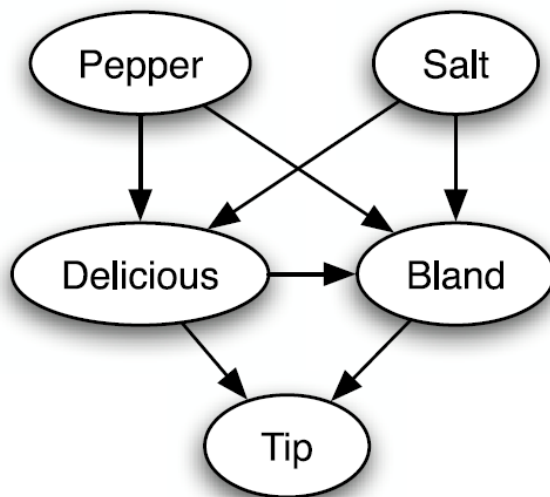


Question 1 (In class)

An ambitious Bongo Burger employee wants to maximize tips, which he thinks depend on the quality of the chef's cooking. When the chef uses more Pepper and more Salt, more people say "Delicious" and fewer people say "how Bland". Better comments come with better tips. He draws the following Bayes' net with 5 Boolean variables to model the scenario.



- a. Which of the following are true of this network? Note that a marginal distribution over a subset of variables in a Bayes' net is the result of summing out all other variables from the full joint distribution.
 (Circle whichever are true, 1 pts each)
 - i. It can represent any joint probability distribution over Salt, Pepper, Delicious, Bland, Tip.
 - ii. For any marginal distribution over Salt, Delicious, Bland, the network can represent at least one joint distribution with that marginal distribution.
 - iii. For any marginal distribution over Salt, Tip, the network can represent at least one joint distribution with that marginal distribution.

- b. How many total probability entries are there in the conditional probability tables for this network. Note: not the number of degrees of freedom, which may be fewer; a coin flip has two probabilities, but only one degree of freedom, and we're asking for the larger number.

- c. List all of the marginal independence assertions made by this graph about pairs of variables.

- d. List all of the conditional independence assertions made by this graph about pairs of variables.

- e. For one of the conditional independence assertions above, briefly explain how you might test its correctness from a set of observations.
- f. “No, no, no,” the chef replies, “People tip because the food is Tasty (influenced by salt and pepper), but comments are more complicated. Customers say Delicious or Bland in part based on whether they like the food, but mostly based on how Polite they are.” Draw a new Bayes' net appropriate to this scenario, which includes variables for Tasty and Polite, but which has as few arcs as possible. You do not need to specify the CPTs.
- g. Given $P(\text{Pepper} = \text{True}) = P(\text{Salt} = \text{True}) = 0.5$, the CPT below and the Bayes' net you just constructed for part (e), can you give $P(\text{Pepper} = \text{True}, \text{Salt} = \text{False} | \text{Tasty} = \text{False})$? If so, compute it. If not, explain why not.

Pepper	Salt	$P(\text{Tasty}=\text{True} \text{pepper}, \text{salt})$
True	True	0.8
True	False	0.6
False	True	0.6
False	False	0.1

- h. What is the (marginal) probability of the food being Tasty?
- i. How might the chef alter her seasoning habits to increase $P(\text{Tasty})$ without using any more salt or pepper overall? What would $P(\text{Tasty})$ be after this change? Would this change alter the Bayes' net you drew for (e)? If so, how? If not, why not?

Question 1 (Homework)

Consider the following network, in which a mouse agent is reasoning about the behavior of a cat. The mouse really wants to know whether the cat will attack (A), which depends on whether the cat is hungry (H) and whether the cat is sleepy (S). The mouse can observe two things, whether the cat is sleepy (S) and whether the cat has a collar (C). The cat is more often sleepy (S) when it's either full (f) or starved (v) than when it is peckish (p) and the collar (C) tends to indicate that the cat is not starved. Note that entries are omitted, such as $P(C = \sim c)$, when their complements are given.

