



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

## Knowledge Checks

### Question 1

1/1 point (graded)

Which of the following is a correct definition of a policy, for a Markov Decision Process (MDP)?

- ☐ The probability of taking an action given a state and the time of being in that state.
- ☒ The probability of taking an action given a state, independent of the time of being in that state. ✓
- ☐ The probability of transitioning to a state from the current state, independent of the time of being in that state.
- ☐ The probability that a state transition will trigger an action, independent of the time of that action.

Submit

You have used 1 of 2 attempts

✓ Correct (1/1 point)

### Question 2

1/1 point (graded)

How can you best describe the Bellman Equations for a Markov Reward Process (MRP)?

- ☐ The value of a state is the reward from that state plus the sum over the product of transition probabilities for the next  $n$  states.

- ☒ The value of a state is the sum over all actions,  $a$ , given the state,  $s$  of the policy, times the sum over the product of transition probabilities from the state to the next state,  $s'$  and the reward from the state plus the discounted value of the next state. ✓

- ☐ The value of a state is the discounted sum over next the product of transition probabilities for next states.

- ☐ The value of a state is the sum over all transition probabilities from the state,  $s$ , to the next state,  $s'$ , times the sum over the product of the policy, and the reward from the state plus the discounted value of the next state.

Submit

You have used 1 of 2 attempts

✓ Correct (1/1 point)

### Question 3

1/1 point (graded)

Which two of the following statements are correct about the use of the Bellman Optimality Equation in Dynamic programming?

- ☒ The transition probabilities  $p(s', r | s, a)$  must be completely known.

- ☐ The transition probabilities  $p(s', r | s, a)$  are computed iteratively.

- ☐ The Bellman equation is solved directly for all state values.

- ☒ The Bellman equation is solved iteratively, as a series of overlapping subproblems on the value function, to find the optimal policy,  $\pi(a | s)$ .



Submit

You have used 1 of 2 attempts

✓ Correct (1/1 point)

## Question 4

1/1 point (graded)

Which two of the following are correct statements about the bootstrapping process?

- ☐ Bootstrapping uses values from states at future time steps ( $t + n$ ) to compute the value,  $v(s)$ , or action value  $q(a, s)$ .
- ☒ Bootstrapping approximates the current state-value estimate based on previously learned estimates.
- ☐ Bootstrapping uses a decay factor at each time step to ensure convergence.
- ☒ Bootstrapping approximates the current state-action value estimate based on previously learned estimates.



Submit

You have used 1 of 2 attempts

✓ Correct (1/1 point)

## Question 5

1/1 point (graded)

Which of the following is a correct statement about the policy improvement theorem?

- ☒ If greedy policy improvement does not improve the policy, then the policy is optimal. ✓
- ☐ If iterative policy evaluation does not improve the policy, then the policy is optimal.
- ☐ If greedy policy improvement does not improve the evaluation of  $v(s)$ , then the policy is optimal.
- ☐ If greedy policy improvement does not improve the probability of an action, then the policy is optimal.

Submit

You have used 1 of 2 attempts

✓ Correct (1/1 point)

## Question 6

1/1 point (graded)

Which of the following is a correct statement about the difference between policy iteration and value iteration?

- ☐ Value iteration requires convergence of the policy evaluation before policy improvement can be performed, whereas, policy iteration performs policy improvement after each sweep of evaluation.
- ☒ Policy iteration requires convergence of the policy evaluation before policy improvement can be performed, whereas, value iteration performs policy improvement after each sweep of evaluation. ✓
- ☐ Policy iteration requires convergence of the policy improvement before policy evaluation can be performed, whereas, value iteration performs policy evaluation after each sweep of policy improvement.
- ☐ Policy iteration requires an approximation of policy evaluation before policy improvement can be performed, whereas, value iteration performs policy improvement after each sweep of policy evaluation.

Submit

You have used 2 of 2 attempts

✓ Correct (1/1 point)

## Question 7

1/1 point (graded)

Which of the following are differences between synchronous and asynchronous (in-place) Dynamic Programming (DP)?

- ☐ At each iteration of synchronous DP,  $v(S)$  is updated, or backed up, in parallel for all states in one step, whereas in asynchronous DP, only states with high probability are backed-up.
- ☐ At each iteration of asynchronous DP,  $v(S)$  is updated, or backed up, one at a time in a sweep, whereas in synchronous DP, states are backed up in parallel for all states in one step.
- ☒ At each iteration of synchronous DP,  $v(S)$  is updated, or backed up, in parallel for all states in one step, whereas in asynchronous DP, states are updated individually in a sweep. ✓
- ☐ At each iteration of synchronous DP,  $v(S)$  is updated, or backed up, in a priority sequence, whereas in asynchronous DP, states are updated individually in a sweep.

Submit

You have used 1 of 2 attempts

✓ Correct (1/1 point)

© All Rights Reserved