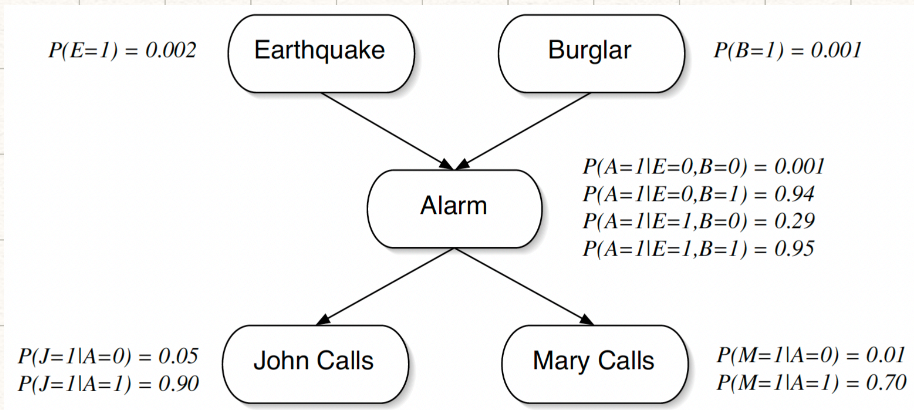
2.1 Probabilistic inference

We recall the alarm BN from class. The DAG and CPTs are shown below:



In this task we will compute numeric values for various probabilities, exploiting relations of marginal and conditional independence as much as possible to simplify our calculations.

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

$$P(E=1 \mid A=1) \overset{\text{Bayes}}{=} \frac{P(A=1 \mid E=1)P(E=1)}{P(A=1)}$$

$$\overset{\text{marginalization}}{=} \frac{\sum_b P(A=1, B=b \mid E=1)P(E=1)}{\sum_b \sum_e P(A=1, B=b, E=e)}$$

We look at the numerator first.

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

$$\overset{\substack{\text{B\&E} \\ \text{marg. indep.}}}{=} \sum_b P(B=b)P(A=1 \mid B=b, E=1)P(E=1)$$

$$= 0.002\,(0.999 \cdot 0.29 + 0.001 \cdot 0.95)$$

Now for the denominator

$$\sum_b \sum_e P(A=1, B=b, E=e) \overset{\text{P.R.}}{=} \sum_b \sum_e P(B=b)P(E=e \mid B=b)P(A \mid B=b, E=e)$$

$$\overset{\substack{\text{B\&E} \\ \text{marg. indep.}}}{=} \sum_b \sum_e P(B=b)P(E=e)P(A \mid B=b, E=e)$$

$$= \sum_e P(E=e)\left[P(B=0)P(A \mid B=0, E=e) + P(B=1)P(A \mid B=1, E=e)\right]$$

$$= P(E=0)\left[P(B=0)P(A \mid B=0, E=0) + P(B=1)P(A \mid B=1, E=0)\right]$$

$$+ P(E=1)\left[P(B=0)P(A \mid B=0, E=1) + P(B=1)P(A \mid B=1, E=1)\right]$$

$$= 0.998\left[0.999 \cdot 0.001 + 0.001 \cdot 0.94\right] + 0.002\left[0.999 \cdot 0.29 + 0.001 \cdot 0.95\right]$$

$$\Rightarrow P(E=1 \mid A=1) = \frac{\sum_b P(A=1, B=b \mid E=1)P(E=1)}{\sum_b \sum_e P(A=1, B=b, E=e)} = \frac{0.002\,(0.999 \cdot 0.29 + 0.001 \cdot 0.95)}{0.998\left[0.999 \cdot 0.001 + 0.001 \cdot 0.94\right] + 0.002\left[0.999 \cdot 0.29 + 0.001 \cdot 0.95\right]} = 0.231$$

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

$$P(E=1 \mid A=1, B=0) \overset{\text{P.R.}}{=} \frac{P(E=1, A=1, B=0)}{P(B=0)\,P(A=1 \mid B=0)}$$

