



[Course](#) > [Bandits](#) > [Knowle...](#) > Knowle...

## Knowledge Checks

### Question 1

1.0/1.0 point (graded)

Which of the following is a fundamental challenge of reinforcement learning?

- ☐ Representation of the states of the system.
- ☐ Generalization from training to other possible states of the system.
- ☐ Temporal credit assignment to determine which actions are important in determining the outcome.
- ☐ Exploration of states or actions that are not optimal.
- ☒ All of the above.



Submit

You have used 1 of 2 attempts

### Question 2

1.0/1.0 point (graded)

Why is it important for a reinforcement agent to take both exploitation and exploratory steps?

- ☐ An agent which takes only exploratory steps gains information which guarantees and optimal exploitation of the state space.
- ☐ An agent which only takes greedy steps will converge quickly using generalized information to optimize policy.
- ☐ An agent that only take greedy steps gather generalized information and converge rapidly to an optimal policy.
- ☒ An agent which takes some exploratory steps along with greedy steps will converge to optimal an policy while acquiring information to improve the policy.

Submit

You have used 1 of 2 attempts

### Question 3

1.0/1.0 point (graded)

Which two of the following cases are "extreme" for the epsilon-greedy?

- ☒ When  $\epsilon = 0$  only greedy steps are taken and regret is minimized.
- ☒ When  $\epsilon = 1$  no greedy steps are taken and regret is maximized.
- ☐ When  $\epsilon = 0$  only greedy steps are taken and regret is maximized.
- ☐ When  $\epsilon = 1$  no greedy steps are taken and regret is minimized.



Submit

You have used 1 of 2 attempts

## Question 4

1.0/1.0 point (graded)

Which three reasons are why generalization is important in reinforcement learning?

- ☒ An agent that generalizes will be able to provide good solutions when faced with states and action options not encountered before.
- ☒ An agent that generalizes will be able to solve complex problems where the number of states is very large.
- ☒ An agent that generalizes will be able to choose actions from a large number of possibilities.
- ☐ An agent which generalizes must have acquired knowledge of all possible states.



Submit

You have used 1 of 2 attempts

## Question 5

1.0/1.0 point (graded)

Which two of the following are examples of exploration AND exploitation?

- ☒ A movie recommender agent that occasionally offers a recommendation to the user in a new category.
- ☒ An advertising display agent which occasionally displays an advertisement for a different class of goods to a user who has never shown interest in that class of goods.
- ☐ A chess playing agent which takes the move with the highest probability of winning at each move of each game.

- ☐ A agent which explores for oil by always choosing a drilling location in a distinctly different area from the last drilling location.



Submit

You have used 1 of 2 attempts

## Question 6

1.0/1.0 point (graded)

Which two of the following are true about regret?

- ☒ Regret is the expected value of the sum over all actions of the difference between optimal value and value of each action.
- ☐ Regret is never linear with time steps for an agent that both explores and exploits.
- ☐ Regret can be computed only from the action value function.
- ☒ Regret is the expected value of the sum over actions for each time step of the opportunity lost.



Submit

You have used 1 of 2 attempts

## Question 7

1.0/1.0 point (graded)

Which of the following statements about the UCB algorithm is correct?

- ☐ The UCB algorithm uniformly samples in order to reduce uncertainty and maximize reward.

- ☒ The UCB algorithm selects the action with the largest uncertainty for the reward in order to reduce that uncertainty.



- ☐ The UCB algorithm selects actions with lowest uncertainty and therefore largest expected reward to ensure rapid convergence.

- ☐ The UCB algorithm selects actions with the lowest uncertainty to minimize regret.

Submit

You have used 1 of 2 attempts

## Question 8

1.0/1.0 point (graded)

What do we mean by contextual learning and how is it different from reinforcement learning (RL)?

- ☐ A contextual bandit takes actions and receives rewards which may depend on state, whereas an RL agent takes actions regardless of state and receive rewards.

- ☐ A contextual bandit takes actions determined by state and receives rewards which may be determined by state, whereas an RL agent takes actions to change state and receive rewards.

- ☒ A contextual bandit takes actions and receives rewards which may depend on state, whereas an RL agent takes actions to change state and receive rewards.



- ☐ A contextual bandit is a type of RL agent.

Submit

You have used 1 of 2 attempts

