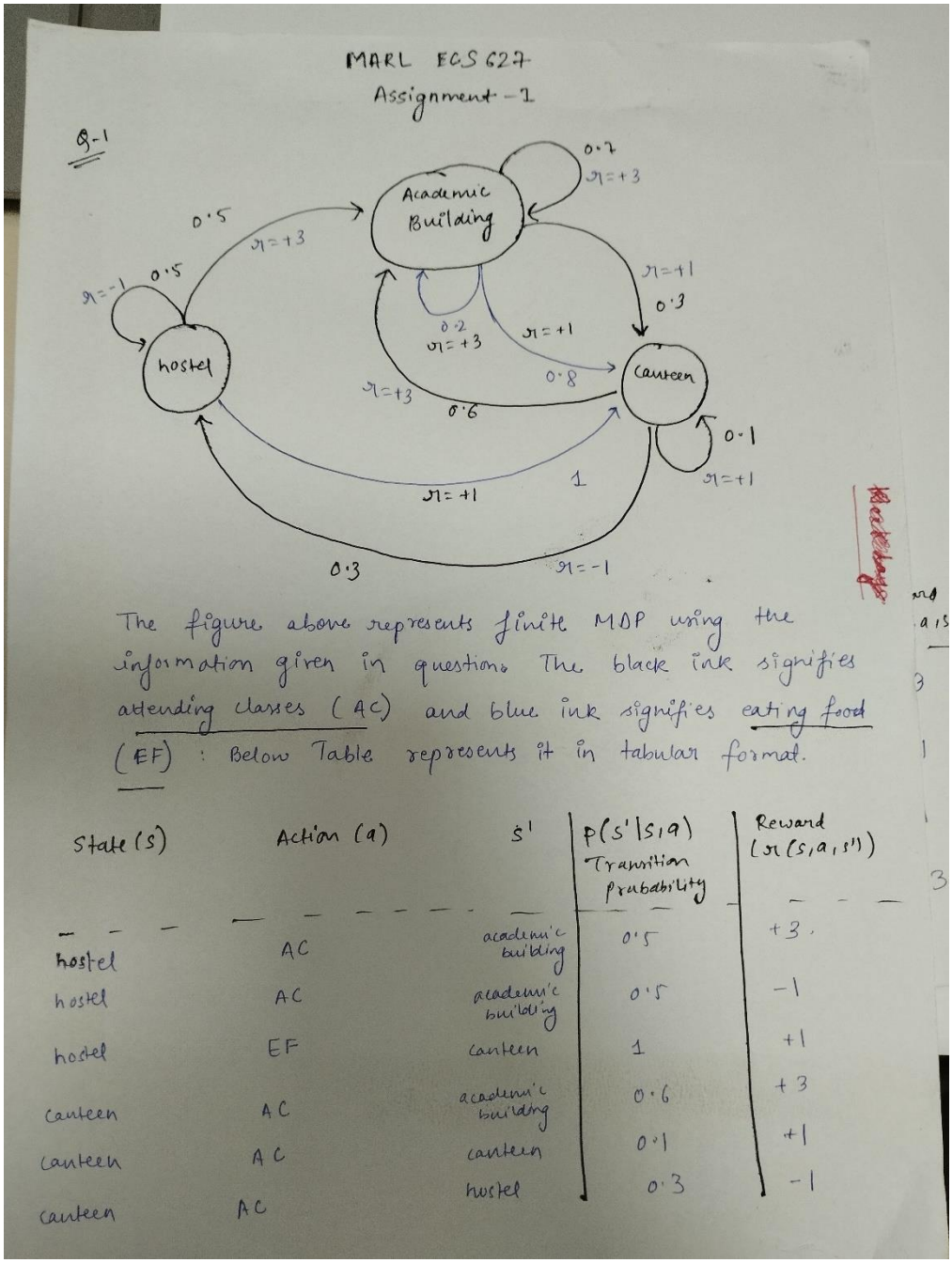
# MARL ECS 627

## Assignment 1

## Name: Shirish Shekhar Jha

## Roll Number : 2410705

Q1:



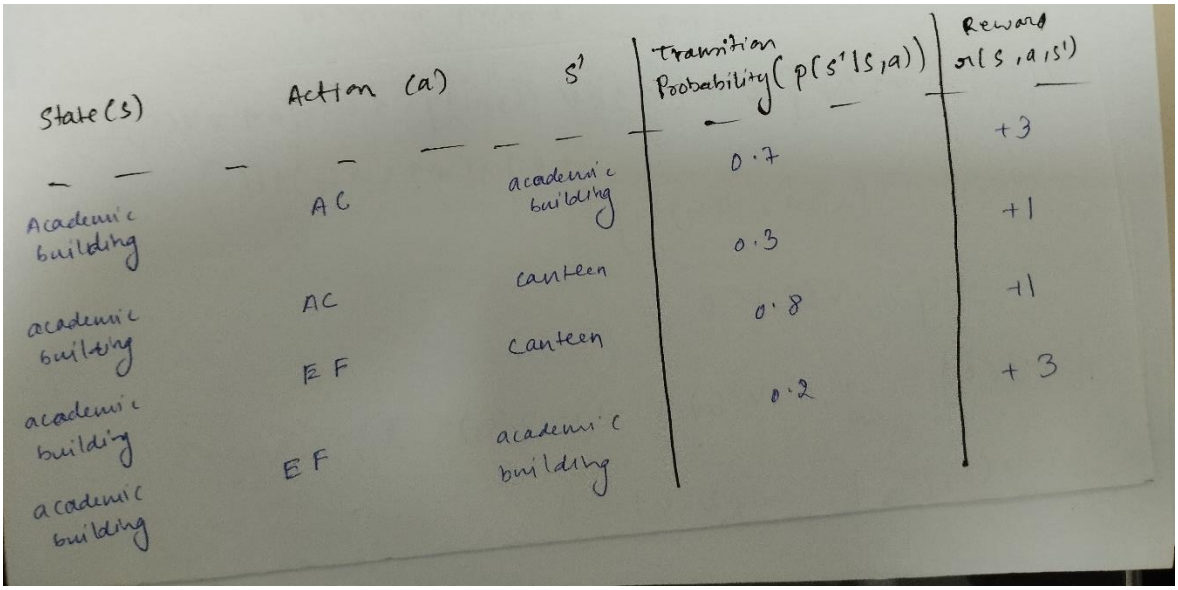
MARL ECS 627
Assignment - 1

Q-1



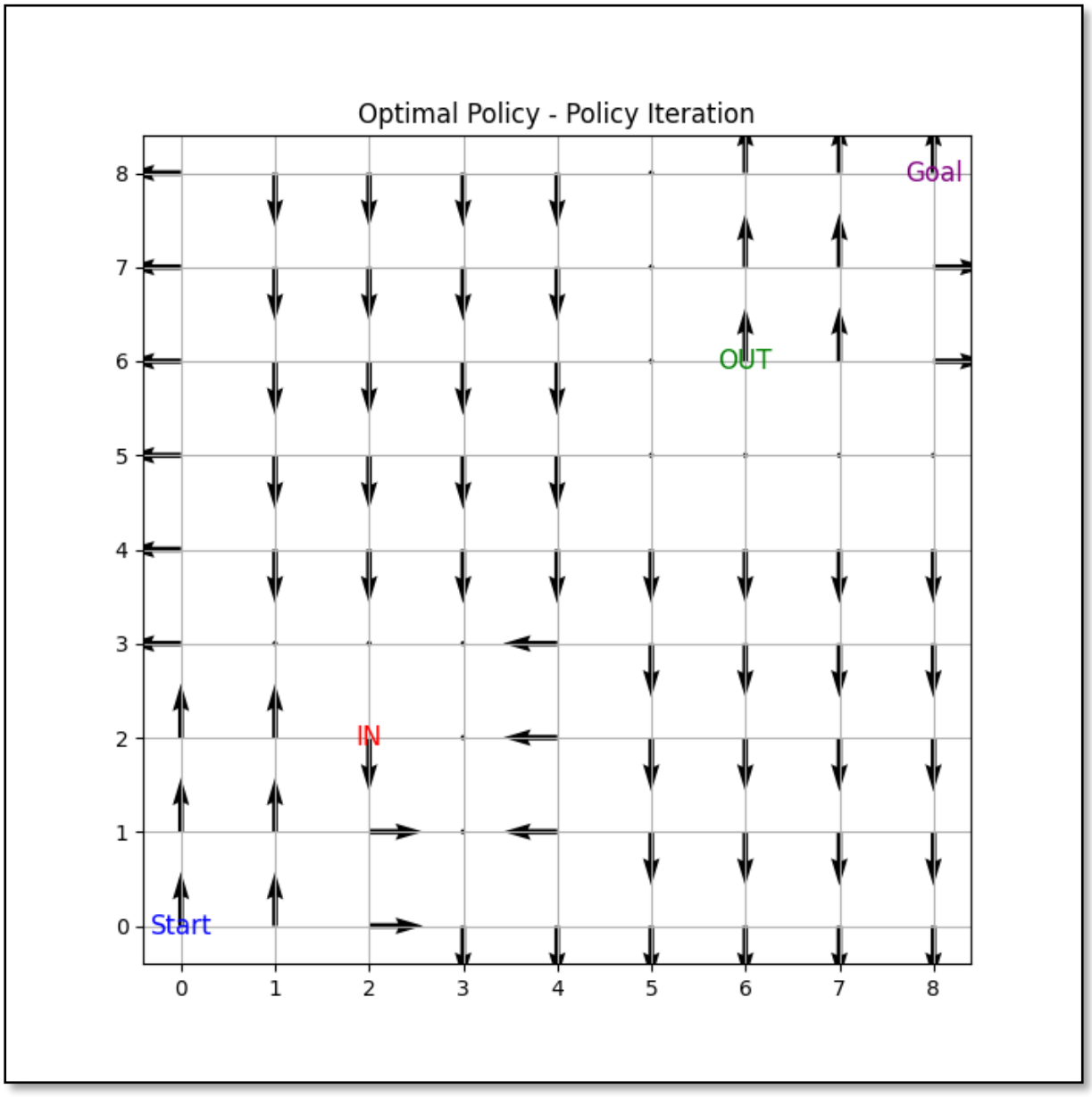
The figure above represents finite MDP using the information given in questions. The black ink signifies attending classes (AC) and blue ink signifies eating food (EF) : Below Table represents it in tabular format.

