CSE 150A / 250A - Homework 2 solutions – *(51 pts total)*

2.1 Probabilistic inference (*12 pts — 2 pts each*)

(a) $\boxed{P(E=1|A=1) = 0.2310}$

From Bayes rule:

$$P(E=1|A=1) \;=\; \frac{P(A=1|E=1)P(E=1)}{P(A=1)}$$

Denominator:

$$\begin{aligned}
P(A=1) &= \sum_{b,e} P(E=e, B=b, A=1) \quad \boxed{\textbf{marginalization}} \\
&= \sum_{b,e} P(E=e)\, P(B=b|E=e)\, P(A=1|B=b, E=e) \quad \boxed{\textbf{product rule}} \\
&= \sum_{b,e} P(E=e)\, P(B=b)\, P(A=1|B=b, E=e) \quad \boxed{\textbf{conditional independence}} \\
&= P(E=0)\, P(B=0)\, P(A=1|E=0, B=0) \;+ \\
&\quad\; P(E=0)\, P(B=1)\, P(A=1|E=0, B=1) \;+ \\
&\quad\; P(E=1)\, P(B=0)\, P(A=1|E=1, B=0) \;+ \\
&\quad\; P(E=1)\, P(B=1)\, P(A=1|E=1, B=1) \\
&= (1-0.002) * ((1-0.001) * 0.001 + 0.001 * 0.94) \;+ \\
&\quad\; 0.002 * ((1-0.001) * 0.29 + 0.001 * 0.95) \\
&= \mathbf{0.002516442}
\end{aligned}$$

First term in numerator:

$$\begin{aligned}
P(A=1|E=1) &= \sum_{b} P(A=1, B=b|E=1) \quad \boxed{\textbf{marginalization}} \\
&= \sum_{b} P(A=1|B=b, E=1)\, P(B=b|E=1) \quad \boxed{\textbf{product rule}} \\
&= \sum_{b} P(A=1|B=b, E=1)\, P(B=b) \quad \boxed{\textbf{conditional independence}} \\
&= P(A=1|E=1, B=1)\, P(B=1) \;+\; P(A=1|E=1, B=0)\, P(B=0)
\end{aligned}$$

$$= \quad 0.95 * 0.001 + 0.29 * (1 - 0.001)$$

$$= \quad \mathbf{0.29066}$$

Substituting the results in bold:

$$P(E=1|A=1) \;=\; \frac{0.29066 * 0.002}{0.002516442} \;=\; \mathbf{0.2310}$$

(b) $\boxed{P(E=1|A=1,B=0) = 0.3676}$

From Bayes rule:

$$\begin{aligned} P(E=1|A=1,B=0) &= \frac{P(A=1|E=1,B=0)P(E=1|B=0)}{P(A=1|B=0)} \\ &= \frac{P(A=1|E=1,B=0)P(E=1)}{P(A=1|B=0)} \quad \boxed{\textbf{conditional independence}} \end{aligned}$$

Denominator:

$$\begin{aligned} P(A=1|B=0) &= \sum_e P(A=1,E=e|B=0) \quad \boxed{\textbf{marginalization}} \\ &= \sum_e P(A=1|E=e,B=0)\, P(E=e|B=0) \quad \boxed{\textbf{product rule}} \\ &= \sum_e P(A=1|E=e,B=0)\, P(E=e) \quad \boxed{\textbf{conditional independence}} \\ &= P(A=1|E=1,B=0)P(E=1) + P(A=1|E=0,B=0)P(E=0) \\ &= 0.29 * 0.002 + 0.001 * (1 - 0.002) \\ &= \mathbf{0.001578} \end{aligned}$$

Substituting the result in bold:

$$P(E=1|A=1,B=0) \;=\; \frac{0.29 * 0.002}{0.001578} \;=\; \mathbf{0.3676}$$

Comparing (a) and (b), we find that $P(E{=}1|A{=}1,B{=}0) > P(E{=}1|A{=}1)$.

This agrees with common sense: if we know that a burglar did not trip the alarm, then we are more likely to believe than an earthquake was responsible.

