1.

State-action value function Graded Quiz • 30 min

coursera



Due Mar 10, 11:59 PM

☞ Congratulations! You passed!

Grade received 100% Latest Submission Grade 100% To pass 80% or higher

Go to next item

State-action value function

Submit your assignment

Due Mar 10, 11:59 PM CET

1/1 point

Which of the following accurately describes the state-action value function Q(s,a)?

Try again

- lacktriangle It is the return if you start from state s, take action a (once), then behave optimally after that.
- \bigcirc It is the return if you start from state s and repeatedly take action a.
- It is the return if you start from state s and behave optimally.
- O It is the immediate reward if you start from state s and take action a (once).
 - Great!

Your grade

100%

View Feedback 2.

1/1 point

We keep your highest score You are controlling a robot that has 3 actions: \leftarrow (left), \rightarrow (right) and STOP. From a given state s, you have computed Q(s, \leftarrow) = -10, Q(s, \rightarrow) = -20, STOP) = 0.

☐ Report an issue

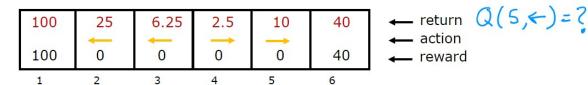
√ Dislike ے Like What is the optimal action to take in state s?

- STOP
- \bigcirc \leftarrow (left)
- \bigcirc \rightarrow (right)
- O Impossible to tell
- **⊘** Correct

Yes, because this has the greatest value.

3. 1/1 point

For this problem, $\gamma=0.25$. The diagram below shows the return and the optimal action from each state. Please compute Q(5, \leftarrow).



- 0.625
- 0.391
- 0 1.25
- 2.5
- Correct

Yes, we get 0 reward in state 5. Then 0*0.25 discounted reward in state 4, since we moved left for our action. Now we behave optimally starting from state 4 onwards. So, we move right to state 5 from state 4 and receive $0*0.25^2$ discounted reward. Finally, we move right in state 5 to state 6 to receive a discounted reward of $40*0.25^3$. Adding these together we get 0.625.

