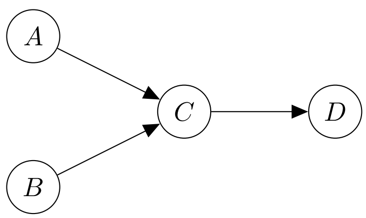
CS188: Exam Practice Session 8 Solutions

Q1. 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?

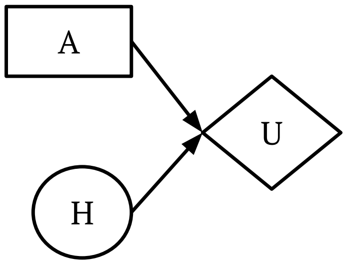
Algorithm	Query
Likelihood Weighting	$P(C\| + a)$

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.

Q2. Decision Networks

After years of battles between the ghosts and Pacman, the ghosts challenge Pacman to a winner-take-all showdown, and the game is a coin flip. Pacman has a decision to make: whether to accept the challenge (*accept*) or decline (*decline*). If the coin comes out heads ($+h$) Pacman wins. If the coin comes out tails ($-h$), the ghosts win. No matter what decision Pacman makes, the outcome of the coin is revealed.



H	$P(H)$
+h	0.5
-h	0.5

H	A	U(H,A)
+h	*accept*	100
-h	*accept*	-100
+h	*decline*	-30
-h	*decline*	50

(a) Maximum Expected Utility

Compute the following quantities:

$EU(accept) = P(+h)U(+h, accept) + P(-h)U(-h, accept) = 0.5 * 100 + 0.5 * -100 = 0$

$EU(decline) = P(+h)U(+h, decline) + P(-h)U(-h, decline) = 0.5 * -30 + 0.5 * 50 = 10$

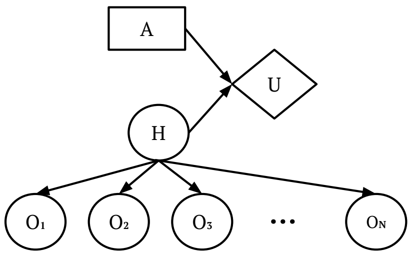
$MEU(\{\}) = max(0, 10) = 10$

Action that achieves $MEU(\{\}) =$ *decline*

(b) VPI relationships When deciding whether to accept the winner-take-all coin flip, Pacman can consult a few fortune tellers that he knows. There are N fortune tellers, and each one provides a prediction O_n for H.

For each of the questions below, circle **all** of the VPI relations that are guaranteed to be true, or select *None of the above.*

(i) In this situation, the fortune tellers give perfect predictions.
Specifically, $P(O_n = +h \mid H = +h) = 1$, $P(O_n = -h \mid H = -h) = 1$, for all n from 1 to N.



○ $\text{VPI}(O_1,\ O_2) \geq \text{VPI}(O_1) + \text{VPI}(O_2)$

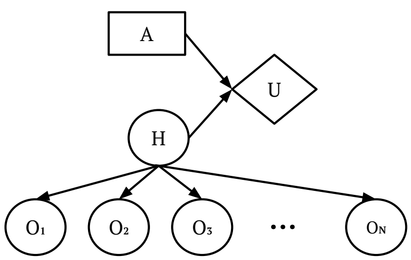
● $\text{VPI}(O_i) = \text{VPI}(O_j)$ where $i \neq j$

○ $\text{VPI}(O_3 \mid O_2,\ O_1) > \text{VPI}(O_2 \mid O_1)$.

○ $\text{VPI}(H) > \text{VPI}(O_1, O_2, \ldots O_N)$

○ None of the above.

(ii) In another situation, the fortune tellers are pretty good, but not perfect.
Specifically, $P(O_n = +h \mid H = +h) = 0.8$, $P(O_n = -h \mid H = -h) = 0.5$, for all n from 1 to N.



○ $\text{VPI}(O_1,\ O_2) \geq \text{VPI}(O_1) + \text{VPI}(O_2)$

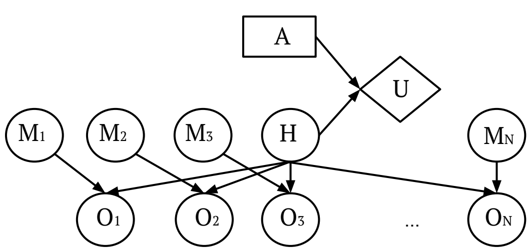
● $\text{VPI}(O_i) = \text{VPI}(O_j)$ where $i \neq j$

○ $\text{VPI}(O_3 \mid O_2,\ O_1) > \text{VPI}(O_2 \mid O_1)$.

● $\text{VPI}(H) > \text{VPI}(O_1, O_2, \ldots O_N)$

○ None of the above.

(iii) In a third situation, each fortune teller's prediction is affected by their mood. If the fortune teller is in a good mood ($+m$), then that fortune teller's prediction is guaranteed to be correct. If the fortune teller is in a bad mood ($-m$), then that teller's prediction is guaranteed to be incorrect. Each fortune teller is happy with probability $P(M_n = +m) = 0.8$.



○ $\text{VPI}(M_1) > 0$

● $\forall i\ \ \text{VPI}(M_i | O_i) > 0$

○ $\text{VPI}(M_1,\ M_2,\ \ldots,\ M_N) > \text{VPI}(M_1)$

● $\forall i\ \ \text{VPI}(H) = \text{VPI}(M_i, O_i)$

○ None of the above.