CSE 415 Spring 2026 Assignment 5

Last name:____________________ First name:________________ UWNetID: _______________

Due Monday night May 11 via Gradescope at 11:59 PM. You may turn in either of the following types of PDFs: (1) Scans of these pages that include your answers (handwriting is OK, if it's clear), or (2) Documents you create with the answers, saved as PDFs. When you upload to Gradescope, you'll be prompted to identify where in your document your answer to each question lies.

Do all four exercises. These are intended to take 30-60 minutes each if you know how to do them. Each is worth 25 points. If any corrections have to be made to this assignment, these will be posted in ED.

This is an *individual-work* assignment. Do not collaborate on this assignment. You may use AI to assist you with this assignment. If you do, you must explain why in your learning diary entry (LDE) for this assignment. Also, read the AI prompting guidelines in ED.

Whether you use AI or not, submit a learning diary entry (LDE) at the end of your PDF. Choose 2 of the 4 A5 questions to focus on within your learning diary entry answers. Identify those two questions in the first section of the LDE ("Goals"). If you used AI, include your reasons in this section, too. In each section (including the Goals section), use two separate paragraphs to represent your answer to that item for each A5 question you selected. Points will be deducted for missing or insubstantial answers, up to half the points for each of the two A5 questions you have chosen to focus on. If you do not use AI, then you can omit including prompts and other LLM details.

Prepare your answers in a neat, easy-to-read PDF. Our grading rubric will be set up such that when a question is not easily readable or not correctly tagged or with pages repeated or out of order, then points will be deducted. However, if all answers are clearly presented, in proper order, and tagged correctly when submitted to Gradescope, we will award a 5-point bonus. (Thus the maximum points for A5 will be 105).

If you choose to typeset your answers in Latex using the template file for this document, please put your answers in blue while leaving the original text black.

Version of 26-05-07. (Q4's sections g-j were previously missing.) Any additional updates to this document will be announced in ED.

1 Expectimax-for-MDPs

Consider the Markov Decision Process from lecture, in which an unreliable elevator is modeled (the April 24 lecture on MDPs, slides 6-10). Assume a person Jim is at level P1 and wants to determine the optimal action, assuming a discount factor of $\gamma = 0.1$. Draw an Expectimax tree that is deep enough for two actions (having 3 levels for state nodes and 2 levels for q-state nodes). Assume the leaf node state values are all 0. Compute the values at the all the other nodes of the tree, and finally identify the policy implied by these values.

(a) (6 points) Draw the tree and label each edge with either the action being taken (below a state node) or the probability and reward (below at q-state node).

(b) (2 points) Write the values of the leaf nodes next to those nodes.

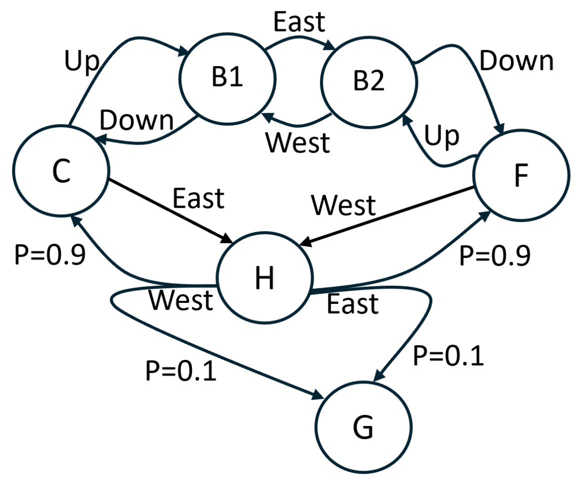
(c) (8 points) Compute the values at all other nodes

(d) (2 points) Determine a policy that maximizes the expected utility according to your answer to (c).

(e) (7 points) Explain the similarity between taking this elevator (as modeled by the MDP) and playing a game against an adversary who plays randomly, and why Expectimax is a reasonable means for determining the value of a state or an action from that state.

2 MDP Values

A bionic chicken starts in the coop (C), and wants to cross the road (actually a Highway H), to get to the Field (F), and it can either go across the Highway or via a Bridge (B1 and B2). If it goes on the Highway, there is a chance of getting run over (0.1 probability), so only a 0.9 probability of getting safely across. If it gets run over, it goes to the Graveyard (G). The situation is described by a Markov Decision Process whose state-transition diagram is shown on the right.



Assume the transitions are deterministic except with the actions East and West from the Highway. Living reward $= -1$. Arrival in F from either B2 or H has reward 20. Arrival in G has reward -100 and ends the episode. (i.e., G is a terminal state).

