Graded Quiz

Graded Assignment • 30 min

← Back

English Due Jun 11, 11:59 PM +07

Your grade: 80%

Your latest: **80%** • Your highest: **80%**

To pass you need at least 80%. We keep your highest score.

Next item →

1. Which approach ensures continual (never-ending) exploration? (**Select all that apply**) 1 / 1 point

- [x] Exploring starts

  Correct
  Correct! Exploring starts guarantee that all state-action pairs are visited an infinite number of times in the limit of an infinite number of episodes.

- [ ] On-policy learning with a **deterministic** policy
- [x] On-policy learning with an ϵ-soft policy

  Correct
  Correct! ϵ-soft policies assign non-zero probabilities to all state-action pairs.

- [x] Off-Policy learning with an ϵ-soft behavior policy and a **deterministic** target policy

  Correct
  Correct! ϵ-soft policies have non-zero probabilities for all actions in all states. The behavior policy is used to generate samples and should be exploratory.

- [ ] Off-Policy learning with an ϵ-soft target policy and a **deterministic** behavior policy

2. When can Monte Carlo methods, as defined in the course, be applied? (Select all that apply) 1 point

- [x] When the problem is **continuing** and given a batch of data containing sequences of states, actions, and rewards

  This should not be selected
  Incorrect, the full return is only available at the end of an episode. Please review Lesson 1 (Video: What is Monte Carlo?)

- [ ] When the problem is **continuing** and there is a model that produces samples of the next state and reward
- [x] When the problem is **episodic** and given a batch of data containing sample episodes (sequences of states, actions, and rewards)

  Correct
  Correct! Well-defined returns are available in episodic tasks.

- [ ] When the problem is **episodic** and there is a model that produces samples of the next state and reward

3. Which of the following learning settings are examples of off-policy learning? (Select all that apply) 1 / 1 point

- [x] Learning the optimal policy while continuing to explore

  Correct
  Correct! An off-policy method with an exploratory behavior policy can assure continual exploration.