

Øving 1 TDT4171 - Uncertainty, Bayesian networks

Problem 1

The probabilities that a person has 0, 1, 2, 3, 4, 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 2 siblings?

$$P(X \leq 2) = \sum_{x=0}^2 P(x) = 0.15 + 0.49 + 0.27 = 0.91 = 91\%$$

Hvor x er "antall søsken".

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

$$\begin{aligned} P(X > 2 | X \geq 1) &= \frac{P(X \geq 1 | X > 2) * P(X > 2)}{P(X \geq 1)} \\ &= \frac{1 - P(X \leq 2)}{1 - P(X \leq 0)} \\ &= \frac{0.09}{0.85} \approx 0.106 \end{aligned}$$

Bruker Bayes' formel. Ser at $P(X \geq 1 | 2) = 1$, så sannsynligheten blir bare ratio mellom $P(X > 2)$ og $P(X \geq 1)$. Vet også at $P(X > x) = 1 - P(X \leq x)$ og bruker dette.

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

Her satt jeg opp alle mulige måter det var mulig å ha tilsammen tre søsken:

- 1 søsken hver: 1 måte, $P(F_1 = 1 \cap F_2 = 1 \cap F_3 = 1)$, hvor F_i er "Friend i ".
- 2 søsken, 1 søsken og 0 søsken: Alle mulige kombinasjoner av $P(F_1 = 2 \cap F_2 = 1 \cap F_3 = 0)$ - 6 måter.
- 3 søsken, 0 søsken og 0 søsken: Alle mulige kombinasjoner av $P(F_1 = 3 \cap F_2 = 0 \cap F_3 = 0)$ - 3 måter.

Dette gir meg regnestykket:

$$3 \cdot (0.06 \cdot 0.15^2) + 6 \cdot (0.27 \cdot 0.49 \cdot 0.15) + 0.49^3 = 0.240769$$

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?

I denne oppgaven antok jeg at sannsynligheten for $P(E + j = 3) = 1$, fordi det er gitt i oppgaveteksten. I tillegg tenkte jeg at sannsynlighetene gitt for 0 -> 5 eller fler søsken var ubrukelige, siden vi har en ny verden hvor $P(X > 3) = 0$.

Vi har da 4 mulige måter Emma og Jacob kan ha 3 søsken til sammen:

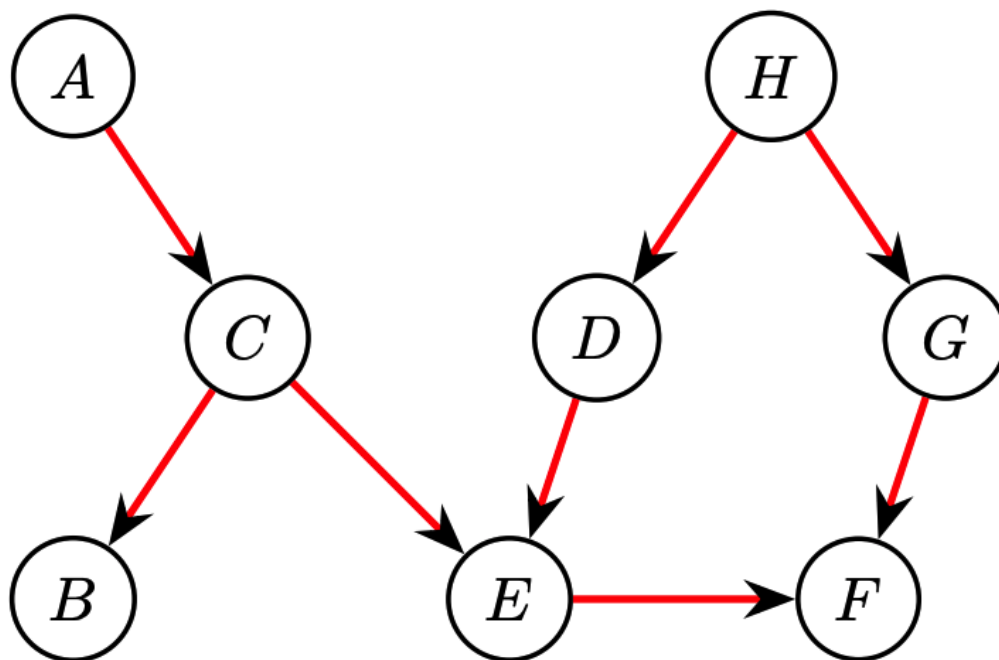
- Emma: 3, Jacob: 0
- Emma: 0, Jacob: 3
- Emma: 1, Jacob: 2
- Emma: 2, Jacob: 1

Av disse fire mulighetene er det én som tilfredsstiller kravet om at Emma skal ha 0, og sannsynligheten blir da $0.25 = 25\%$. Eventuelt kan man sette opp:

$$\begin{aligned} P(E = 0 \mid J + E = 3) &= \frac{P(E = 0 \cap J + E = 3)}{P(J + E = 3)} \\ &= \frac{P(E = 0 \cap J = 3)}{P(J + E = 3)} \\ &= \frac{0.25}{1} = 0.25 \end{aligned}$$

Problem 2

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



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

Hver av nodene i nettverket kan representeres med 2^k tall, der k er antall foreldre.

