CS747 Assignment 2

Yash Salunkhe - 210020156

1 Task 1

My implementations of the algorithms Value Iteration, Linear Programming and Howard's Policy Iteration are in the file `planner.py`. Tiebreaks are done randomly as I have used `np.argmax` in my code. The parameters endstate, mdptype are not required as these algorithms do not need this information to find optimal value functions. One observation was that value iteration is best(for my system) for larger MDPs and a possible reason for that is matrix inversions and operations become computationally expensive for larger sizes.

2 Task 2

The number of states in my MDP problem is 8194. The indexing done is as - 0 indicates a loss, 1 - 8192 indicates intermediate states, 8193 indicates a win. There will be 8192 states as each state is represented by the positions of my 2 players and opponent's 1 player along with possession of the ball.
The required plots are as shown:

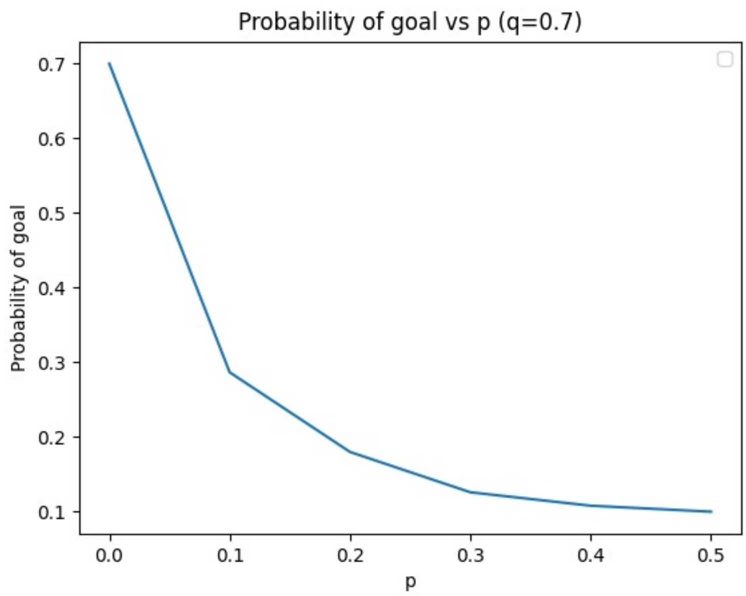

Figure 1: Expected number of goals vs variation in p

Explanation - With increase in the value of p, $1 - 2p$ decreases which is the probability a player will keep possession in the desired direction. The game ends with a probability of $2p$(for a player with the ball) and p(for a player without the ball) and hence the expected number of goals decreases with increase in p.

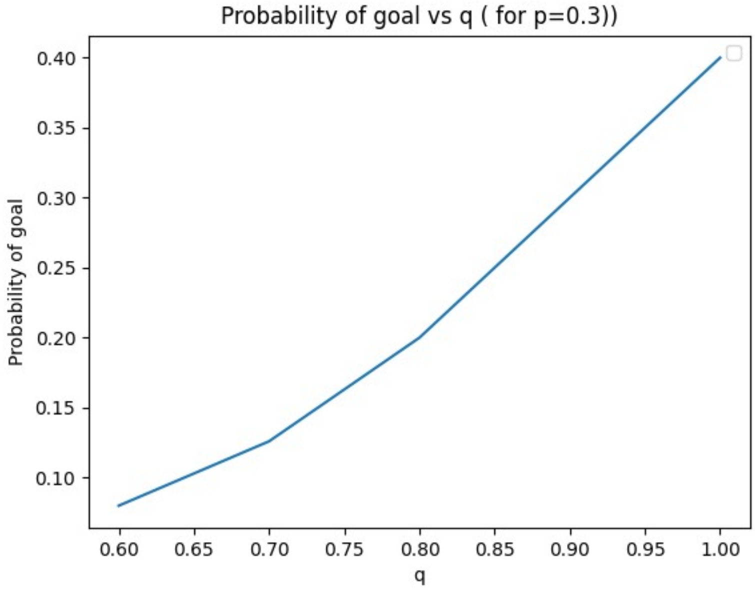

Figure 2: Expected number of goals with variation of q

Explanation - With increase in q, my probability of a successful shoot and a successful pass increases which is proportional to q and hence the expected number of goals increases as successful passes lead to high probabilities of goal scoring.
Probability of successful pass - $q - 0.1 \cdot max(|x1 - x2|, |y1 - y2|)$
Probability of a successful shoot (goal) - $q - 0.2 \cdot (3 - x1)$