$P(C)$	$P(H C)$	$P(S H)$	$P(A H,S)$																																																																	
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr><th>C</th><th>P</th></tr> </thead> <tbody> <tr><td>c</td><td>0.30</td></tr> </tbody> </table>	C	P	c	0.30	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr><th>H</th><th>C</th><th>P</th></tr> </thead> <tbody> <tr><td>f</td><td>c</td><td>0.60</td></tr> <tr><td>v</td><td>c</td><td>0.10</td></tr> <tr><td>p</td><td>c</td><td>0.30</td></tr> <tr><td>f</td><td>$\sim c$</td><td>0.20</td></tr> <tr><td>v</td><td>$\sim c$</td><td>0.40</td></tr> <tr><td>p</td><td>$\sim c$</td><td>0.30</td></tr> </tbody> </table>	H	C	P	f	c	0.60	v	c	0.10	p	c	0.30	f	$\sim c$	0.20	v	$\sim c$	0.40	p	$\sim c$	0.30	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr><th>S</th><th>H</th><th>P</th></tr> </thead> <tbody> <tr><td>s</td><td>f</td><td>0.90</td></tr> <tr><td>s</td><td>v</td><td>0.70</td></tr> <tr><td>s</td><td>p</td><td>0.30</td></tr> </tbody> </table>	S	H	P	s	f	0.90	s	v	0.70	s	p	0.30	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr><th>A</th><th>H</th><th>S</th><th>P</th></tr> </thead> <tbody> <tr><td>a</td><td>f</td><td>s</td><td>0.01</td></tr> <tr><td>a</td><td>f</td><td>$\sim s$</td><td>0.10</td></tr> <tr><td>a</td><td>v</td><td>s</td><td>0.50</td></tr> <tr><td>a</td><td>v</td><td>$\sim s$</td><td>0.90</td></tr> <tr><td>a</td><td>p</td><td>s</td><td>0.20</td></tr> <tr><td>a</td><td>p</td><td>$\sim s$</td><td>0.70</td></tr> </tbody> </table>	A	H	S	P	a	f	s	0.01	a	f	$\sim s$	0.10	a	v	s	0.50	a	v	$\sim s$	0.90	a	p	s	0.20	a	p	$\sim s$	0.70
C	P																																																																			
c	0.30																																																																			
H	C	P																																																																		
f	c	0.60																																																																		
v	c	0.10																																																																		
p	c	0.30																																																																		
f	$\sim c$	0.20																																																																		
v	$\sim c$	0.40																																																																		
p	$\sim c$	0.30																																																																		
S	H	P																																																																		
s	f	0.90																																																																		
s	v	0.70																																																																		
s	p	0.30																																																																		
A	H	S	P																																																																	
a	f	s	0.01																																																																	
a	f	$\sim s$	0.10																																																																	
a	v	s	0.50																																																																	
a	v	$\sim s$	0.90																																																																	
a	p	s	0.20																																																																	
a	p	$\sim s$	0.70																																																																	

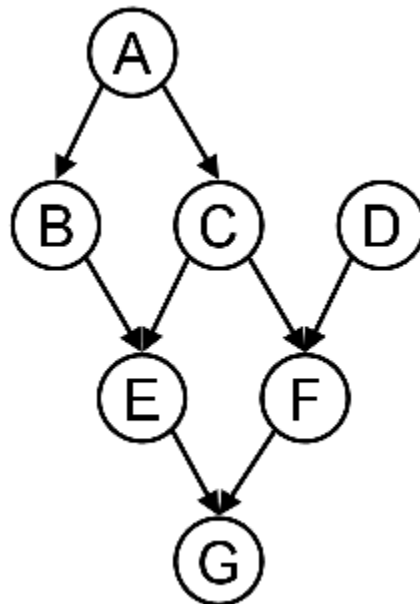
Assume you have the following samples relative to the evidence $C=c, S=s$:

- C=c, H=f, S=s, A= $\sim a$
- C=c, H=f, S=s, A= $\sim a$
- C=c, H=f, S=s, A= $\sim a$
- C=c, H=p, S=s, A=a
- C=c, H=v, S=s, A= $\sim a$

1. Draw the graphical model associated with this problem.
2. Calculate $P(A \mid C=c, S=s)$ using joint inference by enumeration
3. Assume the samples were generated by rejection sampling. What would the estimate of $P(A \mid C = c, S = s)$ be?
4. Assume the samples were generated by likelihood weighting. What would the estimate of $P(A \mid C = c, S = s)$ be?
5. Which sampling scheme is more likely to have generated these samples, rejection sampling or likelihood weighting?
6. It might seem strange that collars "cause" hunger: their relationship is one of correlation, not causation. Propose a new node which allows a more sensible causal interpretation and state how the network connections should change to accommodate it.

Question 2 (Homework)

In the network below, consider which nodes are possibly dependent on the listed node with the listed evidence:



- a. B with no observations
- b. B given G
- c. G given F and C
- d. C given E, F, and B

Question 3 (Homework)

Sometimes we can prune down a network before inference to simplify our computations. In this problem, you will develop a condition under which "dangling" variables can be pruned.

- a. Consider two networks, G and G' . In G , we have nodes X, Y , and Z where X is the parent of Y and Y is the parent of Z , so that $P(x, y, z) = P(x) P(y | x) P(z | y)$. G' is identical to G except, Z and its CPTs have been deleted. Show that $P(x | y)$ is the same whether we compute it from G or G' .

- b. More generally, assume we have a network G in which there are nodes $X_1 \dots X_n$. Assume that we wish to calculate $P(Q_1 = q_1, \dots, Q_k = q_k | E_1 = e_1, \dots, E_m = e_m)$, where each query variable Q and evidence variable E is one of the variables X_i . Let H_1, \dots, H_p be all the remaining (hidden) nodes which are neither observed nor queried. Write an expression for $P(q_1, \dots, q_k | e_1, \dots, e_m)$ in terms of the full joint distribution entries $P(q_1, \dots, q_k, e_1, \dots, e_m, h_1, \dots, h_p)$.

- c. Assume that one of the hidden variables H , say H_1 , is a leaf node, that is, it has no children in G . Let G' be identical to G but with H removed. Show that $P(q_1, \dots, q_k | e_1, \dots, e_m)$ is the same whether calculated in G or G' .

- d. Prove that any node which does not dominate either an evidence or query node ("dangling nodes") may be pruned without effecting the result of a query.