



Assignment 3 - Markov Decision Processes

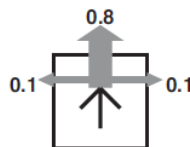
1 Problem Statement

Consider the 3x3 world shown in the following 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. (See Figure 17.1)



1.1 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.



1.2 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 Notes

- You should work on this project in teams of 3 members.
- You must submit a report showing your commented source code, results and required questions answers. You should deliver the report by Thursday 21 December.
- You must submit the report via email with the following naming convention (For the email subject and report file name):
"MDP_ID1_ID2_ID3" to the assigned course email: *ai.csed2018@gmail.com*
for example, if the team members have the following IDs: 12, 22 and 36, they MUST submit a PDF file named:
"MDP_12_22_36.pdf"
- Cheating will be heavily penalized
- You can refer to AIMA 3rd edition: Sections 17.1, 17.2 and 17.3. You might also find Berkely AI Lectures 8 and 9 useful.

Good Luck