

Optimality in RL

测验, 5 个问题

1
point

1。

What are the main sources of randomness in Reinforcement Learning?

- ☒ Randomness of the next state, given state and action
 - ☐ Randomness of expected return given policy and MDP
 - ☐ There is no randomness.
 - ☒ Randomness of the action given state.
 - ☒ Randomness of reward, given state and action.
-

1
point

2。

What is the definition of value function $v_{\pi}(s)$ for policy π ?

- ☐ Maximum reward, that agent can get out from the environment, starting from state s and acting according to optimal policy.
- ☐ Minimum reward, that agent can get out from the environment, starting from state s and acting according to π .
- ☐ Minimum reward, that agent can get out from the environment, starting from state s and acting according to optimal policy.
- ☐ Maximum reward, that agent can get out from the environment, starting from state s and acting according to π .
- ☐ Mean reward, that agent can get out from the environment, starting from state s and acting according to optimal policy.
- ☒ Mean reward, that agent can get out from the environment, starting from state s and acting according to π .

Optimality in RL

测验, 5 个问题

1
point

3.

What is the definition of action-value function $q_{\pi}(s, a)$ for policy π ?

- ☐ Minimum reward, that agent can get out of the environment after making action a in state s and subsequently acting according to optimal policy.
- ☐ Minimum reward, that agent can get out of the environment after making action a in state s and subsequently following current policy.
- ☒ Mean reward, that agent can get out of the environment after making action a in state s and subsequently acting according to π .
- ☐ Mean reward, that agent can get out of the environment after making action a in state s and subsequently acting according to optimal policy.
- ☐ Maximum reward, that agent can get out of the environment after making action a in state s and subsequently acting according to optimal policy.
- ☐ Maximum reward, that agent can get out of the environment after making action a in state s and subsequently acting according to π .

1
point

4.

How many deterministic optimal policies are there in a finite MDP?

- ☐ It depends on the particular MDP: there may be no optimal deterministic policies at all.
 - ☒ One or more.
 - ☐ Infinite.
 - ☐ Only one.
-

Optimality in RL

1 point

测验, 5 个问题

5.

What from the list below allow to conclude an agent follows the optimal policy π^* ? Consider each option in isolation from others.

- ☐ Provided the first state s_0 is fixed, agent plays the policy π that achieves the maximum possible $v_\pi(s_0)$ across all possible policies.
- ☐ In each state s agent makes the action that maximises the value function $v_\pi(s')$ of the next state s' for the current agent's policy π
- ☐ In each state s an agent makes the action maximising the $R(s, a, s')$
- ☐ An agent acts in the way that allow it to come to a state s that has highest $v_\pi(s)$ for the current agent's policy π
- ☒ In each state s an agent makes the action maximising the $q_\pi(s, a)$ for the current agent's policy π



我（伟臣 沈）了解提交不是我自己完成的作业 将永远不会通过此课程或导致我的 Coursera 帐号被关闭。

[了解荣誉准则的更多信息](#)

Submit Quiz

