

**National University of Singapore
College of Design & Engineering – ECE**

**EE4400 Data Engineering and Deep Learning
Tutorial 5 – Reinforcement Learning**

Q1. There are 4 states in a reinforcement learning (RL) task denoted by $\{s_0, s_1, s_2, s_3\}$.

An RL agent takes 3 steps from the start state s_0 at time $t=0$ and reaches the terminal state s_3 at time $t=3$ in the episode shown below and receives rewards 1, 0, 2 after each step, respectively.

$$s_0 \rightarrow s_1 \rightarrow s_2 \rightarrow s_3$$

The discount factor γ is 0.7. The learning rate α is 0.1.

- (a) Determine the discounted return G_t at each time step t .
- (b) If the initial state-values at all states other than the terminal state are 0.1, determine the state-values after this episode if the TD(0) algorithm is used.

Q2. In a different version of the RL task described in Q1, 2 actions, ‘move left’ (L) and ‘move right’ (R), are available in states s_1 and s_2 . In state s_0 , only ‘move right’ (R) is available. State s_3 is the terminal state.

$$s_0 \leftrightarrow s_1 \leftrightarrow s_2 \rightarrow s_3$$

The RL agent followed the same state trajectory as in Q1. The rewards, discount factor and learning rate are the same as in Q1.

The initial action-values at all states other than the terminal state are 0.1.

Determine the action-values after this episode if the Q-Learning algorithm is used.

Q3. Explain why a deep learning neural network is needed and how it is trained in the deep reinforcement learning method called Deep Q-Network (DQN) that can learn policies for playing Atari games at a high level of competency.