Reinforcement Learning

Homework 1

	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	Goal

- Consider the gridworld in the figure.
- The robot agent starts from grid 1 and has to arrive the goal via the shortest path.
- The robot moves move up, down, left or right.
- The robot always has to stay in the grid. For example, when the robot agent takes action left at grid 6, it remains at grid 6.
- The agent can not visit the red cells.
- Define MDP properly for this problem, especially reward.
- Make python code implementing a policy iteration and value iteration for the gridworld in the figure.
- Your code has to generate the optimal value function and corresponding policy.