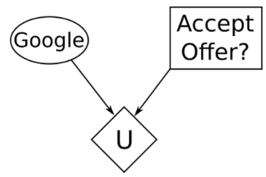
Name: ______________________________ Student ID: ________________

1. (21 points) DECISIONS

   You've been job hunting, and you've narrowed your options to two companies: Acme and Google. You already have an offer from Acme, but it expires today, and you are still waiting for a response from Google. You are faced with the dilemma of whether or not to accept the offer from Acme, which is modeled by the following decision diagram:



   Your prior belief about whether Google will hire you and utility over possible outcomes are as follows:

Google outcome	P(Google outcome)
hired	0.25
not hired	0.75

Action	Google outcome	U
accept Acme offer	hired	2000
accept Acme offer	not hired	8000
reject Acme offer	hired	10000
reject Acme offer	not hired	0

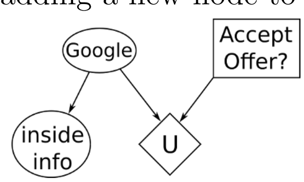
   (a) What is the expected utility of each action? (Note: throughout this problem answers will be evaluated to whole-number precision, so your answer should differ by no more than 1 from the exact answer.)

   Action: accept Acme offer = ____________

   Action: reject Acme offer = ____________

   Which action should you take? ☐ reject ☐ accept

   (b) Suddenly, the phone rings. It's your uncle, who works at Google. Your uncle tells you he has some inside information about the status of your application. Your uncle won't tell you what the information is yet, but he might be willing to divulge it for the right price. You model the new situation by adding a new node to your decision diagram:



   You create a CPT to model the relationship between the inside information and Google's future hiring decision:

info	Google outcome	P(info | Google outcome)
good news	hired	0.7
bad news	hired	0.3
good news	not hired	0.1
bad news	not hired	0.9

We'll help grind through the probabilistic inference. The resulting distributions are:

info	P(info)
good news	0.25
bad news	0.75

Google outcome	info	P(Google outcome | info)
hired	good news	0.7
not hired	good news	0.3
hired	bad news	0.1
not hired	bad news	0.9

Fill in the expected utilities for each action, for each possible type of information we could be given:

EU(accept Acme offer | good news) = ____________

EU(reject Acme offer | good news) = ____________

EU(accept Acme offer | bad news) = ____________

EU(reject Acme offer | bad news) = ____________

What is the maximum expected utility for each type of information we could be given?

MEU(good news) = ____________

MEU(bad news) = ____________

If we are given the inside information, what is the expected value of MEU?

Answer. ____________

What is the value of perfect information of the random variable Inside Info?

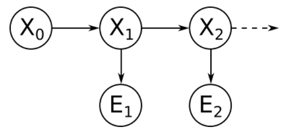
Answer. ____________

2. (16 points) VALUE OF PERFECT INFORMATION

Consider the value of perfect information (VPI) of observing some node in an arbitrary decision network. Which of the following are true statements?

- ☐ VPI is guaranteed to be positive (> 0).
- ☐ VPI is guaranteed to be nonnegative (≥ 0).
- ☐ VPI is guaranteed to be nonzero.
- ☐ The MEU after observing a node could potentially be less than the MEU before observing that node.
- ☐ For any two nodes X and Y, $\text{VPI}(X) + \text{VPI}(Y) \geq \text{VPI}(X, Y)$. That is, the sum of individual VPI's for two nodes is always greater than or equal to the VPI of observing both nodes.
- ☐ VPI is guaranteed to be exactly zero for any node that is conditionally independent (given the evidence so far) of all parents of the utility node.

3. (21 points) HMMs, Part I. Consider the HMM shown below.



The prior probability $P(X_0)$, dynamics model $P(X_{t+1} \mid X_t)$, and sensor model $P(E_t \mid X_t)$ are as follows:

X_0	$P(X_0)$
0	0.15
1	0.85

X_{t+1}	X_t	$P(X_{t+1}\|X_t)$
0	0	0.6
1	0	0.4
0	1	0.9
1	1	0.1

E_t	X_t	$P(E_t\|X_t)$
a	0	0.8
b	0	0.15
c	0	0.05
a	1	0.35
b	1	0.05
c	1	0.6

We perform a first dynamics update, and fill in the resulting belief distribution $B'(X_1)$.

