Introduction to Artificial Intelligence Dieter Fox
CSE 473, Spring 2017

Final

June 5, 2017

Directions

This exam has 6 problems with points shown in the table below, and you have 110 minutes to complete it.

- The exam is closed book. No calculators are needed.
- If you have trouble with a question, by all means move on to the next problem—or the next part of the same problem.
- In answers incurring numbers, feel free to not resolve fractions and sums.
- If a question is unclear, feel free to ask me for clarification!
- **Please do not turn the page until I indicate that it is time to begin.**

Answer Key

Name: ______________________________

Number: ______________________________

1) True/False	/30
2) Reinforcement Learning	/13
3) Hidden Markov Models	/20
4) Probabilities	/12
5) Bayes Net	/17
6) Variable Elimination	/9 ~~15~~
Total	/101 ~~107~~

1. (30 points: 2 pts each, -2 pts if wrong) **True or False** - Circle the correct answer.

   (a) A* search with a heuristic that is not completely admissible cannot find the shortest path to the goal state. . . . . T F

   False

   (b) The heuristic $h(n) = 0$ is admissible for every search problem with non-negative cost. . . . . T F

   True

   (c) Expectimax is better suited than minimax for a stochastic environment. . . . . T F

   True

   (d) The Degree Heuristic for CSPs says to select the variable that imposes the most constraints on the remaining variables. . . . . T F

   True

   (e) If two MDPs have different discount factors (γ) but are otherwise identical, they will still always have the same optimal policy. . . . . T F

   False

   (f) In a Bayes Net, if the nodes for X and Y are D-separated given Z, then the two random variables X and Y are always conditionally independent given only* the random variable Z. . . . . T F

   True (*"Only" was clarified during the exam)

   (g) In a Bayes Net, suppose X is independent of Y given Z. Is it possible the assumption will no longer hold when conditioning on additional evidence to other variables in the network? . . . . T F

   True

   (h) In a Bayes net with at least three variables, if $A \perp\!\!\!\perp B$, there is always a variable C (other than A or B) that also makes $A \perp\!\!\!\perp B|C$. . . . . T F

   False

   (i) Particle filtering can never produce exactly the correct distribution. . . . . T F

   False

   (j) Model-free reinforcement learning is usually more efficient than model-based reinforcement learning, because it does not have to learn the transition model. . . . . T F

   True

(k) Feature-based reinforcement learning converges to the same policy as the non-feature based version and converges faster. . . . . T F

False

(l) A and B are random variables with binary domains. The probability table $P(A|B)$ contains 4 entries that sum to 2. . . . . T F

True

(m) Given random variables A and B and the distributions $P(A)$ and $P(B|A)$, we can conclude $P(A)P(B|A) = P(A, B)$ without any assumptions. . . . . T F

True

(n) The Viterbi algorithm is designed to determine the state with maximum aposteriori probability at the last time step of a sequence. . . . . T F

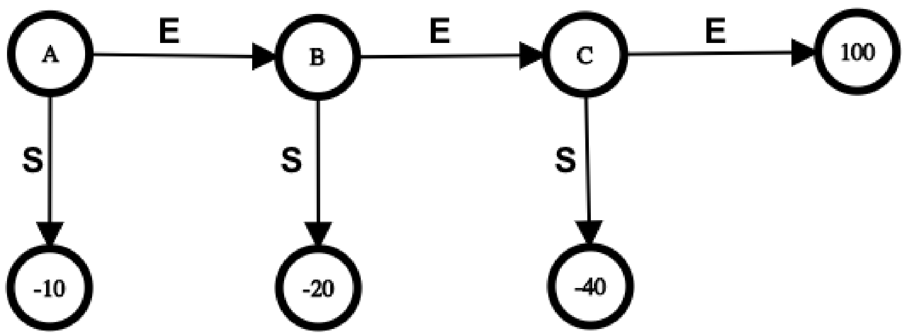
False

(o) The running time of the Viterbi algorithm is linear in the length of a sequence. . T F

