2.1

(a) $P(E=1 \mid A=1)$

$$= \frac{P(A=1 \mid E=1) \cdot P(E=1)}{P(A=1)} \quad [\text{Bayes Rule}]$$

Known $P(E=1)=0.002$, Find $P(A=1 \mid E=1)$, $P(A=1)$

$P(A=1 \mid E=1)$

$$= \sum_B P(A=1, B \mid E=1) \quad [\text{Marginalization}]$$

$$= \sum_B P(A=1 \mid B, E=1) \cdot P(B \mid E=1) \quad [\text{Product Rule}]$$

$$= \sum_B P(A=1 \mid B, E=1) \cdot P(B) \quad [\text{conditional independence}]$$

$$= P(A=1 \mid B=0, E=1) \cdot P(B=0) + P(A=1 \mid B=1, E=1) \cdot P(B=1)$$

$$= 0.29 \times 0.999 + 0.95 \cdot 0.001 = 0.29065$$

$$P(A=1) = \sum_{B,E} P(A=1, B, E) \quad [\text{Marginalization}]$$

$$= \sum_{B,E} P(A=1 \mid B, E) P(B \mid E) P(E)$$

$$= \sum_{B,E} P(A=1 \mid B, E) P(B) P(E)$$

$$= P(A=1 \mid B=0, E=0) P(B=0) P(E=0) + P(A=1 \mid B=0, E=1) P(B=0) P(E=1)$$
$$+ P(A=1 \mid B=1, E=0) P(B=1) P(E=0) + P(A=1 \mid B=1, E=1) P(B=1) P(E=1)$$

$$= 0.00252$$

$$P(E=1 \mid A=1) = (0.29065 \times 0.002) / 0.00252 = 0.230675 \simeq \underline{\underline{0.23}}$$

(b) $P(E=1 \mid A=1, B=0)$

$$= \frac{P(A=1 \mid E=1, B=0) \cdot P(E=1 \mid B=0)}{P(A=1 \mid B=0)} \quad \text{[Bayes Rule]}$$

$P(A=1 \mid E=1, B=0)$ is given

$P(E=1 \mid B=0) = P(E=1)$ [conditional independence]

$P(A=1 \mid B=0)$

$= \sum_E P(A=1, E \mid B=0)$ [Marginalization]

$= \sum_E P(A=1 \mid E, B=0) \cdot P(E \mid B=0)$ [product rule]

$= \sum_E P(A=1 \mid E, B=0) \cdot P(E)$ [conditional independence]

$= P(A=1 \mid E=0, B=0) \cdot P(E=0) + P(A=1 \mid E=1, B=0) \cdot P(E=1)$

$= 0.001 \times 0.998 + 0.29 \times 0.002 = 0.001578$

$P(E=1 \mid A=1, B=0) = (0.29 \times 0.002) / 0.001578 = 0.3675 \approx \underline{\underline{0.37}}$

Compare a to b:

$P(E=1 \mid A=1) = 0.23$

$P(E=1 \mid A=1, B=0) = 0.37$ $\downarrow$ increase

Makes sense, since we know sure no burglar then alarm sound is because of earthquake makes more sense

(c) $P(A=1 \mid M=1)$

$$= \frac{P(M=1 \mid A=1) \cdot P(A=1)}{P(M=1)}$$

$P(M=1)$ Given, $P(A=1)$ from 1.a is 0.00252

$$P(M=1) = \sum_A P(M=1, A)$$

$$= \sum_A P(M=1 \mid A) \cdot P(A)$$

$$= P(M=1 \mid A=0) \cdot P(A=0) + P(M=1 \mid A=1) \cdot P(A=1)$$

$$= 0.01 \times (1-0.00252) + 0.7 \times 0.00252$$

$$= 0.0099748 + 0.001764 = 0.0117$$

$$P(A=1 \mid M=1) = (0.7 \times 0.00252) / 0.0117 = 0.15077 \approx 0.15$$

(d) $P(A=1 \mid M=1, J=0)$

$$= \frac{P(M=1, J=0 \mid A=1) \cdot P(A=1)}{P(M=1, J=0)}$$

$P(M=1, J=0 \mid A=1) = P(M=1 \mid A=1) \cdot P(J=0 \mid A=1)$

[M, J conditional independent on evidence A)