X_1	$B'(X_1)$
0	0.855
1	0.145

We incorporate the evidence $E_1 = c$. We fill in the evidence-weighted distribution $P(E_1 = c \mid X_1)B'(X_1)$, and the (normalized) belief distribution $B(X_1)$.

X_1	$P(E_1 = c\|X_1)B'(X_1)$
0	0.04275
1	0.087

X_1	$B(X_1)$
0	0.329479768786
1	0.670520231214

You get to perform the second dynamics update. Fill in the resulting belief distribution $B'(X_2)$.

$B'(X_2 = 0) =$ ____________

$B'(X_2 = 1) =$ ____________

Now incorporate the evidence $E_2 = c$. Fill in the evidence-weighted distribution $P(E_2 = c \mid X_2)B'(X_2)$, and the (normalized) belief distribution $B(X_2)$.

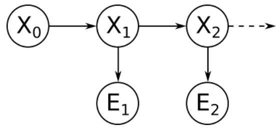
$P(E_2 = c \mid X_2)B'(X_2)$ when $X_2 = 0$ is ____________

$P(E_2 = c \mid X_2)B'(X_2)$ when $X_2 = 1$ is ____________

$B(X_2 = 0) =$ ____________

$B(X_2 = 1) =$ ____________

4. (21 points) HMMs, Part II. Consider the same HMM, but with different probabilities.



The prior probability $P(X_0)$, dynamics model $P(X_{t+1} \mid X_t)$, and sensor model $P(E_t \mid X_t)$ are as follows:

X_0	$P(X_0)$
0	0.2
1	0.8

X_{t+1}	X_t	$P(X_{t+1} \mid X_t)$
0	0	0.3
1	0	0.7
0	1	0.05
1	1	0.95

E_t	X_t	$P(E_t \mid X_t)$
a	0	0.3
b	0	0.15
c	0	0.55
a	1	0.1
b	1	0.45
c	1	0.45

In this question we'll assume the sensor is broken and we get no more evidence readings after E_2. We are forced to rely on dynamics updates only going forward. In the limit as $t \to \infty$, our belief about X_t should converge to a stationary distribution $\tilde{B}(X_\infty)$ defined as follows:

$$\tilde{B}(X_\infty) := \lim_{t\to\infty} P(X_t \mid E_1, E_2)$$

(a) Recall that the stationary distribution satisfies the equation
$\tilde{B}(X_\infty) = \sum_{X_\infty} P(X_{t+1} \mid X_t)\tilde{B}(X_\infty)$
for all values in the domain of X.

In the case of this problem, we can write these relations as a set of linear equations of the form

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \tilde{B}(X_\infty = 0) \\ \tilde{B}(X_\infty = 1) \end{bmatrix} = \begin{bmatrix} \tilde{B}(X_\infty = 0) \\ \tilde{B}(X_\infty = 1) \end{bmatrix}$$

In the spaces below, fill in the coefficients of the linear system.

$a =$ ______________

$b =$ ______________

$c =$ ______________

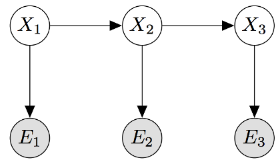
$d =$ ______________

(b) The system you have written has many solutions (consider (0,0), for example), but to get a probability distribution we want the solution that sums to one. Fill in your solution below.

$\tilde{B}(X_\infty = 0) =$ ______________

$\tilde{B}(X_\infty = 1) =$ ______________

5. (21 points) MODIFIED HMM UPDATE EQUATIONS. Consider the HMM graph structure shown below



Recall the Forward algorithm is a two step iterative algorithm used to approximate the probability distribution, $P(X_t \mid e_1, \ldots, e_t)$.

