Assignment 3

Seyedeh Mina Mousavifar
CMPT820 - Machine Learning
University of Saskatchewan

March 16, 2020

Question 1

Part (a)

$$P(S=0|F=0) = \underline{0.82}$$

Part (b)

$$\sum_S P(S|F=0) = P(S=0|F=0) + P(S=1|F=0)$$

$$\Rightarrow \sum_S P(S|F=0) = 0.82 + 0.18 = \underline{1}$$

Part (c)

$$P(F=0) = \underline{0.13}$$

Part (d)

$$\sum_S P(S|F=0)P(F=0) = (P(S=0|F=0)P(F=0)) + (P(S=1|F=0)P(F=0))$$

$$\Rightarrow \sum_S P(S|F=0)P(F=0) = (0.82 * 0.13) + (0.18 * 0.13) = \underline{0.13}$$

So $\sum_S P(S|F=0)P(F=0) = P(F=0)$, which shows that S is a nuisance variable.

Part (e)

$$P(F=0|S=0) = \frac{P(S=0|F=0)P(F=0)}{P(S=0)}$$

Which, $P(S=0)$ is the marginal here.

$$\begin{aligned}
P(S=0) &= \sum_F P(S=0,F) = \sum_F P(S=0|F)P(F) \\
&= P(S=0|F=0)P(F=0) + P(S=0|F=1)P(F=1) \\
&= 0.82 * 0.13 + 0.07 * 0.87 \\
&= 0.1675
\end{aligned}$$

$$\Rightarrow P(F=0|S=0) = \frac{0.82 * 0.13}{0.1675} = \underline{0.6364}$$

Part (f)

$$P(F=1|S=0) = \frac{P(S=0|F=1)P(F=1)}{P(S=0)}$$

Which, $P(S=0)$ is the marginal here, which was calculated above.

$$\Rightarrow P(F=1|S=0) = \frac{0.07 * 0.87}{0.1675} = \underline{0.36358}$$

Question 2

Part (a). $P(A=1|F=1)$

Because T is an ancestor of A, it is relevant to the query, but S is irrelevant.

$$P(A=1|F=1) = \frac{P(A=1,F=1)}{P(F=1)}$$

$$\begin{aligned}
P(A=1,F=1) &= \sum_T P(A=1,F=1,T) \\
&= \sum_T P(A=1|F=1,T)P(T)P(F=1) \\
&= P(F=1)\sum_T P(A=1|F=1,T)P(T) \\
&= 0.87 * [0.994 * 0.99 + 0.75 * 0.01] = 0.8626572
\end{aligned}$$

$$\Rightarrow P(A=1|F=1) = \frac{P(A=1,F=1)}{P(F=1)} = \frac{0.8626572}{0.87} = \underline{0.99156}$$

Part (b) $P(F=0|S=0)$

Both T and A are irrelevant to the query, because they are not in the query or an ancestor. With *Bayes' Rule* we have:

$$P(F=0|S=0) = \frac{P(S=0|F=0)P(F=0)}{P(S=0)}$$

$$\begin{aligned}
P(S=0) &= \sum_F P(S=0,F) = \sum_F P(S=0|F)P(F) \\
&= P(S=0|F=0)P(F=0) + P(S=0|F=1)P(F=1) \\
&= 0.82 * 0.13 + 0.07 * 0.87 = 0.1675
\end{aligned}$$

$$\Rightarrow P(F=0|S=0) = \frac{P(S=0|F=0)P(F=0)}{P(S=0)} = \frac{0.82 * 0.13}{0.1675} = \underline{0.6364}$$

Part (c) $P(A=1|S=1)$

Both T and F are relevant to the query, because of F being on the path and T is the ancestor of A. In the below equation the *commutative property of conjunction* is used.

$$P(A=1|S=1) = \frac{P(A=1,S=1)}{P(S=1)}$$

$$\begin{aligned}
P(A=1,S=1) &= \sum_{F,T} P(A=1,S=1,F,T) \\
&= \sum_{F,T} P(A=1|F,T)P(T)P(F)P(S=1|F) \\
&= \sum_T \sum_F P(A=1|F,T)P(T)P(F)P(S=1|F) \\
&= \sum_F P(F)P(S=1|F) \sum_T P(A=1|F,T)P(T) \\
&= \sum_F P(F)P(S=1|F) * [P(A=1|T=0,F)P(T=0) + P(A=1|T=1,F)P(T=1)] \\
&= P(F=0)P(S=1|F=0)P(A=1|F=0,T=0)P(T=0) \\
&+ P(F=0)P(S=1|F=0)P(A=1|F=0,T=1)P(T=1) \\
&+ P(F=1)P(S=1|F=1)P(A=1|F=1,T=0)P(T=0) \\
&+ P(F=1)P(S=1|F=1)P(A=1|F=1,T=1)P(T=1) \\
&= (0.13 * 0.18 * (0.07 * 0.99 + 0.98 * 0.01)) \\
&+ (0.87 * 0.93 * (0.994 * 0.99 + 0.75 * 0.01)) \\
&= 0.00185094 + 0.802271196 \\
&= 0.8041221356
\end{aligned}$$