$P(J=0 \mid A=1) = 1 - P(J=1 \mid A=1) = 0.1$

$P(A=1)$ from 1.a is 0.00252

$P(M=1, J=0) = \sum_A P(M=1, J=0, A)$

$= \sum_A P(M=1, J=0 \mid A) P(A)$

$= \sum_A P(M=1 \mid A) P(J=0 \mid A) P(A)$

$= P(M=1 \mid A=0) P(J=0 \mid A=0) P(A=0) + P(M=1 \mid A=1) P(J=0 \mid A=1) P(A=1)$

$= 0.01 \times (1-0.05) \times (1-0.00252) + 0.7 \times 0.1 \times 0.00252 = 0.009652$

$P(A=1 \mid M=1, J=0) = (0.7 \times 0.1 \times 0.00252) / 0.009652$

$= 0.01827 \approx 0.018$

Consider c and d:

$P(A=1 \mid M=1) = 0.15$

$P(A=1 \mid M=1, J=0) = 0.018$ $\downarrow$

Makes sense since if there is an alarm, it is more likely to cause both of them to call. However, given mary called but not John, it is less likely to be an alarm

(e)

$P(A=1 \mid M=0)$

$$= \frac{P(M=0 \mid A=1)\, P(A=1)}{P(M=0)}$$

$P(M=0 \mid A=1) = 1 - P(M=1 \mid A=1) = 0.3$

$P(A=1)$ from 1.a is 0.00252

$P(M=1)$ from 1.c is 0.0117, so $P(M=0) = 1 - 0.0117 = 0.9883$

$P(A=1 \mid M=0) = (0.3 \times 0.00252) / 0.9883 = 0.000765$

(f) $P(A=1 \mid M=0, B=1)$

$$= \frac{P(M=0, B=1 \mid A=1) \cdot P(A=1)}{P(M=0, B=1)}$$

$$= \frac{P(M=0 \mid A=1) \cdot \boxed{P(B=1 \mid A=1)} \cdot P(A=1)}{\boxed{P(M=0, B=1)}}$$

[M, B conditional independent on evidence A]

$P(M=0 \mid A=1) = 1 - P(M=1 \mid A=1) = 0.3$

$P(A=1) = 0.00252$ (from a)

$$\boxed{P(B=1 \mid A=1)} = \frac{\boxed{P(A=1 \mid B=1)} \cdot P(B=1)}{P(A=1)}$$

$$P(B=0 \mid A=0) = \frac{P(A=0, B=0) \cdot P(B=0)}{P(A=0)}$$

[From B]

$$= \frac{1 - P(A=1, B=0) \cdot 0.999}{0.99748}$$

$$= 0.99994$$

$$\boxed{P(A=1 \mid B=1)} = \sum_E P(A=1, E \mid B=1)$$

$$= \sum_E P(A=1 \mid E, B=1) P(E \mid B=1)$$

$$= P(A=1 \mid E=0, B=1) P(E=0) + P(A=1 \mid E=1, B=1) \cdot P(E=1)$$

$$= 0.94 \times 0.998 + 0.95 \times 0.002 = 0.94$$

$$\boxed{P(B=1 \mid A=1)} = (0.94 \times 0.001) / 0.00252 = 0.373$$

$$\boxed{P(M=0, B=1)} = \sum_A P(M=0, B=1, A)$$

$$= \sum_A P(M=0, B=1 \mid A) \cdot P(A)$$

$$= \sum_A P(M=0 \mid A) P(B=1 \mid A) \cdot P(A)$$

$$P(M=0 \mid A=0) P(B=1 \mid A=0) P(A=0) + P(M=0 \mid A=1) P(B=1, A=1) P(A=1)$$

$$= 0.99 \times (1 - 0.99994) \times 0.99748 + 0.3 \times 0.373 \times 0.00252$$

$$= 0.00034$$

$P(A=1 \mid M=0, B=1) = (0.3 \times 0.373 \times 0.00252) / 0.00034$

$= 0.829 \approx 0.83$

Consider e and f:

$P(A=1 \mid M=0) = 0.000765$

$P(A=1 \mid M=0, B=1) = 0.83$ ↓

Even the friend mary doesn't call, if we know burglar is true, then the chance of alarm ringing increases

