Heuristic Analysis

In this we report, we measure the relative performance of an AI agent with three custom heuristic evaluations by playing the game of isolation against several other pre-defined agents. We compare the results to the baseline performance of an agent “AB_Improved” which uses iterative deepening alpha beta search and “improved_score” heuristic.

All agents are listed below.

	Algorithm	Heuristic
Random	Random moves	
MM_Open	Minimax search with depth 3	Counts the number of legal moves of the current player
MM_Center	Minimax search with depth 3	Measures the distance between the player's location and the center of the game board
MM_Improved	Minimax search with depth 3	Number of legal moves of the current player - Number of legal moves of the opponent player
AB_Open	Iterative deepening alpha beta	Same as MM_Open
AB_Center	Iterative deepening alpha beta	Same as MM_Center
AB_Improved	Iterative deepening alpha beta	Same as MM_Improved
AB_Custom	Iterative deepening alpha beta	Number of legal moves of the current player - 2 * Number of legal moves of the opponent player
AB_Custom2	Iterative deepening alpha beta	2 * Number of legal moves of the current player - Number of legal moves of the opponent player
AB_Custom3	Iterative deepening alpha beta	Number of legal moves of the current player - Number of legal moves of the opponent player+the distance between the player's location and the center of the game board

Before we present the results, let’s highlight the difference between our custom agents and the baseline agent AB_improved. AB_Custom weights more on limiting opponent’s moves, while AB_Custom2 focuses more on its own moves. AB_Custom3 also values the distance to the center other than the number of legal moves.

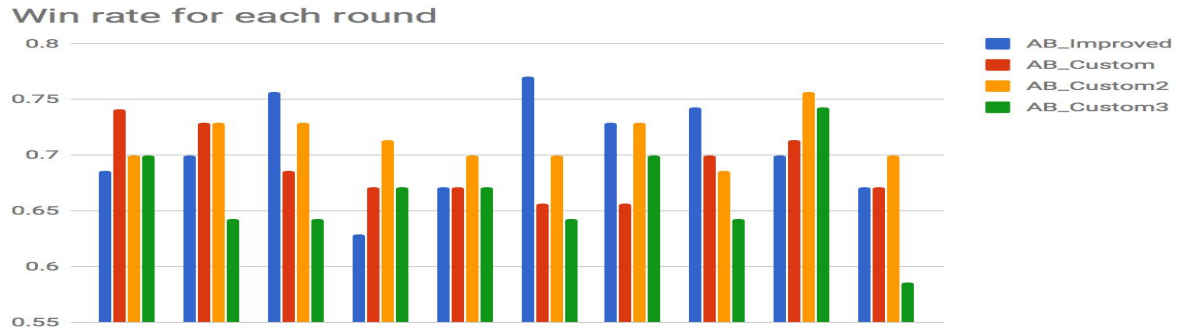
Now let’s take a look at the result. In a single test, each pairs play 10 games of Isolation. We run the test 10 times so each pairs play 100 games against each other. The following table shows the winning record.

	AB_Improved		AB_Custom		AB_Custom_2		AB_Custom_3	
	win	loss	win	loss	win	loss	win	loss
Random	94	6	94	6	95	5	95	5
MM_Open	84	16	69	31	71	29	69	31
MM_Center	87	13	88	12	87	13	90	10
MM_Improved	69	31	72	28	80	20	67	33
AB_Open	54	46	56	44	60	40	45	55
AB_Center	58	42	55	45	59	41	50	50
AB_Improved	48	52	47	53	48	52	49	51
Win Rate	70.57%		68.71%		71.42%		66.42%	
Win Rate vs AB	53.33%		52.66%		55.66%		48%	
WIn Rate vs MM	80%		76.33%		79.33%		75.33%	

First Observations:

1. All custom agents have a winning record.
2. AB_Custom3 performs worse than AB_Improved with a relatively large margin, especially when considering only the records against agents using iterative deepening alpha beta search.
3. Agents using iterative deepening alpha beta consistently beat agents using Minimax with depth 3. This is expected because alpha beta searches deeper by cutting off nodes.
4. The margin is small between AB_Improvd and both of the rest two agents. Detailed analysis is needed.

Now let’s take a look at the breakdown of each test.



Standard Deviation of Win Rate

AB_Improved	AB_Custom	AB_Custom2	AB_Custom3
0.04418911379	0.02995199863	0.02131874501	0.04317676844

Further Observation:

1. The first wining table shows that AB_Custom2 performs only slightly better than AB_Improved with a small sample size. The above chart shows that AB_Custom2 has a winning rate at least 70% in 9 out of 10 tests, while the performance of AB_Improved is very inconsistent, and it beats AB_Custom2 in only 3 tests.

2. AB_Custom3 performs worse than other agents. Not only it is inconsistent, but it also never has the highest winning rate for a single test, while it records some lowest ones.

3. AB_Custom1 neither shows significant improvement nor plays worse.

Analysis and Conclusion:

1. Iterative deepening alpha beta search does do better than minimax with fixed depth, which is consistent with the theory.
2. One shall consider both its own moves and the opponent's move to perform well, though one may consider weights more on its own move.
3. One shall reconsider the weight of distance to the center. As we play in a larger board, the distance have bigger impact on the score.