

CS747 Assignment 2

Roll no: 170100025

Task 1

1. Used PuLP for solving linear equations in Howard Policy Iteration
2. Chose the improvable action with highest action value in HPI

Task 2

1. 4 actions [N, E, W, S] encoded as [0,1,2,3] respectively
2. Each free grid square considered to be a potential state of the agent
3. Transition probability between any adjacent free grid squares is 1
4. We want to minimize total moves, thus reward of -1 for all valid transitions
5. Gamma of 1 as all rewards are equally important and episodic task as it is guaranteed to end.