

# Lab Assignment 3A

CS 330: Artificial Intelligence

Instructor: Dr. Divya Padmanabhan

TAs: Tina, Umang

To be submitted on or before: September 5, 2023

*You are expected to code from scratch in Python or C or C++.*

Consider the standard gambler's problem where a gambler can make bets on the outcomes of a sequence of coin flips. If the coin comes up heads, he wins as many units as he has staked. If it is tails, he loses what he has staked. The coin comes up with heads with a probability  $p$  (which should be tunable). The game ends when the gambler has reached a  $N$  units of capital ( $N$  should again be tunable) or loses by having zero capital. On each flip the gambler has to decide the portion of his capital to stake (in integers). If he makes  $N$  units of capital, he will obtain a one-time reward of  $2N$  units and the game ends. The problem can be modeled as an MDP, where state is the gambler's capital  $s \in \{0, 1, 2, \dots, N\}$ , 0 and  $N$  are terminal states. Actions are stakes  $a \in \{0, 1, \dots, \min(s, N - s)\}$ . Every transition has a zero reward except for when the gambler reaches state  $N$  where a one time reward of  $2N$  is obtained.

1. Write a function to evaluate the value of states under any fixed policy  $\pi$  and explore how various policies you can think of perform. Report the values of policy 1:  $\pi(s) = \min(s, N - s) \forall s$  and policy 2:  $\pi(s) = 1 \forall s$ .
2. Implement value iteration algorithm to find the optimal policy. Report the optimal policy and optimal value function you got.
3. Implement policy iteration algorithm to find the optimal policy. Report the optimal policy and value of the policy you got.

You may take  $N = 10$ ,  $p = 0.4$  and  $\gamma = 0.9$  to answer the questions, however please experiment with different values of  $p$ , larger values of  $N$  (say upto  $N = 100$  at least) and  $0 < \gamma < 1$ . Also experiment with how the optimal policy might change if every transition (other than transition to state  $N$ ) incurs a reward of  $-1$ .

## Files to be submitted:

- Code files (.py or .cpp or .c). Please ensure that all files have the names and roll numbers of group members.
- Brief report with final answers to Qns 1, 2 and 3. If you have observed any interesting phenomena in your experiments, you may report them as well.