ELEC-E8125 - Reinforcement learning D, Lecture, 4.9.2023-29.11.2023

Started on	Saturday, 9 September 2023, 3:05 PM
State	Finished
Completed on	Saturday, 9 September 2023, 3:44 PM
Time taken	38 mins 3 secs
Grade	**10.00** out of 10.00 (**100**%)

Question 1

Flag question Mark 1.00 out of 1.00 Correct

A policy is

Select one:

- a. a mapping from action to state
- b. a mapping from state to action ✔
- c. a mapping from observation to state
- d. a mapping from state to observation
- e. a mapping from state at one time to state at the next instant

Your answer is correct.
The correct answer is: a mapping from state to action

Question 2

Flag question Mark 1.00 out of 1.00 Correct

Let G_t denote discounted returns as in Sutton & Barto chapter 3.
Suppose $\gamma = 0.5$, $R_1 = 9$, $R_2 = 14$, $T = 2$.
What is G_0?

Answer: 16 ✔

The correct answer is: 16.00

Question 3

Flag question Mark 1.00 out of 1.00 Correct

"Snakes and ladders" board game, where move depends entirely on dice, exhibits Markov property.

Select one:

- True ✔
- False

The correct answer is 'True'.

Question 4

Flag question Mark 1.00 out of 1.00 Correct

Consider a 3-element linear world with value function shown in table below.

16 8 4

The agent is in the middle state. Two actions are available, corresponding to left and right motions. Immediate rewards are zero for both actions when taken in the middle state (R(middle,left)=R(middle,right)=0) and discount factor is 0.5

Which action will the agent take?

Select one:

- a. left ✔
- b. right

Your answer is correct.
The correct answer is: left

Question 5

Flag question Mark 1.00 out of 1.00 Correct

Consider a 3-element linear world with value function shown in table below.

-8 ? 12

The agent is in the middle state with unknown value. Two actions are available, corresponding to left and right motions. Immediate rewards are zero for both actions when taken in middle state (R(middle,left)=R(middle,right)=0) and discount factor is 0.5.

What is the value of the middle state if policy is uniformly random (left and right equally probable)?

Answer: 1 ✔

The correct answer is: 1.00

Question 6

Flag question Mark 1.00 out of 1.00 Correct

Consider a 3-element linear world with value function shown in table below.

4 8 16

The agent is in the middle state. Two actions are available, corresponding to left and right motions, which are always successful. Immediate rewards are zero for both actions when taken in middle state (R(middle,left)=R(middle,right)=0) and discount factor is 0.5.

An optimal value function must satisfy the Bellman optimality equation. Does the value of the middle state represent an optimal policy?

Select one:

- a. yes ✔
- b. no

Your answer is correct.
The correct answer is: yes

Question 7

Flag question Mark 1.00 out of 1.00 Correct

Value iteration

Consider a 3-element linear world with value function shown in table below.

-3 ? 8

In the middle state two actions are available U = { a, b }. Immediate rewards are zero for both actions when taken in middle state (R(middle,a)=R(middle,b)=0) and discount factor is 0.5.

Transition probabilities are p(left|middle,a) = p(right|middle,a) = 0.5, p(left|middle,b)=1, p(right|middle,b)=0.

Value iteration has been chosen as the method to find an optimal value function for the middle state. Initial value estimate for the middle state is V(middle)=0.

What is the updated value of the middle state V(middle) after one step of value iteration? Update only the value of the middle state.

Answer: 1.25 ✔

The correct answer is: 1.25

Question 8

Flag question Mark 1.00 out of 1.00 Correct

An optimal state value function defines a unique optimal policy.

Select one:

- True
- False ✔

The correct answer is 'False'.

Question 9

Flag question Mark 1.00 out of 1.00 Correct

Optimal action value function is always unique for a task/environment.

Select one:

- True ✔
- False

The correct answer is 'True'.

Question 10

Flag question Mark 1.00 out of 1.00 Correct

An optimal action value function defines a unique optimal policy.

Select one:

- True
- False ✔

The correct answer is 'False'.

Finish review

Previous activity
◄ Lecture 2: Solving discrete MDPs

Next activity
Quiz 2 ►