Numerator:

$$P(E=1, A=1, B=0) \overset{\text{P.R.}}{=} P(E=1)\,P(B=0 \mid E=1)\,P(A=1 \mid E=1, B=0)$$

$$\overset{\substack{\text{B\&E} \\ \text{indep.}}}{=} P(E=1)\,P(B=0)\,P(A=1 \mid E=1, B=0)$$

$$= \underline{0.002 \cdot 0.999 \cdot 0.29}$$

Denominator:

$$P(B=0)\,P(A=1 \mid B=0) \overset{\text{margin.}}{=} P(B=0) \sum_e P(A=1, E=e \mid B=0)$$

$$\overset{\text{P.R.}}{=} P(B=0) \sum_e P(E=e \mid B=0)\,P(A=1 \mid E=e, B=0)$$

$$\overset{\substack{\text{B\&E} \\ \text{indep.}}}{=} P(B=0)\left[P(E=0)\,P(A=1 \mid E=0, B=0) + P(E=1)\,P(A=1 \mid E=1, B=0)\right]$$

$$= \underline{0.999\,(0.998 \cdot 0.001 + 0.002 \cdot 0.29)}$$

$$\Rightarrow \underline{\underline{P(E=1 \mid A=1, B=0)}} = \frac{P(E=1, A=1, B=0)}{P(B=0)\,P(A=1 \mid B=0)} = \frac{0.002 \cdot 0.999 \cdot 0.29}{0.999\,(0.998 \cdot 0.001 + 0.002 \cdot 0.29)} \approx \underline{\underline{0.368}}$$

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

$$P(A=1 \mid M=1) \overset{\text{Bayes}}{=} \frac{P(M=1 \mid A=1)\,P(A=1)}{P(M=1)}$$

Numerator:

From the denominator in a) we have $P(A=1) = 0.998\,[0.999 \cdot 0.001 + 0.001 \cdot 0.94] + 0.002\,[0.999 \cdot 0.29 + 0.001 \cdot 0.95] = 0.002516442$

Thus, $P(M=1 \mid A=1)\,P(A=1) \overset{\text{CPT}}{=} 0.7 \cdot \left(0.998\,[0.999 \cdot 0.001 + 0.001 \cdot 0.94] + 0.002\,[0.999 \cdot 0.29 + 0.001 \cdot 0.95]\right) = 0.0017615094$

Denominator:

$$P(M=1) \overset{\text{marg.}}{=} \sum_a P(M=1, A=a) \overset{\text{P.R.}}{=} \sum_a P(A=a)\,P(M=1 \mid A=a) = P(A=0)\,P(M=1 \mid A=0) + P(A=1)\,P(M=1 \mid A=1)$$

$$= (1 - 0.002516442) \cdot 0.01 + 0.002516442 \cdot 0.7$$

$$\Rightarrow \underline{\underline{P(A=1 \mid M=1)}} = \frac{P(M=1 \mid A=1)\,P(A=1)}{P(M=1)} = \frac{0.0017615094}{(1 - 0.002516442) \cdot 0.01 + 0.002516442 \cdot 0.7} = \underline{\underline{0.150}}$$

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

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

Numerator:

$$P(M=1, J=0 \mid A=1)P(A=1) \overset{\substack{\text{M\&J}\\ \text{indep}}}{=} P(M=1 \mid A=1)P(J=0 \mid A=1)P(A=1) \overset{\substack{\text{table}\\ \text{and } P(A=1)\\ \text{from a)}}}{=} 0.7 \cdot (1-0.9) \cdot 0.002516442 = 0.00017615094$$

Denominator:

$$P(M=1, J=0) \overset{\text{marg.}}{=} \sum_a P(M=1, J=0, A=a)$$

$$\overset{\substack{\text{P.R.}\\ \text{and M\&J}\\ \text{indep.}}}{=} \sum_a P(A=a)P(M=1 \mid A=a)P(J=0 \mid A=a)$$

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

$$\overset{\substack{\text{tables}\\ \text{and } P(A=1) \text{ from a)}}}{=} (1-0.00251664) \cdot 0.01 \cdot (1-0.05) + 0.00251664 \cdot 0.7 \cdot (1-0.9) = 0.0096 5225672$$

$$\Rightarrow \underline{P(A=1 \mid M=1, J=0)} = \frac{P(M=1, J=0 \mid A=1)P(A=1)}{P(M=1, J=0)} = \frac{0.00017615094}{0.0096 5225672} = \underline{0.0182}$$

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

$$P(A=1 \mid M=0) \overset{\text{Bayes}}{=} \frac{P(M=0 \mid A=1)P(A=1)}{P(M=0)}$$

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

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

$$\overset{\substack{\text{tables and}\\ P(M=1) \text{ from c)}\\ \text{and } P(A=1) \text{ from a)}}}{=} \frac{(1-0.7) \cdot 0.00251664}{1 - ((1-0.002516442) \cdot 0.01 + 0.002516442 \cdot 0.7)}$$

$$\underline{P(A=1 \mid M=0) = 0.000764}$$

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

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

We separate into three calculations.

