F29AI

CourseWork 2: MDPs & Reinforcement Learning; Tic Tac Toe

Value Iteration

Question 2 (1 point): Test your Value Iteration Agent against each of the provided agents 50 times and report on the results – how many games they won, lost & drew against each of the other rule based agents. The rule based agents are: *random*, *aggressive*, *defensive*.

This should take the form of a very short .pdf report named: vi-agent-report.pdf. Commit this together with your code, and push to your fork.

Against Defensive Agent:

Wins: 47 Losses: 0 Draws: 3

Against Aggressive Agent:

Wins: 50 Losses: 0 Draws: 0

Against Random Agent:

Wins: 50 Losses: 0 Draws: 0

iterate():

In the iterate() method, you perform **value iteration** for k iterations. For each state, you evaluate all possible moves, compute the expected value of each move using the Bellman equation, and update the state's value to the maximum expected value.

extractPolicy():

In the extractPolicy() method, you **extract the optimal policy** by selecting the best move for each state based on the current value function. For each state, you evaluate all possible moves, calculate their expected values, and choose the move with the highest expected value.