



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

Knowledge Checks

Question 1

1/1 point (graded)

Which of the following are two characteristics of Monte Carlo (MC) and Temporal Difference (TD) learning?

- ☒ MC methods provide an estimate of $V(s)$ only once an episode terminates, whereas TD provides an estimate of after n steps.
- ☐ MC requires to know the model of the environment i.e. the transition probabilities, whereas TD requires no such model.
- ☒ Both MC and TD are model free methods.
- ☐ Both MC and TD use bootstrapping.



Submit

You have used 1 of 2 attempts

✓ Correct (1/1 point)

Question 2

1/1 point (graded)

Which of the following are two advantages of the TD algorithm compared to the MC algorithms?

- ☐ No Bias in the estimate of value.

- ☒ Works in continuing (non-terminating) environments.
- ☐ Sensitive to initial values.
- ☐ Model free.
- ☒ Exploits the Markov Decision Process properties to gain efficiency.



Submit

You have used 1 of 2 attempts

✓ Correct (1/1 point)

Question 3

1/1 point (graded)

Which of the following statements about sampling are true for Monte Carlo (MC) and temporal difference (TD) algorithms?

- ☐ Monte Carlo algorithms randomly sample all possible state-action pairs.
- ☒ TD and MC sample states, S_{t+n} , from the current state S_t . ✓
- ☐ MC algorithms use deterministic sampling.
- ☐ TD use random sampling of state-action pairs one time step ahead only.

Submit

You have used 1 of 2 attempts

✓ Correct (1/1 point)

Question 4

1/1 point (graded)

Which two of the following describe bias-variance trade-off between MC and TD?

Which two of the following describe bias-variance trade-off between MC and TD?

- ☐ The MC algorithm reduces variance by sampling until the terminal state, leading to higher bias.
- ☒ The MC algorithm reduces bias by sampling until the terminal state, leading to higher variance.
- ☒ The TD algorithm reduces variance by sampling a small number of time steps, leading to higher bias.
- ☐ The TD algorithm reduces bias by sampling a small number of a time steps, leading to higher variance.



Submit

You have used 1 of 2 attempts

✓ Correct (1/1 point)

Question 5

1/1 point (graded)

What is the difference between on-policy and off-policy learning?

- ☐ On-policy learning learns by evaluating the results of a behavior policy to perform policy improvement on a target policy, whereas off-policy learns from experience by evaluating a target policy and performing policy improvement on the target policy.
- ☒ On-policy learning learns from experience by evaluating a target policy and performing policy improvement on the target policy, whereas off-policy learning learns by evaluating the results of a behavior policy to perform policy improvement on a target policy. ✓
- ☐ On-policy learning learns from experience by evaluating a target policy and performing policy improvement on the target policy, whereas off-policy learning learns by evaluating the target policy to perform policy improvement on a behavior policy.

- ☐ On-policy learning learns from experience by evaluating a behavior policy and performing policy improvement on the target policy, whereas off-policy learning learns by evaluating the results of a behavior policy to perform policy improvement on the behavior policy.

Submit

You have used 1 of 2 attempts

✓ Correct (1/1 point)

Question 6

1/1 point (graded)

Which two statements describe eligibility traces?

- ☒ Eligibility traces down weight the contribution of states that are rarely visited to computing average V_s or $Q(s,a)$.
- ☐ Eligibility traces encourage further exploration of the state space.
- ☐ Eligibility traces assign credit to action.
- ☒ Eligibility traces assign credit to both the most frequently visited and last visited states.



Submit

You have used 1 of 2 attempts

✓ Correct (1/1 point)

Question 7

1/1 point (graded)

Which of the following characterizes Q-Learning and SARSA?

- ☐ Q-Learning uses bootstrapping and SARSA does not.
- ☐ SARSA uses bootstrapping and Q-Learning does not.
- ☐ Both SARSA and Q-Learning are control algorithms.
- ☐ Q-Learning is an on-policy algorithm, whereas SARSA is an off-policy algorithm
- ☒ SARSA is an on-policy algorithm, whereas Q-Learning is an off-policy algorithm ✓

Submit

You have used 1 of 2 attempts

✓ Correct (1/1 point)

© All Rights Reserved