

# Exploration

测验, 3 个问题

1  
point

1.

Which of the following is true about regret?

- ☒ Regret estimates how quickly does a given exploration strategy converge to an optimal policy.
  - ☐ Larger regret means better exploration
  - ☒ Smaller regret means better exploration
  - ☐ A better exploration strategy will have better regret for up to any given moment of time
- 

1  
point

2.

Which of the following is true about epsilon-greedy strategy?

- ☐ An epsilon-greedy strategy will reach optimal policy if you set  $\epsilon = 1 - t / 1000$ , where  $t$  is the total number of actions taken
  - ☒ An epsilon-greedy strategy will reach optimal policy if you set  $\epsilon = 1 / t$ , where  $t$  is the total number of actions taken
  - ☒ With constant epsilon, e-greedy exploration has a linearly growing regret
  - ☐ With constant epsilon, e-greedy exploration has a constant regret
- 

1  
point

3.

Which of the following is true about uncertainty-based exploration

## Exploration

测验, 3 个问题

☐

In case of a simple multi-armed bandit, Thompson Sampling has an asymptotically smaller regret than epsilon-greedy strategy with  $\epsilon=0.5$

☐

UCB has a linear regret if the percentile is constant over time

☐

UCB works better than epsilon-greedy strategy in any decision process

☐

In some cases, epsilon-greedy strategy with  $\epsilon=0.2$  can sometimes have smaller regret than Thompson Sampling by 100-th action.



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