(a) (5 points) Determine $V_2(C)$ the expected utility at C with a time horizon of 2 assuming optimal actions and explain how the chicken should act for that. Assume no discounting.

(b) (5 points) Determine $V_3(C)$, and explain how the chicken should act for that.

(c) (5 points) Determine $V^*(C)$, and explain how the chicken should act for that.

(d) (10 points) Now assume discounting with $\gamma = 0.5$. Redo your answers to (a), (b), and (c).

3 Q-Learning

(25 points) Use the Markov Decision Process (MDP) diagram below to help you as you answer the following question Q Learning. The states are A, B, C, D, E, F, G, and an unshown Terminal state.

A	B (-10)	C	D	E	F	G (+10)

Use Q Learning to estimate the true Q values of this MDP given the observed state transitions below. For each state transition, indicate which Q value is changing by giving the state and action associated with that value and giving its new value. You may wish to keep your own diagram of the current values while you work through the given transitions.

You may assume that at states $\{A, C, D, E, F\}$ there are two allowed actions, {Left, Right}, at states $\{B, G\}$, the only allowable action is the Exit action, and at the Terminal state there are no valid actions. Make sure to process the transitions in the order provided. All Q values are initialized to 0. Use a discount factor of $\gamma = 0.8$ and a learning rate of $\alpha = 0.5$.

State Transitions				New Q Values		
State (s)	Action (a)	New State (s')	Reward (r)	State	Action	Q(S,A)
A	Right	B	-1			
B	Exit	Terminal	-10			
E	Right	D	-1			
D	Right	E	-1			
E	Right	F	-1			
F	Right	G	-1			
G	Exit	Terminal	10			

4 Joint Probability Distributions

(Updated with parts g-j on May 7). (25 points) Let C represent the proposition that it is cloudy in Seattle. Let R represent the proposition that it is raining in Seattle. Consider the table given below.

C	R	$P(C,R)$
cloudy	*rain*	0.53
cloudy	*sun*	0.13
clear	*rain*	0.02
clear	*sun*	0.32

(a) (2 points) Compute the marginal distribution $P(C)$ and express it as a table.

(b) (2 points) Similarly, compute the marginal distribution $P(R)$ and express it as a table.

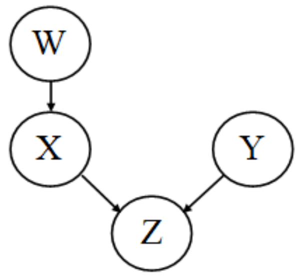
(c) (2 points) Compute the conditional distribution $P(R|C = \mathit{cloudy})$ and express it as a table. Show your work/calculations.

(d) (2 points) Compute the conditional distribution $P(C|R = \mathit{sun})$ and express it as a table. Show your work/calculations.

(e) (3 points) Is it true that $C \perp\!\!\!\perp R$? (i.e., are they statistically independent?) Explain your reasoning.

(f) (4 points) Suppose you decide to track additional weather patterns of Seattle such as temperature (hot/cold), humidity (humid/dry), and wind (windy/calm) denoted as the random variables T, W, H respectively. Is it possible to compute $P(C, R, T, W, H)$ as a product of five terms? If so, show your work. What assumptions need to be made, if any? Otherwise, explain why it is not possible.

(g) (2 points) Consider a Bayes net whose graph is shown below.



Random variable W has a domain with two values $\{w_1, w_2\}$; the domain for X has three values: $\{x_1, x_2, x_3)$; Y's domain has three values: $\{y_1, y_2, y_3\}$; and Z's domain has two values: $\{z_1, z_2\}$. Give a formula for the joint distribution of all four random variables, in terms of the marginals (e.g., $P(W = w_i)$), and conditionals that must be part of the Bayes net (e.g., $P(Z = z_m | X = X_j, Y = y_k)$).

(h) (2 point) How many probability values belong in the (full) joint distribution table for this set of random variables?

(i) (4 points) For each random variable: give the number of probability values in its marginal (for W) or conditional distribution table (for the others).

W:

X:

Y:

Z:

(j) (2 points) For each of the variables $, X, Y, Z$ give the number of "free parameters" (unforced values) in its CPT. Hint: In the marginal distribution for a binary variable, say A, if $P(A = T)$ is p, then $P(A = F)$ is forced to be $1 - p$ by the rules of probability. Therefore there can be only one free parameter in any distribution $P(A)$ for this binary variable. In a conditional distribution, there will be two or more groups of probability values that each sum to 1.