CS 188 Summer 2020 — Introduction to Artificial Intelligence — Midterm 2

- You have approximately 110 minutes.
- The exam is open book, open calculator, and open notes.
- In the interest of fairness, we want everyone to have access to the same information. To that end, we will not be answering questions about the content or making clarifications.
- For multiple choice questions,
  - ☐ means mark **all options** that apply
  - ○ means mark a **single choice**

First name	
Last name	
SID	

For staff use only:

Q1.	Plants vs. Zombies, Cont'd	/26
Q2.	Potpourri	/15
Q3.	Value of Stock Information	/11
Q4.	Secret Tunnel	/21
Q5.	Probability	/27
	Total	/100

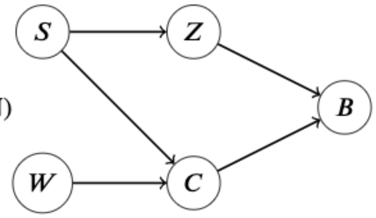
THIS PAGE IS INTENTIONALLY LEFT BLANK

SID: ______________

Q1. [26 pts] Plants vs. Zombies, Cont'd

(a) Zomboss is sending zombies to invade one of the grass lanes on your lawn! There could be a Snow Pea ($S = \pm s$) and/or a Wall-nut ($W = \pm w$) on the lane, and Zomboss may put a regular Zombie ($Z = \pm z$) and/or a Conehead zombie ($C = \pm c$) to try to get your brain ($B = \pm b$).
You come up with the following Bayes net to model the situation:

(Your BN)



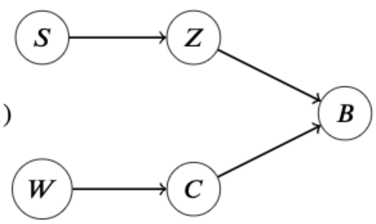
All randomized exams had a different orientation of the same Bayes Net, for both (Your BN) and (1), (2), (3), and (4). Answers for 1a.i, 1a.ii, 1a.iii, 1a.iv, and 1b are the same between versions.

(i) [2 pts] Is $W \perp\!\!\!\perp B|C$ guaranteed?
○ Guaranteed ● Not Guaranteed
Active path: $W - C(shaded) - S - Z - B$.

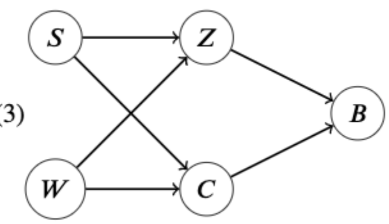
(ii) [2 pts] Is $Z \perp\!\!\!\perp C|S$ guaranteed?
● Guaranteed ○ Not Guaranteed
There is no active path between C and Z when S is shaded.

Pacman also comes up with a bunch of Bayes net representations, seen below:

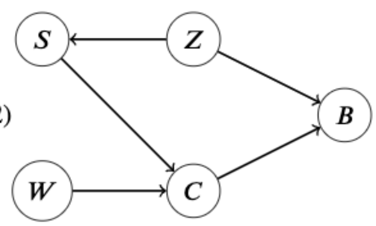
(1)



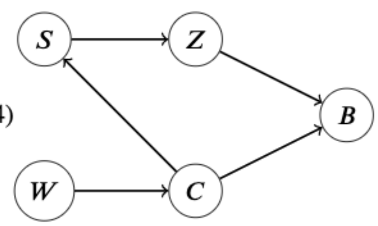
(3)



(2)



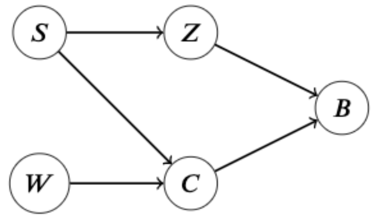
(4)



(iii) [3 pts] Which of Pacman's Bayes nets can represent at least one of the distributions that your Bayes net can represent?
■ (1) ■ (2) ■ (3) ■ (4) ○ None
All the above can represent the fully independent joint distribution.

(iv) [4 pts] Which of Pacman's Bayes nets can represent every distribution that your Bayes net can represent?
□ (1) ■ (2) ■ (3) □ (4) ○ None
If Pacman's Bayes Net has additional independence assumptions compared to yours, then it cannot represent every distribution that your Bayes net can represent.
(1) assumes $S \perp\!\!\!\perp C$, which is not assumed by your Bayes Net.
(2) assumes $S \perp\!\!\!\perp W$, $S \perp\!\!\!\perp B|C, Z$, $Z \perp\!\!\!\perp C|S$, $Z \perp\!\!\!\perp W|S, C$, $Z \perp\!\!\!\perp W|S$,
all of which are assumed by your Bayes Net. Another way to think about this is that, the only change is the edge between S and Z, but the triples that includes the edge, aka $Z - S - C$ and $S - Z - B$ have the same properties as before, either with S (or Z) shaded or not.
(3) has all the directed edges in your Bayes Net. Hence it can represent every distribution that your Bayes net can represent. It's stronger.
(4) assumes $S \perp\!\!\!\perp W|C$, which is not assumed by your Bayes Net.

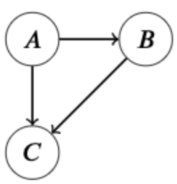
(b) [4 pts] Perform the first step of variable elimination on your Bayes net representation by eliminating variable C. Which probability tables and factors do you have now? Your Bayes net is repeated below for your convenience.