True

2. (13 points) **Reinforcement Learning**

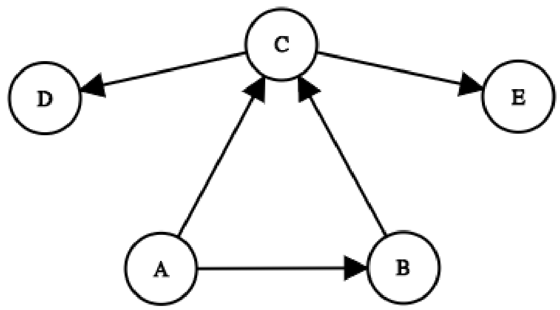
   (a) (4 points) Policy evaluation: You are a treasure hunter, and just found a map toward a fabulous treasure. You are currently in state A. A policy toward the treasure is given by the map, which says to always take the action to the East (E). However, there are three pitfalls along the way. At each state, there is a 50% chance that you will actually go East, and a 50% chance to go South (S) and fall into a pitfall. The penalties of each pitfall and the reward of the treasure is marked in the graph below. You will get the reward and penalty immediately after you reach the state (it corresponds to the familiar R(s, a, s') part). All other transitions have 0 reward. Please evaluate the values of state A, B, and C according to this policy. Use a discount factor of $\gamma = 1$.



Policy π:
$\pi(A) = E$
$\pi(B) = E$
$\pi(C) = E$

	$t=0$	$t=1$	$t=2$	$t=3$	$t=4$
A	0	-5	-10	-2.5	-2.5
B	0	-10	5	5	5
C	0	30	30	30	30

(b) (4 points) Q-Learning: After collecting your treasure, you are surprised to find another map about another island. This map does not mark the position of the treasure. Instead, it contains initial Q-values and two records of experiences a previous treasure hunter made. Each record represents current state, action, resulting state, and the reward, where $X \rightarrow Y$ represents the action from state X to state Y.



Initial Q-values:

i. $Q(A, A \rightarrow B) = Q(B, B \rightarrow C) = Q(A, A \rightarrow C) = 2$

ii. $Q(C, C \rightarrow D) = 4$

iii. $Q(C, C \rightarrow E) = 6$

Records:

i. $A, A \rightarrow B, B$, reward: 4

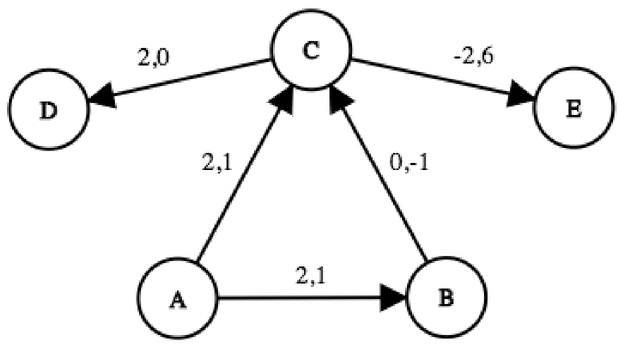
ii. $B, B \rightarrow C, C$, reward: -1

Unfortunately, the previous treasure hunter did not update the Q-values based on these records. Provide the updated Q-values corresponding to each record, using the initial Q-values. Use a learning rate of $\alpha = 0.5$ and discount factor of $\gamma = 1.0$. Please show your work.

$Q(A \rightarrow B) = (1 - 0.5) * 2 + 0.5 * [4 + 1 * 2] = 4$

$Q(B \rightarrow C) = (1 - 0.5) * 2 + 0.5 * [-1 + 1 * 6] = 3.5$

(c) (4 points) Approximate Q-Learning: Suddenly you realize that there's more information on the map: two feature values for each action you could take, which are marked on the map. With these features and the same records, please perform feature-based Q-learning and calculate the two feature weights w_1 and w_2 according to the two provided actions one by one in order. Initially, $w_1 = 1$ and $w_2 = 0$. Use a learning rate of $\alpha = 0.5$ and discount factor of $\gamma = 1.0$.