$$
\begin{aligned}
P(S=1) = \sum_A P(S=1, A) = \sum_A P(A, S=1) &= P(A=0, S=1) + P(A=1, S=1) \\
&= P(F=0)P(S=1|F=0)P(A=0|F=0,T=0)P(T=0) \\
&+ P(F=0)P(S=1|F=0)P(A=0|F=0,T=1)P(T=1) \\
&+ P(F=1)P(S=1|F=1)P(A=0|F=1,T=0)P(T=0) \\
&+ P(F=1)P(S=1|F=1)P(A=0|F=1,T=1)P(T=1) \\
&+ P(F=0)P(S=1|F=0)P(A=1|F=0,T=0)P(T=0) \\
&+ P(F=0)P(S=1|F=0)P(A=1|F=0,T=1)P(T=1) \\
&+ P(F=1)P(S=1|F=1)P(A=1|F=1,T=0)P(T=0) \\
&+ P(F=1)P(S=1|F=1)P(A=1|F=1,T=1)P(T=1) \\
&= (0.13 * 0.18 * (0.93 * 0.99 + 0.02 * 0.01)) \\
&+ (0.87 * 0.93 * (0.006 * 0.99 + 0.25 * 0.01)) \\
&+ (0.13 * 0.18 * (0.07 * 0.99 + 0.98 * 0.01)) \\
&+ (0.87 * 0.93 * (0.994 * 0.99 + 0.75 * 0.01)) \\
&= 0.02154906 + 0.006828804 + 0.00185094 + 0.802271196 \\
&= 0.8325
\end{aligned}
$$

$$
\Rightarrow P(A=0|S=1) = \frac{P(A=0, S=1)}{P(S=1)} = \frac{= 0.8041221356}{0.8325} = \underline{0.96591247}
$$

Question 3

Part (1) Given the empty set, is A conditionally independent of D?

Path 1. A $\rightarrow$ C $\rightarrow$ E $\leftarrow$ D
This path is blocked because E is the convergent of D and C, but has no evidence.
Path 2. A $\rightarrow$ C $\rightarrow$ E $\rightarrow$ F $\rightarrow$ H $\leftarrow$ G $\leftarrow$ D
This path is also blocked because H is the convergent of F and G, but has no evidence.

Consequently, all paths are blocked. So given the empty set, A is conditionally independent of D.

Part (2) Given H, is C conditionally independent of B?

Path 1. C $\rightarrow$ E $\leftarrow$ D $\leftarrow$ B
This path is active because the evidence in H propagates through F to E, which is the convergent of C and D. This evidence in E enables this path.
Path 2. C $\rightarrow$ E $\rightarrow$ F $\rightarrow$ H $\leftarrow$ G $\leftarrow$ D $\leftarrow$ B
This path is also active because H is the convergent of F and G, which has evidence.

Consequently, both paths are active. So Given H, C is <u>conditionally dependent</u> of B.

Part (3) Given F, is H conditionally independent of A?

Path 1. H $\leftarrow$ F $\leftarrow$ E $\leftarrow$ C $\leftarrow$ A
This path is blocked because the evidence in F, which is a casual chain blocks this path.
Path 2. H $\leftarrow$ G $\leftarrow$ D $\rightarrow$ E $\leftarrow$ C $\leftarrow$ A
This path is active because propagating evidence to E from F unblocks this converging node. On the other hand, D is divergent to G and E, and no evidence in this node enables this path.

Consequently, there is one active path from H to A. So given F, H is <u>conditionally dependent</u> of A.

Part (4) Given C, is F conditionally independent of B?

Path 1. F $\rightarrow$ H $\leftarrow$ G $\leftarrow$ D $\leftarrow$ B
This path is blocked because H is the convergent of F and G, but has no evidence.
Path 2. F $\leftarrow$ E $\leftarrow$ D $\leftarrow$ B
This path is active, because there is no evidence in this causal chain.

