INSTRUCTIONS

- **Due: Tuesday, November 21, 2023 at 10:00 PM EST.** Remember that you may use up to 1 slip day for the written making the last day to submit **Wednesday, November 22, 2023 at 10:00 PM EST**.
- **Format:** Write your answers in the `yoursolution.tex` file and compile a pdf (preferred) or you can type directly on the blank pdf. Make sure that your answers are within the dedicated regions for each question/part. If you do not follow this format, we may deduct points. Handwritten solutions are not acceptable and may lead to lost points.
- **How to submit:** Submit a pdf with your answers on Gradescope. Log in and click on our class 15-281, click on the HW9 assignment, and upload your pdf containing your answers.
- **Policy:** See the course website for homework policies and academic integrity.

Name	
Andrew ID	
Hours to complete?	◯ (0, 2] hours ◯ (2, 3] hours ◯ (3, 4] hours ◯ (4, 5] hours ◯ (5, 6] hours ◯ (6, 7] hours ◯ (7, 8] hours ◯ > 8 hours

Q1. [38 pts] Variable Elimination Nation

Suppose you are given a Bayes net with the same variables and structure as the alarm Bayes net from lecture, with the conditional probability tables given below.

B	$P(B)$
$+b$	0.420
$-b$	0.58

E	$P(E)$
$+e$	0.31
$-e$	0.69

A	B	E	$P(A \mid B, E)$
$+a$	$+b$	$+e$	0.4
$-a$	$+b$	$+e$	0.6
$+a$	$+b$	$-e$	0.2
$-a$	$+b$	$-e$	0.8
$+a$	$-b$	$+e$	0.5
$-a$	$-b$	$+e$	0.5
$+a$	$-b$	$-e$	0.9
$-a$	$-b$	$-e$	0.1

J	A	$P(J \mid A)$
$+j$	$+a$	0.7
$-j$	$+a$	0.3
$+j$	$-a$	0.3
$-j$	$-a$	0.7

M	A	$P(M \mid A)$
$+m$	$+a$	0.125
$-m$	$+a$	0.875
$+m$	$-a$	0.25
$-m$	$-a$	0.75

Apply the variable elimination algorithm to the query $P(B \mid +j, +m)$. You will have to eliminate the variables E and A, in that order. For each variable to be eliminated, write:

(i) The variables remaining in the resulting factor (e.g., after eliminating X the resulting factor might be $f_1(Y, Z)$, so the variables remaining are Y, Z). DO NOT include constant values (i.e. the given evidence values).

(ii) The summation to calculate the factor (e.g., $f_1(Y, Z) = \sum_x P(Y)P(x \mid Y)P(Z \mid x)$)

(iii) The numeric values in the factor table. Include an entry for each combination of assignments for the remaining variables, similar to the probability tables given above.

(a) [2 pts] Write the full summation over variables A and E to compute $P(B \mid +j, +m)$. Make sure to use the normalization constant (α or Z):

Summation (use α or Z):

(b) [8 pts] Eliminating E:

(i) Variables Remaining in Factor:

(ii) Summation:

(iii) Factor table:

(c) [8 pts] Eliminating A:

(i) Variables Remaining in Factor:	**(ii) Summation:**
(iii) Factor table:	

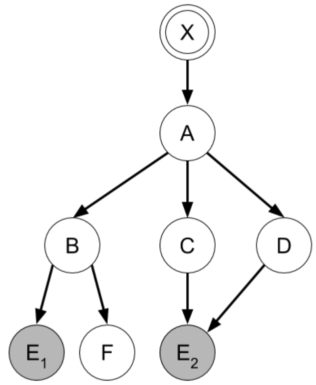
(d) [8 pts] After eliminating E and A above, you must multiply the remaining factors to produce yet another factor table. Specify the variables associated with the resulting factor, and fill out the values in the corresponding factor table.

(i) Variables Remaining in Factor:	**(ii) Factor table:**

Now, find the normalizing constant to make this factor the probability distribution we want, $P(B \mid +j, +m)$. Write out the values of this normalized probability table.

(iii) Constant (α or Z):	**(iv) Probability table for $P(B \mid +j, +m)$:**

Now consider the Bayes net below. Suppose we are trying to compute the query $P(X \mid e_1, e_2)$. Assume all variables are binary.



(e) [2 pts] Suppose we choose to eliminate variables in the order A, C, B, D, F.
Which factor has the most entries in its corresponding table, considering all resulting factors from eliminating each of the variables in this order? Which variable was eliminated to result in this factor? How many entries are in its table? Assume that we have separate entries for pairs of numbers even if we know sum to one (e.g., we would store both $P(X = +x)$ and $P(X = -x)$).

(i) Factor:	**(ii) # entries:**

(f) [10 pts] An optimal variable elimination ordering is one which minimizes the sum of the sizes of factors generated.

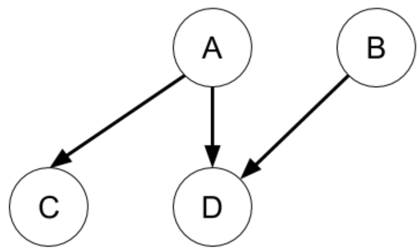
Fill in the table below with an optimal variable elimination ordering. For each variable, include the resulting factor and the number of entries in its table, again assuming that we separately store pairs of numbers which sum to one.