Records:

i. $A, A \to B, B$, reward: 4

ii. $B, B \to C, C$, reward: -1

Provide the updated weights after each record is observed. Please show your work.

$A \to B$:

$w_1 = 1 + 0.5 * \{[4 + 1 * (1 * 0 + 0 * -1)] - (1 * 2 + 0 * 1)\} * 2 = 3$

$w_2 = 0 + 0.5 * \{[4 + 1 * (1 * 0 + 0 * -1)] - (1 * 2 + 0 * 1)\} * 1 = 1$

$B \to C$:

$w_1 = 3 + 0.5 * \{[-1 + 1 * (3 * 2 + 1 * 0)] - (3 * 0 + 1 * -1)\} * 0 = 3$

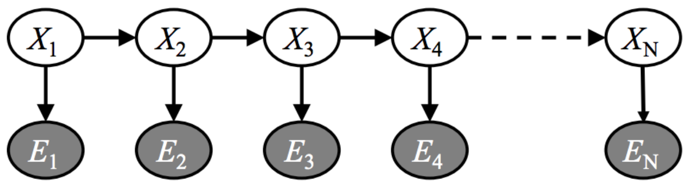
$w_2 = 1 + 0.5 * \{[-1 + 1 * (3 * 2 + 1 * 0)] - (3 * 0 + 1 * -1)\} * -1 = -2$

(d) (1 points) If you want to be more sensitive to each new sample while learning the feature weights, which parameter would you change, and how?

Increase learning rate α

3. (20 points) **Hidden Markov Models**

Consider a hypothetical secret agent working in the government. We will model the agent's political allegiance as a Hidden Markov Model:



Hidden states are binary and represent the secret agent's allegiance to either the United States (U) or Russia (R) on a given day. Between each day there is a 20% chance that his allegiance switches to the opposite country, otherwise his allegiance stays with the same country as the previous day. We know that on day 1 the agent is allied with the United States: $P(X_1 = U) = 1.0$.
Do not consider evidence on questions (a)-(e).

(a) (1 point) Fill in the table with the transition probabilities as defined above, where the row corresponds to X_{t+1} and the column to X_t.

	U	R
U	$P(X_{t+1} = U \mid X_t = U) = 0.8$	$P(X_{t+1} = U \mid X_t = R) = 0.2$
R	$P(X_{t+1} = R \mid X_t = U) = 0.2$	$P(X_{t+1} = R \mid X_t = R) = 0.8$

(b) (2 points) What is the probability of producing the exact sequence U, U, R, R, R on the first 5 days? You do not need to evaluate arithmetic expressions involving only numbers (also, as state above, you should ignore evidence).

$1.0 * 0.8 * 0.2 * 0.8 * 0.8 = 0.1024$

(c) (2 points) Rank the following sequences from most likely to least likely to occur beginning from the first day. 1 means most likely and 3 means least likely. Write one number (1, 2, or 3) in each box, or the word "TIE" if there is a tie.

U, R, R, R, R	U, U, U, U, U	U, R, U, R, U
2	1	3

Switching is less likely than staying constant.

(d) (6 points) Without considering evidence, compute the steps of the Forward algorithm for Markov Models. Fill in the table with the probabilities for the secret agent's political allegiance at each time step. You do not need to simplify numbers.

	$t = 1$	$t = 2$	$t = 3$
$P(X_t = U)$	1.0	0.8	0.68
$P(X_t = R)$	0.0	0.2	0.32

$P(X_1 = U) = 1.0$ (Given)
$P(X_1 = R) = 0.0$ (Given)
$P(X_2 = U) = 1.0 * 0.8 + 0.0 * 0.2 = 0.8$
$P(X_2 = R) = 1.0 * 0.2 + 0.0 * 0.8 = 0.2$
$P(X_3 = U) = 0.8 * 0.8 + 0.2 * 0.2 = 0.68$
$P(X_3 = R) = 0.8 * 0.2 + 0.2 * 0.8 = 0.32$

(e) (2 points) Consider the final day of the secret agent's career, several years in the future. For a large enough t, the probabilities $P(X_t)$ will converge to a stationary distribution. Provide the probabilities to which the stationary distribution will ultimately converge.