- A og H: $2 \cdot 2^0 = 2$
- B, C, D og G: $4 \cdot 2^1 = 8$

- E og F: $2 \cdot 2^2 = 8$

$2 + 8 + 8 = 18$. Oppgave a) er **True**.

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

I oppgave b), c) og d) brukte jeg d-separasjon, med reglene:

1. Dersom det finnes en path og en vertex v på denne pathen slik at det er $\rightarrow v \rightarrow$ eller $\leftarrow v \leftarrow$ og $v \in e$, hvor e er evidence settet, er pathen blokkert og de to nodene er d-separert.
2. Dersom det finnes en path og en vertex v på denne pathen slik at det er $\rightarrow v \leftarrow$ og $v \notin e$ og ingen av barnene til v er i e , er pathen blokkert og de to nodene er d-separert.

(her forstod jeg også at *path* i d-separasjon sammenheng er undirected)

I denne oppgaven ser jeg at det ikke er noen blokkert path fra A til G (A - C - E - F - G) f.eks, som vil si at de ikke er uavhengige. Oppgaven er **False**.

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

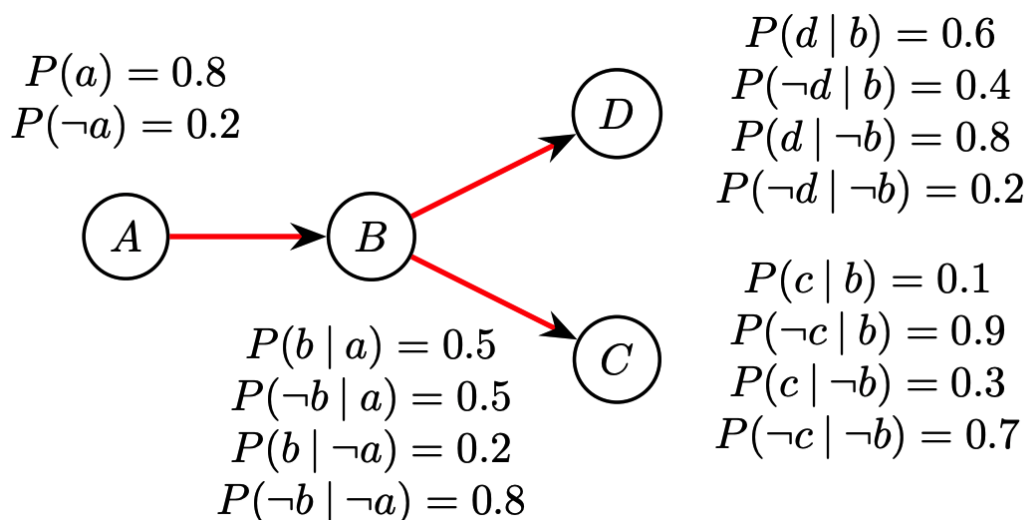
Denne oppgaven er **True**. Bruker regel nummer 1 fra tidligere og ser at både pathen E - D - H er blokkert og E - F - G - H er blokkert, som vil si at E og H er d-separert og dermed også uavhengige.

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

Vi har nå at regel nummer 1 slår ut på pathen H - D - E, men regel nummer 2 blir "motbevist"/invalidert ved at F ligger i evidencen settet vårt. Dette gir at pathen E - F - G - H ikke lenger er blokkert. Påstanden er derfor **False**.

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:



a) $P(b)$

$$\begin{aligned}
 P(b) &= P(b \cap \neg a) + P(b \cap a) \\
 &= P(b|a) \cdot P(a) + P(b|\neg a) \cdot P(\neg a) \\
 &= 0.5 \cdot 0.8 + 0.2^2 \\
 &= \frac{11}{25} \approx 0.44
 \end{aligned}$$

b) $P(d)$

$$\begin{aligned}
 P(d) &= P(d \cap \neg b) + P(d \cap b) \\
 &= P(d|b) \cdot P(b) + P(d|\neg b) \cdot P(\neg b) \\
 &= 0.6 \cdot \frac{11}{25} + 0.8 \cdot (1 - \frac{11}{25}) \\
 &= \frac{89}{125} \approx 0.712
 \end{aligned}$$

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

$$P(c|\neg d) = \frac{P(c \cap \neg d)}{P(\neg d)}$$

Finner $P(c \cap \neg d)$:

$$\begin{aligned}
 P(c \cap \neg d) &= P(b) \cdot P(c|b) \cdot P(\neg d|b) + P(\neg b) \cdot P(c|\neg b) \cdot P(\neg d|\neg b) \\
 &= \frac{11}{25} \cdot 0.1 \cdot 0.4 + (1 - \frac{11}{25}) \cdot 0.3 \cdot 0.2 \\
 &= \frac{32}{625}
 \end{aligned}$$

