CMSC 474 Project 2 Report

Manu Madhu Pillai 117817928

Experimentation

Improved UCT was formed by varying the ratio of exploitation to exploration. Upon countless trials, the variation in the ratio did not improve the performance of the algorithm significantly.

Based on the game knowledge, three additional reward functions were introduced. These were:

1. Inner score reward function: This reward function encourages the player to play towards the center of the board, to reduce the chanced of flipping by the opponent without a counteraction.
2. Greedy penalty function: This penalty function discourages greedy tile flipping at the beginning of the game as it leads to reduction in available moves at subsequent turns.
3. Position penalty function: This penalty function discourages particular edge and diagonal positions next to the corners as this position allows the opponent to perform corner capturing.

Results for Tmax = 1:

UCT Ordinary vs Computer:

Player 1	Player 2	Winner	Point Difference
computer	ordinary	1	30
computer	ordinary	2	6
computer	ordinary	2	4
computer	ordinary	2	2
computer	ordinary	2	12
computer	ordinary	2	2
computer	ordinary	2	12
computer	ordinary	2	28
computer	ordinary	2	16
computer	ordinary	1	36
ordinary	computer	2	30
ordinary	computer	1	17
ordinary	computer	2	26

ordinary	computer	1	6
ordinary	computer	1	8
ordinary	computer	1	1
ordinary	computer	1	13
ordinary	computer	2	16
ordinary	computer	1	1
ordinary	computer	1	3

Win Ratio of UCT Ordinary: 75%

Random vs Computer:

Player 1	Player 2	Winner	Point Difference
computer	random	1	11
computer	random	1	32
computer	random	1	32
computer	random	1	18
computer	random	1	22
computer	random	1	24
computer	random	1	30
computer	random	1	24
computer	random	1	14
computer	random	1	36
random	computer	2	26
random	computer	2	32
random	computer	2	24
random	computer	2	26
random	computer	2	17
random	computer	2	30
random	computer	1	13
random	computer	2	40
random	computer	2	40

random	computer	2	39

Win Ratio of random: 5%

UCT Improved vs Computer:

Player 1	Player 2	Winner	Point Difference
computer	improved	2	4
computer	improved	2	8
computer	improved	2	14
computer	improved	2	4
computer	improved	2	26
computer	improved	2	20
computer	improved	2	4
computer	improved	2	22
computer	improved	2	16
computer	improved	2	12
improved	computer	2	31
improved	computer	1	7
improved	computer	1	11
improved	computer	2	26
improved	computer	1	10
improved	computer	1	3
improved	computer	1	3
improved	computer	2	26
improved	computer	1	9
improved	computer	1	19

Win Ratio of UCT Improved: 85%

UCT Improved vs UCT Ordinary:

Player 1	Player 2	Winner	Point Difference

ordinary	improved	2	4
ordinary	improved	1	27
ordinary	improved	1	2
ordinary	improved	1	38
ordinary	improved	2	4
ordinary	improved	2	12
ordinary	improved	2	4
ordinary	improved	1	4
ordinary	improved	2	6
ordinary	improved	2	19
improved	ordinary	1	9
improved	ordinary	2	28
improved	ordinary	2	24
improved	ordinary	1	3
improved	ordinary	2	30
improved	ordinary	1	15
improved	ordinary	1	13
improved	ordinary	1	3
improved	ordinary	1	19
improved	ordinary	1	1

Win Ratio of UCT Improved: 65%

Results for Tmax = 0.5:

UCT Ordinary vs Computer:

Player 1	Player 2	Winner	Point Difference
computer	ordinary	2	18
computer	ordinary	1	2
computer	ordinary	2	17

computer	ordinary	2	26
computer	ordinary	1	2
computer	ordinary	1	16
computer	ordinary	2	2
computer	ordinary	2	6
computer	ordinary	2	20
computer	ordinary	1	12
ordinary	computer	2	25
ordinary	computer	2	26
ordinary	computer	1	7
ordinary	computer	1	6
ordinary	computer	2	40
ordinary	computer	2	28
ordinary	computer	1	6
ordinary	computer	1	1
ordinary	computer	1	2
ordinary	computer	1	7

Win Ratio of UCT Ordinary: 60%

Random vs Computer:

Player 1	Player 2	Winner	Point Difference
computer	random	1	22
computer	random	1	28
computer	random	1	2
computer	random	1	29
computer	random	1	26
computer	random	1	28
computer	random	1	22
computer	random	1	24
computer	random	1	35

computer	random	1	27
random	computer	2	5
random	computer	2	39
random	computer	2	8
random	computer	2	29
random	computer	2	39
random	computer	2	43
random	computer	2	20
random	computer	2	21
random	computer	2	31
random	computer	2	34