2.2

(a) $$r_k = \frac{P(D=0 \mid S_1=1, S_2=1 \cdots S_k=1)}{P(D=1 \mid S_1=1, S_2=1 \cdots S_k=1)}$$

$$= \frac{P(S_1=1, S_2=1 \cdots S_k=1 \mid D=0)\cdot P(D=0)}{\cancel{P(S_1=1, S_2=1 \cdots S_k=1)}} \Big/ \frac{P(S_1=1, S_2=1 \cdots S_k=1 \mid D=1)\cdot P(D=1)}{\cancel{P(S_1=1, S_2=1 \cdots S_k=1)}}$$

$$= \frac{P(S_1=1, S_2=1 \cdots S_k=1 \mid D=0)\cdot P(D=0)}{P(S_1=1, S_2=1 \cdots S_k=1 \mid D=1)\cdot P(D=1)}$$

Because S_1, S_2, S_3 conditional independent given evidence D,

$$f(a) = P(S_1=1, S_2=1 \cdots S_k=1 \mid D=0) = P(S_1=1 \mid D_0)\times P(S_2=1 \mid D_0)\times \cdots \times P(S_k=1 \mid D_0)$$

$$f(b) = P(S_1=1, S_2=1 \cdots S_k=1 \mid D=1) = P(S_1=1 \mid D_1)\times P(S_2=1 \mid D_1)\times P(S_k=1 \mid D_1)$$

Let $r_k = \frac{f(a)\times P(D=0)}{f(b)\times P(D=1)}$, Given $P(S_k=1 \mid D=0) = \frac{2^{(k-1)}+(-1)^{(k-1)}}{2^k+(-1)^k}$, $P(S_k=1 \mid D=1) = \frac{1}{2}$

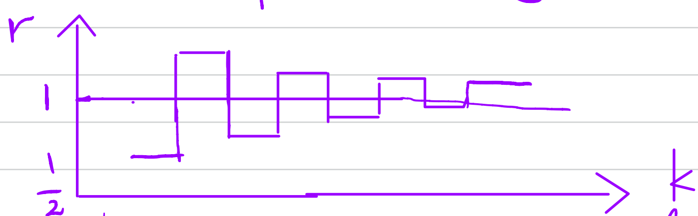
$$f(a) = \prod_{k} \frac{2^{(k-1)}+(-1)^{(k-1)}}{2^k+(-1)^k} = \frac{2^1+(-1)}{\cancel{2^2+(-1)^2}} \times \frac{\cancel{2^2+(-1)^2}}{\cancel{2^3+(-1)^3}} \times \cdots \frac{\cancel{2^{k-1}+(-1)^{k-1}}}{2^k+(-1)^k} = \frac{1}{2^k+(-1)^k}$$

$(k \geq 2)$

$$f(b) = \left(\frac{1}{2}\right)^k = 2^{-k} \qquad r_k = \frac{\frac{1}{2^k+(-1)^k} \times \cancel{\frac{1}{2}}}{2^{-k} \times \cancel{\frac{1}{2}}} = \boxed{\frac{2^k}{2^k+(-1)^k}}$$

from $r = \dfrac{2^k}{2^k + (-1)^k}$ $(k \geq 2)$, we can see if $k \in (3, 5 \cdots 2k-1)$, doctor will diagnose the patient as Disease 0 $(D=0)$ since $r > 1$; if $k \in \{2, 4, 6 \cdots 2k\}$, doctor will diagnose Disease 1 $(D=1)$ since $r < 1$

(b) Sketch of r (k being discrete value, shows)

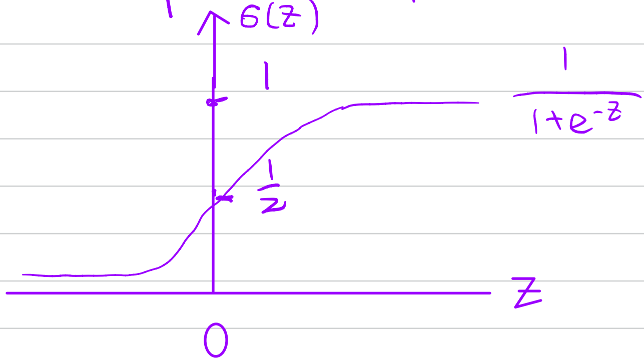


With more symptoms observed (k increase), r's value converge on value 1 [up and down]