Consequently, there is one active path from F to B. So given C, F is <u>conditionally dependent</u> of B.

Part (5) Given F and D, is C conditionally independent of G?

Path 1. C $\rightarrow$ E $\rightarrow$ F $\rightarrow$ H $\leftarrow$ G
This path is blocked because the evidence in F blocks the causal chain. Furthermore, H is the convergent of F and G, but has no evidence.
Path 2. C $\rightarrow$ E $\leftarrow$ D $\rightarrow$ G
This path is also blocked, despite propagating evidence to E from F, which unblocks this converging node, D is divergent to G and E, and having evidence in this node blocks this path.

Consequently, all paths are blocked. So given F and D, C is <u>conditionally independent</u> of G.

Question 4

The query node and the evidence nodes are relevant to the query. All ancestors of relevant nodes are relevant to the query. All other nodes are irrelevant to the query.

Part (1) $P(C|G)$

$$\underline{relevant\_nodes} = [A, B, C, D, G]$$

Part (2) $P(D|E, A)$

$$\underline{relevant\_nodes} = [A, B, C, D, E]$$

Part (3) $P(F)$

$$\underline{relevant\_nodes} = [A, B, C, D, E, F]$$

Part (4) $P(B|D, F)$

$$\underline{relevant\_nodes} = [A, B, C, D, E, F]$$

Part (5) $P(C|F, G)$

$$\underline{relevant\_nodes} = [A, B, C, D, E, F, G]$$

Question 5

Part (1) $P(C)$

Because A is the ancestor of C, only A is relevant to this query, so we use marginalization to eliminate this node.

$$\underline{P(C)} = \sum_{A} P(A, C) = \sum_{A} P(C|A)P(A)$$

So we need JPD of A and C.

Part (2) $P(B|C)$

At first, we investigate the conditional dependency of B on C given an empty set.
Path 1. B $\rightarrow$ D $\rightarrow$ E $\leftarrow$ C is blocked because of lack of evidence in E, which is the convergence node of D and C.
Path 2. B $\rightarrow$ D $\rightarrow$ G $\rightarrow$ H $\leftarrow$ F $\leftarrow$ E $\leftarrow$ C is also blocked because of lack of evidence in H which is the convergence node of G and F.

All paths from B to C are blocked, so B is conditionally independent of C, given an empty set.

$$\underline{P(B|C)} = P(B)$$

Another method is to calculate the equation.

$$P(B|C) = \frac{P(B,C)}{P(C)}$$

$$\begin{aligned} P(B,C) &= \sum_A P(A,B,C) = \sum_A P(C|A)P(A)P(B) \\ &= P(B)\sum_A P(C|A)P(A) = P(B)P(C) \quad \textit{given part}1 \end{aligned}$$

$$\underline{P(B|C)} = \frac{P(B,C)}{P(C)} = \frac{P(B)P(C)}{P(C)} = P(B)$$

So we need probability distribution of B.

Part (3) $P(A|H,C)$

At first, we investigate the conditional dependency of A on H given C.
Path 1. A $\rightarrow$ C $\rightarrow$ E $\rightarrow$ F $\rightarrow$ H, is blocked because of having evidence in C in a causal chain.
Path 2. A $\rightarrow$ C $\rightarrow$ E $\leftarrow$ D $\rightarrow$ G $\rightarrow$ H is also blocked due to lack of evidence in E, which is the convergence node of C and D. Furthermore, evidence in C, which is a causal chain, also blocks this path.

All paths from A to H are blocked, so A is conditionally independent of H given C.

$$P(A|H,C) = P(A|C) \xrightarrow{Bayes' rule} P(A|C) = \frac{P(C|A)P(A)}{P(C)}$$

$$\xrightarrow{Part(1)} P(C) = \sum_A P(C|A)P(A)$$

$$\underline{P(A|H,C)} = \frac{P(C|A)P(A)}{\sum_A P(C|A)P(A)}$$

So we need JPD of A and C.

