Amrita Vishwa Vidyapeetham
Amrita School of Computing, Bangalore
Department of Computer Science and Engineering

21AIE311 - Reinforcement Learning
Lab Worksheet - 1
Tic-Tac-Toe and OpenAI Gym Exploration

1. Explore the demo of the Temporal Difference based RL learning agent to play Tic-Tac-Toe game in the below link and answer the following questions briefly to the point.

Link: https://jinglescode.github.io/2019/06/30/reinforcement-learning-value-function/

a) Agent Vs Agent - Change the hyper-parameters as below and train the agents for 10000 episodes. After each case training, make both the agents play against each other five time in simulate mode. Which player won majority of times and why?, Observe the values of the states during the interaction of the game and report it.

b) Develop a state tree and value table (initialization as discussed in class) for the tic-tac-toe game starting from black state (9 cells empty) and calculate the value for each state as per the next move choice. Show the final state tree and the corresponding value table for 10 episodes of training.

Cases	**Hyper Parameters**	**Agent-1 (O)**	**Agent-2(X)**
Case-1	Exploration Probability	0.1	0.5
	Learning Rate	0.1	0.1
Case-2	Exploration Probability	0.1	0.1
	Learning Rate	0.5	0.1
Case-3	Exploration Probability	0.5	0.25
	Learning Rate	0.5	0.25

2. Visit the Gymnasium webpage given in the link below. Gymnasium is an API standard for reinforcement learning with a diverse collection of reference environments developed by OpenAI. Answer the following questions.

Link: https://gymnasium.farama.org

a) Name all the classic environment available in Gymnasium.
b) Describe the action space, rewards, and terminal condition for every episode for the following environments
    i. Cartpole
    ii. Mountain car continuous
c) In all the environment give in Gymnasium there is an Observation Space. What is significance of it?

3. Identify the problem domain and a brief problem statement that you wish to implement as your course project. Will Gymnasium environments can be used for that? If Yes, name the environment.

Workout this experiments as a team and submit the answers within one week in the teams assignment that will be created shortly.