CS432/632: Reinforcement Learning
Lab Assignment #4

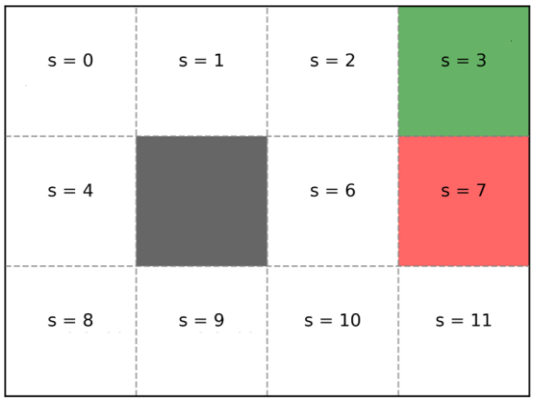
Due Date: April 4, 2024

Instructor: Dr. Surya Prakash (E-mail: surya@iiti.ac.in)

Course TAs:
1. Prasad Kanhegaonkar (E-mail: phd2101201007@iiti.ac.in)
2. Doli Uppal (E-mail: phd2201201001@iiti.ac.in)
3. Prashant Digambar Pathak (E-mail: phd2301101010@iiti.ac.in)
4. Aishwarya Priyadarshini (E-mail: ms2304101003@iiti.ac.in)
5. Aravind Ramagiri (E-mail: phd2301101006@iiti.ac.in)

Policy and Value Iteration Algorithms

Problem 1: Consider a 3 x 4 grid world as shown below where there are four actions: UP, DOWN, LEFT, RIGHT, possible at every state. Further, consider that values of all the states are initialized to zero (that is, initially, v(s)=0 for all s) and the value of $\gamma = 0.5$.



→ The grid world is surrounded by non-passable walls.
→ There are also non-passable walls inside the boundary represented by the black square.
→ The green diamond in the square at the upper right corner represents a finish line.
→ Square with a red poison shows presence of ghosts where if we step into, we lose the game and incur a lot of penalty

Considering r(s) as the reward function, the value of a state, is given as

$$v(s) = r(s) + \gamma \max_a \left[\sum_{s'} p(s'|s,a) v(s') \right]$$

Further, policy and reward are defined as given below.

Policy: Stochastic policy with the probability of actually moving in the intended direction is 0.8. There is a 0.1 probability of moving 90 degrees left to the intended direction and another 0.1 probability of moving 90 degrees right to the intended direction.

Reward: In the grid world, a normal state has a reward of -0.04, a good green ending state has a reward of +1, and a bad red ending state has a reward of -1.

Use value iteration algorithm and obtain the optimum policy.

Problem 2: Do the above problem following the policy iteration algorithm and do comparative analysis in terms of the following

- Final outcome (that is, the optimum policy obtained) and
- The time required for conversion

Submission: The programs written for above problems need to be submitted for evaluation. The submissions will be taken using a Google form and the link for the same is https://forms.gle/BGSuBru3PpTexCfX8. Yon need to also demonstrate your code to one of the TAs with proper explanation.