



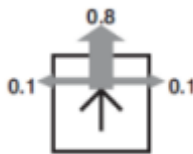
Assignment 4 Markov Decision Process

1 Problem Statement

Consider the following 3x3 world shown in the figure

r	-1	+10
-1	-1	-1
-1	-1	-1

The agent has four actions Up, Down, Right and Left. The transition model is: 80% of the time the agent goes in the direction it selects; the rest of the time it moves at right angles to the intended direction. A collision with a wall results in no movement.



2 Requirements

1. Implement value iteration for this world for each value of r below
 - $r = 100$.
 - $r = 3$
 - $r = 0$
 - $r = -3$
2. Use discounted rewards with a discount factor of 0.99
3. Show the policy obtained in each case.
4. Explain intuitively why the value of r leads to each policy.



2.1 Bonus

Find the optimal policy for each of the previous cases of r using Policy Iteration algorithm. You may start the algorithm with a randomly generated policy.

2.2 Further Notes

- You may use Java, Python or C++ for your implementation.
- Copied assignments will be severely penalized. whether from online or each other
- You can work in groups of 2 or 3.
- You will be evaluated individually in discussion

Good Luck