

Policy Iteration

The gamma discount was set to 0.9 the default. It iterates until convergence due to the nature of policy iteration. The Junit test took a total of 3.468s to run, aggressive took 0.948s, defensive took 1.691s and random took 0.825s. The table below shows the result of the policy iteration agent against each of these agents.

Opponent	Wins	Losses	Draws
Aggressive	50	0	0
Defensive	40	0	10
Random	49	0	1