As a result, smaller and smaller gap between the two diagnosis to be distinguished, so it will be less certain for the diagnosis

2.3

Sketch of $\sigma(z) = \dfrac{1}{1+e^{-z}}$



(a) $\sigma'(z) = \sigma(z)\sigma(-z)$

$$\sigma'(z) = \frac{d}{dz}\sigma(z) = \frac{d}{dz}\frac{1}{1+e^{-z}} = \frac{d}{dz}\left(1+e^{-z}\right)^{-1}$$

$$= -\left(1+e^{-z}\right)^{-2} \times \frac{d}{dz}\left(1+e^{-z}\right) = \frac{e^{-z}}{(1+e^{-z})^2}$$

$$= \frac{1}{1+e^{-z}} \times \frac{e^{-z}}{1+e^{-z}}$$

$$\sigma(z)\cdot\sigma(-z) = \frac{1}{1+e^{-z}} \times \frac{1}{1+e^{z}} = \frac{1}{1+e^{-z}} \times \frac{1}{1+e^{z}} \times \frac{e^{-z}}{e^{-z}}$$

$$= \frac{1}{1+e^{-z}} \times \frac{e^{-z}}{1+e^{-z}}$$

$$= \sigma'(z)$$

(b) $\sigma(-z) + \sigma(z) = 1$

$$\sigma(-z) = \frac{1}{1+e^{z}} = \frac{1}{1+e^{z}} \times \frac{e^{-z}}{e^{-z}} = \frac{e^{-z}}{1+e^{-z}}$$

$$\sigma(-z) + \sigma(z) = \frac{1}{1+e^{-z}} + \frac{e^{-z}}{1+e^{-z}} = 1$$

(c) $L(\sigma(z)) = z$

$$L(\sigma(z)) = \log\left(\frac{\frac{1}{1+e^{-z}}}{1-\frac{1}{1+e^{-z}}}\right) = \log\left(\frac{\frac{1}{1+e^{-z}}}{\frac{e^{-z}}{1+e^{-z}}}\right) = \log\left(\frac{1}{e^{-z}}\right) = \log\left(e^{z}\right)$$

$$= z$$

(d) From $P(Y=1 \mid X = X_1, X_2 = x_2, \cdots X_k = x_k) = \sigma\left(\sum_{i=1}^{k} w_i x_i\right)$

$\Rightarrow P(Y=1 \mid X_i = 1, X_j = 0 \text{ for all } j, j \neq i) = \sigma w_i = p_i$

$L(p_i) = L(\sigma w_i)$

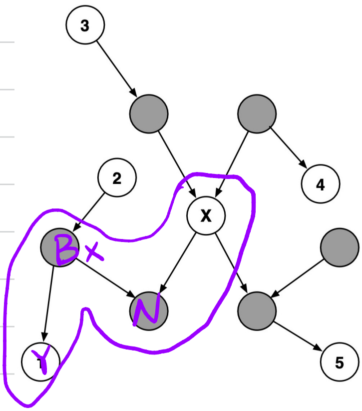
from c, $L(\sigma w_i) = w_i$, so $L(p_i) = w_i$

2.4

1. X,Y {month, fall}, E {puddle}
2. X,Y {rain, fall}, E {puddle}
3. X,Y {sprinkler, fall}, E {puddle}
4. X,Y {sprinkler, rain}, E {month}
5. X,Y {sprinkler, fall}, E {month, puddle}
6. X,Y {rain, fall}, E {month, puddle}
7. X,Y {month, fall}, E {rain, sprinkler}
8. X,Y {month, puddle}, E {rain, sprinkler}
9. X,Y {fall, sprinkler}, E {rain, puddle}
10. X,Y {fall, month}, E {rain, puddle}
11. X,Y {fall, rain}, E {sprinkler, puddle}
12. X,Y {fall, month}, E {sprinkler, puddle}
13. X,Y {fall, sprinkler}, E {rain, puddle, month}
14. X,Y {fall, rain}, E {sprinkler, puddle, month}
15. X,Y {month, puddle}, E {rain, sprinkler, fall}
16. X,Y {month, fall}, E {rain, sprinkler, puddle}

No, there is no with case $E = \emptyset$ since there are no random variable mutually independent in this graph given no evidence