(c) $\boxed{P(A=1|M=1) = 0.1501}$

From Bayes rule:

$$P(A=1|M=1) \;=\; \frac{P(M=1|A=1)P(A=1)}{P(M=1)}$$

Denominator:

$$\begin{aligned}
P(M=1) &= \sum_a P(M=1, A=a) \quad \boxed{\textbf{marginalization}} \\
&= \sum_a P(M=1|A=a)\, P(A=a) \quad \boxed{\textbf{product rule}} \\
&= P(M\!=\!1|A\!=\!0)\, P(A\!=\!0) + P(M\!=\!1|A\!=\!1)\, P(A\!=\!1) \\
&= 0.01 * (1 - 0.002516442) + 0.70 * 0.002516442 \quad \boxed{\textbf{using result from part (a)}} \\
&= \mathbf{0.01173634498}
\end{aligned}$$

Substituting into Bayes rule:

$$P(A=1|M=1) \;=\; \frac{0.7 * 0.002516442}{0.01173634498} \;=\; \mathbf{0.1501}$$

(d) $\boxed{P(A=1|M=1, J=0) = 0.0182}$

From Bayes rule:

$$\begin{aligned}
P(A=1|M=1, J=0) &= \frac{P(M=1, J=0|A=1)\, P(A=1)}{P(M=1, J=0)} \\
&= \frac{P(M=1|A=1)\, P(J=0|A=1)P(A=1)}{P(M=1, J=0)} \quad \boxed{\textbf{conditional independence}}
\end{aligned}$$

Denominator:

$$\begin{aligned}
P(M=1, J=0) &= \sum_a P(M=1, J=0, A=a) \quad \boxed{\textbf{marginalization}} \\
&= \sum_a P(M=1, J=0|A=a)P(A=a) \quad \boxed{\textbf{product rule}} \\
&= \sum_a P(M=1|A=a)P(J=0|A=a)P(A=a) \quad \boxed{\textbf{conditional independence}}
\end{aligned}$$

$$
\begin{aligned}
&= P(M\!=\!1|A\!=\!0)P(J\!=\!0|A\!=\!0)P(A\!=\!0) \;+\; P(M\!=\!1|A\!=\!1)P(J\!=\!0|A\!=\!1)P(A\!=\!1) \\
&= 0.01 * (1-0.05) * (1-0.002516442) + 0.7 * (1-0.9) * 0.002516442 \quad \boxed{\textbf{from part (a)}} \\
&= \mathbf{0.009652244741}
\end{aligned}
$$

Substituting into Bayes rule:

$$
P(A=1|M=1, J=0) \;=\; \frac{0.7 * (1-0.9) * 0.002516442}{0.009652244741} \;=\; 0.0182
$$

Comparing (c) and (d), we find that $P(A\!=\!1|M\!=\!1) > P(A\!=\!1|M\!=\!1, J\!=\!0)$.

This agrees with common sense: we're less likely to believe that the alarm has sounded after learning that John has not called.

(e) $\boxed{P(A=1|M=0) = 0.000764}$

From Bayes rule:

$$
\begin{aligned}
P(A=1|M=0) &= \frac{P(M=0|A=1)P(A=1)}{P(M=0)} \\
&= \frac{(1-0.7) * 0.002516442}{1-0.01173634498} \quad \boxed{\textbf{using results from parts (a,c)}} \\
&= 0.000764
\end{aligned}
$$

(f) $\boxed{P(A=1|M=0, B=1) = 0.826}$

From Bayes rule:

$$
\begin{aligned}
P(A=1|M=0, B=1) &= \frac{P(M=0|A=1, B=1)P(A=1|B=1)}{P(M=0|B=1)} \\
&= \frac{P(M=0|A=1)P(A=1|B=1)}{P(M=0|B=1)} \quad \boxed{\textbf{conditional independence}}
\end{aligned}
$$

Second term in numerator:

$$
\begin{aligned}
P(A=1|B=1) &= \sum_e P(A=1, E=e|B=1) \quad \boxed{\textbf{marginalization}} \\
&= \sum_e P(A=1|E=e, B=1)\, P(E=e|B=1) \quad \boxed{\textbf{product rule}}
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_e P(A=1|E=e,B=1)\,P(E=e) \quad \boxed{\textbf{conditional independence}} \\
&= P(A=1|E=0,B=1)\,P(E=0) \;+\; P(A=1|E=1,B=1)\,P(E=1) \\
&= 0.94 * (1-0.002) + 0.95 * 0.002 \\
&= \mathbf{0.94002}
\end{aligned}
$$

