Dhruv Jain (22M0828) - Report Assignment 1

Task 1)

Design decisions:
Programming Language Used: Python 3.8
Used 2D-3D data-structure for storing MDP details.

Assumptions:
The defender will never go out of the bound or make any illegal move regardless of the policy it uses.

Observations:
hpi: 30.20s user 0.92s system 284% cpu 10.933 total
vi: 34.73s user 1.10s system 231% cpu 15.488 total
lp: 31.83s user 1.05s system 207% cpu 15.863 total
All 3 algorithms took roughly 30 secs to run auto-grader on Mac M1.

Task 2)

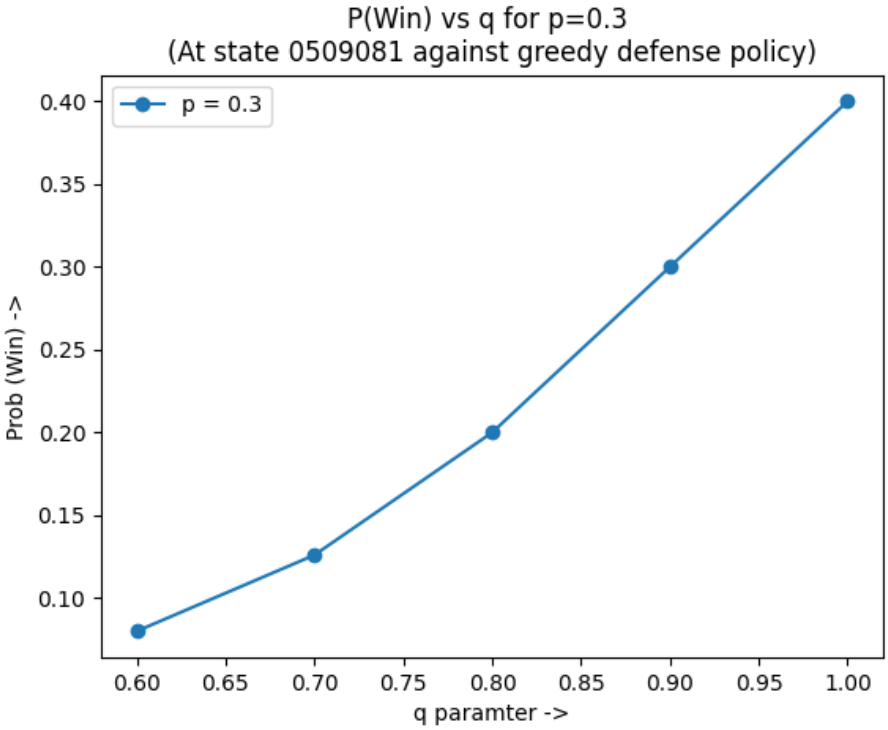
MDP Formulation:
Iterating over the states of the defender policy. For each game state configuration, iterate over all the defend actions possible, so that we move the defender first before the player as suggested by the problem statement. Once, the defender has moved. Iterate over all the 10 actions of our attacking players and for each of the action, check if the new action was valid; if not, end the scene: if yes, make a note of the transition, from the starting state to the new state of the game and mark its relevant transition probabilities. Make a new dummy state (1616163 in case) and mark it as a win state, any transitions into the win state will only have reward as 1, rest all others transitions will have 0 rewards. In this way making sure to handle all movement, tackling, passing and shooting scenarios of our players, we encode the football task2 into a mdp problem.

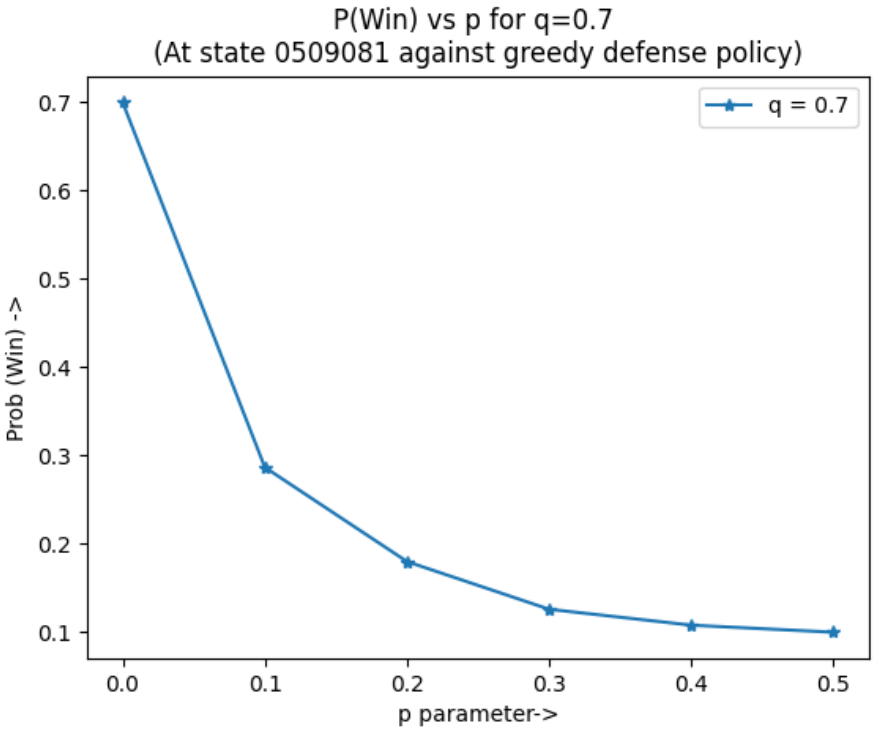
Observations:

At a fixed p, as the q value increases: the probability of winning increases as the chances of a successful pass and successful shoot increases; which is clearly evident from its success equations (Pass: q - 0.1*max(|x1-x2|, |y1-y2|) and shoot: q - 0.2*(3-x1)). So, the trend is justifying out intuitions. $P(Win) \propto q$ (at constant p)

At a fixed q, as the p value increases: the probability of winning decreases as the chances of a successful movement decreases; which is clearly evident from its success equations (Ball holder move: 1-2p, Non-Ball holder move: 1-p). So, the trend is justifying out intuitions. $P(Win) \propto 1/p$ (at constant q).

GRAPHS:





THANK YOU

Report by - Dhruv Jain (22M0828)

Date: 15th October, 2023.