Plugger dette inn i brøken over:

$$\begin{aligned}
 P(c|\neg d) &= \frac{P(c \cap \neg d)}{P(\neg d)} \\
 &= \frac{\frac{32}{625}}{1 - (\frac{89}{125})} \\
 &= \frac{8}{45} \approx 0.1778
 \end{aligned}$$

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

$$P(a|\neg c, d) = \frac{P(a \cap \neg c \cap d)}{P(\neg c \cap d)}$$

Finner teller og nevner ved hjelp av inference by enumeration:

$$\begin{aligned}
P(a \cap \neg c \cap d) &= \sum_b P(a \cap b \cap \neg c \cap d) \\
&= \sum_b P(a) \cdot P(b|a) \cdot P(\neg c|b) \cdot P(d|b) \\
&= P(a) \cdot \sum_b P(b|a) \cdot P(\neg c|b) \cdot P(d|b) \\
&= 0.8 \cdot (P(b|a) \cdot P(\neg c|b) \cdot P(d|b) + P(\neg b|a) \cdot P(\neg c|\neg b) \cdot P(d|\neg b)) \\
&= 0.8 \cdot (0.5 \cdot 0.9 \cdot 0.6 + 0.5 \cdot 0.7 \cdot 0.8) = \frac{11}{25}
\end{aligned}$$

Og nenver:

$$\begin{aligned}
P(\neg c \cap d) &= \sum_b P(b) \cdot P(\neg c|b) \cdot P(d|b) \\
&= P(b) \cdot P(\neg c|b) \cdot P(d|b) + P(\neg b) \cdot P(\neg c|\neg b) \cdot P(d|\neg b) \\
&= \frac{11}{25} \cdot 0.9 \cdot 0.6 + (1 - \frac{11}{25}) \cdot 0.7 \cdot 0.8 \\
&= \frac{689}{1250}
\end{aligned}$$

Setter dette inn i det originale uttrykket:

$$\begin{aligned}
P(a | \neg c, d) &= \frac{P(a \cap \neg c \cap d)}{P(\neg c \cap d)} \\
&= \frac{\frac{11}{25}}{\frac{689}{1250}} \approx 0.79726
\end{aligned}$$

Problem 4

a) Se vedlagt kode b) Se vedlagt kode c) Se vedlagt kode for utregning av Monty Hall. Se under for sannsynlighetsfordelinger og Bayesian network for Monty Hall problemet.

Output fra koden, problem3c og Monty Hall:

Problem 3c:

Probability distribution, P(C | !D)

+-----+-----+		
C(0)	0.1778	
+-----+-----+		
C(1)	0.8222	
+-----+-----+		

Monty Hall:

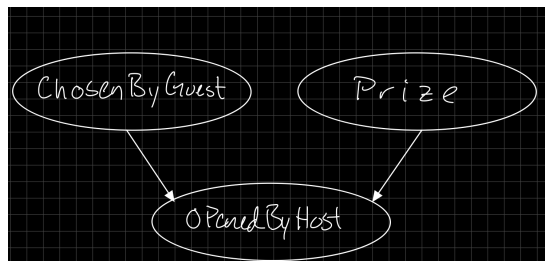
Probability distribution, P(Prize | ChosenByGuest = 1, OpenedByHost = 3)

+-----+-----+		
Prize(0)	0.3333	
+-----+-----+		
Prize(1)	0.6667	
+-----+-----+		

```
| Prize(2) | 0.0000 |
+-----+
```

Is it to my advantage to switch doors, given that I have chosen door 1 and host opens 3?

Yes!



Det er ganske intuitivt at OpenedByHost er gitt av både hvilken dør gjesten åpner, og hvor bilen ligger. (Åpner ikke den gjesten velger, og åpner ikke der bilen er). Derfor peker både ChosenByGuest og Prize på OpenedByHost

Prize 1	Prize 2	Prize 3
$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

Sannsynligheten for at prisen er under dør 1, 2 og 3

Chosen 1	Chosen 2	Chosen 3
$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

Sannsynligheten for at gjest velger dør 1, 2 og 3

	Chosen 1			Chosen 2			Chosen 3		
	P1	P2	P3	P1	P2	P3	P1	P2	P3
Host 1	0	0	0	0	0.5	1	0	1	0.5
Host 2	0.5	0	1	0	0	0	1	0	0.5
Host 3	0.5	1	0	1	0.5	0	0	0	0

Sannsynligheten for at hosten åpner dør 1, 2 og 3, gitt at bilen er under 1, 2 og 3 og gitt at gjesten velger dør 1, 2 og 3