Satwik Murarka
190020101

CS-747 Assignment-2

Task 1:

Value Iteration:

- Initialized the initial value of value iteration as random numbers from a uniform distribution.
- Took a tolerance of 1e-10 for comparing the values in the two iterations
- Some MDPs took longer to run for this and also failed at lower tolerance levels

Linear Programming:

- Used Pulp library to carry out the task
- Maximized the expression by taking negative sign as shown in slides

Howard Policy Iteration:

- Initialized the value of the policy with zero for every state.
- Solved the linear equations using the linalg function in NumPy

Task 2:

The problem can be formulated as an MDP by taking states as given and using the fact that one can only go to a selected number of states from a given state.
For eg. from state 1530, one can go to 1430,1429,1428,1427,1426,1425 and 1424.
Now we assign transition probabilities for each of these transfers and assign rewards accordingly.
The task is an episodic task with end states.