■ P(S)	■ $f(B\|S,Z,W)$	☐ $f(B\|Z,W,C)$	☐ P($B\|Z,C$)	○ None of these
■ P($Z\|S$)	☐ $f(B\|Z)$	☐ $f(C\|S)$	☐ P($B\|Z,W$)	
■ P(W)	☐ $f(B\|Z,C)$	☐ P($B\|S,Z,W$)	☐ P($B\|Z,W,C$)	
☐ P($C\|S,W$)	☐ $f(B\|Z,W)$	☐ P($B\|Z$)	☐ P($C\|S$)	

The original CPTs are $\mathcal{P}(S)$, $\mathcal{P}(W)$, $\mathcal{P}(Z|S)$, $\mathcal{P}(C|S,W)$, and $\mathcal{P}(B|Z,C)$. The ones involving C are $\mathcal{P}(C|S,W)$, $\mathcal{P}(B|Z,C)$. Join and get $B,C|S,W,Z$, sum out C and get $f(B|S,W,Z)$. And we also keep the original CTPs $\mathcal{P}(S)$, $\mathcal{P}(W)$, and $\mathcal{P}(Z|S)$.

(c)



A	P(A)
a_1	0.1
a_2	0.9

B	A	P(B|A)
b_1	a_1	0.3
b_2	a_1	0.7
b_1	a_2	0.8
b_2	a_2	0.2

C	A	B	P(C|A,B)
c_1	a_1	b_1	0.3
c_2	a_1	b_1	0.7
c_1	a_2	b_1	0.6
c_2	a_2	b_1	0.4
c_1	a_1	b_2	0.9
c_2	a_1	b_2	0.1
c_1	a_2	b_2	0.2
c_2	a_2	b_2	0.8

You'd like to try out sampling to find $P(a_1, c_1|b_2)$. Recall from homework that we generate samples by assigning values to variables **(ordering: A, B, C)** with bins and random numbers in the range [0,1].

You generate 200 samples with likelihood sampling.

The number of samples was randomized. Please refer to your own exam and plug in the number of samples for x in the solutions to verify your own answer.

(i) [3 pts] How many random numbers did you use to generate these samples with likelihood sampling? Simplify your answer as much as possible.

$\boxed{2x = 400}$

For each sample, you use a random number to assign a value to A, then enforce B to be b_2 by giving a weight, and use another random number to assign a value to C. So, 2 random numbers for each sample, hence $2x$ random numbers.

(ii) [4 pts] How many random numbers would you approximately need to generate with prior sampling if you wanted to reach a similar accuracy as likelihood sampling from the previous part? Recall that we generated 200 samples with likelihood sampling. Simplify your answer as much as possible.

(Hint: first, think about the ***number of samples*** *you would need to generate with prior sampling if you wanted to reach a similar accuracy.)*

$\boxed{12x = 2400}$

$P(b_2) = 0.1 \cdot 0.7 + 0.9 \cdot 0.2 = 0.25$. So to generate x samples that matches the evidence, we need about $4x$ samples in total using prior sampling.
You need 3 random numbers for each sample because you are assigning all of A, B, and C according to the random numbers. Hence $3 \cdot 4x$.

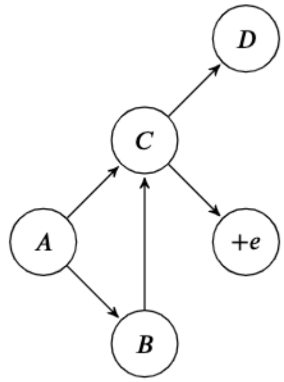
(iii) [4 pts] How many random numbers would you approximately need to generate with rejection sampling if you wanted to reach a similar accuracy as the previous part? Recall that we generated 200 samples with likelihood sampling. Also recall that we assign values to our variables in the order A, B, C. Simplify your answer as much as possible.

$\boxed{9x = 1800}$

For the $\approx x$ samples that matches the evidence, you would need 3 random numbers for each. You would reject $\approx 3x$ samples, and you would use 2 random numbers for each of those. So total random number is about $3 \cdot x + 2 \cdot 3x$.

Q2. [15 pts] Potpourri

(a) [4 pts] (1 pt each) Suppose we run variable elimination on the following Bayes net, where each random variable has domain size 2. We define the size of a factor to be its number of entries. For each variable elimination ordering, what is the size of the largest factor generated?



Domain size was randomized, so refer to your personalized exam to compute the following. For variable domain size x:

- A, B, C, D ____ $x^2 = 4$ ____
- C, B, A, D ____ $x^3 = 8$ ____
- D, C, B, A ____ $x^2 = 4$ ____
- B, C, D, A ____ $x^2 = 4$ ____

(b) [2 pts] Which of the following statements about decision networks are true?

