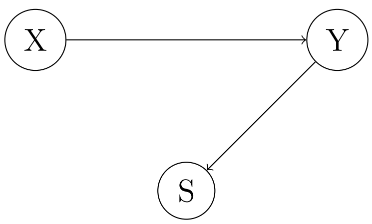
Lemma 1 Proof Take 2

Canyon Foot

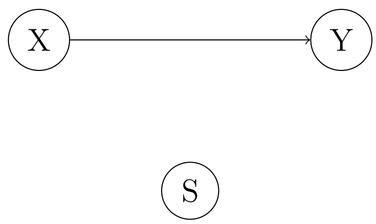
October 27, 2019

For binary random variables X, Y, S, we wish to show that for the two causal graphs, G_1, $G_2 = G_1 \backslash \{Y \to S\}$ we can construct corresponding probability functions P_1, P_2 such that $P_1(Y = y, X = x|S = 1) = P_2(Y = y, X = x|S = 1)$ for all $x, y \in \{0, 1\}$ but $P(Y = y|X = x) \neq P(Y = y|X = x)$ for some $x, y \in \{0, 1\}$. This is exactly the condition that s-recoverability requires, so this example proves that G_1 is not s-recoverable.

G_1



G_2



The graph structure gives us access to the following conditionals, which we have filled in with particular values.

For P_1, we have:

x	$P_1(X = x)$
0	1/2
1	1/2

x	y	$P_1(Y = y\|X = x)$
0	0	1/2
0	1	1/2
1	0	1/2
1	1	1/2

y	s	$P_1(S = 1\|Y = y)$
1	1	3/4
0	1	1/4

And for P_2, we have:

x	$P_2(X = x)$
0	1/2
1	1/2

x	y	$P_2(Y = y\|X = x)$
0	0	1/4
0	1	3/4
1	0	1/4
1	1	3/4

For concision, we will abbreviate $P_1(X = x, Y = 1|S = 1)$ as $P_1(x, y|S = 1)$ (and similar for other quantities). Then, by Bayes rule and the law of total probability we have that:

$$\begin{aligned}
P_1(x, y|S = 1) &= \frac{P_1(S = 1|x, y)P_1(x, y)}{P_1(S = 1)} \\
&= \frac{P_1(S = 1|x, y)P_1(x, y)}{\sum_{x', y' \in \{0,1\}} P_1(S = 1|x', y')P_1(x', y')} \\
&= \frac{P_1(S = 1|x, y)P_1(x, y)}{\sum_{x', y' \in \{0,1\}} P_1(S = 1|x', y')P_1(x', y')}
\end{aligned}$$

So, d-separation (and therefore conditional independence) we get:

$$\begin{aligned}
P_1(x,y|S=1) &= \frac{P_1(S=1|x,y)P(x,y)}{\sum_{x',y'\in\{0,1\}} P_1(S=1|x',y')P_1(x',y')} \\
&= \frac{P_2(S=1|y)P_2(y|x)P_2(x)}{\sum_{x',y'\in\{0,1\}} P_2(S=1|y')P_1(y'|x')P_1(x')}
\end{aligned}$$

Then, since we have defined $P_1(y|x)$ and $P_1(x)$ to be always be $\frac{1}{2}$ we get these terms to cancel, leaving $\sum_{x',y'\in\{0,1\}} P_1(S=1|y')=2$ in the denominator. So,

$$P_1(x,y|S=1)=\frac{P_1(S=1|y)}{2}.$$

The table for $P_1(x,y|S=1)$ is then:

x	y	$P_1(y,x\|S=1)=\frac{P_1(S=1\|y)}{2}$
0	0	1/8
0	1	3/8
1	0	1/8
1	1	3/8

Now we move on to P_2. Since there is no path between S and X or Y in G_2, we have that:

$$P_2(y,x|S=1)=P_2(y,x)=P_2(y|x)P_2(x)=\frac{P_2(y|x)}{2}$$

This gives the table:

x	y	$P_2(y,x\|S=1)=\frac{P_2(y\|x)}{2}$
0	0	1/8
0	1	3/8
1	0	1/8
1	1	3/8

So, $P_1(Y=y|X=x|S=1)=P_2(Y=y|X=x|S=1)$ for all $x,y\in\{0,1\}$. But by consulting the original tables, we can see that $P_1(y|x)\neq P_2(y|x)$. Then G_1 is not s-recoverable.