Lab Assignment 2 : Implementation of value iteration algorithm

Consider a 4 x 4 grid as shown in Figure 1. The start state is located at the bottom left, and the target states *are* (0, 3) and (3,3) with a reward of +2 and -2. All other states have a -1 reward. Possible actions of MDP are to move left, right, up, or down. Assume the environment is stochastic. The transition probabilities are shown in Figure 1.

a. Implement value iteration algorithm to find the optimal policy. consider uniform random policy (0.25 for all four actions)
b. Check the convergence of the algorithm (how many iterations?)
c. Modify the state transition probability matrix by applying the grid transition probabilities shown in Figure 1.

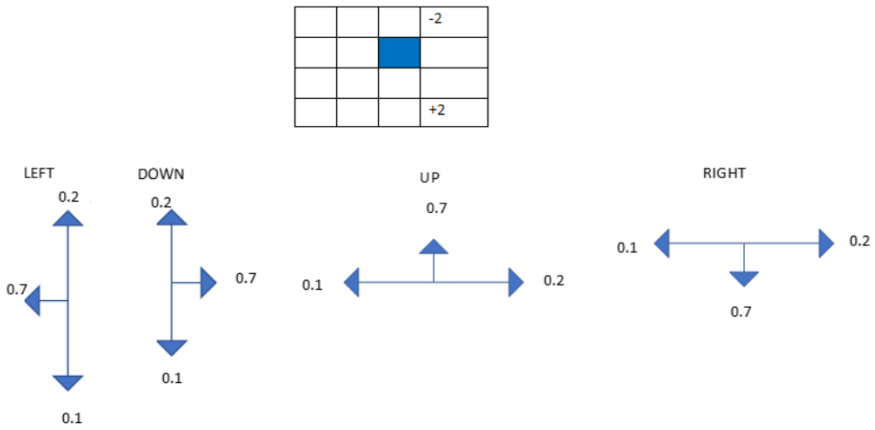

Figure 1. Grid and transition probabilities