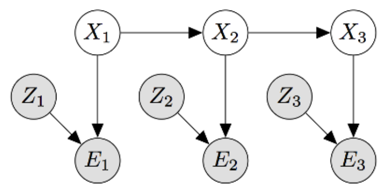
The two steps of the algorithm are as follows:

Elapse Time: $P(X_t \mid e_{1\ldots t-1}) = \Sigma_{x_{t-1}} P(X_t \mid x_{t-1}) P(x_{t-1} \mid e_{1\ldots t-1})$

Observe: $P(X_t \mid e_{1\ldots t}) = \frac{P(e_t|X_t)P(X_t|e_{1\ldots t-1})}{\sum_{x_t} P(e_t|x_t)P(x_t|e_{1\ldots t-1})}$

For this problem we will consider modifying the forward algorithm as the HMM graph structure changes. Our goal will continue to be to create an iterative algorithm which is able to compute the distribution of states, X_t, given all available evidence from time 0 to time t.

(a) Consider the graph below where new observed variables, Z_i, are introduced and influence the evidence.



What will the modified elapse time update be?

$P(X_t \mid e_{1\ldots t-1}, z_{1\ldots t-1}) =$

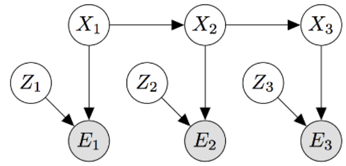
- ☐ $\sum_{x_{t-1}} P(X_t \mid z_{1\ldots t-1}) P(x_{t-1} \mid e_{1\ldots t-1}, z_{1\ldots t-1})$
- ☐ $\sum_{x_{t-1}} P(X_t \mid x_{t-1}) P(x_{t-1} \mid e_{1\ldots t-1}, z_{1\ldots t-1})$
- ☐ $\sum_{x_{t-1}} P(X_t \mid e_{1\ldots t-1}, z_{1\ldots t-1}) P(x_{t-1} \mid x_{t-1}, z_{1\ldots t-1})$
- ☐ $\sum_{x_{t-1}} P(X_t \mid x_{t-1}) P(x_{t-1} \mid e_{1\ldots t-1})$ (no change)

What will the modified observed update be?

$P(X_t \mid e_{1\ldots t}, z_{1\ldots t}) =$

- ☐ $\frac{P(e_t|X_t,z_t)P(X_t|e_{1\ldots t-1},z_{1\ldots t-1})}{\sum_{x_t,z_t} P(e_t|x_t,z_t)P(x_t|e_{1\ldots t-1},z_{1\ldots t-1})}$
- ☐ $\frac{P(e_t|X_t,z_t)P(X_t|e_{1\ldots t-1},z_{1\ldots t-1})}{\sum_{x_t} P(e_t|x_t,z_t)P(x_t|e_{1\ldots t-1},z_{1\ldots t-1})}$
- ☐ $\frac{P(e_t|X_t)P(X_t|e_{1\ldots t-1})}{\sum_{z_t} P(e_t|x_t,z_t)P(x_t|e_{1\ldots t-1},z_{1\ldots t-1})}$
- ☐ $\frac{P(e_t|X_t,z_t)P(X_t|e_{1\ldots t-1},z_{1\ldots t-1})}{\sum_{z_t} P(e_t|x_t,z_t)P(x_t|e_{1\ldots t-1},z_{1\ldots t-1})}$
- ☐ $\frac{P(e_t|X_t)P(X_t|e_{1\ldots t-1})}{\sum_{x_t} P(e_t|x_t)P(x_t|e_{1\ldots t-1})}$ (no change)

(b) Next, consider the graph below where the Z_i variables are unobserved.



What will the modified elapse time update be?

