Assignment 4
Reinforcement Learning

1 Overview

1. For this assignment you will be making a maze solver. Your program will generate a maze of size NxN. Also you should generate barriers at random grid locations. Then you will try to learn the path out of the grid using policy and value iteration.

2 Application

2.1 Algorithm

You will need to implement the NxN maze solver problem using the Policy Iteration and Value Iteration.

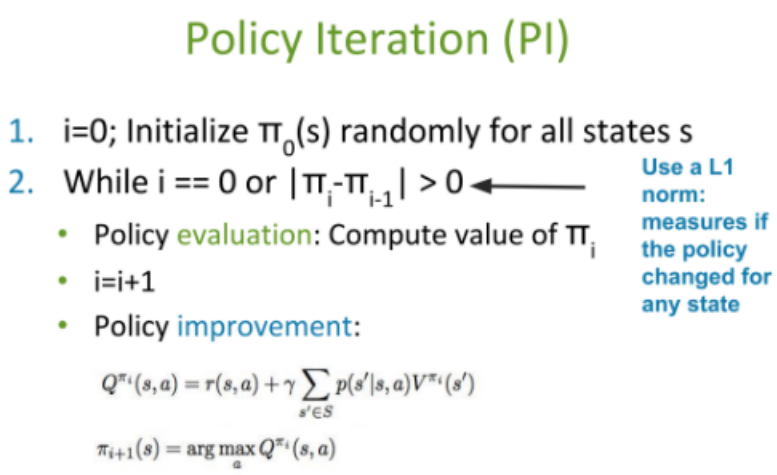
Value Iteration (VI)

1. Initialize $V_0(s)=0$ for all states s
2. Set k=1
3. Loop until [finite horizon, convergence]
   - For each state s

$$V_{k+1}(s) = \max_a R(s,a) + \gamma \sum_{s' \in S} P(s'|s,a) V_k(s')$$

   - View as Bellman backup on value function

$$V_{k+1} = BV_k$$
$$\pi_{k+1}(s) = \arg\max_a R(s,a) + \gamma \sum_{s' \in S} P(s'|s,a) V_k(s')$$



3 Notes

Output should be in a traaceable format as every step should be printed (best visualization will grant a bonus)

3.1 Deliverables

1. well commented code
2. Report should include
   (a) path to goal
   (b) cost of path
   (c) running time
3. Report sbould contain data structures used (if any) and algorithms, Assumptions and details you find necessary to be clarified, sample runs and how algorithms operate

3.2 Further Notes

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

Good Luck