Stationary distribution is uniform because the transitions are symmetric, so $P(X_{stationary} = R) = P(X_{stationary} = U) = 0.5$.

Observing evidence - We will now consider evidence in the form of the secret agent's Twitter posts. When allied with the United States, he is equally likely to make a Tweet (T) or not make a Tweet ($\neg T$) on that day. However, when allied with Russia, there is a 90% chance that there will be a Tweet on that day.

(f) (1 point) Fill in the table with the conditional probabilities of the evidence based on the description above.

	U	R
T	$P(T\|U) = 0.5$	$P(T\|R) = 0.9$
$\neg T$	$P(\neg T\|U) = 0.5$	$P(\neg T\|R) = 0.1$

(g) (6 points) Fully define the conditional distribution $P(X_2|E_2)$ for the initial belief of $P(X_1 = U) = 1.0$. This initial belief is after having observed E_1, that is, you can ignore E_1 in your computation. Consider the network only up to time step 2. Provide the probabilities for each possible setting of the variables in $P(X_2|E_2)$. You do not need to simplify values, but you do need to give all terms necessary for each equation.

$P(X_2 = U|E_2 = T) \propto 1.0 * 0.8 * 0.5 = 0.4$
$P(X_2 = R|E_2 = T) \propto 1.0 * 0.2 * 0.9 = 0.18$
$P(X_2 = U|E_2 = \neg T) \propto 1.0 * 0.8 * 0.5 = 0.4$
$P(X_2 = R|E_2 = \neg T) \propto 1.0 * 0.2 * 0.1 = 0.02$

$P(X_2 = U|E_2 = T) = (0.8 * 0.5)/(0.8 * 0.5 + 0.2 * 0.9) = 0.4/0.58$
$P(X_2 = R|E_2 = T) = (0.2 * 0.9)/(0.8 * 0.5 + 0.2 * 0.9) = 0.18/0.58$
$P(X_2 = U|E_2 = \neg T) = (0.8 * 0.5)/(0.8 * 0.5 + 0.2 * 0.1) = 0.4/0.42$
$P(X_2 = R|E_2 = \neg T) = (0.2 * 0.1)/(0.8 * 0.5 + 0.2 * 0.1) = 0.02/0.42$

4. (12 points) **Probabilities**

Consider a distributed computer system of three nodes. Each node may be up or down at any given time, denoted U for Up and D for Down. Their joint probability distribution is given below. For all parts of this question, you should try to combine like terms but you do not need to simplify your fractions.

N1	**N2**	**N3**	**P(N1,N2,N3)**
U	U	U	0.5
U	U	D	0.1
U	D	U	0.1
U	D	D	0.05
D	U	U	0.1
D	U	D	0.05
D	D	U	0.05
D	D	D	0.05

(a) (2 points) What is $P(N3 = U)$? Provide the individual terms involved in this probability.

$0.5 + 0.1 + 0.1 + 0.05 = 0.75$

(b) (2 points) What is $P(N3 = U \mid N2 = U)$? Provide the individual terms involved in this probability.

$$\frac{P(N3 = U, N2 = U)}{P(N2 = U)} = \frac{0.5 + 0.1}{0.5 + 0.1 + 0.1 + 0.05} = \frac{0.6}{0.75} = 0.8$$

(c) (2 points) If the system functions correctly as long as at least two nodes are up, what is the probability that the system is functioning correctly?

0.5 + 0.1 + 0.1 + 0.1 = 0.8

(d) (2 points) Are $N1$ and $N3$ independent, that is, $N1 \perp\!\!\!\perp N3$? Justify your answer.

