Indian Institute Of Technology Patna
Department of Computer Science and Engineering

CS-603 Reinforcement Learning
Mid Semester Examination(2023)

26th September 2023

Full Marks: 60
Duration: 2 hours

Instruction
Make reasonable assumptions as and whenever necessary. You can answer the questions in any sequence. However, answers of all the parts to any particular question should appear together. Markings will be based on the correctness and soundness of the outputs. Proper indentation and appropriate comments (if necessary) are mandatory.

1.
   a. What are the different learning paradigms? How is reinforcement learning different from other learning paradigms?
   b. Can Rock-Paper-Scissor be designed as a RL problem ?

   (3 + 2 + 2)

2.
   a. Define Markov Decision Process.
   b. A pole is attached by a free joint to a cart, which moves along a frictionless track. The pendulum is placed upright on the cart and the goal is to balance the pole by applying forces in the left and right direction on the cart.

   Design an episodic game environment and a Markov Decision Problem for the above game.

   (3 + 5)

3.
   a. Define dialogue system. What are the different components of a dialogue system?
   b. How can reinforcement learning be used to improve a dialogue system?

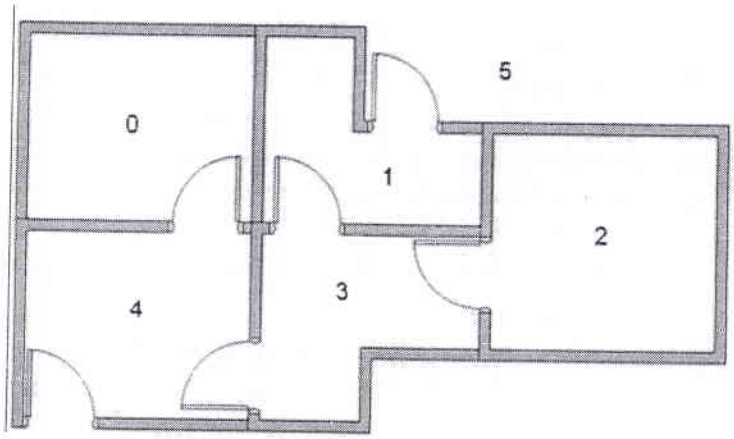
   (2 + 3 + 3)

4.
   a. Define Bellman equation.
   b. What is cumulative reward? Why is cumulative reward more important than immediate reward ?
   c. Explain the effect of adding a constant c to all the rewards in (i)Episodic environment and (ii) Continual environment

   (3 + 3 + 3)

5.

a. Consider five rooms in the apartments connected by **doors**. The goal is to leave the apartment. The outside of the apartment can be considered as one big room (marked as room no. 5) for ease of computation. Reaching outside the apartment should give an award of 100. Take epsilon = 0.5. Use tabular Q-Learning for three episodes.



(3)

b. Give an intuitive explanation of why tabular Q-learning tends to converge to a solution.

(3)

6. Explain any four ML/AI biases with examples. How can you prevent them?

(4 + 2)

7.

a. Explain the process of
   i. DQN
   ii. Double DQN
   iii. Duelling DQN

(5 + 5 + 5)

8. How does Prioritised Experience Replay Memory improve over Conventional Replay Memory?

(3)