Exploration

Quiz, 3 questions

1 point

1.

Which of the following is true about regret?

- [x] Regret estimates how quickly does a given exploration strategy converge to an optimal policy.
- [ ] Larger regret means better exploration
- [x] 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?

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

---

1 point

3.

Which of the following is true about uncertainty-based exploration

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

Exploration

Quiz, 3 questions

- [ ] UCB has a linear regret if the percentile is constant over time
- [ ] UCB works better than epsilon-greedy strategy in any decision process
- [x] In some cases, epsilon-greedy strategy with epsilon=0.2 can sometimes have smaller regret than Thompson Sampling by 100-th action.

- [ ] I, **Jiadai Zhao**, understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account.

Learn more about Coursera's Honor Code

Submit Quiz