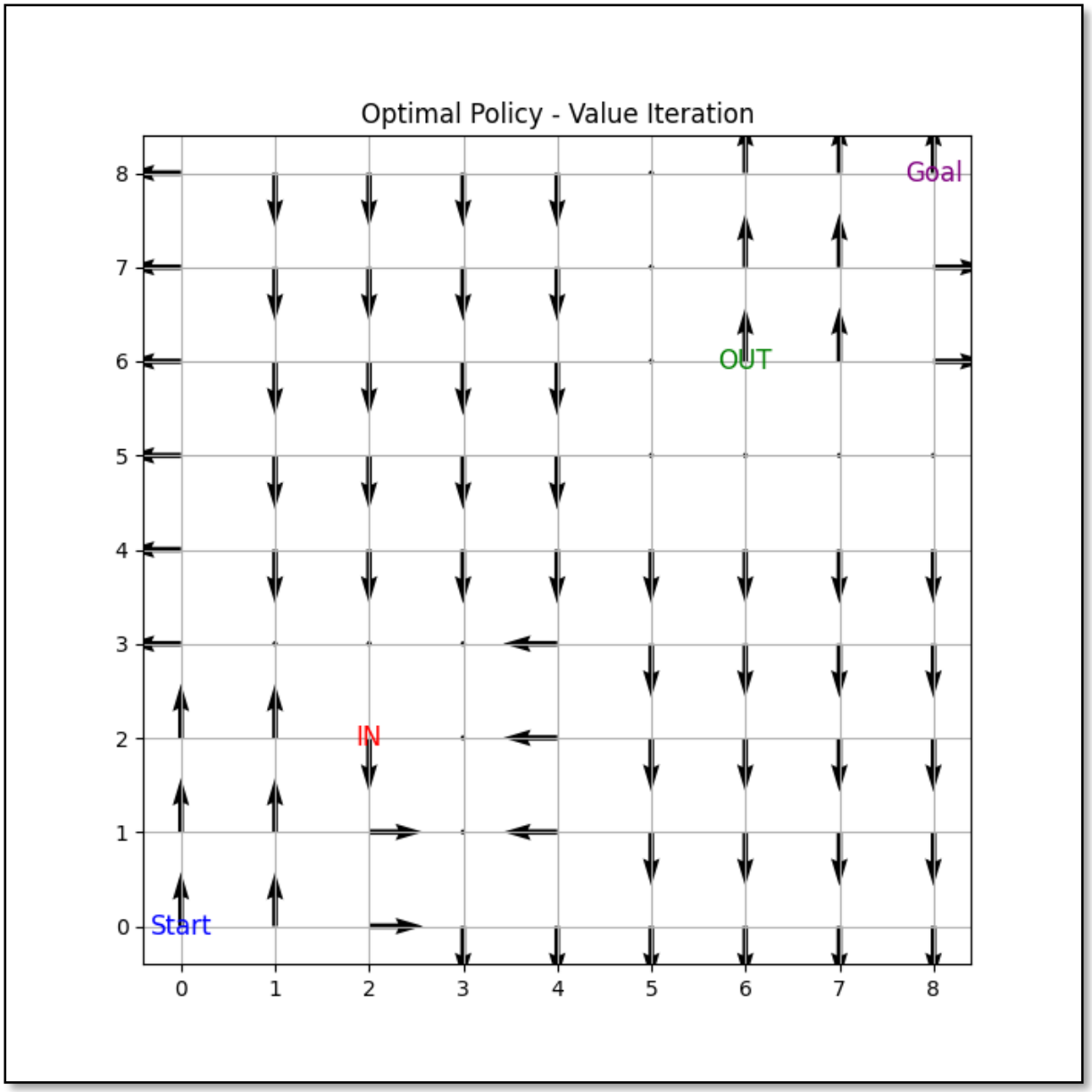
| State (s) | Action (a) | s' | p(s'\|s,a) Transition Probability | Reward (r(s,a,s')) |
|---|---|---|---|---|
| hostel | AC | academic building | 0.5 | +3 |
| hostel | AC | academic building | 0.5 | -1 |
| hostel | EF | canteen | 1 | +1 |
| canteen | AC | academic building | 0.6 | +3 |
| canteen | AC | canteen | 0.1 | +1 |
| canteen | AC | hostel | 0.3 | -1 |




Both Value Iteration and Policy Iteration suggest the same optimal policy for the student: attending class at each location (Hostel, Academic Building, Canteen). This outcome makes sense given the higher rewards associated with attending class, especially in the Academic Building where the reward is +3. The uniformity of the policy across all states suggests that, regardless of where the student is on campus, attending class maximizes long-term rewards.

Q2 :


Optimal Policy - Policy Iteration
Goal
OUT
IN
Start
0
1
2
3
4
5
6
7
8


Optimal Policy - Value Iteration
Goal
OUT
IN
Start
0
1
2
3
4
5
6
7
8