

AI

Coursework 2, Part 1

Tic-Tac-Toe: Markov Decision Processes and Reinforcement Learning

Sam Fay-Hunt sf52@hw.ac.uk

1 Policy iteration agent

Results from 50 games played against Random, Aggressive, and Defensive agents. Using the Polity iteration agent.

This Agent used the default local reward and living reward, the convergence delta was set to 0.1.

Agent	Wins	Losses	Draws
Random	50	0	0
Aggressive	50	0	0
Defensive	44	0	6