

Washington State University
School of Electrical Engineering and Computer Science
Fall 2021

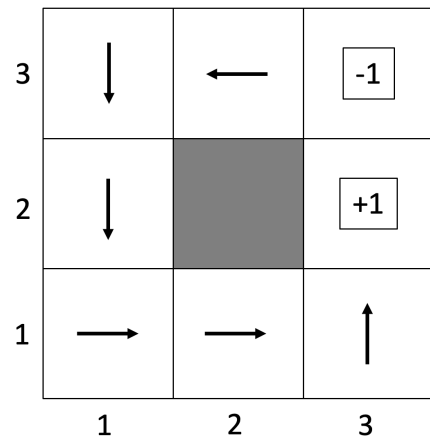
CptS 440/540 Artificial Intelligence

Homework 11

Due: December 2, 2021 (11:59pm)

General Instructions: Put your answers to the following problems into a PDF document and upload the document as your submission for Homework 11 for the course CptS 440 Pullman (all sections of CptS 440 and 540 are merged under the CptS 440 Pullman section) on the Canvas system by the above deadline. Note that you may submit multiple times, but we will only grade the most recent entry submitted before the deadline.

1. Consider a 3x3 version of the reinforcement learning task from the Learning lecture (shown on the right). Other than being a smaller world, the details are the same. The agent can execute actions: up, down, left, or right. When executing an action, the agent has an 80% chance of actually moving in that direction, a 10% chance of moving in the -90 degrees direction, and a 10% chance of moving in the +90 degrees direction. For example, if the agent executes the “right” action, then there is an 80% chance the agent moves right, a 10% chance the agent moves up, and a 10% chance the agent moves down. If the agent attempts to move into a wall or obstacle (gray area), then the agent stays in the same location. If the agent moves into location (3,2), it receives a +1 reward and the task is over. If the agent moves into location (3,3), it receives a -1 reward and the task is over. For all other actions, the agent receives a -0.04 reward.



- a. Compute the utility $U(s)$ of each non-terminal state s given the policy shown above using utility estimation, as described on slide 57 of the Learning lecture. Perform five iterations of utility updates. Each iteration should use the utilities from the previous iteration, and each utility calculation should round the utility to 2 decimal precision. The first iteration is provided below. Note that (3,2) and (3,3) are terminal states, where $U(3,2) = +1$ and $U(3,3) = -1$. You may assume $\gamma = 0.9$.

Iteration	U(1,1)	U(1,2)	U(1,3)	U(2,1)	U(2,3)	U(3,1)
0	0	0	0	0	0	0
1	-0.04	-0.04	-0.04	-0.04	-0.04	0.68
2						
3						
4						
5						

- b. Using temporal difference Q-learning, as described on slide 63 of the Learning lecture, compute the Q values for $Q(1,1,\text{right})$, $Q(2,1,\text{right})$, $Q(3,1,\text{up})$, after each of seven iterations of the action sequence: right, right, up (starting from (1,1) for each sequence). You may assume $\alpha = 0.9$, $\gamma = 0.9$, and all Q values for non-terminal states are initially zero. The first iteration is provided below. Round values to 2 decimal precision.

Iteration	$Q(1,1,\text{right})$	$Q(2,1,\text{right})$	$Q(3,1,\text{up})$
0	0	0	0
1	-0.04	-0.04	0.77
2			
3			
4			
5			
6			
7			

2. Given the following bigram model, compute the probability of the two sentences below. Show your work. Answers should be rounded to 2-decimal precision.

- a. “the wumpus smells the gold”
b. “the wumpus is dead”

Word 1	Word 2	Frequency
agent	is	5,000
is	dead	2,000
smells	the	4,000
smells	you	2,000
the	agent	6,000
the	gold	2,000
the	wumpus	8,000
wumpus	is	6,000
wumpus	smells	5,000

3. Using the lexicon and grammar on slides 23 and 24 of the Natural Language lecture, show all possible parse trees for each of the following sentences. Or indicate that there is no parse.
- “the wumpus smells the gold”
 - “the wumpus smells the gold in 2 3”
 - “the wumpus is dead”
4. *CptS 540 Students Only*: The Stanford Parser is available at nlp.stanford.edu:8080/parser/. For each of the sentences in problem 3, show the parse tree obtained by the Stanford Parser. You may just copy-and-paste the Parse result into your homework submission; no need to draw the parse tree. But make sure the indentation is preserved.
5. *Extra Credit (6 points)*. Continue updating the utilities in 1(a) until they converge. Show the final utility values for the 6 non-terminal states. *Hint*: Utilities should converge after 12 iterations.