$P(M=0 \mid A=1, B=1) = P(M=0 \mid A=1)$ because d-separation holds as there is a direct path $B \to A \to M$

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

$$P(A=1 \mid B=1) \overset{\text{marg.}}{=} \sum_e P(A=1, E=e \mid B=1) \overset{\text{P.R.}}{=} \sum_e P(E=e)P(A=1 \mid E=e, B=1) = P(E=0)P(A=1 \mid E=0, B=1) + P(E=1)P(A=1 \mid E=1, B=1) \overset{\text{tables}}{=} 0.998 \cdot 0.94 + 0.002 \cdot 0.95 = 0.94002$$

$$P(M=0 \mid B=1) \overset{\text{marg.}}{=} \sum_a P(M=0, A=a \mid B=1) \overset{\text{P.R.}}{=} \sum_a P(A=a \mid B=1)P(M=0 \mid A=a, B=1) \overset{\substack{\text{d-separation}\\ B \to A \to M}}{=} \sum_a P(A=a \mid B=1)P(M=0 \mid A=a)$$

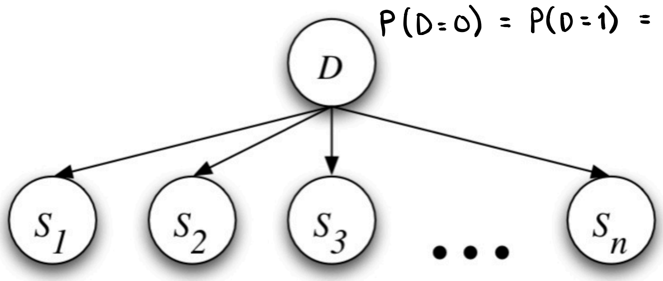
$$= P(A=0 \mid B=1)P(M=0 \mid A=0) + P(A=1 \mid B=1)P(M=0 \mid A=1)$$

$$\overset{\substack{\text{tables}\\ \text{and } P(A=1 \mid B=1)\\ \text{from above}}}{=} (1-0.94002) \cdot (1-0.01) + 0.94002 \cdot (1-0.7) = 0.3413862$$

$$\Rightarrow P(A=1 \mid M=0, B=1) = \frac{P(M=0 \mid A=1, B=1)P(A=1 \mid B=1)}{P(M=0 \mid B=1)} = \frac{0.3 \cdot 0.94002}{0.3413862} = \underline{0.826}$$

2.2 Probabilistic reasoning

A patient is known to have a rare disease which comes in two forms, represented by the values of the random variable $D \in \{0,1\}$. Symptoms of the disease are represented by the random variables $S_k \in \{0,1\}$, and knowledge about the disease is summarized by the belief network:



$P(D=0) = P(D=1) = \frac{1}{2}$ $\qquad P(S_1=1 \mid D=0) = 1$

$P(S_k=1 \mid D=0) = \frac{f(k-1)}{f(k)}$, $k \geq 2$, where $f(k) = 2^k + (-1)^k$

$P(S_k=1 \mid D=1) = \frac{1}{2} \quad \forall k$

We suppose a test is done on the k^{th} day of the month to determine if k^{th} symptom is active (and it is). Thus, on the k^{th} day, the doctor observes $\{S_1=1, S_2=1, \dots, S_k=1\}$ and makes a diagnosis by computing

$r_k = \frac{P(D=0 \mid S_1=1, S_2=1, \dots, S_k=1)}{P(D=1 \mid S_1=1, S_2=1, \dots, S_k=1)}$. If $r_k > 1$: diagnosis is $D=0$. Else: diagnosis is $D=1$.

a) Compute r_k as a function of k. How does the diagnosis depend on the day of the month?

$r_k = \frac{P(D=0 \mid S_1=1, S_2=1, \dots, S_k=1)}{P(D=1 \mid S_1=1, S_2=1, \dots, S_k=1)}$ We use Bayes rule on both the numerator and the denominator.

