

Assignment 4: Value Iteration

In a 10x10 grid world, do 30 times value iteration.

Requirements:

1. A 10x10 grid world
2. Noise=0.15 (if you go north, 0.85 goes North, 0.075 goes to East or West, 0 goes south.), discount = 0.91
3. There are four actions west, east, north, south. The actions are all “north” in the initial policy. After each iteration, the actions in the policy need to be updated to the direction that points to the adjacent grid that has the highest value.
4. There are two terminal grids, one has reward +2 and the other one is -2. The location should be given set by using the given variables at the beginning.
5. One grid is occupied by a stone whose location is specified by the giving variables at the beginning.
6. The grid world is surrounded by walls.
7. The agent will stay at the same position if it moves towards a stone or a wall.
8. The output should contain the number of iteration statuses of the grid world in the following format. The following is one status. STONE stands for a stone. 0.00 stands for the current value. The actions of the current policy are indicated as w, e, n, and s.

0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n)

0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n)

0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),2.00(n),0.00(n),0.00(n),0.00(n)

0.00(n),0.00(n),STONE,0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n)

0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),-2.00(n),0.00(n),0.00(n),0.00(n)

0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n)

0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n)

0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n)

0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n)

0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n),0.00(n)

9. You need to write a program that use value iteration to update this table. If the iteration amount is set to twenty, you need to print out the status of the grid world after each iteration.
10. The example the end of the MDP slides could be helpful.
11. The given ipynb file must be used in this assignment. And you must follow the guideline in the ipynb file.