

Heuristic analysis

This brief report describes what kind of heuristics I used in this project, their evaluation results in `tournament.py` and overall decision of taking one for final submission.

Heuristics used

`score_moves`

First and very obvious heuristic is `score_moves` which just counts available moves of the player and the opponent. This one is used already in `improved_score` function by default, so I took an idea from lessons to not just use raw difference, but add some bigger multiplier for opponent moves: `score = player_moves - opponent_scaling_factor * opponent_moves`. Values used for tests are 1, 1.5 and 2.

`score_center`


Second heuristic `score_center` is based on the idea that cells closer to center are more preferable than edge ones -- there is potentially more space to go since no board-limits exist. I used Manhattan distance from the very central cell to count the distance. Because best position gives 0 distance and worst gives 6 and we want heuristic to return higher values for better moves, I subtract distance from maximum value. Thus, worst value is 0 and the best is 6. As before, I added scaling factor for opponent moves. Noteworthy is that *I used scaling for opponent moves for every custom heuristics*.

`score_freecells`

Third idea is to count all available free cells in a radius of 2 (since knight move range). If given cell has more open spaces around, consider it better. This is a bit correlated with number of moves, but is different. Consider a cell which has only 1 move available (e.g. one near to the corner). But there is about 5 other cells available so within this little space there might be more possible moves. Another cell can have 2 moves and 2 free cells so after moving once game might be also over. This seems better to combine number of moves and number of space. I make it below.

`score_moves_center`

This one is a combination of `score_moves` and `score_center`. Scaling factor applied for both heuristics independently and with same value, i.e.

```
 def score_moves_center(game, player, opponent_scaling_factor=1):  
    if game.is_loser(player):
```

```

        return float("-inf")

    if game.is_winner(player):
        return float("inf")

    move_score = score_moves(game, player, opponent_scaling_factor)
    center_score = score_center(game, player, opponent_scaling_factor)
    score = move_score + center_score
    return score

```

score_moves_freecells

Combination of score_moves and score_freecells. Scaling factor policy is the same.

score_moves_center_freecells

Ultimate combination of score_moves, score_center and score_freecells. Scaling factor policy is the same.

score_combined

Ultimate combination of score_moves, score_center and score_freecells. Scaling factor policy is based on previous evaluations of listed 3 heuristics. Best values used:

- x1.5 for score_moves -- 82.86% win ratio (see table below)
- x2 for score_center -- 75.00% win ratio (see table below)
- default for score_freecells -- 77.86% win ratio (see table below)

Evaluation results

Default column means player_score - opponent_score for custom heuristics. Scaling factors 1.5 and 2 noted in corresponding columns **x1.5** and **x2**. Raw results of tournament.py represented in Appendix section.

Heuristic	Default	x1.5	x2	Comments
ID_improved -- baseline	70.00%			
score_moves	72.86%	82.86%	77.14%	82.86%: minimal is 14-6 vs AB_Improved
score_center	72.14%	72.14%	75.00%	Not that much good heuristic
score_freecells	77.86%	75.71%	75.71%	Almost no difference with scaling

Heuristic	Default	x1.5	x2	Comments
score_moves_center	70.00%	75.00%	80.71%	80.71%: minimal is 12-8 vs AB_Improved
score_moves_freecells	72.86%	73.57%	72.14%	Kinda useless combination :(
score_moves_center_freecells	80.00%	82.14%	80.71%	Stablest heuristic -- always about 80%!
score_combined	72.86%			Much worse than separate ones

Every custom heuristic is better than baseline, but because of stochastic nature I consider results less than 80% as poor ones. In only 5 cases 80% threshold was passed, 3 times of them by last heuristic -- score_moves_center_freecells. It seems to be best one because of its stability.

Resulting heuristic and recommendation

Final heuristic score_moves_center_freecells with scaling factor 1.5 is chosen for submission based on its results. Highest value is given by simple score_moves, but it is much less stable with different scaling factors. This high score might be caused by randomness, but score_moves_center_freecells seems to be more stable.

