Heuristic Evaluation Function Analysis

Three different evaluation functions are implemented, and each one is based on the previous one. The first one is based on the *improved_score*, with an extra penalty. The penalty term is the euclidean distance between the players current position and the board center. Reason for adding this penalty is to try to keep the player's position prone to stay in the mid area, because a player in margin area are more likely to be stuck and lose the game. The first evaluation function is called *Center_Dist_ID_Improved*.

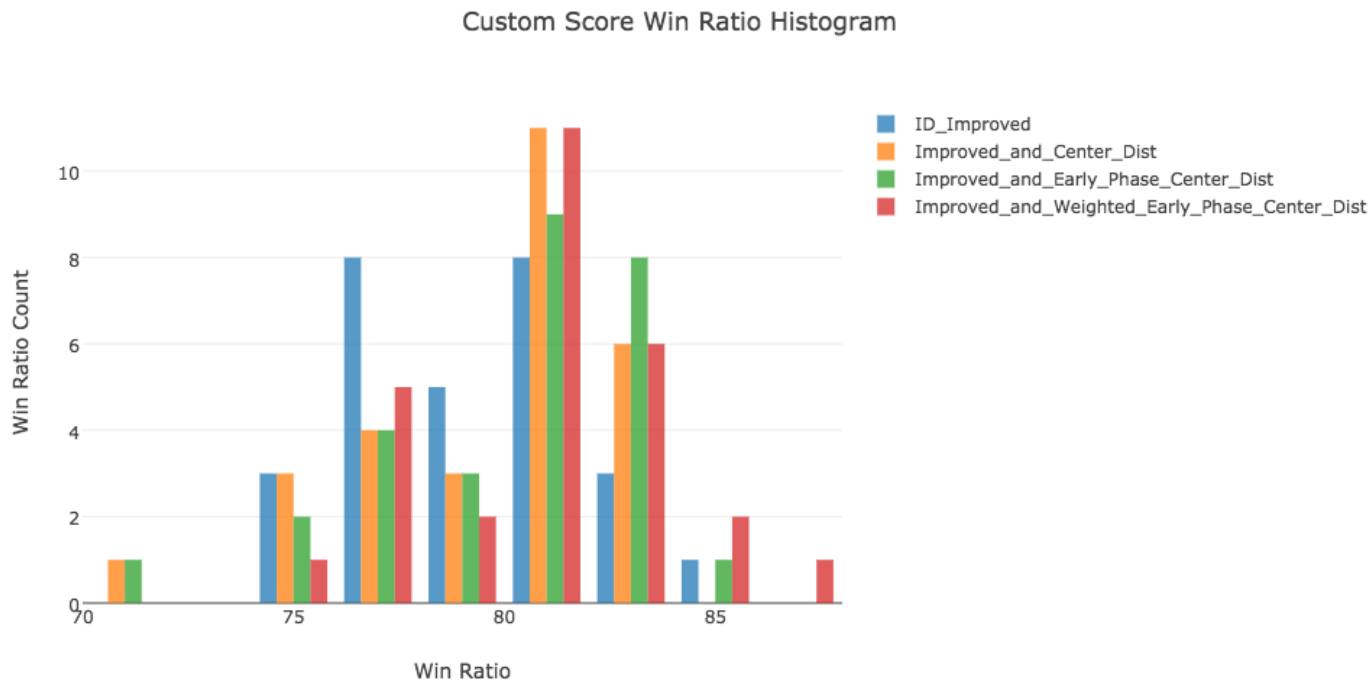
Experiment is held by running 5 matches (20 games) against each opponent (*Random*, *MM_Null*, *MM_Open*, *MM_Improved*, *AB_Null*, *AB_Open*, and *AB_Improved*). The related win ratio is recorded. 28 experiments are run for *ID_Improved* and the three evaluation functions implemented. T-test is used to evaluate the performance between each one and *ID_Improved*. *Center_Dist_ID_Improved* is slightly better than *ID_Improved*, as shown in Table.1 but not significant at all.

The second evaluation function, is called *Early_Stage_Center_Dist_ID_Improved*, which is based on *Center_Dist_ID_Improved*, but the penalty only work in the early phase of the game, when the blank space is still more than 31. At later phase of the game, the tree search can go more deeper and the penalty term may not be so necessary. *Early_Stage_Center_Dist_ID_Improved* performs better than *Center_Dist_ID_Improved*, as shown in Table.1. But its win ratio distribution is still not significantly better than *ID_Improved*, according to the t-test result shown in Table.1.

The third evaluation function is called *Early_Stage_Weighted_Center_Dist_ID_Improved*. It is based on *Early_Stage_Center_Dist_ID_Improved*, but with its penalty term weighted by 1/2. Different weight can bring different combinational effect to the evaluation function, and this makes *Early_Stage_Weighted_Center_Dist_ID_Improved* finally performs significantly better than *ID_Improved* according to the t-test result shown in Table.1. And of course, the finally introduced evalucation function, *Early_Stage_Weighted_Center_Dist_ID_Improved* is recommended to play the game, according to the experimental data.

Evaluation Function Name	t-statistic	p-value
Center_Dist_ID_Improved	0.245	0.807
Early_Stage_Center_Dist_ID_Improved	0.955	0.344
Early_Stage_Weighted_Center_Dist_ID_Improved	2.023	0.047

Table 1. Win ratio t-test with ID_Improved.



Custom Score Win Ratio Histogram