Sols HW1

290C W06

MV

2.6 #1

Prove Weak Union & Contraction

20

Weak Union: Prove that $(X \perp Y, W \mid Z) \Rightarrow (X \perp Y \mid Z, W)$

(1) the def. of cond. prob. gives: $(X \perp Y, W \mid Z) \Leftrightarrow P(X \mid Y, W, Z) = P(X \mid Z)$

(2) by decomposition, $(X \perp Y, W \mid Z) \Rightarrow (X \perp W \mid Z)$.

(3) def. of cond. prob $(X \perp W \mid Z) \Leftrightarrow P(X \mid W, Z) = P(X \mid Z)$

(4) combine (1) and (3): $P(X \mid Y, W, Z) = P(X \mid W, Z)$

(5) def. of cond. prob. $P(X \mid Y, W, Z) = P(X \mid W, Z) \Leftrightarrow (X \perp Y \mid W, Z)$ ✓

proved!

Contraction: Prove that $(X \perp W \mid Z, Y)$ & $(X \perp Y \mid Z) \Rightarrow (X \perp Y, W \mid Z)$

(1) Def. of cond. prob: $(X \perp W \mid Z, Y) \Leftrightarrow P(X \mid W, Y, Z) = P(X \mid Y, Z)$

(2) Def. of cond. prob: $(X \perp Y \mid Z) \Leftrightarrow P(X \mid Y, Z) = P(X \mid Z)$

(3) Combine (1) & (2) $P(X \mid W, Y, Z) = P(X \mid Z)$

(4) Def. of cond. prob. $P(X \mid W, Y, Z) = P(X \mid Z) \Leftrightarrow (X \perp Y, W \mid Z)$

2.20 $I_P(X;Y) \leq I_P(X;Y,Z)$

20/20

$$\begin{aligned} I_P(X;Y,Z) &= H(X) - H(X|Y,Z) \\ &\geq H(X) - H(X|Y) \text{ since } H(X|Y) \geq H(X|Y,Z) \\ &\geq I_P(X;Y) \end{aligned}$$

or by chain rule:

$$I_p(X;Y,Z) = I_p(X;Y) + I_p(X;Z|Y)$$

$$I_p(X;Y) = I_p(X;Y,Z) - \underbrace{I_p(X;Z|Y)}_{\geq 0}$$

Mutual inf. is a relative entropy
=> always ≥ 0

& $I_p(X;Y) \leq I_p(X;Y,Z)$

2.9

20/20

1.

$$P(H|E_1, E_2) = \frac{P(H, E_1, E_2)}{P(E_1, E_2)} = \frac{P(H)P(E_1, E_2|H)}{P(E_1, E_2)}$$

a) $P(E_1, E_2), P(H), P(E_1|H)$, and $P(E_2|H)$.

Not enough: we do not know if $(E_1 \perp E_2|H)$

b) $P(E_1, E_2), P(H)$, and $(E_1, E_2|H)$. It can be evaluated:

$$\frac{P(H)P(E_1, E_2|H)}{P(E_1, E_2)}$$

c) $P(E_1|H), P(E_2|H)$ and $P(H)$.

Not enough: It is a subset of the first case.

2

a) $P(E_1, E_2), P(H), P(E_1|H)$, and $P(E_2|H)$. It can be evaluated:

$$\frac{P(H)P(E_1|H)P(E_2|H)}{P(E_1, E_2)}$$

b) $P(E_1, E_2), P(H)$, and $(E_1, E_2|H)$. It can be evaluated:

$$\frac{P(H)P(E_1, E_2|H)}{P(E_1, E_2)}$$

c) $P(E_1|H), P(E_2|H)$ and $P(H)$. It can be evaluated:

$$\frac{P(H)P(E_1|H)P(E_2|H)}{\underbrace{\sum_{h \in H} P(h)P(E_1|h)P(E_2|h)}_{P(E_1, E_2)}}$$

3.3 20

1) Constraints for explaining away

$P(B) \approx P(B \mid E, A) < P(B \mid A)$

explains away alarm by earthquake

$P(E) \approx P(E \mid B, A) < P(E \mid A)$

explains away alarm by burglary

2

$\frac{10}{10}$

$$
\begin{aligned}
P(b^1|a^1,e^1) &= \frac{P(b^1,a^1,e^1)}{P(a^1,e^1)} \\
&= \frac{P(b^1)P(e^1)P(a^1|b^1,e^1)}{P(a^1,e^1)} \\
&= \frac{P(b^1)P(e^1)}{P(a^1,e^1)} \\
&= \frac{P(b^1)P(e^1)}{P(a^1,e^1,b^1)+P(a^1,e^1,b^0)} \text{ since } P(a^1|b^1,e^1)=1 \\
&= \frac{P(b^1)P(e^1)}{P(b^1)P(e^1)P(a^1|b^1,e^1)+P(b^0)P(e^1)P(a^1|b^0,e^1)} \\
&= \frac{P(b^1)P(e^1)}{P(b^1)P(e^1)+P(b^0)P(e^1)} \\
&= \frac{P(b^1)P(e^1)}{(P(b^1)+P(b^0))P(e^1)} \\
&= P(b^1)
\end{aligned}
$$

20

EC: 3.1.

uniform distribution on all tupels s.t. $x_1 \oplus x_2 \oplus x_3$

x_1	x_2	x_3	
0	0	0	$\frac{1}{4}$
1	1	0	$\frac{1}{4}$
1	0	1	$\frac{1}{4}$
0	1	1	$\frac{1}{4}$
rest			0

$P(x_i) = \frac{1}{2}$

$P(x_i, x_j) = \frac{1}{4}$
$i \neq j$

$$\left.\begin{aligned} &P(x_1=1, x_2=1, x_3=1) = 0 \\ &\underbrace{P(x_1=1, x_2=1)}_{\frac{1}{4}}\ \underbrace{P(x_3=1)}_{\frac{1}{2}} = \frac{1}{8} \end{aligned}\right] \neq$$

3.9 Show that all paths P from X to non-desc. of X are blocked

I

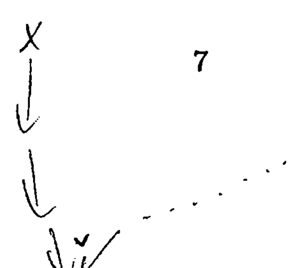
Proof. ~~Every~~ path P from a node X to one of its non descendants ~~must have~~ has one of the vertices in Pa_X say Y as the first node in the path. The edge connecting them is from Y to X ($Y \rightarrow X$) since Y is an ancestor of X. The next node in the path say Z can be connected to node Y in two different ways:

-The first way is from $Z \rightarrow Y$. Therefore there is a structure $X \leftarrow Y \leftarrow Z$ and since Y is an evidence node then Y blocks every path from X to its non-descendants.

-The second way is from $Y \rightarrow Z$. Therefore there is a structure $X \leftarrow Y \rightarrow Z$ and since Y is an evidence node then Y blocks every path from X to its non-descendants.

Because every path is blocked then X is d-separated from its non-descendants given its parents.

II Path P goes first to a # of descendents (may be just X itself) of X and then branches off via v structure



All those paths are blocked because of v.