MARL - Assignment 1

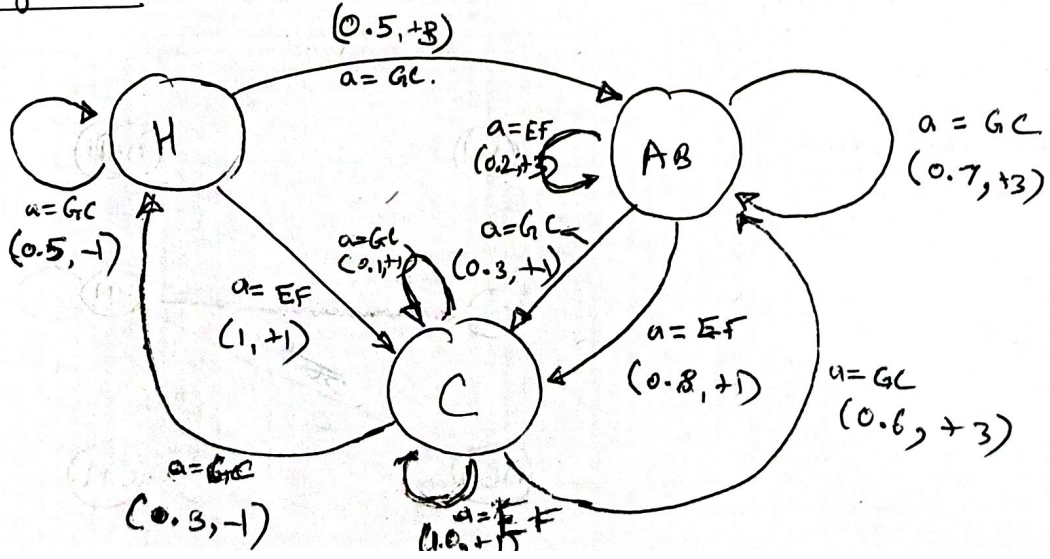
Ananya Gandhi (20319)

- States: Hostel (H)
  Acad Building (AB)
  Canteen (C)

- Actions: (a) Eat Food when Hungry (EF)
  Go to Class when not Hungry (GC)

Well, the student seems pretty ideal to me at this point though.

- MDP Diagram:



• Transition Matrix.

Current State (s)	Action. (a)	Next State. (s')	Transition. Probability. (P_T)	Rewards. (r)
H	GC.	AB	0.5	+3.
H.	GC	H	0.5	-1.
H	EF	C	1	+1
AB	GC	AB	0.7	+3
AB	EF	AB	0.2.	+3.
AB.	EF	C	0.8	+1
AB	GC	C	0.3	+1
C	GC	H	0.3	-1
C	GC	AB	0.6	+3
C	GC	C	0.1	+1.
C	~~GC~~ EF	C	1	+1.

- Value Iteration.

» We will use Bellman Optimality Equation.

We take: Discounting Factor (γ) = 0.9.

& small threshold (ϵ) = 10^{-5}. [for convergence]

- Optimal Values for Value Iteration:

Hostel = 16.056171

Acad. Building = 21.84645

Canteen = 18.82664

- Optimal ~~Values~~ Policy for Value Iteration:

'Hostel' : 'Class'

'Academic Building' : 'Class'.

'Canteen' : 'Class'

- ~~Optimal~~ Policy Iteration : $\gamma = 0.9$

- Optimal Policy for Policy Iteration:

'Hostel' : 'Class'.

'Academic Building' : 'Class'.

'Canteen' : 'Class'.