Variable	Factor	# Entries

Q2. [12 pts] Irrelevance Criteria

A variable in a Bayes net is said to be *irrelevant* to a given query if we could remove the variable from the Bayes net without changing the answer to the query. For example, in the Bayes net below, the variable D is irrelevant to the query $\mathbf{P}(B)$, because we already have the $\mathbf{P}(B)$ table. On the other hand, B is *not* irrelevant to the query $\mathbf{P}(D)$, because $\mathbf{P}(B)$ is needed to compute $\mathbf{P}(D)$.

Consider the Bayes net below:



(a) [6 pts] Suppose we are making the query $\mathbf{P}(A \mid D = d)$. Prove that C is irrelevant to this query using the following steps:

(i) Write the full joint distribution as a product of the CPTs in the Bayes net.

(ii) Sum over this product and normalize to get $\mathbf{P}(A \mid d)$. Make sure you write out the full equations (i.e. don't use any normalizing constants).

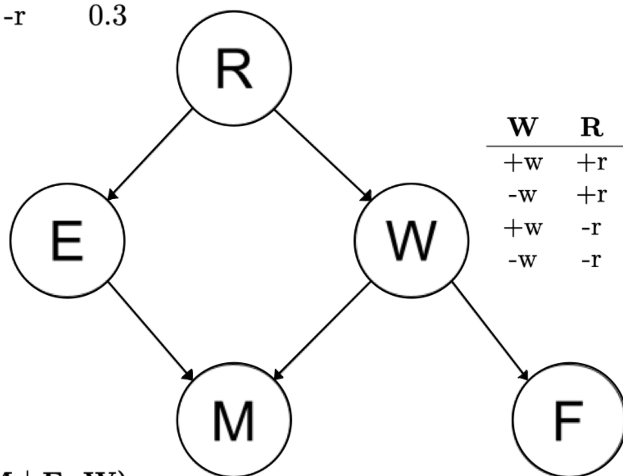
(iii) Explain why this expression does not depend on the variable C.

(b) [6 pts] Now suppose we are making the query $\mathbf{P}(C \mid D = d)$. Execute the first two steps in part (a) for this query, and then argue why B is *not* irrelevant.

Q3. [20 pts] Bayes Nets Sampling

Consider the following Bayes Net and corresponding probability tables.

R	**P(R)**
+r	0.7
-r	0.3



W	**R**	**P(W | R)**
+w	+r	0.9
-w	+r	0.1
+w	-r	0.3
-w	-r	0.7

E	**R**	**P(E | R)**
+e	+r	0.6
-e	+r	0.4
+e	-r	0.2
-e	-r	0.8

M	**E**	**W**	**P(M | E, W)**
+m	+e	+w	0.25
-m	+e	+w	0.75
+m	+e	-w	0.35
-m	+e	-w	0.65
+m	-e	+w	0.45
-m	-e	+w	0.55
+m	-e	-w	0.1
-m	-e	-w	0.9

F	**W**	**P(F | W)**
+f	+w	0.25
-f	+w	0.75
+f	-w	0.85
-f	-w	0.15

Consider the case where we are sampling to approximate the query $P(R \mid +f, +m)$.

(a) [12 pts] Fill in the following table with the probabilities of ***drawing AND using*** each respective sample given that we are using each of the following sampling techniques.

Please show your calculation for partial credit (please don't change the sizes of the boxes - your submission will become misaligned).

Method	$< +r, +e, -w, +m, +f >$	$< +r, -e, +w, -m, +f >$
Prior sampling		
Rejection sampling		
Likelihood weighting		

(b) [8 pts]

"Sampling is important." – Gibby Gibson, *iCarly*, S2 Ep. 81

We are going to use Gibbs sampling to estimate the probability of getting the sample $< +r, +e, -w, +m, +f >$. We will start from the sample $< -r, -e, -w, +m, +f >$ and resample twice. First we'll resample E. Second we'll resample R.

Please show your work for partial credit (please don't change the sizes of the boxes - your submission will become misaligned).

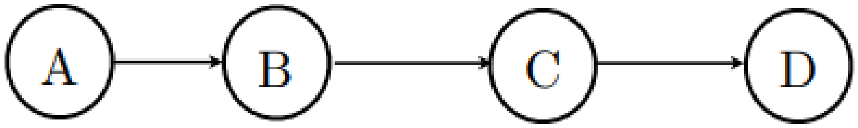
(i) [3 pts] Resample E $(+e)$. Compute the probability of the first resampling. That is, calculate the probability of $+e$ given the values of the remaining variables *in the starting sample*:

(ii) [3 pts] Resample R $(+r)$. Compute the probability of the second resampling. That is, compute the probability of $+r$ given the values of the remaining variables *after the first resampling*:

(iii) [2 pts] What is the probability of drawing sample $< +r, +e, -w, +m, +f >$?

Q4. [24 pts] More Sampling

Consider the following Bayes Net and corresponding probability tables. Show work for partial credit.