Bayes

$$= \frac{\dfrac{P(S_1=1,S_2=1,\dots,S_k=1 \mid D=0)\cancel{P(D=0)}}{\cancel{P(S_1=1,S_2=1,\dots,S_k=1)}}}{\dfrac{P(S_1=1,S_2=1,\dots,S_k=1 \mid D=1)\cancel{P(D=1)}}{\cancel{P(S_1=1,S_2=1,\dots,S_k=1)}}}$$

(Because $P(D=0)=P(D=1)$ we can remove them)

Cancelling terms

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

We compute the numerator and denominator separately

$$P(S_1=1,S_2=1,\dots,S_k=1 \mid D=0) = \prod_{j=1}^{k} P(S_j=1 \mid D=0) = P(S_1=1 \mid D=0)\prod_{j=2}^{k} P(S_j=1 \mid D=0)$$

(S_i & S_j are marg. indep for $i \neq j$)

$$= 1 \cdot \prod_{j=2}^{k} P(S_j=1 \mid D=0) = 1 \cdot \frac{f(1)}{\cancel{f(2)}} \cdot \frac{\cancel{f(2)}}{\cancel{f(3)}} \cdot \frac{\cancel{f(3)}}{\cancel{f(4)}} \cdots \frac{\cancel{f(k-2)}}{\cancel{f(k-1)}} \frac{\cancel{f(k-1)}}{f(k)}$$

$$= \frac{f(1)}{f(k)} = \frac{1}{2^k + (-1)^k}$$

$$P(S_1=1,S_2=1,\dots,S_k=1 \mid D=1) = \prod_{i=1}^{k} P(S_i=1 \mid D=1) = \left(\frac{1}{2}\right)^k \Rightarrow \underline{\underline{r_k = \frac{\frac{1}{2^k+(-1)^k}}{\left(\frac{1}{2}\right)^k} = \frac{2^k}{2^k+(-1)^k}}}$$

For $k = 2p-1$, $p=\{1,2,\dots\}$, $r_k > 1 \Rightarrow$ Diagnosis is $D=0$ (Odd days of the month)

For $k = 2p$, $p=\{1,2,\dots\}$, $r_k < 1 \Rightarrow$ Diagnosis is $D=1$ (Even days of the month)

b) Does the diagnosis become more or less certain as more symptoms are observed?

As more symptoms are observed, k becomes larger, and r_k tends towards 1 as $\lim_{k \to \infty} r_k = 1$.

This is the point where the diagnosis shifts, so the diagnosis becomes less certain as more symptoms are observed

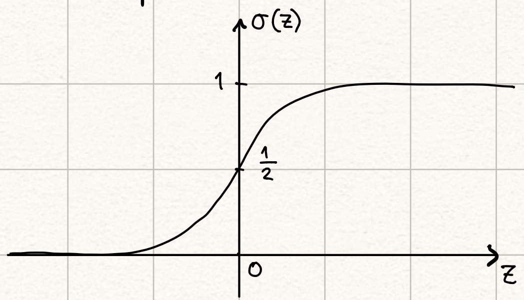
2.3 Sigmoid function

We let $Y \in \{0,1\}$ denote a binary random variable that depends on k other RVs X_i as:

$$P(Y=1 \mid X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k) = \sigma\left(\sum_{i=1}^{k} \omega_i x_i\right) \quad \text{where} \quad \sigma(z) = \frac{1}{1+e^{-z}}$$

The real-valued parameters ω_i are known as weights.

In this problem we will sketch the function $\sigma(z)$ and verify some properties.



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

$$\sigma'(z) = \frac{0 - (-e^{-z})}{(1+e^{-z})^2} = \frac{e^{-z}}{e^{-2z} + 2e^{-z} + 1} \cdot \frac{\frac{1}{e^{-z}}}{\frac{1}{e^{-z}}} = \frac{1}{e^{-z} + 2 + e^{z}}$$

$$e^{z} + e^{-z} + 2 = (1+e^{z})(1+e^{-z})$$

$$\Rightarrow \sigma'(z) = \frac{1}{(1+e^{z})(1+e^{-z})} = \sigma(z)\sigma(-z) \qquad \text{QED}$$

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

$$\sigma(-z) + \sigma(z) = \frac{1}{1+e^{-z}} + \frac{1}{1+e^{z}} = \frac{1+e^{z}+1+e^{-z}}{(1+e^{-z})(1+e^{z})} = \frac{\cancel{2+e^{z}+e^{-z}}}{\cancel{2+e^{z}+e^{-z}}} = 1 \qquad \text{QED}$$

c) Verify $L(\sigma(z)) = z$, where $L(p) = \log\left(\frac{p}{1-p}\right)$ is the log-odds function.

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

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

$$\Rightarrow L(\sigma(z)) = \log\left(\frac{\sigma(z)}{1-\sigma(z)}\right) = \log(e^{z}) = z \qquad \text{QED.}$$

d) Verify $\omega_i = L(p_i)$, where $p_i = P(Y=1 \mid X_i = 1, X_j = 0 \ \forall j \neq i)$.

We have $P(Y=1 \mid X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k) = \sigma\left(\sum_{i=1}^{k} \omega_i x_i\right)$, which here becomes

$$P = P(Y=1 \mid X_i = 1, X_j = 0 \ \forall j \neq i) = \sigma\left(\sum_{i=1}^{k} \omega_i \cdot x_i\right)$$

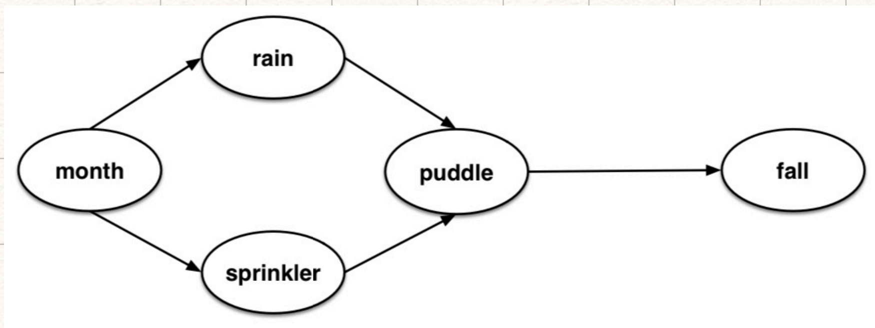
$$\Rightarrow p_i = \sigma(\omega_i x_i) = \sigma(\omega_i) \quad \text{as } x_i = 1 \ \forall i$$

From c) we know that $L(\sigma(z)) = z$, so if we take the log-odds on both sides, we get

$$L(p_i) = L(\sigma(\omega_i)) \Leftrightarrow \omega_i = L(p_i) \qquad \text{QED}$$

2.4 Conditional independence

We consider the DAG shown below:



We have the following:

$X, Y \in \{month, rain, sprinkler, puddle, fall\}$

$E \subseteq \{month, rain, sprinkler, puddle, fall\}$

$X \neq Y$

$X, Y \notin E$

List all tuples $\{X, Y, E\}$ such that $P(X, Y \mid E) = P(X \mid E) P(Y \mid E)$.

What we want is all events X and Y that are independent given the event or events E.

$\{X,Y,E\} = \{X = rain, Y = sprinkler, E = month\}$

$\{X,Y,E\} = \{X = fall, Y = sprinkler, E = puddle\}$

$\{X,Y,E\} = \{X = fall, Y = rain, E = puddle\}$

$\{X,Y,E\} = \{X = fall, Y = month, E = puddle\}$

$\{X,Y,E\} = \{X = fall, Y = rain, E = \{puddle, month\}\}$

$\{X,Y,E\} = \{X = fall, Y = sprinkler, E = \{puddle, month\}\}$

$\{X,Y,E\} = \{X = fall, Y = sprinkler, E = \{puddle, rain\}\}$

$\{X,Y,E\} = \{X = fall, Y = month, E = \{puddle, rain\}\}$

$\{X,Y,E\} = \{X = fall, Y = month, E = \{puddle, sprinkler\}\}$

$\{X,Y,E\} = \{X = fall, Y = rain, E = \{puddle, sprinkler\}\}$

$\{X,Y,E\} = \{X = fall, Y = month, E = \{rain, sprinkler\}\}$

$\{X,Y,E\} = \{X = puddle, Y = month, E = \{rain, sprinkler\}\}$

$\{X,Y,E\} = \{X = puddle, Y = month, E = \{rain, sprinkler, fall\}\}$

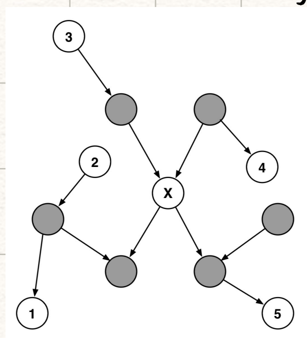
$\{X,Y,E\} = \{X = fall, Y = month, E = \{rain, sprinkler, puddle\}\}$

$\{X,Y,E\} = \{X = fall, Y = rain, E = \{month, sprinkler, puddle\}\}$

$\{X,Y,E\} = \{X = fall, Y = sprinkler, E = \{month, rain, puddle\}\}$

2.5 Markov blankets

We have the following picture:

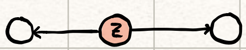


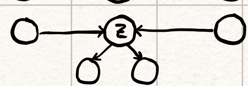
The Markov blanket B_x of a node X in a belief network includes its parents, its children and the parents of its children (excluding the node X itself).