Another method is to calculate the equation.

$$P(A|H,C) = \frac{P(A,C,H)}{P(C,H)}$$

$$\begin{aligned}
P(A,C,H) &= \sum_{B,D,E,F,G} P(A,B,C,D,E,F,G,H) \\
&= \sum_{B,D,E,F,G} P(H|G,F)P(F|E)P(E|C,D)P(C|A)P(A)P(G|D)P(D|B)P(B) \\
&= P(C|A)P(A) \sum_{B,D,E,F,G} P(H|G,F)P(F|E)P(E|C,D)P(G|D)P(D|B)P(B)
\end{aligned}$$

$$\begin{aligned}
P(C,H) &= \sum_{B,D,E,F,G,H} P(A,B,C,D,E,F,G,H) \\
&= \sum_{A,B,D,E,F,G} P(H|G,F)P(F|E)P(E|C,D)P(C|A)P(A)P(G|D)P(D|B)P(B) \\
&= \sum_{A} P(C|A)P(A) \sum_{B,D,E,F,G} P(H|G,F)P(F|E)P(E|C,D)P(G|D)P(D|B)P(B)
\end{aligned}$$

$$\begin{aligned}
P(A|H,C) &= \frac{P(A,C,H)}{P(C,H)} \\
&= \frac{P(C|A)P(A)\sum_{B,D,E,F,G} P(H|G,F)P(F|E)P(E|C,D)P(G|D)P(D|B)P(B)}{\sum_A P(C|A)P(A)\sum_{B,D,E,F,G} P(H|G,F)P(F|E)P(E|C,D)P(G|D)P(D|B)P(B)} \\
&= \frac{P(C|A)P(A)}{\sum_A P(C|A)P(A)}
\end{aligned}$$

$$\underline{P(A|H,C)} = \frac{P(C|A)P(A)}{\sum_A P(C|A)P(A)}$$

So we need JPD of A and C.

Part (4) $P(C|D,F)$

At first, we investigate the conditional dependency of C on D given F.
Path 1. C $\rightarrow$ E $\leftarrow$ D is active because the evidence in F propagates to E, which is the convergent node of C and D, and enables this path.
Path 2. C $\rightarrow$ E $\rightarrow$ F $\rightarrow$ H $\leftarrow$ G $\leftarrow$ D is blocked because of lack of evidence in H, which is the convergent node of F and G. Furthermore, evidence in F, which is a causal chain, also blocks this path.
Secondly, we investigate the conditional dependency of C on F given D.
Path 1. C $\rightarrow$ E $\rightarrow$ F is active because nothing blocks this causal chain.
Path 2. C $\rightarrow$ E $\leftarrow$ D $\rightarrow$ G $\rightarrow$ H $\leftarrow$ F is blocked because of having evidence in D, which is the divergent node to E and G.
Consequently, C is conditionally dependent on D and F.

$$relevant\_nodes = [A, B, C, D, E, F]$$

$$P(C|D,F) = \frac{P(C,D,F)}{P(D,F)}$$

$$\begin{aligned}
P(C,D,F) &= \sum_{A,B,E} P(F|E)P(E|C,D)P(D|B)P(B)P(C|A)P(A) \\
&= \sum_{B,E} P(F|E)P(E|C,D)P(D|B)P(B) \sum_{A} P(C|A)P(A) \\
&= \sum_{B} P(D|B)P(B) \sum_{E} P(F|E)P(E|C,D) \sum_{A} P(C|A)P(A) \\
&= \sum_{B} P(D|B)P(B) \sum_{E} P(F|E)P(E|C,D)(P(C)) \quad given\ part1 \\
&= P(C) \sum_{B} P(D|B)P(B) \sum_{E} P(F|E)P(E|C,D)
\end{aligned}$$

$$\begin{aligned}
P(D,F) &= \sum_{A,B,C,E} P(F|E)P(E|C,D)P(D|B)P(B)P(C|A)P(A) \\
&= \sum_{B} P(D|B)P(B) \sum_{E} P(F|E) \sum_{C} P(E|C,D) \sum_{A} P(C|A)P(A) \quad given\ part1 \\
&= \sum_{B} P(D|B)P(B) \sum_{E} P(F|E) \sum_{C} P(E|C,D)P(C)
\end{aligned}$$

$$
\begin{aligned}
P(C|D,F) &= \frac{P(C,D,F)}{P(D,F)} \\
&= \frac{P(C)\sum_B P(D|B)P(B)\sum_E P(F|E)P(E|C,D)}{\sum_B P(D|B)P(B)\sum_E P(F|E)\sum_C P(E|C,D)P(C)} \\
&= \frac{P(C)\sum_E P(F|E)P(E|C,D)}{\sum_E P(F|E)\sum_C P(E|C,D)P(C)}
\end{aligned}
$$

$$
\underline{P(C|D,F)} = \frac{P(C)\sum_E P(F|E)P(E|C,D)}{\sum_E P(F|E)\sum_C P(E|C,D)P(C)}
$$

So we need JPD of C, D, E, F.