2.5



Case 1

Case 1, Y is the Children of X 's spouse

Let the spouse of X be Bx

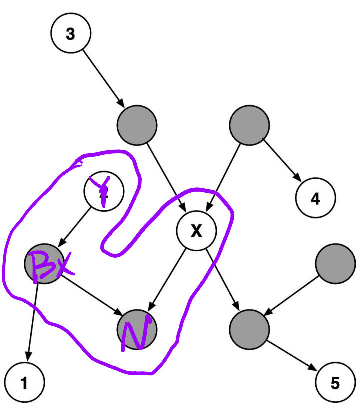
Prove $P(X, Y | B_x) = P(X | B_x) P(Y | B_x)$
Which means X, Y conditional independent given Bx

From X to Y, there is only 1 path (given polytree), which is:

$$Y \leftarrow B_x \rightarrow N \leftarrow X$$

Apply case 2 of d-separation,

X, Y conditional independent on Bx



Case 2

Case 2, Y is the parent of X 's spouse

Let the spouse of X be Bx

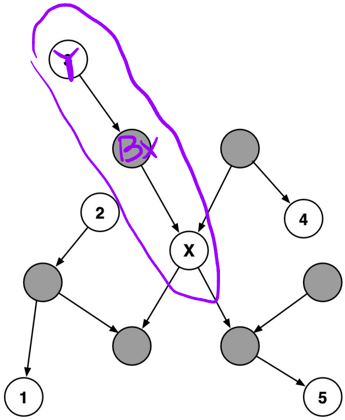
Prove $P(X, Y | B_x) = P(X | B_x) P(Y | B_x)$
Which means X, Y conditional independent given Bx

From X to Y, there is only 1 path (given polytree), which is:

$$Y \rightarrow B_x \rightarrow N \leftarrow X$$

Apply case 1 of d-separation,

X, Y conditional independent on Bx



Case 3

Case 3, Y is the parent of X's parent

Let the parent of X be Bx

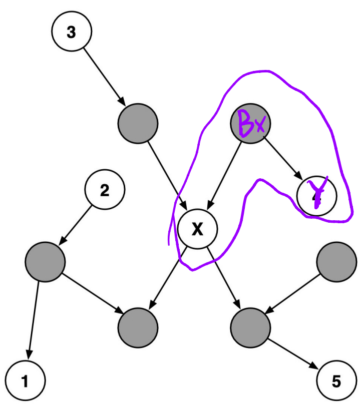
Prove $P(X,Y|B_x) = P(X|B_x)P(Y|B_x)$
which means X, Y conditional independent given Bx

From X to Y, there is only 1 path (given polytree), which is:

$$Y \rightarrow B_x \rightarrow X$$

Apply case 1 of d-separation,

X, Y conditional independent on Bx



Case 4

Case 4, Y is the children of X's parent, (Y is spouse of X) Let the parent of X be Bx

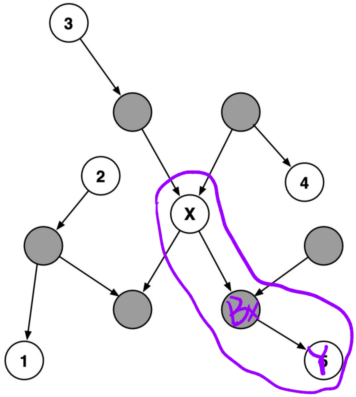
Prove $P(X,Y|B_x) = P(X|B_x)P(Y|B_x)$
which means X, Y conditional independent given Bx

From X to Y, there is only 1 path (given polytree), which is:

$$Y \leftarrow B_x \rightarrow X$$

Apply case 2 of d-separation,

X, Y conditional independent on Bx



Case 5

Case 5, Y is the Children of X 's children

Let the children of X be Bx

Prove $P(X, Y|Bx) = P(X|Bx)P(Y|Bx)$
which means X, Y conditional independent given Bx

From X to Y, there is only 1 path (given polytree), which is:

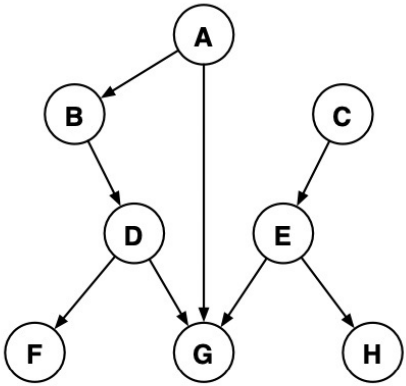
$$X \rightarrow Bx \rightarrow Y$$

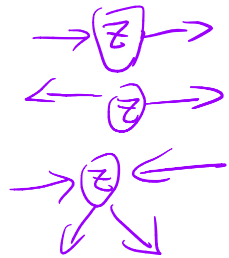
Apply case 1 of d-separation,

X, Y conditional independent on Bx

2.6 True or false

For the belief network shown below, indicate whether the following statements of marginal or conditional independence are **true (T)** or **false (F)**.





F $\qquad P(B|G,C) = P(B|G)$

T $\qquad P(F,G|D) = P(F|D)\,P(G|D)$

T $\qquad P(A,C) = P(A)\,P(C)$

F $\qquad P(D|B,F,G) = P(D|B,F,G,A)$

T $\qquad P(F,H) = P(F)\,P(H)$

T $\qquad P(D,E|F,H) = P(D|F)\,P(E|H)$

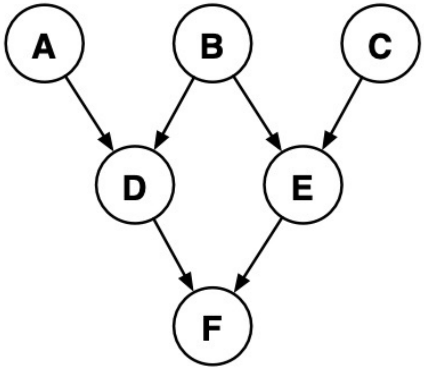
F $\qquad P(F,C|G) = P(F|G)\,P(C|G)$

T $\qquad P(D,E,G) = P(D)\,P(E)\,P(G|D,E)$

T $\qquad P(H|C) = P(H|A,B,C,D,F)$

F $\qquad P(H|A,C) = P(H|A,C,G)$

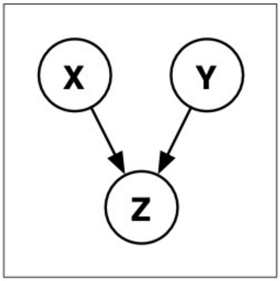
2.7 Subsets



For the belief network shown above, consider the following statements of conditional independence. Indicate the largest subset of nodes $\mathcal{S} \subset \{A, B, C, D, E, F\}$ for which each statement is true. Note that one possible answer is the empty set $\mathcal{S} = \emptyset$ or $\mathcal{S} = \{\}$ (whichever notation you prefer). The first one is done as an example.

$P(A)$	=	$P(A\|\mathcal{S})$	$\mathcal{S} = \{B, C, E\}$
$P(A\|D)$	=	$P(A\|\mathcal{S})$	S = {D, C}
$P(A\|B, D)$	=	$P(A\|\mathcal{S})$	S = {B, D, C, E, F}
$P(B\|D, E)$	=	$P(B\|\mathcal{S})$	S = {D, E, F}
$P(E)$	=	$P(E\|\mathcal{S})$	S = {A}
$P(E\|F)$	=	$P(E\|\mathcal{S})$	S = {F}
$P(E\|D, F)$	=	$P(E\|\mathcal{S})$	S = {D, F}
$P(E\|B, C)$	=	$P(E\|\mathcal{S})$	S = {B, C, A, D}
$P(F)$	=	$P(F\|\mathcal{S})$	S = { }
$P(F\|D)$	=	$P(F\|\mathcal{S})$	S = {D}
$P(F\|D, E)$	=	$P(F\|\mathcal{S})$	S = {D, E, A, B, C}

2.8 Noisy-OR



Nodes: $X \in \{0,1\}, Y \in \{0,1\}, Z \in \{0,1\}$

Noisy-OR CPT: $P(Z=1|X,Y) = 1 - (1-p_x)^X (1-p_y)^Y$

Parameters: $p_x \in [0,1], p_y \in [0,1], p_x < p_y$

Suppose that the nodes in this network represent binary random variables and that the CPT for $P(Z|X,Y)$ is parameterized by a noisy-OR model, as shown above. Suppose also that

$$0 < P(X=1) < 1,$$
$$0 < P(Y=1) < 1,$$

while the parameters of the noisy-OR model satisfy:

$$0 < p_x < p_y < 1.$$

Consider the following pairs of probabilities. In each case, indicate whether the probability on the left is equal (=), greater than (>), or less than (<) the probability on the right. The first one has been filled in for you as an example. (You should use your intuition for these problems; you are **not** required to show work.)