They are not independent

$P(N1 = U) = 0.75$

$P(N3 = U) = 0.75$

$P(N1 = U, N3 = U) = 0.5 + 0.1 = \frac{6}{10}$

$P(N1 = U)P(N3 = U) = 0.75 * 0.75 = \frac{9}{16} \neq \frac{6}{10}$

(e) (2 points) Are $N1$ and $N3$ conditionally independent given $N2$? Justify your answer.

They are not conditionally independent.

$P(N1 = U|N2 = U, N3 = U) = \frac{0.5}{0.5+0.1} = 0.5/0.6 = \frac{5}{6}$

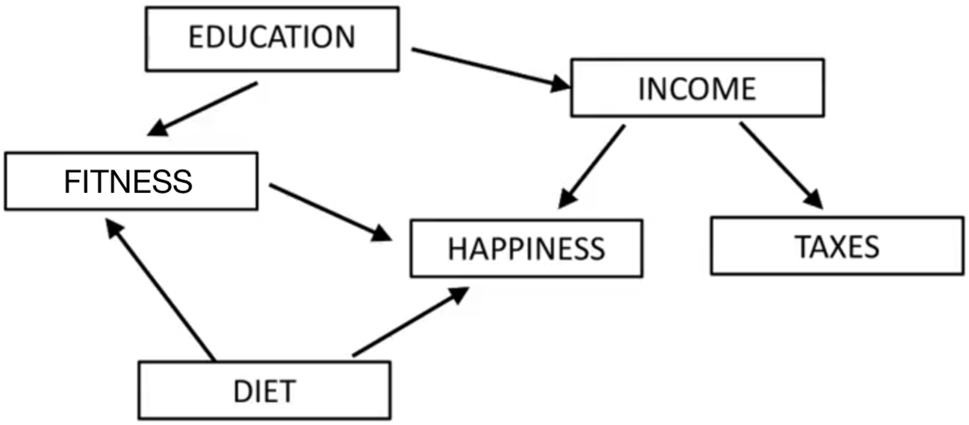
$P(N1|N2) = 0.8 = \frac{4}{5} \neq \frac{5}{6}$

(f) (2 points) Now consider a single remote node. At any moment, the probability it has a disk failure is 0.2. The probability it fails to respond to requests over the network given it has a disk failure is 1.0, and the probability it fails to respond given it does NOT have a disk failure is 0.5. Use Bayes' rule to find the probability that there is a disk failure given the node does not respond to a request. Show your math using D for "Disk failure", N for "No response", and $\neg D, \neg N$ for the opposite cases.

$$D = \text{There is a disk failure}$$
$$N = \text{Fails to respond}$$
$$P(D \mid N) = \frac{P(N \mid D)P(D)}{P(N)}$$
$$P(N) = P(N \mid D)P(D) + P(N \mid \neg D)P(\neg D)$$
$$P(N) = (1.0)(0.2) + (0.5)(0.8) = 0.6$$
$$P(D \mid N) = \frac{1.0 * 0.2}{0.6} = \frac{1}{3}$$

5. (17 points) **Bayes Net**

Consider the Bayes Net shown below. Each variable can be either high or low.



(a) (2 points each, -1 point if wrong) By independent we mean whether they are independent for any setting of the CPTs.

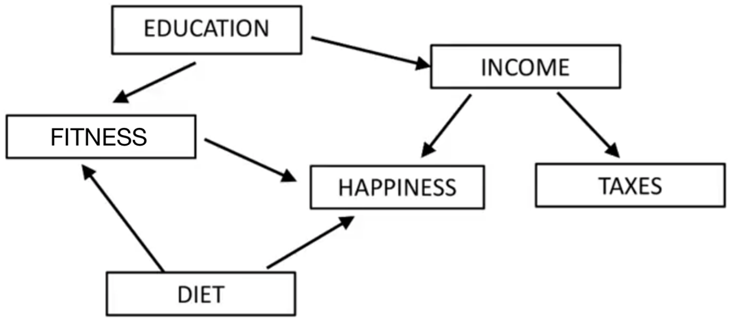
i. Are Diet and Taxes independent? .... T F

ii. Are Diet and Taxes independent given Happiness? .... T F

iii. Are Fitness and Taxes independent given Diet and Education? .... T F

iv. Are Education and Diet independent given Happiness? .... T F

i) T
ii) F ($D->H<-I->T$ active)
iii) T
iv) F ($E->I->H<-D$ active)