$P(A)$	
$-a$	3/4
$+a$	1/4

$P(B\|A)$		
$-a$	$-b$	2/3
$-a$	$+b$	1/3
$+a$	$-b$	4/5
$+a$	$+b$	1/5

$P(C\|B)$		
$-b$	$-c$	1/4
$-b$	$+c$	3/4
$+b$	$-c$	1/2
$+b$	$+c$	1/2

$P(D\|C)$		
$-c$	$-d$	1/8
$-c$	$+d$	7/8
$+c$	$-d$	5/6
$+c$	$+d$	1/6

(a) [6 pts] You are given the following samples:

$(-a, +b, +c, -d)$	$(-a, -b, +c, -d)$	$(-a, +b, +c, +d)$	$(-a, -b, -c, -d)$
$(-a, -b, -c, +d)$	$(-a, -b, +c, -d)$	$(+a, +b, -c, +d)$	$(+a, +b, +c, -d)$

(i) [1 pt] Estimate $P(-c)$. Assume that these samples came from performing **prior sampling**.

P(-c):

(ii) [5 pts] Now we will query $P(-c \mid -a, -d)$ using **rejection sampling**. List out the samples from above that would not be used when doing rejection sampling for this task, and write the estimate for $P(-c \mid -a, -d)$.

Rejected Samples:

P(-c|-a,-d):

(b) [4 pts] We will now use **likelihood weighted sampling** to estimate $P(-c \mid -a, -d)$. Give the weights for the following samples.

(i) [1 pt] $(-a, +b, +c, -d)$

Weight:

(ii) [1 pt] $(-a, +b, -c, -d)$

Weight:

(iii) [1 pt] $(-a, -b, +c, -d)$

Weight:

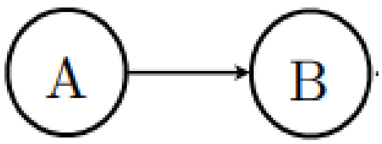
(iv) [1 pt] $(-a, -b, -c, -d)$

Weight:

(c) [2 pts] Estimate $P(-c \mid -a, -d)$ via **likelihood weighted sampling**. Use the four samples and their calculated weights from part **(b)**. Assume a count of 1 for each sample. **DO NOT** use the eight samples from part **(a)**. Round your answer to 4 decimal places.

P(-c|-a,-d):

(d) [12 pts] Consider the following Bayes net with binary variables A, B.



A	$P(A)$
$+a$	.01
$-a$	.99

A	B	$P(B\|A)$
$+a$	$+b$	.99
$+a$	$-b$	.01
$-a$	$+b$	.0001
$-a$	$-b$	.9999

(i) [2 pts] Use **inference** to find the actual probability distribution $P(A \mid +b)$.

P(+a|+b): | **P(-a|+b):**

(ii) [2 pts] We are trying to estimate $P(A \mid +b)$ using **likelihood weighted sampling**. What is the first sample most likely to be, and what is the corresponding weight of the sample?

Sample: | **Weight:**

(iii) [3 pts] Following **likelihood weighted sampling**, suppose we only draw 5 samples, how many of them will we expect to be $(-a, +b)$? (Round to the nearest integer). What distribution $P(A \mid +b)$ will be estimated from using these 5 expected samples?

Number $(-a, +b)$**:** | **P(+a|+b):** | **P(-a|+b):**

(iv) [3 pts] How about 1000 samples?

Number $(-a, +b)$**:** | **P(+a|+b):** | **P(-a|+b):**

(v) [2 pts] In some Bayes net $A \to B \to C \to D \to E$, will **likelihood weighted sampling** provide more accurate distributions for $P(A \mid E)$ or $P(E \mid A)$ with a relatively low sample count? Explain why (you may use previous parts as an example).

Hint: what happens if the evidence can be reached only by unlikely values of *upstream* (opposite direction of Bayes Net arrows) variables?

Answer:

Q5. [6 pts] Bayes Nets IRL

Draw a Bayes Net representing some aspect of your life. For example, you could try to relate different parts of your life as a student like doing homework, attending lecture, laughing at Alex's top-tier memes, getting sleep, playing sports, cooking, partying, etc. Another approach is to think about how a Bayes Net could be applied to a favorite hobby of yours. Feel free to deviate from this, we want you to get as creative as you can. :)

TO GET FULL POINTS, you must include at least one set of variables A, B, C such that $A \perp\!\!\!\perp B \mid C$, and one set of variables X, Y, Z such that $X \cancel{\perp\!\!\!\perp} Y \mid Z$. You don't have to name the variables those names, of course.

(a) [2 pts] Draw your Bayes Net. **Please include at least 5 nodes.**

Answer:

(b) [4 pts]

(i) [2 pts] Write one conditional independence relationship between variables (ex. $A \perp\!\!\!\perp B \mid C$) :

Answer:

(ii) [2 pts] Identify three variables such that two variables that are NOT conditionally independent given a third: (ex. $X \cancel{\perp\!\!\!\perp} Y \mid Z$). Use `\cancel{\bigCI}` to cross out $\perp\!\!\!\perp$.

Answer: