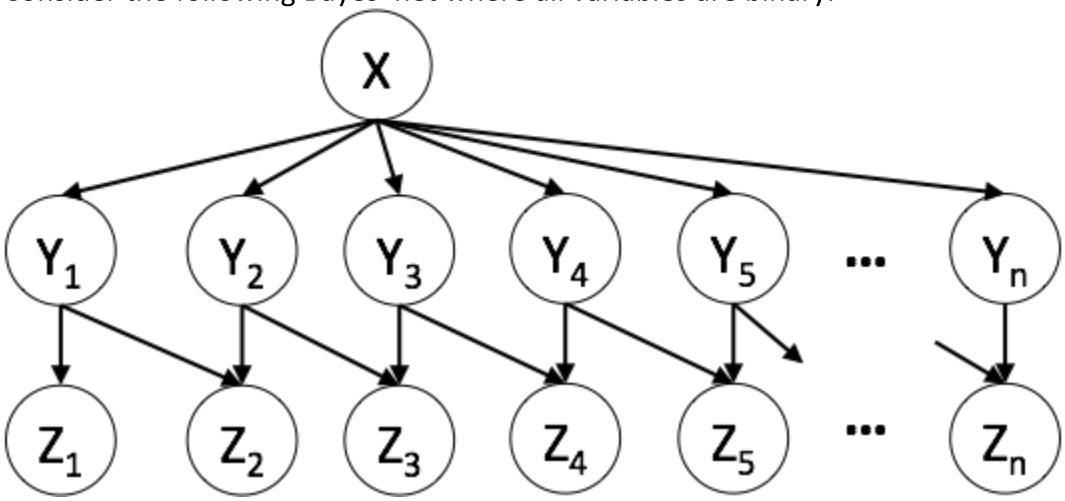
UC Berkeley – Computer Science
CS188: Introduction to Artificial Intelligence

Practice Midterm2, Fall 2017

1. Bayes Net Inference

Consider the following Bayes' net where all variables are binary.



I. Assume that we would like to perform inference to obtain $P(Y_n|Z_1 = z_1, Z_2 = z_2, \cdots, Z_n = z_n)$.

(a) What is the number of rows in the largest factor generated by *inference by enumeration*, for this query $P(Y_n|Z_1 = z_1, Z_2 = z_2, \cdots, Z_n = z_n)$?

____

(b) Suppose we decide to perform variable elimination to calculate the query $P(Y_n|Z_1 = z_1, Z_2 = z_2, \cdots, Z_n = z_n)$, and we eliminate the variables in the ordering $Y_1, Y_2, \cdots, Y_{n-1}, X$. Write out the form and the size of factor generated by the elimination of Y_i and X, where $i \in \{2, \cdots, n-1\}$.

	The new factor generated	The size of the new factor
After Y_i is eliminated	__________________	__________________
After X is eliminated	__________________	__________________

(c) Find the best and the worst variable elimination orderings for this query. The goodness is measured by the sum of the sizes of the factors that are generated.

Best ordering: __________________

Worst ordering: __________________

Ii. Assume now we want to use variable elimination to calculate *another* query $P(Z_n|Y_1, Y_2, \cdots, Y_{n-1})$.

(a) Mark all of the following variables that produce a *constant* factor after being eliminated for all possible elimination orderings.

☐ Z_1 ☐ Z_2 ☐ Z_{n-1} ☐ X

(b) List all the variables that can be ignored (i.e. the conditional probability tables of the variables can be ruled out from the initial factor set), in performing the query $P(Z_n|Y_1, Y_2, \cdots, Y_{n-1})$. Briefly present the reason why these variables can be ignored. (Hint: you can use the results in previous part or the conditional independencies encoded in the given BN).

Variables can be ignored when computing the query $P(Z_n|Y_1, Y_2, \cdots, Y_{n-1})$:

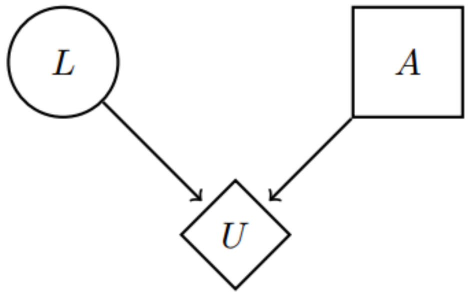
____

Reason:

____

2. Decision Network

In this problem, we will model a lottery as a decision network, where we are deciding whether or not it is worth buying a lottery ticket and playing the lottery.
We can consider the outcome of a lottery to be a node L, with A as the decision to play the lottery or not and U as our utility. The resulting decision network may look something like this:



For all parts of this question, assume that we have an utility function of $U(x) = x^3$, where x is the amount of money we pay and then (hopefully) win in the lottery, and that the price of a lottery ticket is $4.

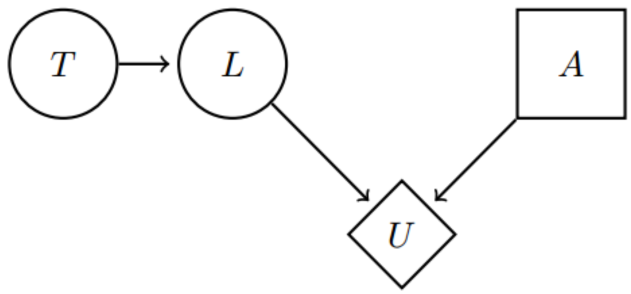
Q1: For this first part, assume that the lottery is $[0,9/10; 14,1/10]$. Fill in the following table for $U|L, A$.