Appealing to the conditions of d-sep, prove that we have $P(X,Y|B_x) = P(X|B_x)P(Y|B_x)$ for any node Y outside B_x.

We have the conditions of d-separations:

1. Edges align: $Z \in E$ 

2. Edges diverge: $Z \in E$ 

3. Edges converge: $Z \notin E$  $\text{descendants}(Z) \cap E = \emptyset$

① For Y being node 1, the child of a parent of a child of X, we have the following path:

B_x

We observe that the way from X to A through B is not blocked as the edges on each side of B converge and B is part of the evidence (d-separation condition 3 not applied). If we continue, we see that the edges on each side of A diverge and $A \in E$. Thus, by d-sep condition 1, the equality holds.

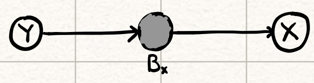
② For Y being node 2, the parent of a parent of a child of X, we have the following path:

and we need to show that $P(X,Y|A,B) = P(X|A,B)P(Y|A,B)$.

From Y to B through A, we observe that the edges align and $A \in E$. Thus, the path is blocked, and the equality $P(X,Y|B_x) = P(X|B_x)P(Y|B_x)$ holds.

③

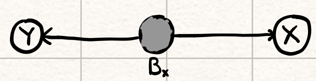
For Y being node 3, the parent of a parent of X, we have the following path:



In this case, the edges align, and condition 1 is therefore applied, and $P(X,Y|B_x) = P(X|B_x)P(Y|B_x)$.

④

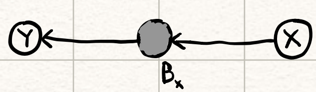
For Y being node 4, the child of a parent of X, we have the following path:



We observe that the edges diverge, and condition 2 is therefore applied, and $P(X,Y|B_x) = P(X|B_x)P(Y|B_x)$.

⑤

For Y being node 5, the child of a child of X, we have the following path:



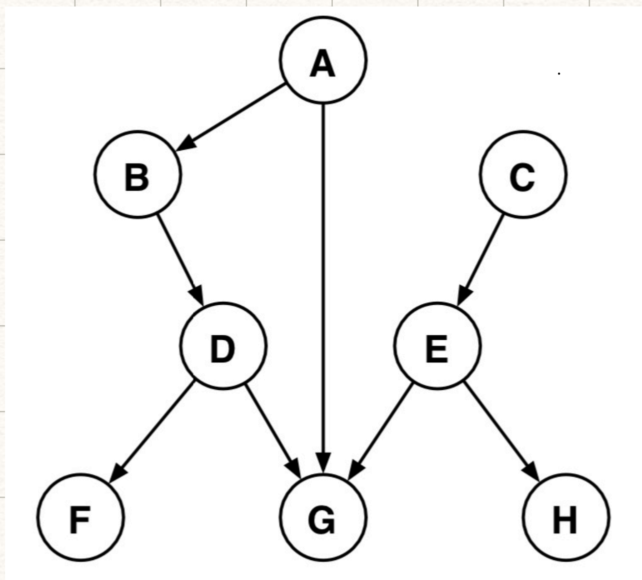
In this case, the edges align, and condition 1 is therefore applied, and $P(X,Y|B_x) = P(X|B_x)P(Y|B_x)$.

These five cases cover the nodes out of the nodes of the Markov Blanket, i.e. the nodes out of the parents of X, children of X and the parents of children of X excluding X. Thus, no other path than these five can exist, and we have shown that Y and X are cond. indep. given B_x for all five cases.

