Reward design

测验, 4 个问题

1 point

1。
What is the function of reward?

- ☑ It defines the value of each (state, action, next state) triple

- ☐ It defines the value of each (state, action) pair

- ☐ It tells agent how to achieve the stated objectives

- ☐ It defines the value of each state

- ☑ It tells agent what we want it to do

---

1 point

2。
What are the typical problems with optimization of return?

- ☑ The length of an episode can be infinite

- ☐ Reward can be discounted

- ☐ The length of the episode can be finite

- ☑ Sparse reward signal

- ☐ Negative feedback loop

- ☑ Positive feedback loop problem

---

1 point

3。

Reward design

By "correct" we mean a way that does not change the optimal policy.

测验, 4 个问题

- ☑ Scale the reward signal - divide it by 10, for example

- ☐ Reshape the rewards - add action-potential shaping function to all of the rewards.

- ☐ Standardise the rewards. That is, for example, subtract mean and divide by standard deviation.

- ☑ Reshape the rewards - add state-potential shaping function to all of the rewards

---

1 point

4。

What does the reward discounting means for the agent point of view?

- ☑ It focuses agent's attention more on close rewards and reduce the value of distance ones.

- ☐ It focuses agent's attention more on distant rewards and reduce the value of close ones.

- ☑ It reduces the variance of the return estimator by decreasing the contribution of distant rewards.

- ☐ It reduces the bias of the return estimator by increasing the contribution of close rewards.

- ✔ 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

👍 👎 🚩

Reward design

测验, 4 个问题