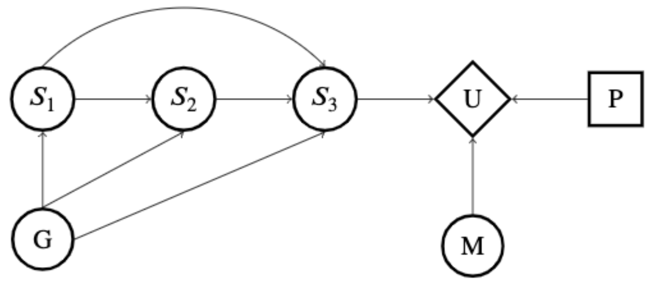
- [ ] If the VPI of observing a variable is 0, then that variable is conditionally independent from the parents of the utility given current evidence.
- [x] If a variable is conditionally independent from the parents of the utility given current evidence, then the VPI of observing that variable is 0.
- [x] A Markov decision process can be represented as a decision network.
- [ ] The value of observing two evidence nodes is always greater than the value of observing one.
- ○ None of the above

- Option 1 is false - for example consider a decision network where the utility is 0 no matter what actions are taken.
- Option 2 is true (see lecture notes).
- Option 3 is true - rewards can be represented by utilities.
- Option 4 is false - the value of information is always non-negative, and can also be 0. Revealing additional variables could leave the MEU the same.

(c) Recall the conventions from the lecture notes:

action nodes as rectangles, chance nodes as ovals, and utility nodes as diamonds.

Consider the following decision network.



There were different randomized orientations of the DN and slightly randomized VPI expressions, but all answers remain the same. Refer to the Alt sections of each subquestion to verify your version.

For each expression, select the choices that are possible:

(i) [1 pt] $VPI(G|S_3)$ — □ < 0 ■ $= 0$ □ > 0

Alt: $VPI(S_2|S_3)$, $VPI(S_1|S_3)$

Since $G \perp\!\!\!\perp U|S_3$, $VPI(G|S_3) = 0$.

(ii) [1 pt] $VPI(S_1|S_2)$ — □ < 0 ■ $= 0$ ■ > 0

Alt: $VPI(G|S_2)$

Since $S_1 \not\perp\!\!\!\perp U|S_2$ (active path: $S_1 - S_3 - U$), $VPI(S_1|S_2) \geq 0$.

(iii) [1 pt] $VPI(G) - VPI(S_2)$ — ■ < 0 ■ $= 0$ ■ > 0

Alt: $VPI(S_1) - VPI(S_2)$

We cannot guarantee any relation between $VPI(G)$ and $VPI(S_2)$.

(iv) [1 pt] $VPI(G) - VPI(M)$ — ■ < 0 ■ $= 0$ ■ > 0

Alt: $VPI(M) - VPI(S_2)$

We cannot guarantee any relation between $VPI(G)$ and $VPI(M)$.

(v) [1 pt] $VPI(S_2) - VPI(S_1, S_2)$ — ■ < 0 ■ $= 0$ □ > 0

Alt: $VPI(S_2) - VPI(S_2, G)$

Recall that according to (ii), $VPI(S_1|S_2) \geq 0$.
$VPI(S_2) - VPI(S_1, S_2) = -(VPI(S_1, S_2) - VPI(S_2)) = -VPI(S_1|S_2) \leq 0$.

(vi) [2 pts] $VPI(M, S_3) - VPI(M) - VPI(S_3)$ — ■ < 0 ■ $= 0$ ■ > 0

Alt: $VPI(S_2, M) - VPI(S_2) - VPI(M)$, $VPI(M, G) - VPI(M) - VPI(G)$

We cannot guarantee any relation between $VPI(M, S_3)$ and $VPI(M) + VPI(S_3)$.
Think about the case of XOR for $VPI(A) + VPI(B) < VPI(A, B)$.

(vii) [2 pts] $VPI(S_2, S_3) - VPI(S_2) - VPI(S_3)$ — ■ < 0 ■ $= 0$ □ > 0

Alt: $VPI(S_3, G) - VPI(S_3) - VPI(G)$

Since $S_2 \perp\!\!\!\perp U|S_3$, $VPI(S_2, S_3) = VPI(S_2)$. So $VPI(S_2, S_3) - VPI(S_2) - VPI(S_3) = -VPI(S_3) \leq 0$.

Q3. [11 pts] Value of Stock Information

A UC Berkeley professor is trying to maximize the value of her stock holdings. The situation is described as follows:

(1) The professor (P) can either sell (s), hold (h), or buy (b) more stock every day.
(2) A market gremlin (G) works for the professor's trading company, and has its own opinion on what its customers should be doing. The gremlin takes one of three positions: sell (s), hold (h), buy (b).
(3) The professor **always writes** (W) on a media platform. She varies between posting her decision on Reddit (r) or Twitter (t).
(4) There is a random internal variable representing the market (M), which either trends down (m_0), trends flat (m_1), or trends up (m_2).

Utility function: The professor only makes money (gains a utility (U) of \$100) when the gremlin matches her order, she posts on twitter (t), **and** the market is up or flat (m_1, m_2). If she selects hold (h) and the other conditions are not met, she gains a utility of \$0. For all other cases, her utility is $-$\$50.
Mathematically, this can be expressed as:

$$U(n) = \begin{cases} 100 & \text{if } (P = G) \wedge (W = t) \wedge (M = m_1 \vee M = m_2) \\ 0 & \text{if } (P = h) \wedge \neg((P = G) \wedge (W = t) \wedge (M = m_1 \vee M = m_2)) \\ -50 & \text{else.} \end{cases}$$

The Utility function was randomized for different versions of the exam. Please refer to your original exam and compute the appropriate solutions by substituting your values. We will refer to the Utility function with x instead of 100 and y instead of -50 when giving the equation for the general case.

