

CS432/632: Reinforcement Learning

Lab Assignment #3

Due Date: March 21, 2024

Instructor: Dr. Surya Prakash (E-mail: surya@iiti.ac.in)

Course TAs:

1. Prasad Kanhegaonkar (E-mail: phd2101201007@iiti.ac.in)
 2. Doli Uppal (E-mail: phd2201201001@iiti.ac.in)
 3. Prashant Digambar Pathak (E-mail: phd2301101010@iiti.ac.in)
 4. Aishwarya Priyadarshini (E-mail: ms2304101003@iiti.ac.in)
 5. Aravind Ramagiri (E-mail: phd2301101006@iiti.ac.in)
-

Iterative Policy Evaluations

Problem 1: Consider an undiscounted (that is, $\gamma=1$), episodic task in a 4 x 4 grid world with following policy and reward mechanism. The objective is to reach the terminal states (denoted by T) from any given state.

Policy: Equiprobable random policy (that is, actions UP, DOWN, LEFT, RIGHT, all have the same probability) and actions are deterministic.

Reward: $R_{s,s'}^a = -1$ for actions a and transition from states s, s'

Consider that $V(s)=0$ for all the states initially.

T	1	2	3
4	5	6	7
8	9	10	11
12	13	14	T

Problem 2: Do the above problem considering that actions are stochastic (desired action has 0.4 probability whereas the rest have 0.2 each) and the task has discounted reward collection with $\gamma=0.1$.

Submission: The programs written for above problems need to be submitted for evaluation. The submissions will be taken using a Google form and the link for the same is <https://forms.gle/k8vZ5Uc9WjSTSQ5n6>. You need to also demonstrate your code to one of the TAs with proper explanation.