Quiz 7

Friday, May 3, 2024 7:44 AM

1 of 10

A policy $\pi : S \rightarrow A$ gives an action for each state. If $|S| = n$, and $|A| = k$, the total number of policies is?

A n^k

B k^n

C kn

D $k(n^2)$

SUBMIT ANSWER

2 of 10

Policy π_1 is said to be better than (or equal to) policy π_2 if...

A Following policy π_1 always return better (or equal) utility values compared to following policy π_2.

B At every state, the expected utility received by following policy π_1 is better than (or equal to) the expected utility received by following policy π_2.

C The sum of rewards received by following policy π_1 is guaranteed to be more than (or equal to) the one received by following policy π_2.

D If the environment is stochastic, we can't say whether one policy is better than the other or not.

SUBMIT ANSWER

3 of 10

$$V^{\pi}(s) = \sum_{s'} P(s'|s,\pi(s))[R(s,\pi(s),s') + \gamma V^{\pi}(s')]$$

Zoom

What is the name of the V^{π} function?

A. The action-value function.

B. The state-value function.

C. The action-value function for policy π.

D. The state-value function for policy π.

SUBMIT ANSWER

4 of 10

$$V_t^{\pi}(s) \leftarrow \sum_{s'} P(s'|s,\pi(s))[R(s,\pi(s),s') + \gamma V_{t-1}^{\pi}(s')]$$

Zoom

S is the set of states. What is the complexity of iterative Policy Evaluation?

A. $O(|S|^2)$

B. $O(|S|^2)$ for each iteration

C. $O(|S|)$

D. $O(|S|)$ for each iteration

SUBMIT ANSWER

5 of 10

$$V^{\pi}(s) = \sum_{s'} P(s'|s, \pi(s))[R(s, \pi(s), s') + \gamma V^{\pi}(s')]$$

Zoom

How can we determine the expected utility received at each state by following π?

A. Solving a system of linear equations, where each equation is a state-value function for policy π when starting at a state.

B. Using iterative policy evaluation to update V^{π} until acceptable convergence.

C. Both A and B can be used.

D. None can be used.

SUBMIT ANSWER

6 of 10

$$Q^{*}(s, a) = \sum_{s'} P(s'|s, a)[R(s, a, s') + \gamma V^{*}(s')]$$

Zoom

What is the name of the Q^{*} function?

A. The optimal action-value function.

B. The optimal state-value function.

C. The state-value function.

D. The action-value function.

SUBMIT ANSWER

7 of 10

$$Q^*(s,a) = \sum_{s'} P(s'|s,a)[R(s,a,s') + \gamma V^*(s')]$$
$$\pi^*(s) = \underset{a}{\operatorname{argmax}}\, Q^*(s,a)$$
$$V^*(s) = \max_a Q^*(s,a)$$

Zoom

Can there be more than one optimal policy?

A Yes

B No

SUBMIT ANSWER

8 of 10

$$Q^*(s,a) = \sum_{s'} P(s'|s,a)[R(s,a,s') + \gamma V^*(s')]$$
$$\pi^*(s) = \underset{a}{\operatorname{argmax}}\, Q^*(s,a)$$
$$V^*(s) = \max_a Q^*(s,a)$$

Zoom

Can there be more than one optimal state-value function?

A Yes

B No

SUBMIT ANSWER

9 of 10

$$V^*(s) \leftarrow \max_a \sum_{s'} P(s'|s,a)[R(s,a,s') + \gamma V^*(s')]$$

Zoom

S is the set of states. A is the set of actions. What is the complexity of Value Iteration?

A. $O(|S| \times |A|)$

B. $O(|S| \times |A|^2)$

C. $O(|S| \times |A|^2)$ for each iteration

D. $O(|A| \times |S|^2)$ for each iteration

SUBMIT ANSWER

10 of 10

$$V^*(s) \leftarrow \max_a \sum_{s'} P(s'|s,a)[R(s,a,s') + \gamma V^*(s')]$$

Zoom

What are the direct results of Value Iteration?

A. The optimal state-value function.

B. The expected utility received at each state by following an optimal policy.

C. The action that should be taken at each state to act optimally.

D. Both A and B are correct.

SUBMIT ANSWER