(a) Before selecting your actions, suppose that someone could tell you the value of M or G. Calculate the **maximum expected utility (MEU)** given M, given G, and given M and G together. Use the following probability tables to compute your answers.

M	$P(M)$
m_0	1/4
m_1	1/4
m_2	1/2

G	$P(G)$
s	1/3
h	1/3
b	1/3

(i) [3 pts] What is $MEU(M)$? Simplify your answer as much as possible and round your answer to the nearest **tenth** (i.e.: 1.234 would round to 1.2).

25

Note: can definitely answer this with intuition (and no math).
$= \frac{1}{4} MEU(M = m_0) + \frac{1}{4} MEU(M = m_1) + \frac{1}{2} MEU(M = m_2)$
$= \frac{1}{4}(max_P(EU(M = m_0, P = s), EU(M = m_0, P = h), EU(M = m_0, P = b)))$
$+ \frac{1}{4}(max_P(EU(M = m_1, P = s), EU(M = m_1, P = h), EU(M = m_1, P = b)))$
$+ \frac{1}{2}(max_P(EU(M = m_2, P = s), EU(M = m_2, P = h), EU(M = m_2, P = b)))$
$= \frac{1}{4}(max(-50, 0, -50)) + \frac{1}{4}(max(0, 33.33, 0)) + \frac{1}{2}(max(0, 33.33, 0))$
$= 25$
Note, these terms primarily come from two factors: a) the professor can make no money when the market is down and b) the professor only has a random guess of getting the gremlin position (G) correct. The $\frac{1}{3}$ chance of guessing the gremlin correct for \$100 is equally weighted with the $\frac{2}{3}$ loss of $\$-50$ for the sell and buy cases, but the $\$-50$ is negated when hold is chosen.

FOR COMPUTING THE RANDOMIZED ANSWER, the final result should be $\frac{x}{4}$.

(ii) [3 pts] What is $MEU(G)$? Simplify your answer as much as possible and round your answer to the nearest **tenth** (i.e.: 1.234 would round to 1.2).

66.7

$$
\begin{aligned}
&\frac{1}{3}MEU(G=s)+\frac{1}{3}MEU(G=h)+\frac{1}{3}MEU(G=b)\\
=&\frac{1}{3}(max_{w,p}EU(G=s,W=w,P=p))+\frac{1}{3}(max_{w,p}EU(G=h,W=w,P=p))\\
&+\frac{1}{3}(max_{w,p}EU(G=b,W=w,P=p))\\
=&\frac{1}{3}EU(G=s,W=t,P=s)+\frac{1}{3}EU(G=h,W=t,P=h)+\frac{1}{3}EU(G=b,W=t,P=b)\\
=&\frac{1}{3}\cdot(0.5\cdot 100+0.25\cdot 100+0.25\cdot -50)+\frac{1}{3}\cdot(0.5\cdot 100+0.25\cdot 100+0.25\cdot 0)\\
&+\frac{1}{3}\cdot(0.5\cdot 100+0.25\cdot 100+0.25\cdot -50)\\
=&66.7
\end{aligned}
$$

FOR COMPUTING THE RANDOMIZED ANSWER, the final result should be $\frac{2x}{3}$.

(iii) [3 pts] What is $MEU(M, G)$? Simplify your answer as much as possible and round your answer to the nearest **tenth** (i.e.: 1.234 would round to 1.2).

75

75, because you have enough information to definitely gain $100 when the market is up or flat (75% of the time).

FOR COMPUTING THE RANDOMIZED ANSWER, the final result should be $\frac{3x}{4}$.

(iv) [2 pts] Would you prefer to be told G, M, or $either$? ● G ○ M ○ $either$

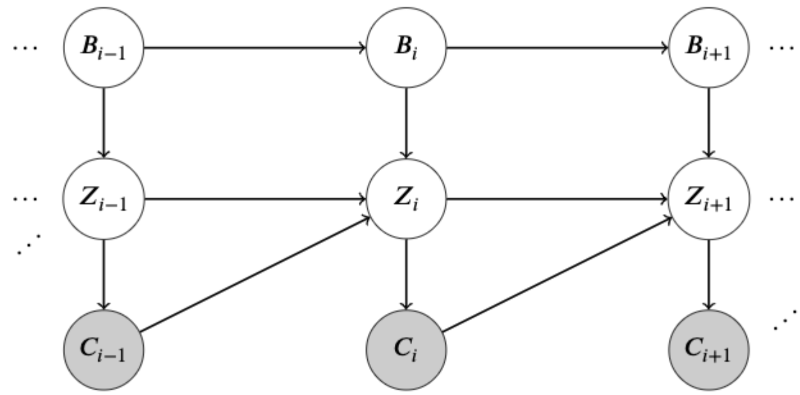
G since it has a higher MEU (and therefore a higher VPI). This is true across all versions

Q4. [21 pts] Secret Tunnel

Aang and his friends are traveling to the Earth Kingdom, when they find themselves trapped in a secret maze of tunnels through the mountains!

They begin in tunnel Z_1 and estimate that any tunnel Z_i will collapse with some probability that they can't observe directly; however, Aang's friend Toph can partially sense information about the tunnels, which directly influences her choices C_i. Aang also senses the existence of an enormous badgermole B_i roaming around the mountains, shaking the tunnel Aang is currently in.

