

Assignment 1

Vasu Dhull

August 28, 2024

Question 1

Initial State	Action	Next State	Probability	Reward
Hostel	Attend Classes	Hostel	0.5	-1
Hostel	Attend Classes	Academic Building	0.5	3
Hostel	Hungry	Canteen	1	1
Academic Building	Attend Classes	Academic Building	0.7	3
Academic Building	Attend Classes	Canteen	0.3	1
Academic Building	Hungry	Academic Building	0.8	3
Academic Building	Hungry	Canteen	0.2	1
Canteen	Attend Classes	Academic Building	0.6	3
Canteen	Attend Classes	Hostel	0.3	-1

Table 1: Student MDP Table

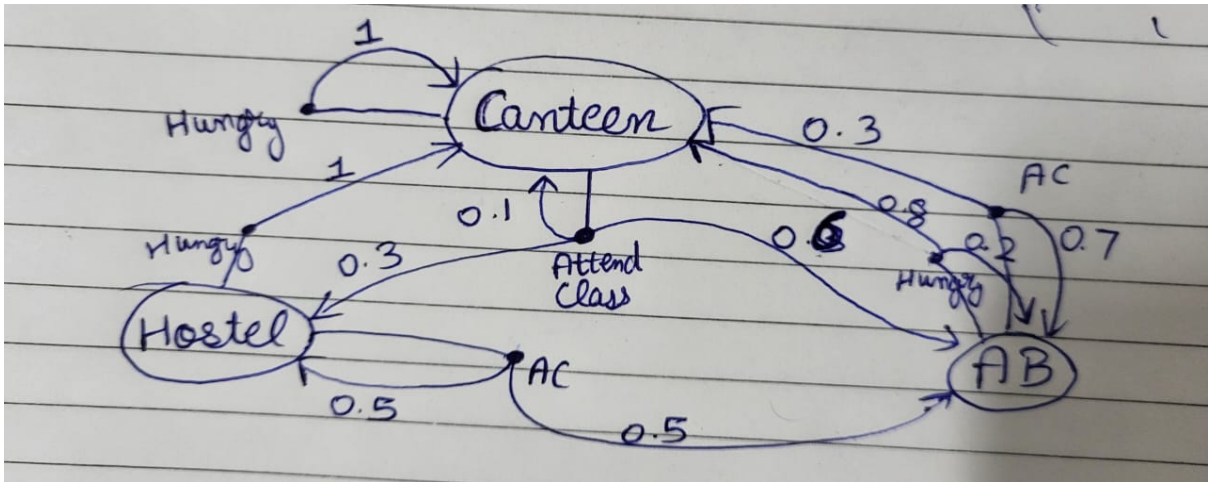


Figure 1: Graphical representation of the Student MDP

Table 1 illustrates the Markov Decision Process (MDP) of a student, while **Figure 1** visually represents this MDP through a graph.

Results

In both the Value Iteration and Policy Iteration methods, the outcomes are consistent. The optimal policy derived from the computations is as follows:

Optimal Policy: {canteen: attend classes, hostel: attend classes, academic building: attend classes}

Optimal Value Function: {canteen: 17.96, hostel: 15.19, academic building: 20.98}

Conclusion

Both Value Iteration and Policy Iteration techniques yield the same optimal policy, where the action "attend classes" is selected in all states. The value function converges similarly in both methods, with the "Academic Building" holding the highest value among all states, indicating its significance in the decision-making process.

Question 2

The solutions and results for this question are documented in the accompanying code file (Assignment-1.ipynb). All the variables are appropriately labeled, and the visualization is enhanced using AI tools. The core logic from Question 1 is reused with minor modifications to suit the requirements of Question 2.

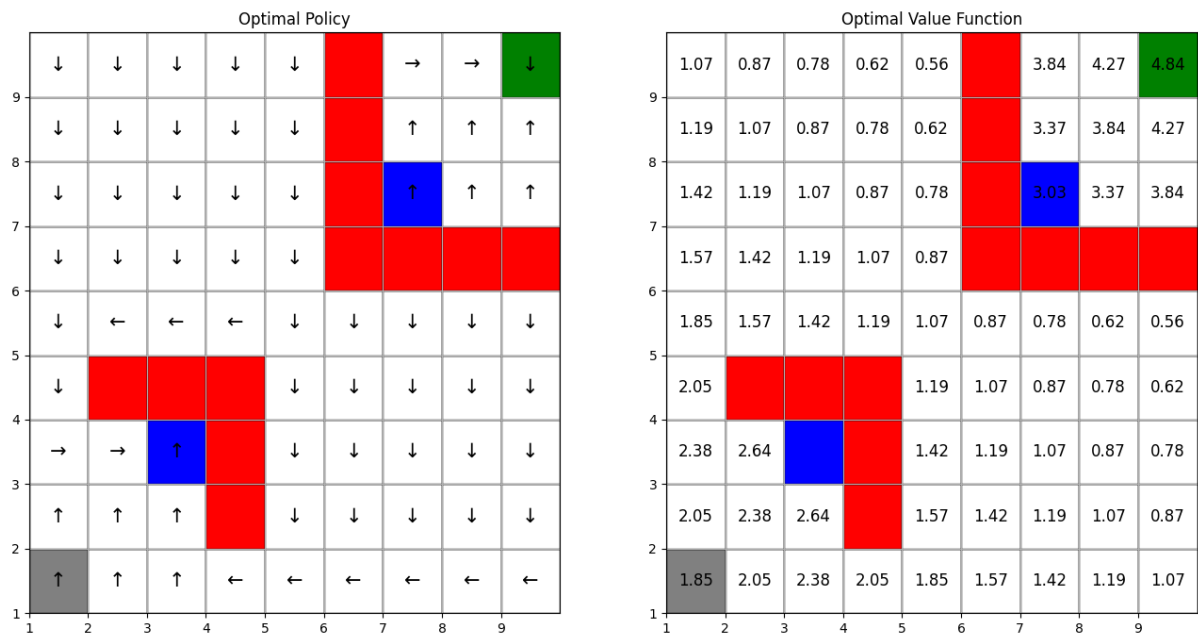


Figure 2: Both give the same output