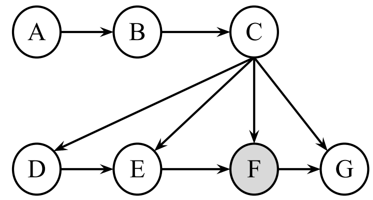
CS188: Exam Practice Session 6 Solutions

Q1. Variable Elimination



For the Bayes' net shown on the right, we are given the query $P(B, D \mid +f)$. All variables have binary domains. Assume we run variable elimination to compute the answer to this query, with the following variable elimination ordering: A, C, E, G.

(a) Complete the following description of the factors generated in this process:

After inserting evidence, we have the following factors to start out with:

$$P(A), P(B|A), P(C|B), P(D|C), P(E|C,D), P(+f|C,E), P(G|C,+f)$$

When eliminating A we generate a new factor f_1 as follows:

$$f_1(B) = \sum_a P(a)P(B|a)$$

This leaves us with the factors:

$$P(C|B), P(D|C), P(E|C,D), P(+f|C,E), P(G|C,+f), f_1(B)$$

(i) When eliminating C we generate a new factor f_2 as follows:

$$f_2(B, D, E, +f, G) = \sum_c P(c|B)P(D|c)P(E|c,D)P(+f|c,E)P(G|c,+f)$$

This leaves us with the factors:

$$f_1(B), f_2(B, D, E, +f, G)$$

(ii) When eliminating E we generate a new factor f_3 as follows:

$$f_3(B, D, +f, G) = \sum_e f_2(B, D, E, +f, G)$$

This leaves us with the factors:

$$f_1(B), f_3(B, D, +f, G)$$

(iii) When eliminating G we generate a new factor f_4 as follows:

$$f_4(B, D, +f) = \sum_g f_3(B, D, +f, g)$$

This leaves us with the factors:

$$f_1(B), f_4(B, D, +f)$$

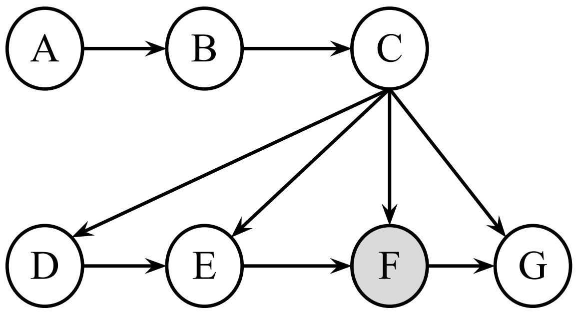
(b) Explain in one sentence how $P(B, D| + f)$ can be computed from the factors left in part (iii) of (a)?

Join $f_1 f_4$ to obtain $P(B, D, +f)$ and normalize it to get $P(B, D \mid f)$ Concretely, $P(b, d \mid +f) = \frac{f_1(b) f_4(b,d,+f)}{\sum_{b',d'} f_1(b) f_4(b,d,+f)}$.

(c) Among $f_1, f_2, \ldots, f_4$, which is the largest factor generated, and how large is it? Assume all variables have binary domains and measure the size of each factor by the number of rows in the table that would represent the factor.

$f_2(B, D, E, +f, G)$ is the largest factor generated. It has 4 variables, hence $2^4 = 16$ entries.

For your convenience, the Bayes' net from the previous page is shown again below.



(d) Find a variable elimination ordering for the same query, i.e., for $P(B, D \mid +f)$, for which the maximum size factor generated along the way is smallest. Hint: the maximum size factor generated in your solution should have only 2 variables, for a table size of $2^2 = 4$. Fill in the variable elimination ordering and the factors generated into the table below.

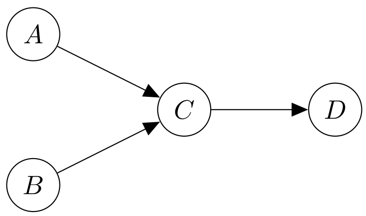
Variable Eliminated	Factor Generated
A	$f_1(B)$
G	$f_2(C, +f)$
E	$f_3(C, D)$
C	$f_4(B, D, +f)$

For example, in the naive ordering we used earlier, the first line in this table would have had the following two entries: A, $f_1(B)$. For this question there is no need to include how each factor is computed, i.e., no need to include expressions of the type $= \sum_a P(a)P(B|a)$.

Note: multiple orderings are possible. In particular in this case all orderings with E and G before C are correct.

Q2. Bayes' Net Sampling

Assume you are given the following Bayes' net and the corresponding distributions over the variables in the Bayes' net.



$P(A)$	
+a	0.1
-a	0.9

$P(B)$	
+b	.7
-b	.3

$P(C\|A,B)$			
+c	+a	+b	.25
-c	+a	+b	.75
+c	-a	+b	.6
-c	-a	+b	.4
+c	+a	-b	.5
-c	+a	-b	.5
+c	-a	-b	.2
-c	-a	-b	.8

$P(D\|C)$		
+d	+c	.5
-d	+c	.5
+d	-c	.8
-d	-c	.2

(a) Assume we receive evidence that $A = +a$. If we were to draw samples using rejection sampling, on expectation what percentage of the samples will be **rejected**?

Since $P(+a) = \frac{1}{10}$, we would expect that only 10% of the samples could be saved. Therefore, expected 90% of the samples will be rejected.

(b) Next, assume we observed both $A = +a$ and $D = +d$. What are the weights for the following samples under likelihood weighting sampling?

Sample	Weight
$(+a, -b, +c, +d)$	$P(+a) \cdot P(+d\| + c) = 0.1 * 0.5 = 0.05$
$(+a, -b, -c, +d)$	$P(+a) \cdot P(+d\| - c) = 0.1 * 0.8 = 0.08$
$(+a, +b, -c, +d)$	$P(+a) \cdot P(+d\| - c) = 0.1 * 0.8 = 0.08$

(c) Given the samples in the previous question, estimate $P(-b| + a, +d)$.

$$P(-b| + a, +d) = \frac{P(+a) \cdot P(+d| + c) + P(+a) \cdot P(+d| - c)}{P(+a) \cdot P(+d| + c) + 2 \cdot P(+a) \cdot P(+d| - c)} = \frac{0.05 + 0.08}{0.05 + 2 \cdot 0.08} = \frac{13}{21}$$

(d) Assume we need to (approximately) answer two different inference queries for this graph: $P(C| + a)$ and $P(C| + d)$. You are required to answer one query using likelihood weighting and one query using Gibbs sampling. In each case you can only collect a relatively small amount of samples, so for maximal accuracy you need to make sure you cleverly assign algorithm to query based on how well the algorithm fits the query. Which query would you answer with each algorithm?

Algorithm	Query
Likelihood Weighting	$P(C\| + a)$

Algorithm	Query
Gibbs Sampling	$P(C\| + d)$

Justify your answer:
You should use Gibbs sampling to find the query answer $P(C| + d)$. This is because likelihood weighting only takes upstream evidence into account when sampling. Therefore, Gibbs, which utilizes both upstream and downstream evidence, is more suited to the query $P(C| + d)$ which has downstream evidence.