L	A	$U\|L, A$
0	buy	
14	buy	
0	no buy	
14	no buy	

What is $MEU(\{\})$?

What action achieves $MEU(\{\})$?

Now, the organization running the lottery has announced a change in policy: there are two lotteries, and when someone buys a lottery ticket, the ticket is randomly for one of the two lotteries.
We can incorporate this into our model by adding a node T that indicates which ticket gets bought: the resulting decision network looks like this:



In addition, we have the following conditional probability tables, that we will be using for the remaining parts of this problem:

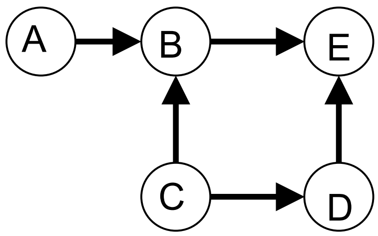
T	$P(T)$
$lottery1$	1/2
$lottery2$	1/2

T	L	$P(L\|T)$
$lottery1$	0	9/10
$lottery1$	14	1/10
$lottery2$	0	7/10
$lottery2$	5	3/10

Q2: What is the new MEU?
What action achieves $MEU(\{\})$?

Q3: What is VPI(T)?

3. Sampling



A	P(A)
0	0.6
1	0.4

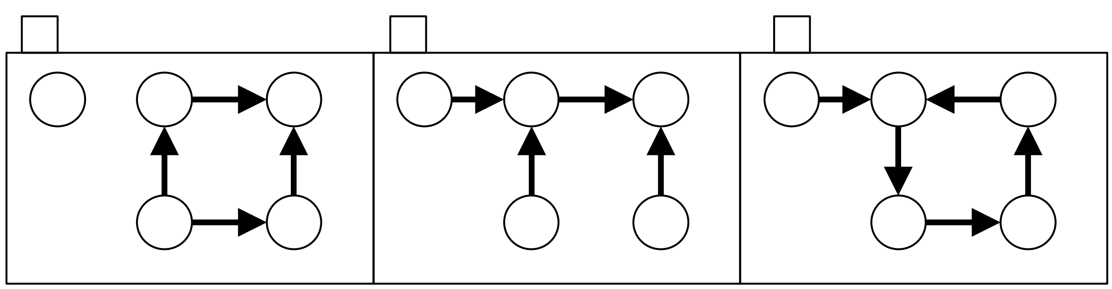
C	D	P(D|C)
0	0	0.2
0	1	0.8
1	0	0.3
1	1	0.7

A	C	B	P(B|A,C)
0	0	0	0.2
0	0	1	0.8
0	1	0	0.4
0	1	1	0.6
1	0	0	0.2
1	0	1	0.8
1	1	0	0.4
1	1	1	0.6

B	D	E	P(E|B,D)
0	0	0	0.75
0	0	1	0.25
0	1	0	0.5
0	1	1	0.5
1	0	0	0.85
1	0	1	0.15
1	1	0	0.4
1	1	1	0.6

C	P(C)
0	0.4
1	0.6

i. (3 pts) Check the boxes above the Bayes' nets below that could also be valid for the above probability tables.



ii. (2 pts) Caryn wants to compute the distribution P(A,C|E=1) using ***prior sampling*** on Model-Q1 (given at the top of this page). She draws a bunch of samples. The first of these is (0, 0, 1, 1, 0), given in (A, B, C, D, E) format. What's the probability of drawing this sample?

iii. (2 pts) Give an example of an inference query for Model-Q1 with **one query variable and one evidence variable** that could be estimated more efficiently (in terms of runtime) using ***rejection sampling*** than by using ***prior sampling***. If none exist, state "not possible".

iv. (2 pts) Give an example of an inference query for Model-Q1 with **one query variable and one evidence variable** for which ***rejection sampling*** provides no efficiency advantage (in terms of runtime) over using ***prior sampling***. If none exist, state "not possible".

v. (2 pts) Now Caryn wants to determine P(A,C|E=1) for Model-Q1 using likelihood weighting. She draws the five samples shown below, which are given in (A, B, C, D, E) format, where the leftmost sample is "Sample 1" and the rightmost is "Sample 5". What are the weights of the samples S1 and S3?

weight(S1): ________________________ weight(S3): ________________________

S1: (0, 0, 1, 1, 1) S2: (0, 0, 1, 1, 1) S3: (1, 0, 1, 1, 1) S4: (0, 1, 0, 0, 1) S5: (0, 1, 0, 0, 1)

vi. (1 pt) For the same samples as in part v, compute P(A=1,C=1|E=1) for Model-Q1. Express your answer as a simplified fraction (e.g. 2 / 3 instead of 4 / 6).

vii. (2 pts) Select True or False for each of the following:

True	False	
○	○	When there is no evidence, prior sampling is guaranteed to yield the exact same answer as inference by enumeration.
		When collecting a sample during likelihood weighting, evidence variables are not sampled.
		When collecting a sample during rejection sampling, variables can be sampled in any order.
		Gibbs sampling is a technique for performing approximate inference, not exact inference.