Multi-Agent Reinforcement Learning - Assignment 1

Name and Roll number:	**Ankur kumar(21043)**
Institute/University Name:	IISER Bhopal
Program/Stream:	DSE
Date of Submission:	August 28, 2024

1 Question 1

The Following tables show the state transition probabilities for both the possible actions in the give problem - "attend" and "hungry" :-

		To state		
	ATTEND	Hostel	Academic- Building	Canteen
From State	Hostel	0.5	0.5	
	Academic- Building	0.7		0.3
	Canteen	0.6	0.3	0.1

Table 1: State Transition probabilities for the action attend

		To state		
	HUNGRY	Hostel	Academic- Building	Canteen
From State	Hostel			1
	Academic- Building		0.2	0.8
	Canteen			1

Table 2: State Transition probabilities for the action hungry

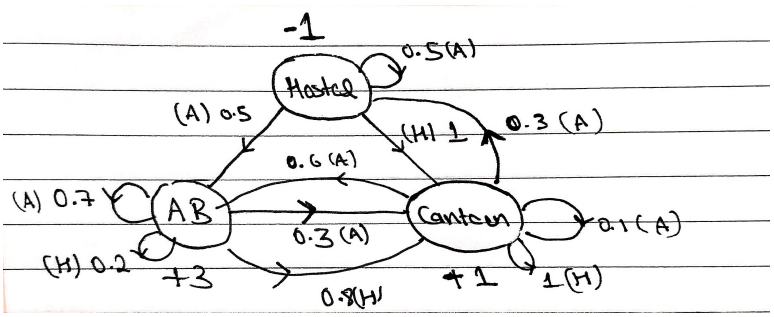

Figure 1: Diagram illustrating the MDP

- The implementations for policy iteration and value iteration can be found on the GitHub repository.
- The results of policy iteration and value iteration are identical because both algorithms converge to the correct value function by satisfying the Bellman equations. They terminate once the optimal policy is achieved. According to both these methods taking the action "attend" in all the states is the optimal policy.

2 Question 2

The implementations for policy iteration and value iteration can be found on the GitHub repository. In this question a maze was given and we were supposed to find an optimal path through the maze, this question can be modeled as an MDP and can thus be solved by methods like policy iteration and value iteration. The problem can be formulated as a Markov Decision Process (MDP) as follows:

- Each cell of the maze represents a distinct state in the MDP.
- State transitions are deterministic, meaning that each action leads to a specific next state with a probability of 1.
- The maze contains walls; any action that would lead to a state occupied by a wall results in the agent remaining in its current state.
- There is a teleportation cell in the maze; any action taken in this cell instantly transports the agent to a different, predetermined cell. This teleportation works only one way.
- There is a reward of +1 on the final state and 0 for all other state.