

Analysis of three different evaluation heuristics

In this project, three custom evaluation functions were implemented. The details of implementation and performance of each custom evaluation function are described below.

Evaluation function one

The first evaluation function is implemented as `custom_score1` in `game_agent.py`. For any board state it returns the difference in the weighted sum of the number of moves available to the first player (max player) and the weighted sum of the number of moves available to the second player (min player). For each player and each move, the weight of each move is defined to be the number of “legal” moves the player has after making that move.

Table 1 shows the overall performance, over 100 tournaments, using this custom evaluation function. Evaluation function one has a higher win percentage compared to `Id_improved` 71% of the time. Table 2 shows the performance against the seven pre-implemented evaluation functions.

Table 1: Overall Performance Using Evaluation Function One

Tournament Number	Id_Improved win %	Agent using evaluation function one win %
1	62.14	65.00
2	67.14	65.00
3	63.57	60.00
4	66.43	73.57
5	62.14	67.86
6	60.00	63.57
7	64.29	63.57
8	63.57	65.71
9	65.71	67.14
10	61.43	60.71
11	58.57	75.00
12	70.00	65.71
13	63.57	70.00
14	67.86	67.14
15	61.43	65.71
16	66.43	65.00
17	65.71	70.71
18	61.43	64.29
19	68.57	65.00
20	59.29	64.29
21	67.14	68.57
22	65.71	67.14
23	64.29	63.57
24	62.86	65.00
25	64.29	67.14
26	70.71	59.29
27	67.14	65.00
28	64.29	72.14
29	61.43	67.14
30	65.00	68.57

Continuation of Table 1		
Tournament Number	Id_Improved win %	Agent using evaluation function one win %
31	60.71	64.29
32	63.57	72.86
33	64.29	70.71
34	65.00	64.29
35	60.00	66.43
36	67.86	69.29
37	60.71	65.71
38	65.00	67.86
39	62.86	63.57
40	60.71	67.86
41	64.29	72.14
42	69.29	60.71
43	60.71	62.14
44	67.14	70.00
45	60.71	69.29
46	64.29	67.14
47	62.14	65.00
48	58.57	66.43
49	65.00	72.14
50	67.14	65.71
51	60.00	68.57
52	66.43	68.57
53	63.57	67.86
54	61.43	65.00
55	60.71	65.71
56	65.00	67.14
57	62.86	65.71
58	60.71	63.57
59	64.29	70.00
60	61.43	64.29
61	63.57	64.29
62	61.43	70.00
63	65.00	71.43
64	62.86	68.57
65	66.43	67.86
66	67.14	64.29
67	63.57	71.43
68	64.29	70.00
69	65.00	70.71
70	65.71	70.71
71	63.57	63.57
72	62.86	62.86
73	68.57	61.43
74	66.43	65.00
75	66.43	68.57
76	65.71	67.14
77	63.57	63.57

Continuation of Table 1		
Tournament Number	Id_Improved win %	Agent using evaluation function one win %
78	67.86	68.57
79	59.29	67.14
80	67.14	72.86
81	61.43	64.29
82	66.43	63.57
83	65.00	62.86
84	63.57	64.29
85	63.57	62.86
86	64.29	70.00
87	59.29	74.29
88	64.29	62.14
89	60.00	67.14
90	62.86	63.57
91	68.57	67.86
92	62.86	62.86
93	62.86	65.00
94	59.29	67.14
95	62.86	62.14
96	64.29	65.00
97	62.86	70.00
98	67.86	66.43
99	65.71	62.14
100	69.29	70.71

Table 2: Performance of Evaluation function one Against the 7 Pre-Implemented Evaluation Functions

Agent Eval Function	ID.improved win %	Agent with Custom Score Eval Function
Random	85.8	87.9
MM_Null	71.65	72.0
MM_open	54.15	57.85
MM_improved	50.0	55.7
AB_Null	66.65	68.55
AB_open	61.05	62.6
AB_improved	58.85	61.8

Evaluation function two

The second evaluation function is implemented as `custom_score2` in `game_agent.py`. For any board state it returns the difference in the weighted sum of the number of moves available to the first player (max player) and the weighted sum of the number of moves available to the second player (min player). For each player and each move, the weight of each move is defined to be the number of moves (not discarding moves into spots that were previously occupied) the player has after making that move.