(b) We would like to generate samples from the joint distribution defined by this network. To do so, you can assume that you have access to a function called *sample*(). If you would like to draw a sample x_s from $P(X|Y, Z)$, the sample function can be used as follows: $x_s = sample(X|Y = y_s, Z = z_s)$. You can use the first letter of each node in your description (e.g. d is diet).

i. (3 points) Describe how to generate samples from the joint distribution using prior sampling. Name a partial sample using the following notation: $\langle \_, \_, \underline{y_s}, \_ \rangle$, where in this case we have a partial sample over four variables and only the third variable has been assigned a value y_s. In your answer, please order the variables in alphabetical order. Provide the individual steps required to generate a complete sample. **After each step, write a definition for every new variable in terms of the *sample*() function calls.**

Step 1. $< \_\_, \_\_, \_\_, \_\_, \_\_, \_\_ >$ $< d_s, \_, \_, \_, \_, \_ >$
$d_s = sample(d)$

Step 2. $< \_\_, \_\_, \_\_, \_\_, \_\_, \_\_ >$ $< d_s, e_s, \_, \_, \_, \_ >$
$e_s = sample(e)$

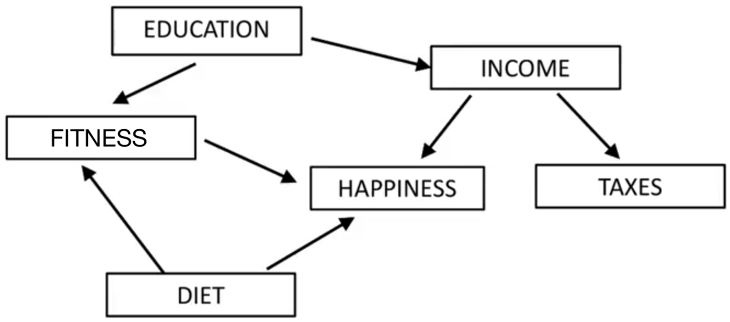
Step 3. $< \_\_, \_\_, \_\_, \_\_, \_\_, \_\_ >$ $< d_s, e_s, f_s, \_, \_, \_ >$
$f_s = sample(f|d = d_s, e = e_s)$

Step 4. $< \_\_, \_\_, \_\_, \_\_, \_\_, \_\_ >$ $< d_s, e_s, f_s, h_s, \_, \_ >$
$h_s = sample(h|d = d_s, e = e_s, f = f_s)$

Step 5. $< \_\_, \_\_, \_\_, \_\_, \_\_, \_\_ >$ $< d_s, e_s, f_s, h_s, i_s, \_ >$
$i_s = sample(i|e = e_s)$

Step 6. $< d_s, e_s, f_s, h_s, i_s, t_s >$
$t_s = sample(t|i = i_s)$

We can start from Diet, generate $d_s = sample(d)$, then generate $e_s = sample(e)$, then generate $f_s = sample(f|d = d_s, e = e_s)$, then generate $h_s = sample(h|d = d_s, e = e_s, f = f_s)$, then generate $i_s = sample(i|e = e_s)$ and finally generate $t_s = sample(t|i = i_s)$.
$d_s, e_s, f_s, h_s, i_s, t_s$ is a complete sample.



ii. (2 points) Given N samples drawn from the joint distribution, how would you determine $P(Taxes = high)$? No need to specify actual values, just provide the important quantities and their relationship.