$\Rightarrow P(X,Y|B_x) = P(X|B_x)P(Y|B_x)$ QED

2.6 True or False

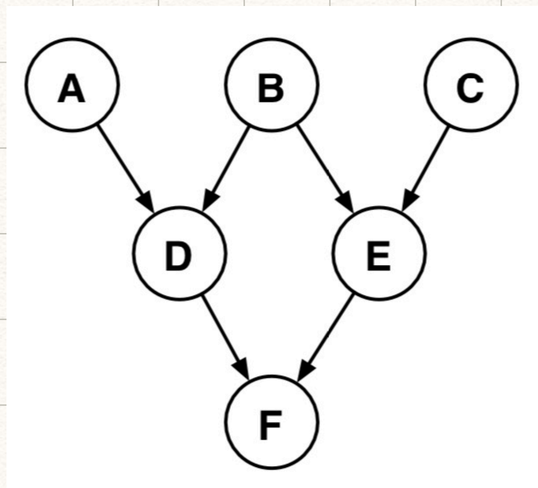
We consider the following belief network:



Indicate whether the following statements of marginal or conditional independence are True or False.

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

2.7 Subsets

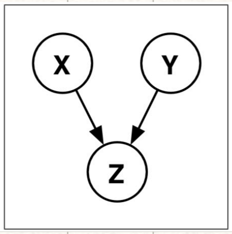


For the belief network above, we consider the following statements of marginal or conditional independence.

Indicate the largest subset of nodes $S \subset \{A,B,C,D,E,F\}$ for which each statement is true.

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

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

We suppose that the nodes in the network represent binary RVs and that the CPT for $P(Z|X,Y)$ is parametrized by a noisy-OR model, as shown above. We also suppose that

$$0 < P(X=1) < 1 \quad \text{and} \quad 0 < P(Y=1) < 1$$

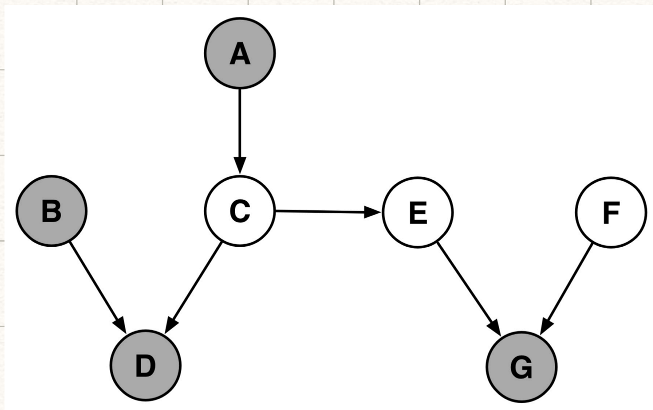
while the parameters of the noisy-OR model satisfy

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

Consider the following pairs of probabilities. In each case, indicate $<$, $>$ or $=$.

	$P(X=1)$	$=$	$P(X=1)$
(a)	$P(Z=1\|X=0,Y=0)$	$<$	$P(Z=1\|X=0,Y=1)$
(b)	$P(Z=1\|X=1,Y=0)$	$<$	$P(Z=1\|X=0,Y=1)$
(c)	$P(Z=1\|X=1,Y=0)$	$<$	$P(Z=1\|X=1,Y=1)$
(d)	$P(X=1)$	$<$	$P(X=1\|Z=1)$
(e)	$P(X=1)$	$=$	$P(X=1\|Y=1)$
(f)	$P(X=1\|Z=1)$	$>$	$P(X=1\|Y=1,Z=1)$
(g)	$P(X=1)\,P(Y=1)\,P(Z=1)$	$<$	$P(X=1,Y=1,Z=1)$

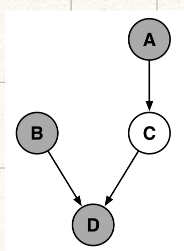
2.9 Polytree inference



We consider the belief network shown above. In this problem, we will be guided through an effective computation of the posterior probability $P(F \mid A, B, D, G)$. We should perform these computations efficiently – that is, by exploiting the structure of the DAG and not marginalizing over more variables than necessary. A, B, D, G are evidence.

a) **Bayes rule:** Show how to compute $P(C \mid A,B,D)$ in terms of $P(A)$, $P(B)$, $P(C \mid A)$ and $P(D \mid B,C)$.

