Started on	Friday, 6 October 2023, 7:27 PM
State	Finished
Completed on	Friday, 6 October 2023, 7:58 PM
Time taken	30 mins 19 secs
Grade	**10.00** out of 10.00 (**100**%)

Question 1

Flag question Mark 1.00 out of 1.00 Correct

Neural network approximation of action-value function is guaranteed to converge to global optimum with MC approach.

Select one:

- True
- False ✔

The correct answer is 'False'.

Question 2

Flag question Mark 1.00 out of 1.00 Correct

Neural network approximation of action-value function is guaranteed to converge to global optimum with SARSA.

Select one:

- True
- False ✔

The correct answer is 'False'.

Question 3

Flag question Mark 1.00 out of 1.00 Correct

Neural networks can be used as policy representations with policy gradient methods.

Select one:

- True ✔
- False

The correct answer is 'True'.

Question 4

Flag question Mark 1.00 out of 1.00 Correct

REINFORCE method is guaranteed to converge to a globally optimal policy.

Select one:

- True
- False ✔

The correct answer is 'False'.

Question 5

Flag question Mark 1.00 out of 1.00 Correct

True or false: Policy gradient methods can be used with continuous action spaces.

Select one:

- True ✔
- False

The correct answer is 'True'.

Question 6

Flag question Mark 1.00 out of 1.00 Correct

Consider a problem with 1-dimensional continuous state space $s \in S = [0, 100]$ and discrete action space with three actions $A = \{a_1, a_2, a_3\}$.
A policy has been defined using parametrized numerical preferences (Sutton & Barto ch. 13.1). The preference function is

$$h(s, a_1) = s$$
$$h(s, a_2) = 100 - s$$
$$h(s, a_3) = 60.$$

Actions are chosen based on soft-max.
Which of the actions may be chosen in state $s = 0$? You can choose multiple if applicable.

Select one or more:

- a. a_1 ✔
- b. a_2 ✔
- c. a_3 ✔

Your answer is correct.
The correct answers are: a_1

, a_2

, a_3

Question 7

Flag question Mark 1.00 out of 1.00 Correct

Consider the setting of the previous question.
Which of the actions is most likely for $s = 50$?

Select one:

- a. a_1
- b. a_2
- c. a_3 ✔
- d. They are equally likely.

Your answer is correct.
The correct answer is: a_3

Question 8

Flag question Mark 1.00 out of 1.00 Correct

With discrete action spaces where action-value functions are appropriate, it is always easier to approximate the action-value functions than policy functions.

Select one:

- True
- False ✔

The correct answer is 'False'.

Question 9

Flag question Mark 1.00 out of 1.00 Correct

True or false: Policy gradient methods are on-policy methods. That is, they are not applicable for off-policy learning.

Select one:

- True
- False ✔

The correct answer is 'False'.

Question 10

Flag question Mark 1.00 out of 1.00 Correct

True or false: Experience replay can be used with policy gradient approaches if properly handled.

Select one:

- True ✔
- False

The correct answer is 'True'.

Finish review

Previous activity
◄ Quiz 4

Next activity
Quiz 6 ►