(a) Aang comes up with the following Hidden Markov Model (HMM), where shaded nodes are observable:



(i) [6 pts] Which of the following independence assumptions are enforced by Aang's model? Assume in this subpart that observable nodes **aren't observed** unless given in the independence assumption.

☐ $Z_{i-1} \perp\!\!\!\perp Z_{i+1} \mid Z_i$
☐ $C_{i-1} \perp\!\!\!\perp C_{i+1} \mid C_i$
■ $B_{i-1} \perp\!\!\!\perp B_{i+1} \mid B_i$
☐ $C_{i-1} \perp\!\!\!\perp Z_{i+1} \mid Z_i$
■ $C_{i-1} \perp\!\!\!\perp Z_{i+1} \mid Z_i, B_i$
■ $C_{i-1} \perp\!\!\!\perp B_{i+1} \mid B_i$
☐ $Z_{i-1} \perp\!\!\!\perp B_{i+1} \mid B_{i-1}, C_i$

1. There is an active path from $Z_{i-1} \to B_{i-1} \leftarrow B_i \to B_{i+1} \to Z_{i+1}$, regardless of whether or not Z_i is given.
2. There is an active path from $C_{i-1} \to Z_i \to Z_{i+1} \to C_{i+1}$, regardless of whether or not C_i is given.
3. Given B_i, every path going right from B_{i-1} is blocked. If we go down to Z_{i-1}, we can see that any path back up to B_{i+1} must somehow go through a common effect triple, which is naturally inactive. Therefore, there are no active paths, and the nodes are guaranteed to be independent.
4. There is an active path from $C_{i-1} \leftarrow Z_{i-1} \leftarrow B_{i-1} \to B_i \to B_{i+1} \to Z_{i+1}$, regardless of whether or not Z_i is given.
5. Given B_i, the path that we used in (4) is now inactive. We can see that since we can no longer traverse upward, the only possible remaining active paths are through Z_i. Out of the triples formed with Z_i, only $C_{i-1} \to Z_i \leftarrow B_i$ is active; however, the subsequent triple $Z_i \leftarrow B_i \to B_{i+1}$ is inactive given B_i. Therefore, there are no active paths between the nodes, and they are guaranteed to be independent.
6. All paths through B_i are inactive given B_i, and all remaining paths must go through a common effect triple, which is also inactive. Therefore, the nodes are guaranteed to be independent.
7. There is an active path from $Z_{i-1} \to Z_i \leftarrow B_i \to B_{i+1}$. The triple $Z_{i-1} \to Z_i \leftarrow B_i$ is active given C_i.

Aang now wants to derive an algorithm to compute the belief distribution, but he's not sure how to get started! You decide to help him out by providing him with the time elapse update equation.

(ii) [8 pts] Derive the time elapse update equation using the available terms going from tunnel $i-1$ to tunnel i. Select one choice per blank.

Time Elapse: **(1)** = **(2)** **(3)** **(4)** **(5)**

(1) ○ $P(Z_i, B_i)$ ○ $P\left(Z_{i-1}, B_{i-1}\right)$ ○ $P\left(Z_{i-1}, B_{i-1}|c_{1:i-1}\right)$ ● $P\left(Z_i, B_i|c_{1:i-1}\right)$ ○ $P\left(Z_i, B_i|c_{1:i}\right)$

(2) ○ $\Sigma_{z_{i-1}}$ ○ $\Sigma_{b_{i-1}}$ ● $\Sigma_{z_{i-1}}\Sigma_{b_{i-1}}$ ○ $\Sigma_{z_i}\Sigma_{b_i}$ ○ $\max_{z_i}\max_{b_i}$ ○ $\max_{z_{i-1}}$ ○ $\max_{b_{i-1}}$ ○ $\max_{z_{i-1}}\max_{b_{i-1}}$ ○ 1

(3) ● $P(B_i|b_{i-1})$ ○ $P(B_{i+1}|b_i, z_i)$ ○ $P(Z_i, B_i|b_{i-1})$ ○ $P(b_i|B_{i-1})$ ○ $P(z_i, b_i|B_{i-1})$ ○ $P(B_{i+1}|B_i, Z_i)$ ○ 1

(4) ○ $P(Z_i|z_{i-1}, B_i)$ ● $P(Z_i|z_{i-1}, B_i, c_{i-1})$ ○ $P(Z_i|z_{i-1}, b_{i-1}, c_{i-1})$ ○ $P(z_i|b_i, c_{i-1})$ ○ $P(z_i|Z_{i-1}, b_i)$ ○ $P(z_i|Z_{i-1}, b_i, c_{i-1})$ ○ 1

(5) ○ $P(z_{i-1}, b_{i-1})$ ● $P(z_{i-1}, b_{i-1}|c_{1:i-1})$ ○ $P(z_{i-1}, b_{i-1}|c_{1:i})$ ○ $P(z_i, b_i|c_{1:i-1})$ ○ $P(z_i, b_i|c_{1:i})$ ○ $P(z_i, b_i)$ ○ 1

The Time Elapse update going from $i-1$ to i corresponds to the belief distribution at $i-1$ after incorporating the transition model, which is $P(Z_i, B_i|c_{1:i-1})$. We can now break down the Time Elapse update into parts, considering both the recursive step and the transition model, while using the CPTs that are available to us.