From the N samples, X samples satisfy $T_s = high$, and N-X samples satisfy $T_s = low$. We can normalize to get $P(Taxes = high) = X/N$

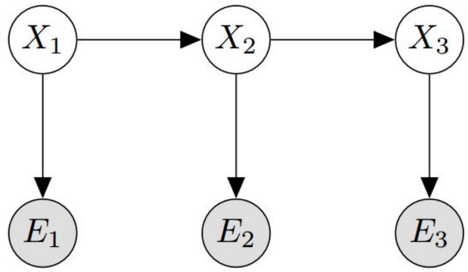
iii. (2 points) We now want to generate samples for $P(Education|Taxes = high)$. Describe in 1-2 sentences how you would do this using rejection sampling.

Do prior sampling like Q(i), and only keep samples that satisfy $T_s = high$.

iv. (2 points) Describe in 1-2 sentences how you would generate samples using likelihood weighting instead of rejection sampling. Specify how you would determine weights.

Fix the evidence variable(Taxes) and only sample the rest of variables, weighting each sample by the likelihood it matches the evidence.

6. (~~15 points~~ 9 pts) **Variable Elimination**



(a) (2 points, 1 pt each) Consider doing variable elimination on this HMM. In general, the ordering of variable elimination can greatly affect efficiency. For the query $P(X_3|e_1, e_2, e_3)$, consider the two possible orderings: (X_1, X_2) and (X_2, X_1).

You can assume all variables are binary for part (a).

i. Using the ordering (X_2, X_1), what is the size of the maximum factor generated during variable elimination (i.e., how many entries are in the table defining the factor)?

$\boxed{2^2 = 4}$

ii. Using the ordering (X_1, X_2), what is the size of the maximum factor generated during variable elimination?

$\boxed{2^1 = 2}$

(b) (~~13 pts~~ 7 pts) Using the ordering (X_1, X_2), compute $P(X_3 \mid e_1, e_2, e_3)$. You can assume that all variables are discrete. In each step, specify the variable you are eliminating, state the corresponding equation, and list all factors still under consideration.

Please use notation such as this one to define factors:
$f_2(X, y) = \sum_z f_1(X, y, Z = z)P(X|Z = z)$.

Initial factors: ______________________________________

Step 1: Eliminate the first variable by introducing the factor f_1

Which variable are you eliminating? ____________________ X_1

Equation for f_1: ______________________________________

New factors: ______________________________________

Step 2: Eliminate the second variable by introducing the factor f_2

Which variable are you eliminating? ____________________

Equation for f_2: ______________________________________

New factors: ______________________________________

Step 3: No hidden variable left, join the remaining factors

Join to get f_3: ______________________________________

Normalize to get: $P(X_3 \mid e_1, e_2, e_3)$: ____________________

1 point for each correct equation

If students correctly specifies what variable to eliminate in both part 1 and 2, they will get 1 (free) point

Step 1

$P(X_3 \mid e_1, e_2, e_3) \propto P(X_1)P(e_1|X_1)P(X_2|X_1)P(e_2|X_2)P(X_3|X_2)P(e_3|X_3)$

Eliminate X_1: $f_1(X_2, e_1) = \sum_{X_1} P(X_1)P(e_1|X_1)P(X_2|X_1)$

New factors: $f_1(X_2, e_1)P(e_2|X_2)P(X_3|X_2)P(e_3|X_3)$

Step 2

Eliminate X_2: $f_2(e_1, e_2, X_3) = \sum_{X_2} f_1(X_2, e_1)P(e_2|X_2)P(X_3|X_2)$

New factors: $f_2(X_3, e_1, e_2)P(e_3|X_3)$

Step 3

Join to get: $f_3(X_3, e_1, e_2) = f_2(X_3, e_1, e_2)P(e_3|X_3)$

Normalize: $P(X_3 \mid e_1, e_2, e_3) = \frac{f_3(X_3, e_1, e_2)}{\sum_{x_3} f_3(X_3, e_1, e_2)}$

Additional space for notes.