HW10

Started: Nov 27 at 8:08pm

Quiz Instructions

While submitting numerical solutions, the accepted margin of error is 0.

Reinforcement Learning

Consider the following Markov Decision Process (MDP), describing a simple robot grid world. Here, the arrows indicate legal actions from each state, and they result in transporting the robot to the adjacent state. The value of the immediate reward $r(s, a)$ which is obtained by taking that action a from that state s is written next to the arrow. Suppose the robot is initially in state 1 and needs to find a path to reach G. Your job is to come up with the an optimal path that the robot can follow. Here, $V^*(s)$ represents value function, and discount factor $\gamma = 0.9$.

Recall that for worlds like this where actions have deterministic outcomes, $V^*(s)$ refers to the value of each state, which can be written:

$$V^*(s) = r(s, a) + \gamma V^*(s')$$

where $r(s, a)$ denotes the immediate reward received when taking action a from state s, and where action a is the optimal action (that is, $a = \pi^*(s)$ is the action recommended by the optimal policy π^*). Here, s' denotes the new state of the robot after taking action a.
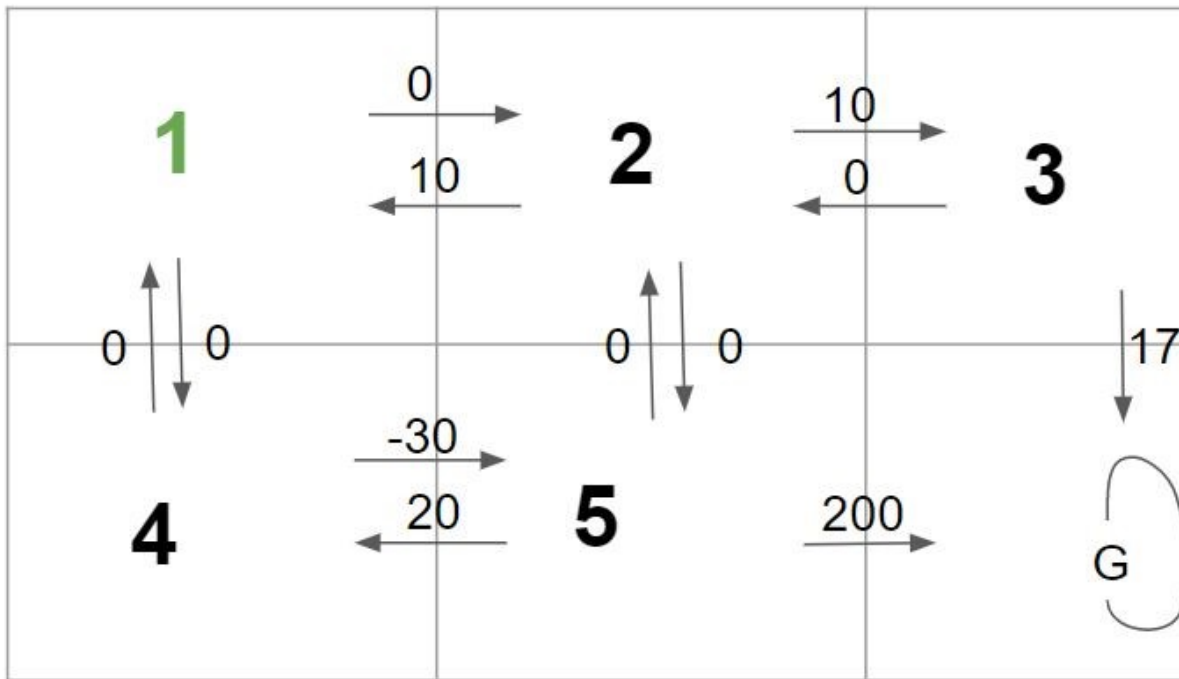
Question 1　　　　　　　　　　　　　　　　　　　　　　　　3.5 pts

What is the $V^*(s)$ value for state 5 ?

200.0000

Question 2　　　　　　　　　　　　　　　　　　　　　　　　3.5 pts

What is the $V^*(s)$ value for state 3 ?

170.0000

Question 3   3.5 pts

What is the $V^*(s)$ value for state 2 ?

> 180.0000

Question 4   3.5 pts

What is the $V^*(s)$ value for state 4 ?

> 150.0000

Question 5   3.5 pts

What is the $V^*(s)$ value for state 1 ?

> 162.0000

Question 6   3.5 pts

Which one of the following can be an optimal policy for the robot?

○ 1-4-5-G

⦿ 1-2-5-G

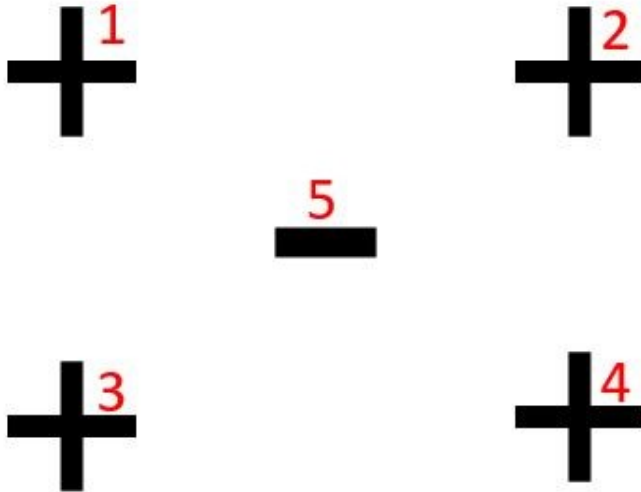○ 1-2-5-4-1-2-5-G

Boosting

Consider training a boosting classifier using decision stumps on the following data set :

Question 7   **3 pts**

Which examples will have their weights increased at the end of the first iteration?