Denominator:

$$
\begin{aligned}
P(M=0|B=1) &= \sum_a P(M=0,A=a|B=1) \quad \boxed{\textbf{marginalization}} \\
&= \sum_a P(M=0|A=a,B=1)\,P(A=a|B=1) \quad \boxed{\textbf{product rule}} \\
&= \sum_a P(M=0|A=a)\,P(A=a|B=1) \quad \boxed{\textbf{conditional independence}} \\
&= P(M=0|A=0)\,P(A=0|B=1) \;+\; P(M=0|A=1)\,P(A=1|B=1) \\
&= (1-0.01)*(1-0.94002) \;+\; (1-0.7)*0.94002 \quad \boxed{\textbf{use result from above}} \\
&= \mathbf{0.3413862}
\end{aligned}
$$

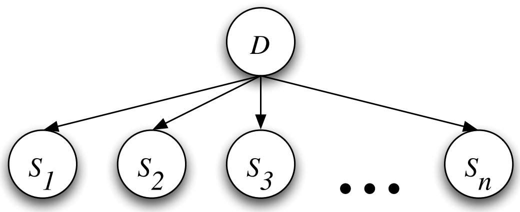
Substituting into Bayes rule:

$$
P(A=1|M=0,B=1) \;=\; \frac{(1-0.7)*0.94002}{0.3413862} \;=\; 0.826
$$

Comparing (e) and (f), we find that $P(A{=}1|M{=}0,B{=}1) \gg P(A{=}1|M{=}0)$.

This agrees with common sense: we're much more likely to believe that the alarm has sounded after learning that a burglary has occurred.

2.2 Probabilistic reasoning *(4 pts)*



Conditional probability tables (CPTs):

$$
\begin{aligned}
P(D\!=\!1) &= \frac{1}{2} \\
P(S_1\!=\!1|D\!=\!0) &= 1 \\
P(S_k\!=\!1|D\!=\!0) &= \frac{f(k-1)}{f(k)} \quad \textbf{where} \quad f(k) = 2^k + (-1)^k \quad \textbf{for} \quad k \geq 2 \\
P(S_k\!=\!1|D\!=\!1) &= \frac{1}{2}
\end{aligned}
$$

(a) **Likelihood ratio** *(3 pts)*

From Bayes rule: *(1 pt)*

$$
\begin{aligned}
P(D\!=\!0|S_1\!=\!1, S_2\!=\!1, \ldots, S_k\!=\!1) &= \frac{P(S_1\!=\!1, S_2\!=\!1, \ldots, S_k\!=\!1|D\!=\!0)\, P(D\!=\!0)}{P(S_1\!=\!1, S_2\!=\!1, \ldots, S_k\!=\!1)} \\
P(D\!=\!1|S_1\!=\!1, S_2\!=\!1, \ldots, S_k\!=\!1) &= \frac{P(S_1\!=\!1, S_2\!=\!1, \ldots, S_k\!=\!1|D\!=\!1)\, P(D\!=\!1)}{P(S_1\!=\!1, S_2\!=\!1, \ldots, S_k\!=\!1)}
\end{aligned}
$$

Taking the ratio: *(1 pt)*

$$
\begin{aligned}
r_k &= \frac{P(D\!=\!0|S_1\!=\!1, S_2\!=\!1, \ldots, S_k\!=\!1)}{P(D\!=\!1|S_1\!=\!1, S_2\!=\!1, \ldots, S_k\!=\!1)} \\
&= \frac{P(S_1\!=\!1, S_2\!=\!1, \ldots, S_k\!=\!1|D\!=\!0)\, P(D\!=\!0)}{P(S_1\!=\!1, S_2\!=\!1, \ldots, S_k\!=\!1|D\!=\!1)\, P(D\!=\!1)} \quad \boxed{\textbf{substituting from above}} \\
&= \frac{P(D\!=\!0)\, \prod_{i=1}^{k} P(S_i\!=\!1|D\!=\!0)}{P(D\!=\!1)\, \prod_{i=1}^{k} P(S_i\!=\!1|D\!=\!1)} \quad \boxed{\textbf{conditional independence}}
\end{aligned}
$$

