Heuristics Analysis

Requirement: Have at least three (3) evaluation heuristics besides null_score(), open_move_score(), and improved_score() been implemented and analyzed? Described the performance of agents against the testing agents.

Answer: In this experiment, we implemented three computer runs (40 matches, 100 matches and 400 matches) between our agent (Min/Max and AB pruning game tree search method capability) and a basic testing agents. Our goal is to see how our agent will perform in heuristic functions settings: null_score(), open_move_score(), and improved_score() when compete with a basic testing agents.

We use parameters AB_ARGS = {"search_depth": 5, "method": 'alphabeta', "iterative": False} and
MM_ARGS = {"search_depth": 3, "method": 'minimax', "iterative": False} which has different depth search levels for our 3 evaluations.
Interestingly, on a summary level of the 3 evaluations, our agent can achieve about 70% winning rate in any group of matches against a basic testing agent. When comparing MM with AB alone in each evaluation, the following table summarized our observations.

Evaluating: ID_Improved			
**********			
Matches Played:	40	100	400
Match 1: ID_Improved vs Random	Result: 34 to 6	Result: 86 to 14	Result: 335 to 65
Match 2: ID_Improved vs MM_Null	Result: 29 to 11	Result: 80 to 20	Result: 317 to 83
Match 3: ID_Improved vs MM_Open	Result: 25 to 15	Result: 61 to 39	Result: 267 to 133
Match 4: ID_Improved vs MM_Improved	Result: 23 to 17	Result: 59 to 41	Result: 247 to 153
Match 5: ID_Improved vs AB_Null	Result: 31 to 9	Result: 69 to 31	Result: 290 to 110
Match 6: ID_Improved vs AB_Open	Result: 29 to 11	Result: 64 to 36	Result: 261 to 139
Match 7: ID_Improved vs AB_Improved	Result: 29 to 11	Result: 58 to 42	Result: 240 to 160
Results:			
ID_Improved	71.43%	68.14%	69.89%

Table: ID_Improved Result

Requirement: Has the performance of agents against the testing agents been adequately described?

A brief report lists (using a table and any appropriate visualizations) and verbally describes the performance of agents using the implemented evaluation functions. Performance data includes results from tournament.py comparing (at a minimum) the best performing student heuristic against the ID_Improved agent.

Answer: The above table shows that in the first evaluation (40 matches); AB has higher winnings score than MM in Null (31 : 29), in Open (29 : 25) , and in Improved 29 :23) heuristics.
In second evaluation (100 matches), where depth parameters remain the same, AB has more winning scores than MM in Open(64: 61); but less winning scores in Null (69 : 80), and in Improved (58 :59) heuristics.
In third evaluation (400 matches) where depth parameters remain the same, AB has less winning scores than MM in Open (261: 267); in Null (290 : 317), and Improved (240 :247) heuristics.
We can look at the `ID_Improved` agent provides a baseline by measuring the performance of a basic agent using Iterative Deepening and the "improved" heuristic) on the computer hardware.
The `Student` agent then measures the performance of Iterative Deepening and the custom heuristic against the same opponents.
Since the parameters in the ID improved and the student agent are the same.

The student agent result below can be used to compare to the 'ID_improved' result above.

**********			
Evaluating: Student			
**********			
Matches Played:	40	100	400
Match 1: ID_Improved vs Random	Result: 33 to 7	Result: 81 to 19	Result: 328 to 72
Match 2: ID_Improved vs MM_Null	Result: 35 to 5	Result: 77 to 23	Result: 313 to 87
Match 3: ID_Improved vs MM_Open	Result: 25 to 15	Result: 64 to 36	Result: 271 to 129
Match 4: ID_Improved vs MM_Improved	Result: 29 to 11	Result: 59 to 41	Result: 260 to 140
Match 5: ID_Improved vs AB_Null	Result: 34 to 6	Result: 74 to 26	Result: 302 to 98
Match 6: ID_Improved vs AB_Open	Result: 29 to 11	Result: 61 to 39	Result: 249 to 151
Match 7: ID_Improved vs AB_Improved	Result: 27 to 13	Result: 65 to 35	Result: 243 to 157
Results:			
Student	75.71%	68.71%	70.21%

Table: Student Agent Result

Requirement: Does the report make a recommendation about the best evaluation function, and is this recommendation adequately justified? Does the report makes a recommendation about which evaluation function should be used and justifies the recommendation with at least three reasons supported by the data.

Answer: It seems that the report cannot make a recommendation about the evaluation function for the following three reasons

- By comparing the student's agent performance with ID_Improved, student's agent has only a slight higher winning percentage than ID_improved (in 40 matches group, student vs ID has 75.71% and 71.43% respectively); (in 100 matches group, student vs ID has 68.71%Vs 68.14%); (in 400 matches group, student vs ID has 70.21% and 69.89% respectively); There's no any evaluations has more than 5% pervasive difference.

- Looking at individual heuristics of the student agent result, as shown in the table below, there's no clear winner when comparing AB to MM by different group of matches.

AB: MM	Null	Open	Improved
40 Matches	34:35	29:25	27:29
100 Matches	74:77	61:64	65:59
400 Matches	302:313	249:271	243:260

- Not enough data to make the conclusion. Due to limitation of computer resource, the maximum group of matches is only 400, which is not enough to assure randomness in this experiment.