Table 3 shows the overall performance, over 100 tournaments, using this custom evaluation function. Evaluation function two has a higher win percentage compared to Id.improved 75% of the time. Table 4 shows the performance against the seven pre-implemented evaluation functions.

Table 3: Overall Performance Using Evaluation Function Two

Tournament Number	Id. Improved win %	Agent using evaluation function two win %
1	64.29	65.71
2	69.29	67.14
3	65.00	65.00
4	67.86	70.71
5	59.29	73.57
6	63.57	68.57
7	66.43	69.29
8	62.14	75.71
9	68.57	70.00
10	65.71	65.71
11	66.43	65.71
12	65.00	68.57
13	61.43	72.14
14	65.00	71.43
15	57.86	68.57
16	61.43	66.43
17	68.57	70.00
18	61.43	72.86
19	60.00	63.57
20	60.71	71.43
21	65.00	65.00
22	67.14	68.57
23	70.71	69.29
24	60.71	62.14
25	59.29	72.14
26	65.00	67.14
27	60.00	68.57
28	63.57	69.29
29	55.00	69.29
30	68.57	74.29
31	60.71	66.43
32	62.14	62.86
33	67.86	69.29
34	61.43	62.86
35	60.71	68.57
36	62.14	72.86
37	65.71	71.43
38	62.86	71.43
39	63.57	70.71
40	64.29	68.57
41	57.86	67.86
42	61.43	75.00
43	69.29	72.86
44	65.00	60.71
45	64.29	72.86

Continuation of Table 3		
Tournament Number	Id. Improved win %	Agent using evaluation function two win %
46	64.29	70.71
47	67.86	68.57
48	62.14	73.57
49	66.43	68.57
50	64.29	68.57
51	64.29	67.14
52	65.71	63.57
53	66.43	72.86
54	68.57	64.29
55	65.00	70.00
56	65.71	67.14
57	62.86	60.71
58	62.86	70.71
59	67.14	63.57
60	60.71	70.71
61	65.00	65.71
62	65.71	61.43
63	65.71	72.86
64	65.00	71.43
65	67.86	63.57
66	69.29	62.14
67	63.57	62.86
68	54.29	67.14
69	65.71	66.43
70	68.57	63.57
71	62.86	64.29
72	65.71	66.43
73	63.57	69.29
74	62.14	59.29
75	65.00	62.86
76	61.43	67.14
77	60.00	66.43
78	63.57	65.00
79	58.57	67.86
80	65.00	68.57
81	60.00	70.71
82	68.57	66.43
83	62.14	67.86
84	66.43	65.71
85	64.29	70.71
86	62.86	70.71
87	69.29	63.57
88	63.57	66.43
89	66.43	63.57
90	63.57	69.29
91	60.71	63.57
92	64.29	67.14

Continuation of Table 3		
Tournament Number	Id_Improved win %	Agent using evaluation function two win %
93	63.57	68.57
94	68.57	67.14
95	69.29	70.00
96	66.43	67.14
97	62.14	68.57
98	63.57	71.43
99	66.43	64.29
100	67.86	62.14

Table 4: Performance of Evaluation function two Against the 7 Pre-Implemented Evaluation Functions

Agent Eval Function	ID_improved win %	Agent using evaluation function two win %
Random	85.65	89.55
MM_Null	71.3	74.1
MM_open	54.35	57.85
MM_improved	49.5	55.95
AB_Null	67.35	73.0
AB_open	62.8	62.95
AB_improved	58.25	61.35

Evaluation function three

The third evaluation function is implemented as `custom_score` in `game_agent.py`. It is a modified version of evaluation function two. We note that the weighted sum computed in evaluation function two would usually contain some double counting of moves. Evaluation function three eliminates double counting by only taking unique subsequent moves.

