

**CS3018 Reinforcement Learning**fast-logo

**Quiz 3**

Date: 3rd March, 2025

Q1: Consider a Markov Decision Process with three states S={s1,s2,s3} and two actions A={a1,a2}. The state transition probabilities along with the values of rewards for each transition are given in the following table:

| Current State (s) | Action (a) | Next state (s’) | Transition Probability p(s’|s,a) | Value of reward  (r) |
| --- | --- | --- | --- | --- |
| s1 | a1 | s2 | 0.8 | 5 |
| s1 | a1 | s3 | 0.2 | 2 |
| s2 | a1 | s1 | 0.6 | 3 |
| s2 | a1 | s3 | 0.4 | 4 |
| s3 | a2 | s3 | 1 | 0 |

1. Write the Bellman’s equations for the state values v(s) for each state s considering the discounting factor of 0.9.
2. Solve the system of equations to get the value of each state.

Q2: Consider a Markov Decision Process with four states S={s1,s2,s3,s4} and two actions A={a1,a2}. The state transition probabilities along with the values of rewards for each transition are given in the following table:

| Current State (s) | Action (a) | Next state (s’) | Transition Probability p(s’|s,a) | Value of reward  (r) |
| --- | --- | --- | --- | --- |
| s1 | a1 | s2 | 0.7 | 3 |
| s1 | a1 | s3 | 0.3 | 2 |
| s1 | a2 | s4 | 1 | 4 |
| s2 | a1 | s1 | 0.8 | 1 |
| s2 | a1 | s4 | 0.2 | 5 |
| s3 | a1 | s3 | 1 | 0 |
| s4 | a2 | s1 | 1 | 6 |

1. Consider equiprobable random policy for states with more than one action. Write the Bellman’s equations for the state action values q(s,a) for each state action value considering the discounting factor of 0.7.

Q3: Write an equation that maps v\*(s) on left side of equation to q\*(s,a) on the right side of the equation. Explain the equation.