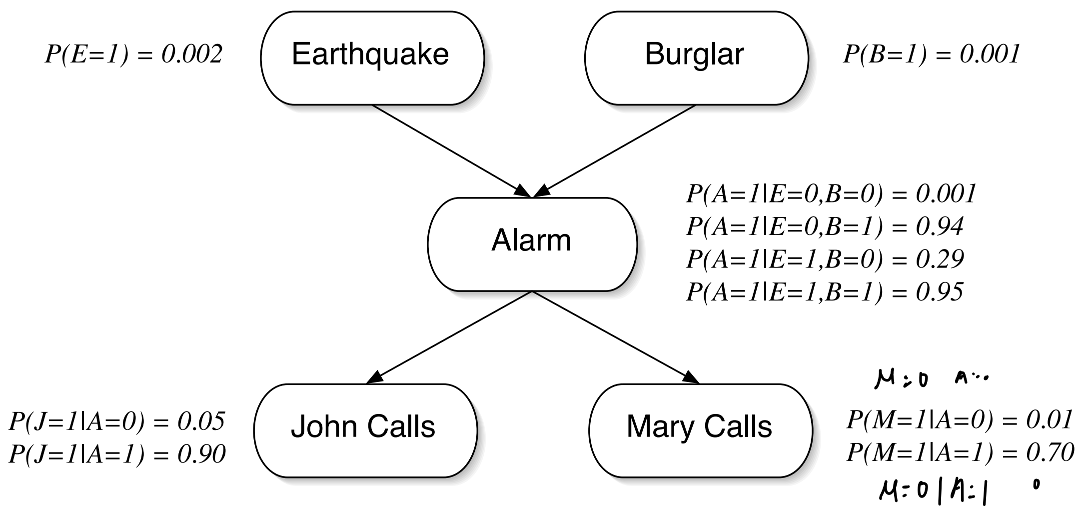
CSE 150A / 250A - Homework 2 *(51 pts total)*

Out: *Tue Oct 8*
Due: *Mon Oct 14* (by 11:59 PM, Pacific Time, via gradescope)
Grace period: 24 hours

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

Recall the alarm belief network described in class. The directed acyclic graph (DAG) and conditional probability tables (CPTs) are shown below:



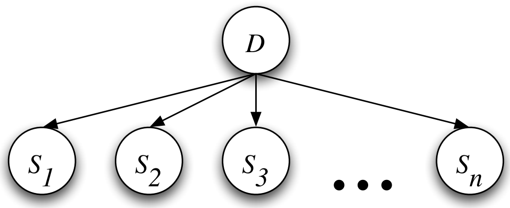
Compute numeric values for the following probabilities, exploiting relations of marginal and conditional independence as much as possible to simplify your calculations. You may re-use numerical results from lecture, but otherwise *show your work*. Be careful not to drop significant digits in your answer.

(a) $P(E\!=\!1\vert A\!=\!1)$	(c) $P(A\!=\!1\vert M\!=\!1)$	(e) $P(A\!=\!1\vert M\!=\!0)$
(b) $P(E\!=\!1\vert A\!=\!1, B\!=\!0)$	(d) $P(A\!=\!1\vert M\!=\!1, J\!=\!0)$	(f) $P(A\!=\!1\vert M\!=\!0, B\!=\!1)$

Consider your results in (b) versus (a), (d) versus (c), and (f) versus (e). Do they seem consistent with commonsense patterns of reasoning?

2.2 Probabilistic reasoning *(4 pts)*

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



The conditional probability tables (CPTs) for this belief network are as follows. In the absence of evidence, both forms of the disease are equally likely, with prior probabilities:

$$P(D\!=\!0) = P(D\!=\!1) = \frac{1}{2}.$$

In one form of the disease $(D\!=\!0)$, the first symptom occurs with probability one,

$$P(S_1\!=\!1|D\!=\!0) = 1,$$

while the k^{th} symptom (with $k\!\geq\!2$) occurs with probability

$$P(S_k\!=\!1|D\!=\!0) = \frac{f(k-1)}{f(k)},$$

where the function $f(k)$ is defined by

$$f(k) = 2^k + (-1)^k.$$

By contrast, in the other form of the disease $(D\!=\!1)$, all the symptoms are uniformly likely to be observed, with

$$P(S_k\!=\!1|D\!=\!1) = \frac{1}{2}$$

for all k. Suppose that on the k^{th} day of the month, a test is done to determine whether the patient is exhibiting the k^{th} symptom, and that each such test returns a positive result. Thus, on the k^{th} day, the doctor observes the patient with symptoms $\{S_1 = 1, S_2 = 1, \ldots, S_k = 1\}$. Based on the cumulative evidence, the doctor makes a new diagnosis each day by computing the ratio:

$$r_k = \frac{P(D\!=\!0|S_1\!=\!1, S_2\!=\!1, \ldots, S_k\!=\!1)}{P(D\!=\!1|S_1\!=\!1, S_2\!=\!1, \ldots, S_k\!=\!1)}.$$

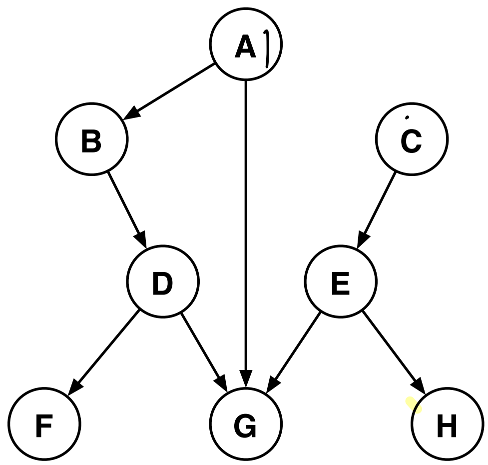
If this ratio is greater than 1, the doctor diagnoses the patient with the $D\!=\!0$ form of the disease; otherwise, with the $D\!=\!1$ form.