	$P(X=1)$	$=$	$P(X=1)$
(a)	$\underbrace{P(Z=1\vert X=0,Y=0)}$ 0	$<$	$P(Z=1\vert X=0,Y=1)$ p_y
(b)	$P(Z=1\vert X=1,Y=0)$ p_x	$<$	$P(Z=1\vert X=0,Y=1)$ p_y
(c)	$P(Z=1\vert X=1,Y=0)$ p_x	$<$	$P(Z=1\vert X=1,Y=1)$ $p_y+p_x-p_xp_y$
(d)	$P(X=1)$	$<$	$P(X=1\vert Z=1)$
(e)	$P(X=1)$	$=$	$P(X=1\vert Y=1)$
(f)	$P(X=1\vert Z=1)$	$>$	$P(X=1\vert Y=1,Z=1)$
(g)	$P(X=1)\,P(Y=1)\,P(Z=1)$	$<$	$P(X=1,Y=1,Z=1)$

X, Y=1, Z=1
P(X)·P(Y) P(X,Y|Z)

2.9

(a) $P(C|A,B,D)$

$$= \frac{P(D|A,B,C)\cdot P(C|A,B)\cdot \cancel{P(A|B)}\cdot \cancel{P(B)}}{P(D|A,B)\cdot \cancel{P(A|B)}\cdot \cancel{P(B)}}$$

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

[apply d-sep, A, D conditional independent on C, B, C conditional independent on D not in Evidence]

Both terms in numerator given, so solve $P(D|A,B)$:

$$P(D|A,B) = \sum_C P(D,C|A,B)$$

$$= \sum_C \frac{P(D|A,B,C)\cdot P(C|A,B)\cdot \cancel{P(A|B)\cdot P(B)}}{\cancel{P(A,B)}}$$

$$= \sum_C P(D|A,B)\cdot P(C|A)$$

$$P(C|A,B,D) = \frac{P(D|B,C)\cdot P(C|A)}{\sum_C P(D|B,C)\cdot P(C|A)}$$

(b) $P(E \mid A, B, D)$

$$= \sum_{C} P(E, C \mid A, B, D)$$

$$= \sum_{C} P(E \mid A, B, C, D) \cdot P(C \mid A, B, D)$$

$$= \sum_{C} P(E \mid C) \cdot P(C \mid A, B, D)$$

[E and A, B, D conditional independent given C]

(c) $P(G \mid A, B, D)$

$$= \sum_{E} P(G, E \mid A, B, D)$$

$$= \sum_{E} P(G \mid E, A, B, D) \cdot P(E \mid A, B, D)$$

$$= \sum_{E} P(G \mid E) \cdot P(E \mid A, B, D)$$

[G and A, B, D conditional independent given E]

(d)

$$P(F \mid A,B,D,G)$$

$$= \frac{P(G \mid A,B,D,F)\cdot P(A,B,D,F)}{P(A,B,D,G)}$$

$$= \frac{P(G \mid A,B,D,F)\cdot \cancel{P(A,B,D)}\cdot P(F)}{P(G \mid A,B,D)\cdot \cancel{P(A,B,D)}}$$

A, B, D conditionally independent with F Given G not in evidence set

$$= \frac{P(G \mid A,B,D,F)\cdot P(F)}{P(G \mid A,B,D)}$$

from (c)

$$= \frac{\sum_E P(G,E \mid A,B,D,F)\cdot P(F)}{P(G \mid A,B,D)}$$

$$= \frac{\sum_E P(G \mid A,B,D,E,F)\cdot P(E \mid A,B,D,F)\cdot P(F)}{P(G \mid A,B,D)}$$

$$= \frac{\sum_E P(G \mid E,F)\cdot P(E \mid A,B,D)\cdot P(F)}{P(G \mid A,B,D)}$$

[G and A, B, D conditional independent given E,
E and F conditional independent given G not in E]