Substituting the CPTs: *(1 pt)*

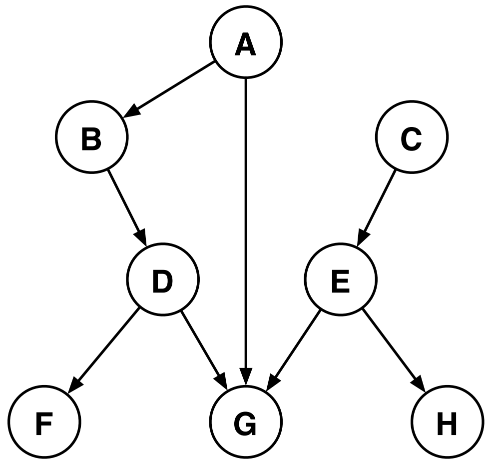
$$
\begin{aligned}
r_k &= \frac{\left(\frac{1}{2}\right)\ (1)\ \prod_{i=2}^{k}\left[\frac{f(i-1)}{f(i)}\right]}{\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^k} \\
&= 2^k\left[\frac{f(1)}{f(2)}\cdot\frac{f(2)}{f(3)}\ \cdots\ \frac{f(k-1)}{f(k)}\right] \\
&= 2^k\left[\frac{f(1)}{f(k)}\right] \\
&= \frac{2^k}{2^k+(-1)^k} \\
&= \frac{1}{1+\left(-\frac{1}{2}\right)^k}
\end{aligned}
$$

(b) **Explanation:** *(1 pt)*

The ratio r_k is less than one for even k and greater than one for odd k. Thus the diagnosis vacillates: on even days, the $D=1$ form of the disease seems more likely; on odd days, the opposite. In addition, the ratio r_k approaches unity as $k \to \infty$. Thus the diagnosis becomes more uncertain with each day. Note that conditioning on more evidence does not always reduce the amount of uncertainty.

2.3 True or false *(10 pts)*

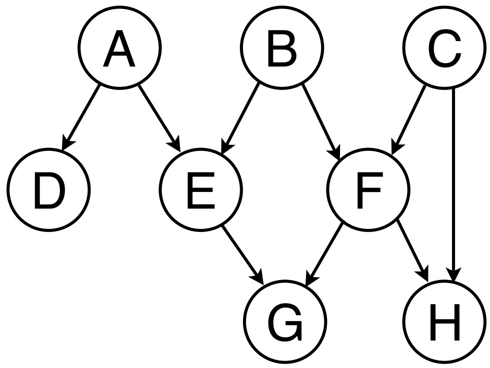
For the belief network shown below, indicate whether the following statements of marginal or conditional independence are **true (T)** or **false (F)**.



false	$P(B\|G,C) = P(B\|G)$
true	$P(F,G\|D) = P(F\|D)\,P(G\|D)$
true	$P(A,C) = P(A)\,P(C)$
false	$P(D\|B,F,G) = P(D\|B,F,G,A)$
true	$P(F,H) = P(F)\,P(H)$
true	$P(D,E\|F,H) = P(D\|F)\,P(E\|H)$
false	$P(F,C\|G) = P(F\|G)\,P(C\|G)$
true	$P(D,E,G) = P(D)\,P(E)\,P(G\|D,E)$
true	$P(H\|C) = P(H\|A,B,C,D,F)$
false	$P(H\|A,C) = P(H\|A,C,G)$

