Assignment 1 Report

ECS 427: Multi-Agent Reinforcement Learning

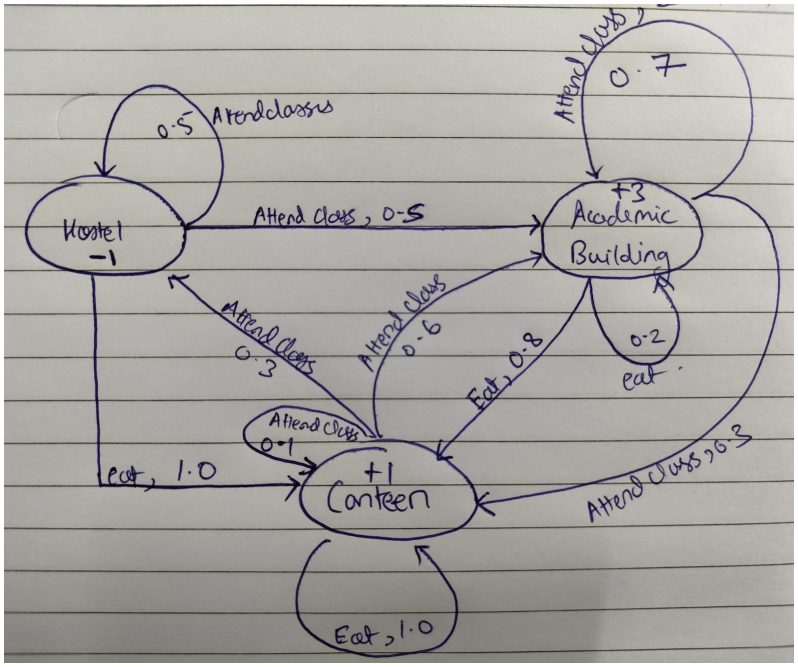
Name: Rahul Kulkarni
Roll No: 21150

Question 1

i. Design a finite MDP by writing down the possible combinations of states, actions, transition probability from one state to another for a given action, and rewards in a tabular form.

State (s)	Action (a)	Next State (s')	Transition Probability P(s'|s,a)	Rewards (s,a,s')
Hostel	Study	Hostel	0.5	-1
Hostel	Study	Academic Building	0.5	3
Academic Building	Eat	Academic Building	0.2	3
Academic Building	Study	Academic Building	0.7	3
Academic Building	Study	Canteen	0.3	1
Academic Building	Eat	Canteen	0.8	1
Canteen	Study	Canteen	0.1	1
Canteen	Study	Academic Building	0.6	3
Canteen	Eat	Canteen	1	1
Canteen	Study	Hostel	0.3	-1
Hostel	Eat	Canteen	1	1

ii. Also, draw a diagram of the MDP from the information mentioning the probability and rewards.



Discussions:

Value Iteration

After Value Iteration, the optimal value of each state:

V(Hostel) = 16.0556977

V(Academic_Building) = 21.84597336

V(Canteen) = 18.82616452

Policy Iteration

The Optimal Policy after Policy iterations:

π(Hostel) = Attend Classes

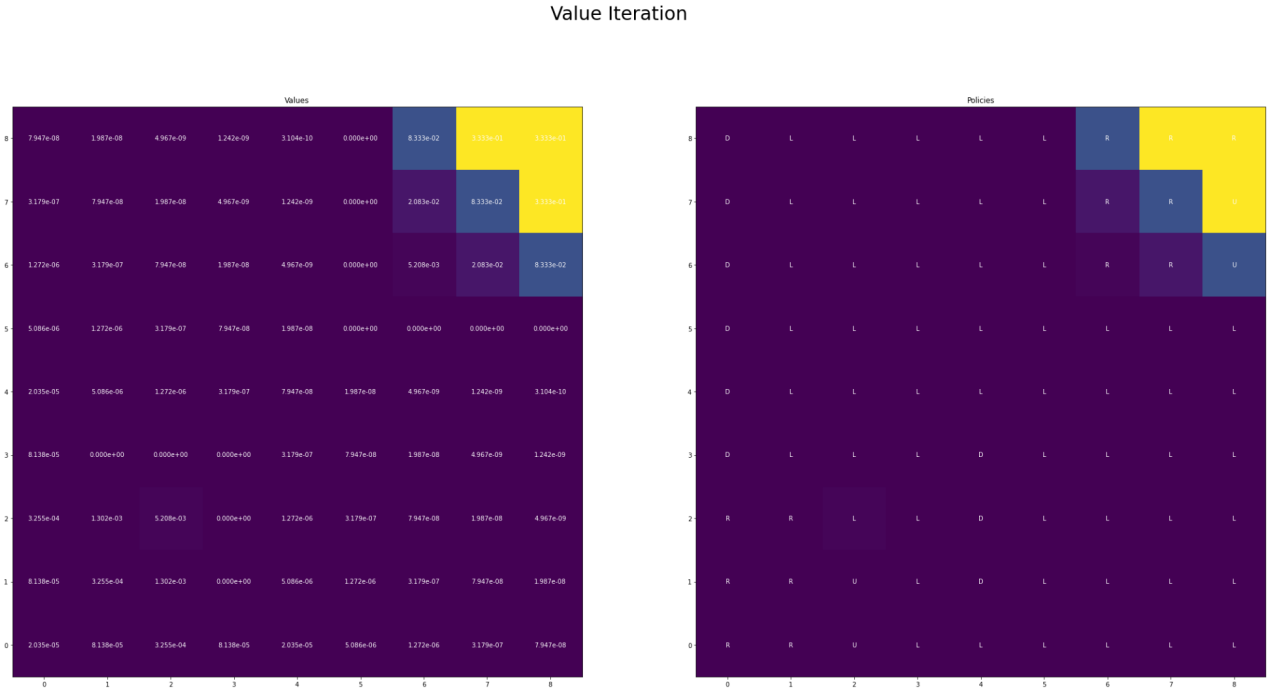
π(Academic_Building) = Attend Classes

π(Canteen) = Attend Classes

Action Set = (Attend Classes, Eat Food)

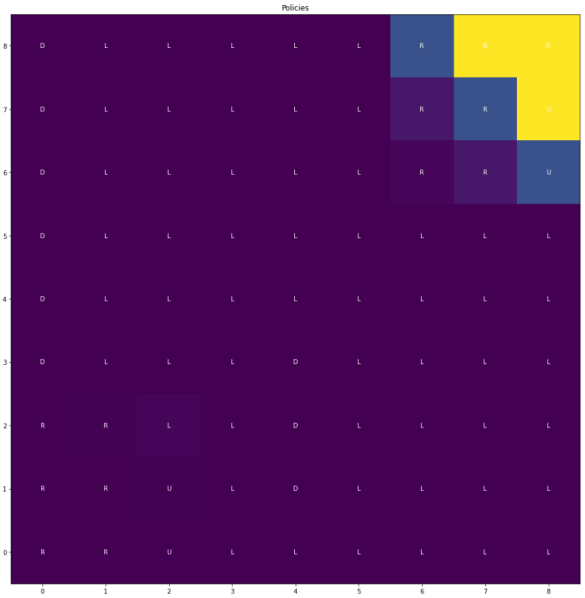
For both value iteration and policy iteration, there is convergence at gamma=0.9. The optimal policy at hostel is to attend class, at academic building is to attend class and at canteen is to attend class.

Question 2




```
Policy Actions:
U = Up
D = Down
L = Left
R = Right
NOTE:
Where ever the Value of a State is 0, it is the obstacle/wall where the robot agent can't travel
```

Policy Iteration



Policy Actions:
U = Up
D = Down
L = Left
R = Right
NOTE:
Where ever the Value of a State is 0, it is the obstacle/wall where the robot agent can't travel