

# AI

## Coursework 2, Part 1

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

Sam Fay-Hunt sf52@hw.ac.uk

See Page 2 for extended tests.

#### 1 Basic Value Iteration Tests

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

This Agent had the default local reward and living reward, but performed  $k = 20$  step expectimax.

Agent	Wins	Losses	Draws
Random	50	0	0
Aggressive	50	0	0
Defensive	45	0	5

#### 2 Basic Policy Iteration Tests

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

### 3 Extended Tests (just for fun)

The next table shows 100,000 games played for each of the agents. The convergence delta for the policy iterator here is 0.0001. The value iterator has  $k = 20$ . We can see that the Value iterator agent takes almost twice as long but over 100,000 games only manages to win 61 more times.

Agent	Wins	Losses	Draws	Time to complete
Value Iterator	87633	0	12367	23h 31m
Policy Iterator	87572	0	12428	12h 42m

Using the same settings for the Policy iterator and Value iterator as in the previous sections. The table below shows a comparison of them both performing 1000 games against the defensive agent only.

Agent	Wins	Losses	Draws
Value Iterator	878	0	122
Policy Iterator	875	0	125

This table shows a comparison between  $k = 5$  and  $k = 20$  for the Value iteration agent vs the defensive agent in 100 games.

K value	Wins	Losses	Draws
$k = 5$	89	0	11
$k = 20$	90	0	10

The table below compares changing the convergence delta between 0.1, 0.0001 and 0.00001 of the policy iterator. I ran each test 3 times against the defensive agent

Run no.	Convergence delta	Wins	Losses	Draws
1	0.1	87	0	13
2	0.1	92	0	8
3	0.1	91	0	9
1	0.0001	89	0	11
2	0.0001	82	0	18
3	0.0001	87	0	13
1	0.00001	90	0	10
2	0.00001	90	0	10
3	0.00001	86	0	14

In 1000 games with a convergence delta of 0.00001, 0.1 and 0.5 respectively:

Convergence delta	Wins	Losses	Draws
0.5	881	0	119
0.1	890	0	110
0.00001	876	0	124