Win Ratio of random: 0%

UCT Improved vs Computer:

Player 1	Player 2	Winner	Point Difference
computer	improved	1	12
computer	improved	2	8
computer	improved	1	10
computer	improved	1	15
computer	improved	2	2
computer	improved	1	24
computer	improved	2	6
computer	improved	2	10
computer	improved	2	4
computer	improved	2	28
improved	computer	2	21
improved	computer	1	1
improved	computer	1	12
improved	computer	1	3
improved	computer	2	22
improved	computer	1	12

improved	computer	2	12
improved	computer	2	15
improved	computer	2	22
improved	computer	2	28

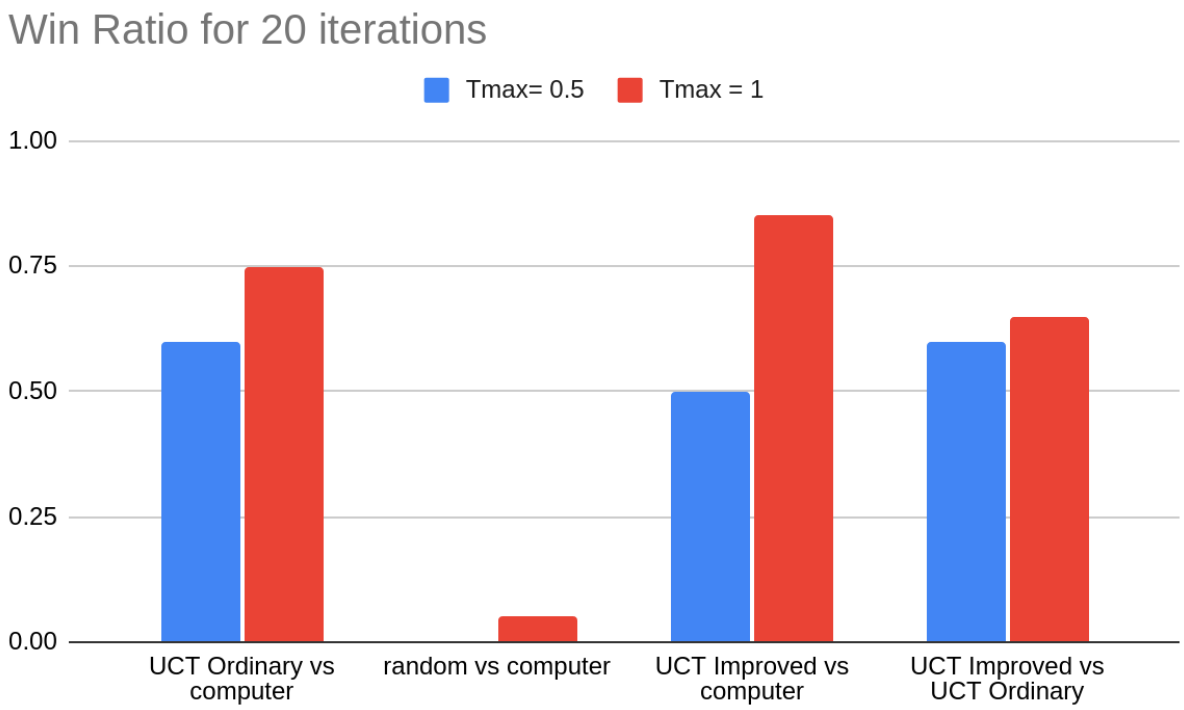
Win Ratio of UCT Improved: 50%

UCT Improved vs UCT Ordinary:

Player 1	Player 2	Winner	Point Difference
ordinary	improved	2	17
ordinary	improved	2	18
ordinary	improved	2	6
ordinary	improved	2	18
ordinary	improved	2	6
ordinary	improved	1	5
ordinary	improved	1	5
ordinary	improved	1	10
ordinary	improved	2	20
ordinary	improved	1	22
improved	ordinary	2	4
improved	ordinary	1	10
improved	ordinary	1	12
improved	ordinary	1	13
improved	ordinary	1	4
improved	ordinary	2	14
improved	ordinary	1	18
improved	ordinary	2	11
improved	ordinary	1	5
improved	ordinary	2	10

Win Ratio of UCT Improved: 60%

Inference:



While both algorithms win against computer, their win ratio is can be improved. Further improvements can be made in terms of limited depth search due to timeout constraints and tuning of the weights.

Furthermore, mobility of the move could be introduced as a reward function where mobility measures the number of available move in all the subsequent moves.

The algorithm can be further improved to perform in more time constraint scenarios by saving common states across moves, and avoiding UCT strategy in the beginning of the game when search space is large.