○ 1

○ 2

○ 3

○ 4

◉ 5

Question 8

3 pts

How many iterations will it take to achieve zero training error ?

◉ at least 3

○ at least 2

○ at least 1

Question 9

3 pts

Why do we want to use weak learners when boosting ?

◉ To prevent overfitting

○ To prevent bias

○ It doesn't matter whether we choose weak learners or not

Weighted Majority Algorithm

Suppose we have results from a study conducted in Amazon Mechanical Turk (a crowdsourcing of data to make predictions). In the worker pool we have all sorts of people. We have experts who have taken Introduction to Machine Learning, some statisticians, and the rest of the people. We have all these people giving predictions for our binary problem. We would want to find a perfect expert among this crowd who never makes an incorrect prediction, but there may not be such perfect expert.Therefore we use the weighted majority algorithm to introduce weights for each person. Let us consider we

have ten such workers and a binary classification problem.

Ten Expert Predictions	Ground Truth
1,1,1,1,0,0,0,1,1,0	1
1,0,1,0,1,0,1,1,0,1	0
1,1,1,1,1,1,1,0,1,0	1
1,0,1,1,0,1,0,1,0,0	0

Consider we use a learning rate(β) of 0.5 and if a_i denotes the ith pool worker and w_i denotes the weight associated with a_i, q_0 and q_1 are the cumulative weights for decision 0 and 1. Answer these questions.

Question 10   3 pts

What will be the value of q_0 on the first example?

○ 0.4

○ 0.6

○ 3

◉ 4

○ 6

Question 11   3 pts

What would the prediction be for the first example?

◉ 1

○ 0

○ Cannot determine with the given data

Question 12

3 pts

What are the updated weight w_1 and w_{10} after training on the first example?

○ 1 and 1

○ 1 and 0

○ 0.5 and 1

◉ 1 and 0.5

Question 13

3 pts

Train the algorithm for all the given examples and report the value of
w_9 ?

○ 0.0625

○ 0.125

○ 0.25

○ 0.5

◉ 1

Question 14

3 pts

After training the Weighted Majority Algorithm on the above four training examples, the experts are given a fifth example. Their individual predictions are 1,0,1,0,0,1,0,1,1,0. What is the prediction output by the Weighted Majority Algorithm?

○ 1

◉ 0

○ Either.

Question 15

3.5 pts

We happen to know that at least one of the ten experts is perfect (they never make an incorrect prediction), but we don't which of the ten they are. What value of β should we use, to guarantee that at any point during training, the weighted majority algorithm output will be influenced only by experts that have been infallible so far?

○ 1

◉ 0

○ 0.5

Question 16

0.5 pts

Collaboration Policy Questions: Complete these question after finishing **all** parts of this assignment (canvas, gradescope, and autolab).

Did you receive any help whatsoever from anyone in solving this assignment?

- Please answer *yes* or *no*.
- If you answer *yes*, please give full details below (e.g. Tyrion Lannister explained to me what is asked in Question 2.1)

HTML Editor

B   *I*   U   A ▾   A ▾ T_x ≡   ≡   ≡   ≡   ≡ x^2 x_2 ⋮≡   ⅓≡

⊞ ▾   ▣   🔗   ⸖   🖼   √x   ▷   ❡   ¶   12pt   ▾   Paragraph   ▾

No

p

Question 17

0.5 pts

Collaboration Policy Questions: Complete these question after finishing **all** parts of this assignment (canvas, gradescope, and autolab).

Did you give any help whatsoever to anyone in solving this assignment?

- Please answer *yes* or *no*.
- If you answer *yes*, please give full details below (e.g. I pointed Jon Snow to section 2.1 since he didn't know how to proceed with Question 2).

HTML Editor

B I U̲ A ▾ A ▾ T̲ₓ ≡ ≡ ≡ ≡ ≡ x² x₂ ☰ ☷

⊞▾ ▣ 🔗 ⋰ 🖼 √x   ▷ ¶ ¶ 12pt ▾ Paragraph ▾

No

p

Question 18

0.5 pts

Collaboration Policy Questions: Complete these question after finishing **all** parts of this assignment (canvas, gradescope, and autolab).

Did you find or come across code that implements any part of this assignment?

- Please answer *yes* or *no*.
- If you answer *yes*, please give full details below (book and page, URL and location, movies and scene, etc).

HTML Editor

B *I* U̲ A ▾ A ▾ T̲ₓ ≡ ≡ ≡ ≡ ≡ x² x₂ ⋮≡ ⅟≡

⊞▾ ▣ 🔗 ✂ 🖼 √x   ▷ ¶ ¶ 12pt ▾ Paragraph ▾

No

p

No new data to save. Last checked at 6:17pm   Submit Quiz