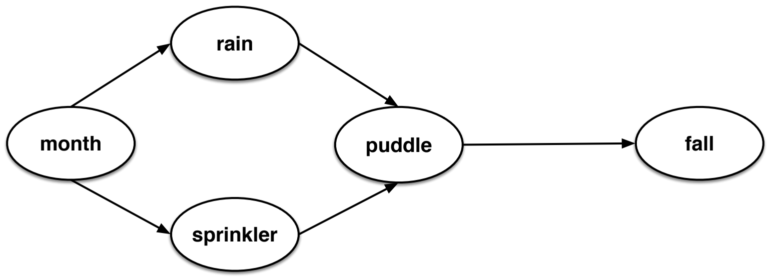
2.4 More on Belief Networks *(10 pts)*

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



false	$P(F\|H)$	=	$P(F\|C,H)$
true	$P(E\|A,B)$	=	$P(E\|A,B,F)$
false	$P(E,F\|B,G)$	=	$P(E\|B,G)\,P(F\|B,G)$
false	$P(F\|B,C,G,H)$	=	$P(F\|B,C,E,G,H)$
false	$P(A,B\|D,E,F)$	=	$P(A,B\|D,E,F,G,H)$
false	$P(D,E,F)$	=	$P(D)\,P(E\|D)\,P(F\|E)$
true	$P(A\|F)$	=	$P(A)$
false	$P(E,F)$	=	$P(E)\,P(F)$
true	$P(D\|A)$	=	$P(D\|A,E)$
true	$P(B,C)$	=	$P(B)\,P(C)$

2.5 Conditional independence *(8 pts)*



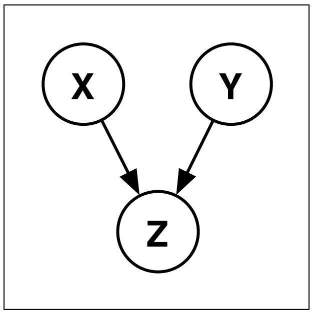
List all the conditional independence relations that must hold in any probability distribution represented by this DAG. More specifically, list all tuples $\{X, Y, E\}$ such that $P(X, Y|E) = P(X|E)P(Y|E)$, where:

$$\begin{aligned} X, Y &\in \{\texttt{month}, \texttt{rain}, \texttt{sprinkler}, \texttt{puddle}, \texttt{fall}\}, \\ E &\subseteq \{\texttt{month}, \texttt{rain}, \texttt{sprinkler}, \texttt{puddle}, \texttt{fall}\}, \\ X &\neq Y, \\ X, Y &\notin E. \end{aligned}$$

X	Y	E
month	puddle	{rain,sprinkler}
month	puddle	{rain,sprinkler,fall}
month	fall	{sprinkler,rain}
month	fall	{sprinkler,rain,puddle}
month	fall	{sprinkler,puddle}
month	fall	{rain,puddle}
month	fall	{puddle}
sprinkler	rain	{month}
sprinkler	fall	{puddle}
sprinkler	fall	{puddle,rain}
sprinkler	fall	{puddle,month}
sprinkler	fall	{puddle,rain,month}
rain	fall	{puddle}
rain	fall	{puddle,sprinkler}
rain	fall	{puddle,month}
rain	fall	{puddle,sprinkler,month}

Grading: +1/2 pt for each correct tuple, -1/2 pt for each incorrect one.

2.6 Noisy-OR *(7 pts)*



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.

(a)	$P(Z=1\|X=0,Y=0)$	$<$	$P(Z=1\|X=0,Y=1)$	**because** $p_y > 0$
(b)	$P(Z=1\|X=1,Y=0)$	$<$	$P(Z=1\|X=0,Y=1)$	**because** $p_x < p_y$
(c)	$P(Z=1\|X=1,Y=0)$	$<$	$P(Z=1\|X=1,Y=1)$	**because** $p_x < p_x + p_x(1-p_y)$
(d)	$P(X=1)$	$=$	$P(X=1\|Y=1)$	**independence**
(e)	$P(X=1)$	$<$	$P(X=1\|Z=1)$	
(f)	$P(X=1\|Z=1)$	$>$	$P(X=1\|Y=1,Z=1)$	**explaining away**
(g)	$P(X=1)\,P(Y=1)\,P(Z=1)$	$<$	$P(X=1,Y=1,Z=1)$	

For CSE 250A ONLY *(5 pts)*:

(e) Consider the term on the right hand side:

$$P(X\!=\!1|Z\!=\!1) = P(X\!=\!1)\left[\frac{P(Z\!=\!1|X\!=\!1)}{P(Z\!=\!1)}\right] \quad \boxed{\textbf{Bayes rule}}$$