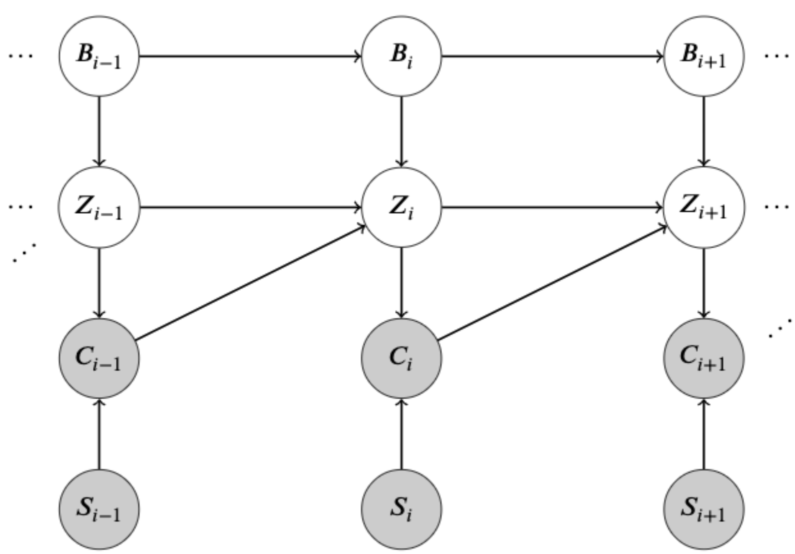
Recursive Step: This is the formula for the belief distribution at $i-1$, which is equivalent to $P(Z_{i-1}, B_{i-1}, |c_{1:i-1})$.

Transition Model: These are the CPTs that correspond to the update of the hidden variables from one timestep to the next. Based on the structure of the HMM, these are $P(B_i|B_{i-1})$ and $P(Z_i|Z_{i-1}, B_i, c_{i-1})$.

Now we need to sum out the hidden variables from the previous timestep, which are Z_{i-1} and B_{i-1}. This leads us to the following final equation:

$$P(Z_i, B_i|c_{1:i-1}) = \Sigma_{z_{i-1}}\Sigma_{b_{i-1}} P(B_i|b_{i-1})P(Z_i|z_{i-1}, B_i, c_{i-1})P(z_{i-1}, b_{i-1}, |c_{1:i-1})$$

(b) Some wandering nomads have joined Aang's group of friends in the tunnels; however, their constant singing S_i annoys Toph, influencing her choices C_i. Aang quickly revises his HMM to incorporate this disturbance, where shaded nodes are observable:



(i) [4 pts] Starting from scratch, Aang wants to derive an algorithm to recompute the belief distribution $P(Z_i, B_i|c_{1:i}, s_{1:i})$ from his revised model. He succeeds at getting the time elapse update equation, but can't remember how to proceed from there.

Complete the observation update equation using the available terms to incorporate evidence for tunnel i. Select one choice per row for each blank.

Observation: $P(Z_i, B_i \mid c_{1:i}, s_{1:i}) \propto$ **(1)** **(2)** **(Time Elapse Update)**

(1) ○ $\Sigma_{z_{i-1}}$ ○ $\Sigma_{b_{i-1}}$ ○ $\Sigma_{z_{i-1}} \Sigma_{b_{i-1}}$ ○ $\Sigma_{z_i} \Sigma_{b_i}$ ○ $\max_{c_i}$
○ $\max_{s_i}$ ○ $\max_{s_{1:i}} \max_{c_i}$ ○ $\max_{s_{1:i}}$ ● 1

(2) ○ $P(c_i)$ ○ $P(c_i|Z_i)$ ○ $P(c_i|Z_i, s_{1:i})$ ● $P(c_i|Z_i, s_i)$ ○ 1

The observation update involves incorporating the evidence at i, which means you just have to multiply by the CPT associated with the evidence with the time elapse update. Therefore, according to this model, we incorporate c_i by multiplying by $P(c_i|Z_i, s_i$. Thus, the final equation looks like the following:

$$P(Z_i, B_i \mid c_{1:i}, s_{1:i}) \propto P(c_i|Z_i, s_i)\textbf{(Time Elapse Update)}$$

Note that we don't need to include $P(S_i)$ because it is a constant.

(ii) [3 pts] Aang's friend Katara advises Aang to use particle filtering, but her brother Sokka disagrees with her. For which of the following cases would particle filtering be a strictly better option over the forward algorithm? Select all that apply, or 'None of the above.'

- ☐ There is an infinite number of tunnels, so $i \to \infty$.
- ☐ All variables are observable.
- ■ There is an exponentially large amount of unobservable states.
- ■ The domains of some random variables are infinite.
- ☐ We want the most accurate possible result.
- ☐ None of the above

1. Infinite tunnels (timesteps) do not necessarily make particle filtering better. If the domain sizes of the random variables are small, then keeping track of particles instead of the full distribution is pointless.
2. All variables being observable negates the necessity to compute a belief distribution in the first place, since we can just compute any distribution directly, but this does not make particle filtering better than the forward algorithm.
3. Particle filtering excels when there are many unobservable hidden variables that we must compute the belief distribution for. An exponential number of unobservable states would force us to save a greater than exponential sized CPT for the full belief distribution if we use the forward algorithm, but we can get away with using particles instead to save much less.
4. Particle filtering excels when domains are infinite, since particles are just samples, taking into account possibilities for the belief distribution. While it is impossible to store the full belief distribution if there exist random variables with infinite domains, we can still use particle filtering to approximate the distribution in those cases.
5. Particle filtering is an approximation, similar to sampling, that we use to approximate the belief distribution. It is not more accurate than computing the exact belief distribution using the forward algorithm.

