	0	0.5	0.5	1	0	1	0.5
Door2	0.5	0	1	0	0	0	1	0	0.5
Door3	0.5	1	0	0.5	0.5	0	0	0	0

Table 3: Probability distribution for $ChosenByGuest$.

$$P(Prize|ChosenByGuest = 1, OpenedByHost = 3) = \langle 0.3333, 0.6667, 0.0000 \rangle$$

Based on this it is of the interest of the guest to switch their choice.