Now consider the numerator of the ratio:

$$\begin{aligned}
P(Z\!=\!1|X\!=\!1) &= \sum_y P(Z\!=\!1, Y\!=\!y|X\!=\!1) \quad \boxed{\textbf{marginalization}}\\
&= \sum_y P(Y\!=\!y|X\!=\!1)\, P(Z\!=\!1|X\!=\!1, Y\!=\!y) \quad \boxed{\textbf{product rule}}\\
&= \sum_y P(Y\!=\!y)\, P(Z\!=\!1|X\!=\!1, Y\!=\!y) \quad \boxed{\textbf{independence}}
\end{aligned}$$

Similarly, for the denominator:

$$\begin{aligned}
P(Z\!=\!1) &= \sum_{x,y} P(X\!=\!x, Y\!=\!y, Z\!=\!1) \quad \boxed{\textbf{marginalization}}\\
&= \sum_{x,y} P(X\!=\!x)\, P(Y\!=\!y|X\!=\!1)\, P(Z\!=\!1|X\!=\!x, Y\!=\!y) \quad \boxed{\textbf{product rule}}\\
&= \sum_{x,y} P(X\!=\!x)\, P(Y\!=\!y)\, P(Z\!=\!1|X\!=\!x, Y\!=\!y) \quad \boxed{\textbf{independence}}\\
&= \sum_{y} P(Y\!=\!y) \sum_x P(X\!=\!x) P(Z\!=\!1|X\!=\!x, Y\!=\!y) \quad \boxed{\textbf{grouping terms}}\\
&< \sum_{y} P(Y\!=\!y)\, \max_x P(Z\!=\!1|X\!=\!x, Y\!=\!y) \quad \boxed{\textbf{average is less than max}}\\
&< \sum_{y} P(Y\!=\!y)\, P(Z\!=\!1|X\!=\!1, Y\!=\!y) \quad \boxed{\textbf{noisy-OR}}\\
&= \sum_{y} P(Z\!=\!1|X\!=\!1) \quad \boxed{\textbf{substituting above}}
\end{aligned}$$

Hence the numerator in this ratio is always larger than the denominator. It follows that the ratio is always greater than one, and that $P(X\!=\!1|Z\!=\!1) > P(X\!=\!1)$.

(f) Consider the term on the right hand side:

$$\begin{aligned}
P(X\!=\!1|Z\!=\!1) &= \sum_y P(X\!=\!1, Y\!=\!y|Z\!=\!1) \quad \boxed{\textbf{marginalization}}\\
&= \sum_y P(Y\!=\!y|Z\!=\!1) P(X\!=\!1|Y\!=\!y, Z\!=\!1) \quad \boxed{\textbf{product rule}}\\
&> \min_y P(X\!=\!1|Y\!=\!y, Z\!=\!1) \quad \boxed{\textbf{average is greater than min}}\\
&= \min\{P(X\!=\!1|Y\!=\!0, Z\!=\!1), P(X\!=\!1|Y\!=\!1, Z\!=\!1)\} \quad \boxed{\textbf{substituting } y \in \{0,1\}}\\
&= \min\{1,\ P(X\!=\!1|Y\!=\!1, Z\!=\!1)\}\\
&= P(X\!=\!1|Y\!=\!1, Z\!=\!1)
\end{aligned}$$

(g) Consider the term on the left hand side:

$$\begin{aligned} P(X=1, Y=1, Z=1) &= P(X=1)\, P(Y=1|X=1)\, P(Z=1|X=1, Y=1) \quad \boxed{\textbf{product rule}} \\ &= P(X=1)\, P(Y=1)\, P(Z=1|X=1, Y=1) \quad \boxed{\textbf{independence}} \end{aligned}$$

Consider the last conditional probability on the right hand side:

$$\begin{aligned} P(Z=1|X=1, Y=1) &= \max_{x,y} \Big[P(Z=1|X=x, Y=y) \Big] \quad \boxed{\textbf{monotonicity of noisy-OR}} \\ &> \sum_{x,y} P(X=x, Y=y) P(Z=1|X=x, Y=y) \quad \boxed{\textbf{max is greater than average}} \\ &= \sum_{x,y} P(X=x, Y=y, Z=1) \quad \boxed{\textbf{product rule}} \\ &= P(Z=1) \quad \boxed{\textbf{marginalization}} \end{aligned}$$

We obtain the desired result by substituting this inequality for $P(Z=1|X=1, Y=1)$ into the right hand side of the equality for $P(X=1, Y=1, Z=1)$.