(a) Compute the ratio r_k as a function of k. How does the doctor's diagnosis depend on the day of the month? Show your work. *(3 pts)*

(b) Does the diagnosis become more or less certain as more symptoms are observed? Explain. (1 pts)

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)**. Your answer will be **only** graded based on correctness. No justifications required.



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

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)**. Your answer will be **only** graded based on correctness. No justifications required.

(a) ____________ $P(F|H) = P(F|C,H)$

(b) ____________ $P(E|A,B) = P(E|A,B,F)$

(c) ____________ $P(E,F|B,G) = P(E|B,G)\,P(F|B,G)$

(d) ____________ $P(F|B,C,G,H) = P(F|B,C,E,G,H)$

(e) ____________ $P(A,B|D,E,F) = P(A,B|D,E,F,G,H)$

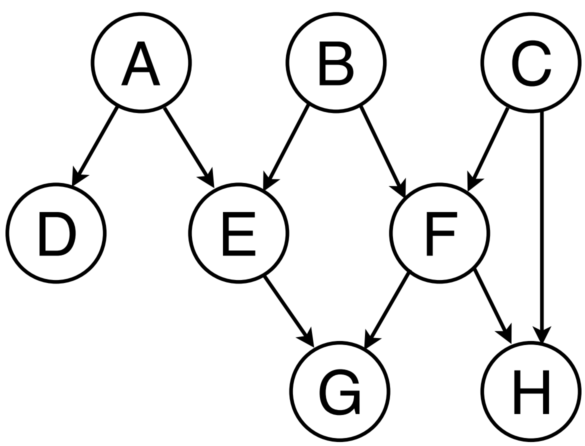
(f) ____________ $P(D,E,F) = P(D)\,P(E|D)\,P(F|E)$

(g) ____________ $P(A|F) = P(A)$

(h) ____________ $P(E,F) = P(E)\,P(F)$

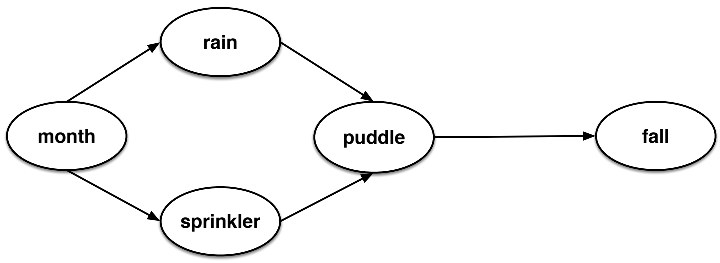
(i) ____________ $P(D|A) = P(D|A,E)$

(j) ____________ $P(B,C) = P(B)\,P(C)$



2.5 Conditional independence *(8 pts)*

Consider the DAG shown below, describing the following domain. Given the `month` of the year, there is some probability of `rain`, and also some probability that the `sprinkler` is turned on. Either of these events leads to some probability that a `puddle` forms on the sidewalk, which in turn leads to some probability that someone has a `fall`.

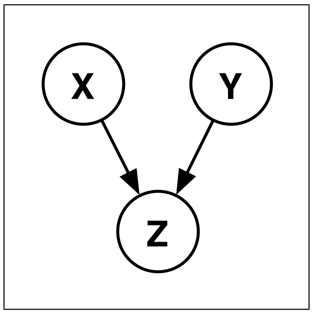


List all the conditional independence relations that must hold in any probability distribution represented by this DAG. More specifically, list all tuples $\{X, Y, E\}$ such that $P(X, Y|E) = P(X|E)P(Y|E)$, where

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

Hint: There are sixteen such tuples, not counting those that are equivalent up to exchange of X and Y. Do any of the tuples contain the case $E = \emptyset$?

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. 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.)

	$P(X=1)$	$\boxed{=}$	$P(X=1)$
(a)	$P(Z=1\|X=0,Y=0)$	$\square$	$P(Z=1\|X=0,Y=1)$
(b)	$P(Z=1\|X=1,Y=0)$	$\square$	$P(Z=1\|X=0,Y=1)$
(c)	$P(Z=1\|X=1,Y=0)$	$\square$	$P(Z=1\|X=1,Y=1)$
(d)	$P(X=1)$	$\square$	$P(X=1\|Z=1)$
(e)	$P(X=1)$	$\square$	$P(X=1\|Y=1)$
(f)	$P(X=1\|Z=1)$	$\square$	$P(X=1\|Y=1,Z=1)$
(g)	$P(X=1)\,P(Y=1)\,P(Z=1)$	$\square$	$P(X=1,Y=1,Z=1)$

Challenge (**optional**)*:* for each case, prove rigorously the correctness of your answer. You will receive partial credits for your efforts on proofs.