

Assignment 1. : student MRP example (10 points)

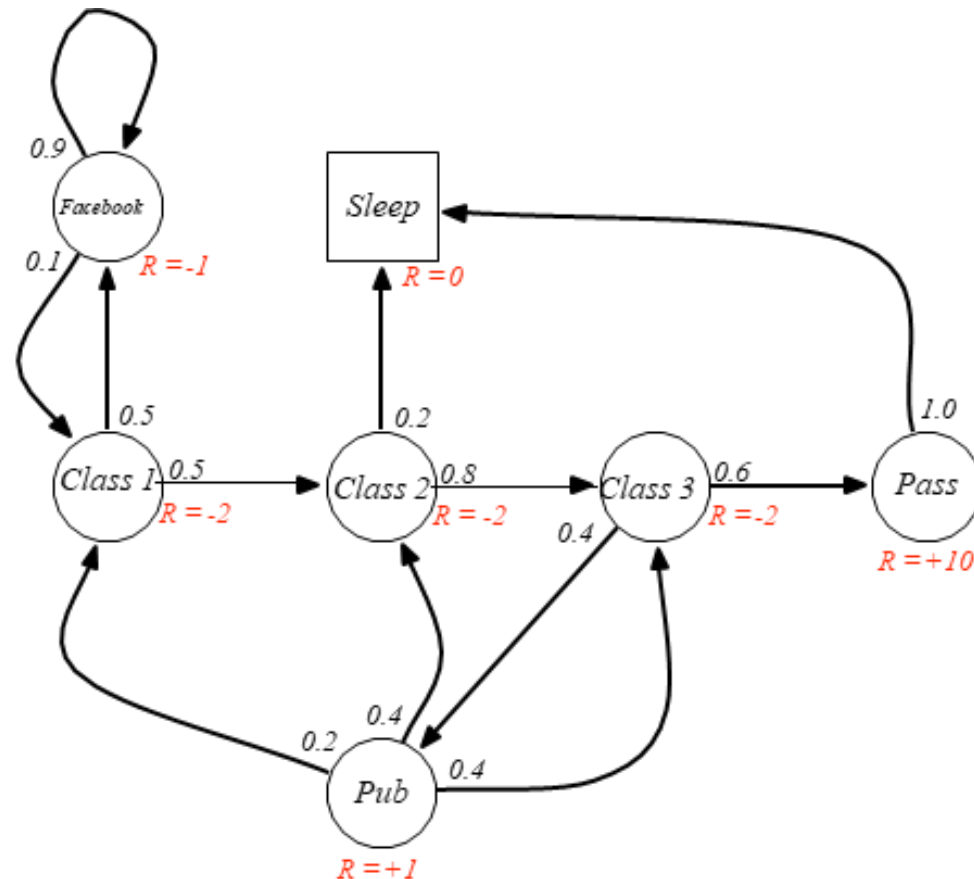
1. Evaluate State Value Evaluation ($V(s)$)

Compute the value of each state $V(s)$ under the given environment, assuming a discount factor of $\gamma = 0.85$.

2. Action Value Evaluation ($Q(s, a)$)

Calculate the action-value function $Q(s, a)$ for each state-action pair, with the same discount factor $\gamma = 0.85$. Convergence (Hint: use Bellman eq.)

3. Analysis Determine how many iterations of value iteration are required to reach the final (converged) value function.



Assignment 2. : Value iteration in random walk (10 points)

Q2.1. State Value Evaluation ($V(s)$)

Consider a random walk environment with the following settings:
There are two terminal states with rewards **100** and **50**, respectively.
All other states have zero rewards.
The policy is random: the probability of moving left or right is equal.
The discount factor is set to $\gamma = 1.0$

Q2.2. Action Value Evaluation ($Q(s, a)$)

Compute the action-value function $Q(s, a)$ for each state–action pair, with the same discount factor $\gamma = 1.0$.

State	1	2	3	4	5	6
Reward	100	0	0	0	0	50
	Terminal state					Terminal state