We consider just the following part of the belief network:

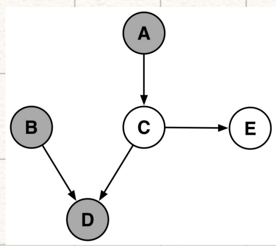


$$
\begin{aligned}
P(C \mid A,B,D) &\overset{\text{P.R.}}{=} \frac{P(A,B,C,D)}{P(A,B,D)} \\
&\overset{\text{P.R.}}{=} \frac{\cancel{P(A)}\,\cancel{P(B \mid A)}\,P(C \mid A,B)\,P(D \mid A,B,C)}{\cancel{P(A)}\,\cancel{P(B \mid A)}\,P(D \mid A,B)} \quad \Bigg\}\ \text{We are left with the conditional Bayes from HW1.} \\
\text{d-sep}_1\ A \to C \to D \quad &= \frac{P(D \mid B,C)\,P(C \mid A,B)}{P(D \mid A,B)} \\
\text{d-sep}_3\ \text{B\&C indep. given A} \quad &= \frac{P(D \mid B,C)\,P(C \mid A)}{P(D \mid A,B)} \\
&\overset{\text{marg.}}{=} \frac{P(D \mid B,C)\,P(C \mid A)}{\sum_c P(D, C=c \mid A,B)} \\
&\overset{\text{P.R.}}{=} \frac{P(D \mid B,C)\,P(C \mid A)}{\sum_c P(C=c \mid A,B)\,P(D \mid A,B,C=c)} \\
\text{d-sep}_3\ \text{B\&C indep. given A} \quad &= \frac{P(D \mid B,C)\,P(C \mid A)}{\sum_c P(C=c \mid A)\,P(D \mid A,B,C=c)} \\
\text{d-sep}_1\ A \to C \to D \quad &= \frac{P(D \mid B,C)\,P(C \mid A)}{\sum_c P(C=c \mid A)\,P(D \mid B,C=c)}
\end{aligned}
$$

We have expressed $P(C \mid A,B,D)$ in terms of $P(A)$, $P(B)$, $P(C \mid A)$ and $P(D \mid B,C)$. QED.

b) **Marginalization:** Show how to compute $P(E|A,B,D)$ in terms of the answer in a) and the CPTs of the belief network.

We consider just the following part of the belief network:



$$P(E|A,B,D) \overset{\text{marg.}}{=} \sum_c P(E, C=c|A,B,D)$$

$$\overset{\text{P.R.}}{=} \sum_c P(C=c|A,B,D)P(E|A,B,C=c,D)$$

d-sep 1 & d-sep 2
E indep. of A,B,D

$$= \underbrace{\sum_c P(C=c|A,B,D)}_{\text{Answer from a)}} \underbrace{P(E|C=c)}_{\text{table}}$$

c) **Marginalization:** Show how to compute $P(G|A,B,D)$ in terms of the answer in b) and the CPTs of the belief network.

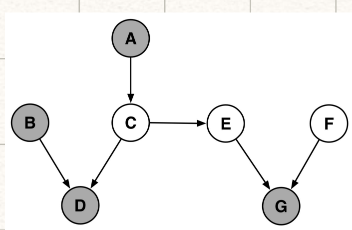
$$P(G|A,B,D) \overset{\text{marg.}}{=} \sum_e P(E=e, G|A,B,D)$$

$$\overset{\text{P.R.}}{=} \sum_e P(E=e|A,B,D)P(G|A,B,D,E=e)$$

d-sep 1 & 3
G cond. indep. of A,B,D given E

$$= \underbrace{\sum_e P(E=e|A,B,D)}_{\text{answer from b)}} \underbrace{P(G|E=e)}_{\text{table}}$$

d) **Explaining away:** Show how to compute $P(F|A,B,D,G)$ in terms of the answers in b,c) and the CPTs of the belief network.



$$P(F|A,B,D,G) = \frac{P(G|A,B,D,F)P(F|A,B,D)}{\underbrace{P(G|A,B,D)}_{\text{answer from c)}}}$$

We look at the numerator, starting with the right one.

$P(F|A,B,D) = P(F)$ because G blocks F from A,B,D by cond. 3 of d-sep as G∉Evidence blocks.

Now with the left term:

$$P(G|A,B,D,F) \overset{\text{marg.}}{=} \sum_e P(E=e, G|A,B,D,F) = \sum_e P(E=e|A,B,D,F)P(G|A,B,D,E=e,F)$$

G blocks E from F d-sep 3
E blocks G from A,B,D d-sep 1

$$= \underbrace{\sum_e P(E=e|A,B,D)}_{\text{answer b)}} \underbrace{P(G|E=e,F)}_{\text{table}}$$

$$\Rightarrow \quad P(F|A,B,D,G) = \frac{\overbrace{\sum_e P(E=e|A,B,D)}^{2.9b)} \overbrace{P(G|E=e,F)P(F)}^{\text{CPTs}}}{\underbrace{P(G|A,B,D)}_{2.9c)}}$$