Table 5 shows the overall performance, over 100 tournaments, using this custom evaluation function. Evaluation function one has a higher win percentage compared to Id_improved 83% of the time. Table 6 shows the performance against the seven pre-implemented evaluation functions.

Table 5: Overall Performance Using Evaluation Function Three

Tournament Number	Id_Improved win %	Agent using evaluation function three win %
1	62.86	63.57
2	65.71	72.14
3	67.86	74.29
4	63.57	71.43
5	60.00	63.57
6	67.14	72.14
7	60.00	70.71
8	65.00	69.29
9	62.14	65.00
10	62.14	70.00
11	70.71	67.14
12	63.57	72.14

Continuation of Table 5		
Tournament Number	Id_Improved win %	Agent using evaluation function three win %
13	62.14	63.57
14	65.71	71.43
15	63.57	76.43
16	64.29	72.86
17	62.86	69.29
18	62.86	71.43
19	67.86	70.71
20	65.00	71.43
21	67.14	65.00
22	62.86	72.14
23	64.29	70.71
24	66.43	71.43
25	60.00	63.57
26	67.86	70.00
27	57.14	67.86
28	62.14	65.71
29	65.71	70.71
30	67.86	72.14
31	70.71	65.00
32	57.86	63.57
33	62.86	64.29
34	62.14	69.29
35	69.29	65.00
36	58.57	71.43
37	65.71	67.14
38	67.14	64.29
39	64.29	71.43
40	63.57	66.43
41	65.71	67.14
42	67.14	68.57
43	68.57	67.14
44	62.14	69.29
45	65.00	70.00
46	60.71	63.57
47	67.86	70.00
48	68.57	65.71
49	60.00	71.43
50	64.29	71.43
51	64.29	66.43
52	67.14	71.43
53	59.29	68.57
54	60.71	65.71
55	69.29	67.86
56	69.29	64.29
57	69.29	64.29
58	65.00	66.43
59	62.86	65.71

Continuation of Table 5		
Tournament Number	Id_Improved win %	Agent using evaluation function three win %
60	60.71	66.43
61	66.43	70.00
62	63.57	64.29
63	63.57	67.14
64	61.43	69.29
65	65.00	68.57
66	66.43	65.71
67	67.86	65.71
68	62.86	68.57
69	70.00	68.57
70	62.86	67.14
71	64.29	66.43
72	59.29	68.57
73	67.86	71.43
74	63.57	72.14
75	65.00	63.57
76	64.29	67.86
77	65.71	65.00
78	59.29	71.43
79	65.00	67.14
80	60.71	65.00
81	61.43	75.71
82	70.71	61.43
83	60.71	65.71
84	63.57	70.00
85	62.14	62.86
86	61.43	75.00
87	67.14	72.86
88	63.57	67.86
89	62.14	66.43
90	65.71	62.86
91	65.00	70.71
92	61.43	64.29
93	64.29	67.86
94	65.00	72.14
95	67.14	70.71
96	62.86	70.00
97	65.00	67.86
98	64.29	67.86
99	69.29	70.00
100	66.43	71.43

In the code submission, evaluation function three was used as the custom evaluation function based on its slightly superior performance as shown by its percentage of wins over 100 games.

Table 6: Performance of Evaluation function three Against the 7 Pre-Implemented Evaluation Functions

Agent Eval Function	ID_improved win %	Agent using evaluation function three win %
Random	85.85	89.1
MM_Null	72.55	76.35
MM_open	54.55	59.15
MM_improved	50.05	56.65
AB_Null	68.7	72.6
AB_open	60.4	62.15
AB_improved	58.75	62.4