Q5. [27 pts] Probability

(a) For each statement below about distributions over W, X, Y and Z:

- If the statement is **always true** for any distribution, then write **'Always true.'**
- If the statement is **not always true**, write **EXACTLY ONE** conditional independence assumption which makes it true. For example, if the necessary assumption is that W must be independent of X given Y, please write $W \perp\!\!\!\perp X | Y$.
- If the statement is **never true**, or is only true with **TWO OR MORE** conditional independence assumptions then write **'No solution.'**

Assume that the distributions are **nontrivial**, meaning that you should not assume them to be always true or always false.

The following questions make use of Chain Rule, Bayes Rule, and the Law of Total Probability to try to prove or disprove an equality.

(i) [2 pts] $P(W, X|Y) = \frac{P(Y,W|X)P(X)}{P(Y)}$ Always true

(ii) [2 pts] $P(Y|X, Z) = \frac{P(X)P(Y|X)P(Z|X,Y)}{\sum_y P(X)P(Y|X)P(Z|Y,X)}$ Always true

(iii) [2 pts] $P(X, Z) = P(X|Y)P(Z)$ No solution

(iv) [2 pts] $P(X, Y, Z) = P(X, Z)P(Y|X, Z)$ Always true

(v) [2 pts] $P(W|X, Y) = \frac{P(W)P(X|W)P(Y|W)}{P(X|Y)P(Y)}$ True if $X \perp\!\!\!\perp Y|W$

(vi) [2 pts] $P(W, X, Y) = P(Y)P(W|Y)P(X|Y)$ True if $W \perp\!\!\!\perp X|Y$

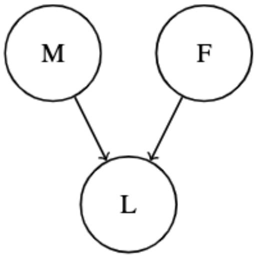
(vii) [2 pts] $P(W, X|Y, Z) = P(W|Y, Z)P(X|W, Y, Z)$ Always true

(viii) [2 pts] $P(X|Y, Z) = \frac{P(X,Y)}{P(Y)P(Z)}$ No solution

(b) You recently flew back to Berkeley for the Online 2020 semester and have decided to use a simple Bayes nets to analyze the probability that you will enjoy your new college experience. However, this depends on whether your friends come back to Berkeley and if you have any early morning classes. You know that the probability of your friends returning (event $F = +f$) is 0.8 and the probability of you having to enroll in a morning class (event $M = +m$) is 0.7. Let random variables F and M represent each of these events and let L take the value $+l$ if you like the new format and $-l$ if you don't.

The values of $P(F = +f)$ and $P(M = +m)$ were randomized between versions of the exam. See the last page after all subparts for the solutions for each different version, and compare your answer to this. Use the equations set up in each subpart for the procedure.

The relationship between these variables is given by the following Bayes Net:



M	F	$P(L=+l \mid M,F)$
$+m$	$+f$	0.05
$+m$	$-f$	0.01
$-m$	$+f$	0.9
$-m$	$-f$	0.02

(i) [3 pts] Give the formula for the joint probability distribution induced by the above Bayes Net.

$P(M)P(F)P(L|M,F)$

(ii) [2 pts] Compute the probability that your friends return and you don't have to take any morning classes. Simplify your answer as much as possible and round your answer to the nearest **hundredth** (i.e.: 1.234 would round to 1.23).

0.24

$P(-m)P(+f) = 0.3 * 0.8 = 0.24$ by independence

(iii) [2 pts] Compute the probability that your friends return *or* you don't have to take any morning classes. Simplify your answer as much as possible and round your answer to the nearest **hundredth** (i.e.: 1.234 would round to 1.23).

0.86

$P(-m) + P(+f) - P(-m,+f) = 0.3 + 0.8 - 0.24 = 0.86$

(iv) [2 pts] Compute the probability that your friends came back given that you like the semester. Simplify your answer as much as possible and round your answer to the nearest **hundredth** (i.e.: 1.234 would round to 1.23).

0.99

$P(+f \mid +l) = \frac{P(+f,+l)}{P(+l)} = \frac{P(+m,+f,+l)+P(-m,+f,+l)}{\sum_{f,m} P(f,m,+l)} = \frac{0.7*0.8*0.05+0.3*0.8*0.9}{0.7*0.8*0.05+0.3*0.8*0.9+0.7*0.2*0.01+0.3*0.2*0.02} = 0.99$

(v) [2 pts] Compute the probability that you had to take a morning class and none of your friends came back given that you do not like the semester. Simplify your answer as much as possible and round your answer to the nearest **hundredth** (i.e.: 1.234 would round to 1.23).