$P(X_t \mid e_{1\ldots t-1}) =$

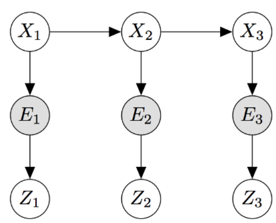
- ☐ $\sum_{x_{t-1}} P(X_t \mid z_{1\ldots t-1}) P(x_{t-1} \mid e_{1\ldots t-1}, z_{1\ldots t-1})$
- ☐ $\sum_{x_{t-1}} P(X_t \mid x_{t-1}) P(x_{t-1} \mid e_{1\ldots t-1}, z_{1\ldots t-1})$
- ☐ $\sum_{x_{t-1}} P(X_t \mid e_{1\ldots t-1}, z_{1\ldots t-1}) P(x_{t-1} \mid x_{t-1}, z_{1\ldots t-1})$
- ☐ $\sum_{x_{t-1}} P(X_t \mid x_{t-1}) P(x_{t-1} \mid e_{1\ldots t-1})$ (no change)

What will the modified observed update be?

$P(X_t \mid e_{1\ldots t}) =$

- ☐ $\frac{P(X_t|e_{1\ldots t-1})P(z_t)P(e_t|X_t,z_t)}{\sum_{z_t} P(x_t|e_{1\ldots t-1})P(e_t|x_t,z_t)P(z_t)}$
- ☐ $\frac{P(X_t|e_{1\ldots t-1})\sum_{z_t} P(z_t)P(e_t|X_t,z_t)}{P(x_t|e_{1\ldots t-1})\sum_{z_t} P(e_t|x_t,z_t)P(z_t)}$
- ☐ $\frac{P(X_t|e_{1\ldots t-1})P(z_t)P(e_t|X_t,z_t)}{\sum_{x_t} P(x_t|e_{1\ldots t-1})P(e_t|x_t,z_t)P(z_t)}$
- ☐ $\frac{P(X_t|e_{1\ldots t-1})\sum_{z_t} P(z_t)P(e_t|X_t,z_t)}{\sum_{x_t} P(x_t|e_{1\ldots t-1})\sum_{z_t} P(e_t|x_t,z_t)P(z_t)}$
- ☐ $\frac{P(e_t|X_t)P(X_t|e_{1\ldots t-1})}{\sum_{x_t} P(e_t|x_t)P(x_t|e_{1\ldots t-1})}$ (no change)

(c) Finally, consider a graph where the newly introduced variables are unobserved and influenced by the evidence nodes.



What will the modified elapse time update be?

$P(X_t \mid e_{1\ldots t-1}) =$

- ☐ $\sum_{x_{t-1}} P(X_t \mid z_{1\ldots t-1}) P(x_{t-1} \mid e_{1\ldots t-1}, z_{1\ldots t-1})$
- ☐ $\sum_{x_{t-1}} P(X_t \mid x_{t-1}) P(x_{t-1} \mid e_{1\ldots t-1}, z_{1\ldots t-1})$
- ☐ $\sum_{x_{t-1}} P(X_t \mid e_{1\ldots t-1}, z_{1\ldots t-1}) P(x_{t-1} \mid x_{t-1}, z_{1\ldots t-1})$
- ☐ $\sum_{x_{t-1}} P(X_t \mid x_{t-1}) P(x_{t-1} \mid e_{1\ldots t-1})$ (no change)

What will the modified observed update be?

$P(X_t \mid e_{1\ldots t}) =$

- ☐ $\frac{P(e_t|X_t,z_t)P(X_t|e_{1\ldots t-1},z_{1\ldots t-1})}{\sum_{z_t} P(e_t|x_t,z_t)P(x_t|e_{1\ldots t-1},z_{1\ldots t-1})}$
- ☐ $\frac{P(e_t|X_t,z_t)P(X_t|e_{1\ldots t-1},z_{1\ldots t-1})}{\sum_{x_t} P(e_t|x_t,z_t)P(x_t|e_{1\ldots t-1},z_{1\ldots t-1})}$
- ☐ $\frac{P(e_t|X_t)P(X_t|e_{1\ldots t-1})}{\sum_{x_t,z_t} P(e_t|x_t,z_t)P(x_t|e_{1\ldots t-1},z_{1\ldots t-1})}$
- ☐ $\frac{P(e_t|X_t)P(X_t|e_{1\ldots t-1})}{\sum_{x_t} P(e_t|x_t)P(x_t|e_{1\ldots t-1})}$ (no change)