0.18

$P(+m,-f \mid -l) = \frac{P(+m)P(-f)P(-l|+m,-f)}{P(-l)} = \frac{P(+m)P(-f)P(-l|+m,-f)}{\sum_{f,m} P(f,m,-l)} = \frac{0.7*0.2*0.99}{0.7*0.8*0.95+0.3*0.8*0.1+0.7*0.2*0.99+0.2*0.3*0.98} =$
0.18

SOLUTIONS TO RANDOMIZED $P(F = +f)$ and $P(M = +m)$:

Column 1: $P(+m)$
Column 2: $P(+f)$
Column 3: Answer to Q5b.ii
Column 4: Answer to Q5b.iii
Column 5: Answer to Q5b.iv
Column 6: Answer to Q5b.v

```
[0.7, 0.7, 0.21, 0.79, 0.98, 0.27]
[0.7, 0.1, 0.03, 0.37, 0.72, 0.65]
[0.7, 0.2, 0.06, 0.44, 0.85, 0.6]
[0.7, 0.3, 0.09, 0.51, 0.91, 0.54]
[0.7, 0.4, 0.12, 0.58, 0.94, 0.48]
[0.7, 0.5, 0.15, 0.65, 0.96, 0.41]
[0.7, 0.6, 0.18, 0.72, 0.97, 0.34]
[0.7, 0.8, 0.24, 0.86, 0.99, 0.18]
[0.1, 0.7, 0.63, 0.97, 0.99, 0.07]
[0.1, 0.1, 0.09, 0.91, 0.83, 0.1]
[0.1, 0.2, 0.18, 0.92, 0.91, 0.1]
[0.1, 0.3, 0.27, 0.93, 0.95, 0.09]
[0.1, 0.4, 0.36, 0.94, 0.97, 0.09]
[0.1, 0.5, 0.45, 0.95, 0.98, 0.08]
[0.1, 0.6, 0.54, 0.96, 0.98, 0.08]
[0.1, 0.8, 0.72, 0.98, 0.99, 0.06]
[0.2, 0.7, 0.56, 0.94, 0.99, 0.12]
[0.2, 0.1, 0.08, 0.82, 0.82, 0.2]
[0.2, 0.2, 0.16, 0.84, 0.91, 0.19]
[0.2, 0.3, 0.24, 0.86, 0.95, 0.18]
[0.2, 0.4, 0.32, 0.88, 0.96, 0.17]
[0.2, 0.5, 0.4, 0.9, 0.98, 0.16]
[0.2, 0.6, 0.48, 0.92, 0.98, 0.14]
[0.2, 0.8, 0.64, 0.96, 0.99, 0.1]
[0.3, 0.7, 0.49, 0.91, 0.99, 0.16]
[0.3, 0.1, 0.07, 0.73, 0.81, 0.29]
[0.3, 0.2, 0.14, 0.76, 0.9, 0.28]
[0.3, 0.3, 0.21, 0.79, 0.94, 0.26]
[0.3, 0.4, 0.28, 0.82, 0.96, 0.24]
[0.3, 0.5, 0.35, 0.85, 0.97, 0.22]
[0.3, 0.6, 0.42, 0.88, 0.98, 0.2]
[0.3, 0.8, 0.56, 0.94, 0.99, 0.12]
[0.4, 0.7, 0.42, 0.88, 0.99, 0.2]
[0.4, 0.1, 0.06, 0.64, 0.8, 0.38]
[0.4, 0.2, 0.12, 0.68, 0.9, 0.36]
[0.4, 0.3, 0.18, 0.72, 0.94, 0.34]
[0.4, 0.4, 0.24, 0.76, 0.96, 0.31]
[0.4, 0.5, 0.3, 0.8, 0.97, 0.28]
[0.4, 0.6, 0.36, 0.84, 0.98, 0.24]
[0.4, 0.8, 0.48, 0.92, 0.99, 0.14]
[0.5, 0.7, 0.35, 0.85, 0.99, 0.22]
[0.5, 0.1, 0.05, 0.55, 0.78, 0.47]
[0.5, 0.2, 0.1, 0.6, 0.89, 0.44]
[0.5, 0.3, 0.15, 0.65, 0.93, 0.41]
[0.5, 0.4, 0.2, 0.7, 0.95, 0.37]
[0.5, 0.5, 0.25, 0.75, 0.97, 0.33]
[0.5, 0.6, 0.3, 0.8, 0.98, 0.28]
[0.5, 0.8, 0.4, 0.9, 0.99, 0.16]
[0.6, 0.7, 0.28, 0.82, 0.98, 0.25]
[0.6, 0.1, 0.04, 0.46, 0.76, 0.56]
[0.6, 0.2, 0.08, 0.52, 0.87, 0.52]
[0.6, 0.3, 0.12, 0.58, 0.92, 0.48]
[0.6, 0.4, 0.16, 0.64, 0.95, 0.43]
[0.6, 0.5, 0.2, 0.7, 0.97, 0.37]
[0.6, 0.6, 0.24, 0.76, 0.98, 0.31]
[0.6, 0.8, 0.32, 0.88, 0.99, 0.17]
[0.8, 0.7, 0.14, 0.76, 0.98, 0.28]
[0.8, 0.1, 0.02, 0.28, 0.67, 0.74]
[0.8, 0.2, 0.04, 0.36, 0.82, 0.67]
[0.8, 0.3, 0.06, 0.44, 0.89, 0.6]
[0.8, 0.4, 0.08, 0.52, 0.92, 0.53]
[0.8, 0.5, 0.1, 0.6, 0.95, 0.45]
[0.8, 0.6, 0.12, 0.68, 0.96, 0.37]
[0.8, 0.8, 0.16, 0.84, 0.99, 0.19]
```