Recommendations about evaluation functions usage:

- combine different approaches together, this provides better results as for score_moves_center_freecells;
- not implemented here, but is good to test different weighting coefficients for such heuristics within a combination;
- based on results, it seems that pushing to choose central cells (as with score_center) is useful, other evaluation functions are likely to try.
- without scaling factors, score_freecells is performing better than score_moves -- so counting open area is also important. More sophisticated functions, which for example can count open cells within Γ -distance from each other, should be tried.

Appendix -- Raw evaluation results

```

*****
Evaluating: ID_Improved
*****

Playing Matches:
-----
Match 1: ID_Improved vs Random Result: 16 to 4

```

Match 2: ID_Improved vs MM_Null Result: 16 to 4
Match 3: ID_Improved vs MM_Open Result: 17 to 3
Match 4: ID_Improved vs MM_Improved Result: 11 to 9
Match 5: ID_Improved vs AB_Null Result: 15 to 5
Match 6: ID_Improved vs AB_Open Result: 13 to 7
Match 7: ID_Improved vs AB_Improved Result: 10 to 10

Results:

ID_Improved 70.00%

Evaluating: Student: score_moves

Playing Matches:

Match 1: Student: score_moves vs Random Result: 19 to 1
Match 2: Student: score_moves vs MM_Null Result: 16 to 4
Match 3: Student: score_moves vs MM_Open Result: 14 to 6
Match 4: Student: score_moves vs MM_Improved Result: 11 to 9
Match 5: Student: score_moves vs AB_Null Result: 14 to 6
Match 6: Student: score_moves vs AB_Open Result: 15 to 5
Match 7: Student: score_moves vs AB_Improved Result: 13 to 7

Results:

Student: score_moves 72.86%

Evaluating: Student: score_moves p-1.5o

Playing Matches:

Match 1: Student: score_moves p-1.5o vs Random Result: 17 to 3
Match 2: Student: score_moves p-1.5o vs MM_Null Result: 16 to 4
Match 3: Student: score_moves p-1.5o vs MM_Open Result: 19 to 1
Match 4: Student: score_moves p-1.5o vs MM_Improved Result: 16 to 4
Match 5: Student: score_moves p-1.5o vs AB_Null Result: 17 to 3
Match 6: Student: score_moves p-1.5o vs AB_Open Result: 17 to 3
Match 7: Student: score_moves p-1.5o vs AB_Improved Result: 14 to 6

Results:

Student: score_moves p-1.5o 82.86%

Evaluating: Student: score_moves p-2o

Playing Matches:

Match 1:	Student: score_moves p-2o vs	Random	Result: 20 to 0
Match 2:	Student: score_moves p-2o vs	MM_Null	Result: 16 to 4
Match 3:	Student: score_moves p-2o vs	MM_Open	Result: 16 to 4
Match 4:	Student: score_moves p-2o vs	MM_Improved	Result: 14 to 6
Match 5:	Student: score_moves p-2o vs	AB_Null	Result: 16 to 4
Match 6:	Student: score_moves p-2o vs	AB_Open	Result: 13 to 7
Match 7:	Student: score_moves p-2o vs	AB_Improved	Result: 13 to 7

Results:

Student: score_moves p-2o 77.14%

Evaluating: Student: score_center

Playing Matches:

Match 1:	Student: score_center vs	Random	Result: 19 to 1
Match 2:	Student: score_center vs	MM_Null	Result: 18 to 2
Match 3:	Student: score_center vs	MM_Open	Result: 13 to 7
Match 4:	Student: score_center vs	MM_Improved	Result: 10 to 10
Match 5:	Student: score_center vs	AB_Null	Result: 18 to 2
Match 6:	Student: score_center vs	AB_Open	Result: 15 to 5
Match 7:	Student: score_center vs	AB_Improved	Result: 8 to 12

Results:

Student: score_center 72.14%

Evaluating: Student: score_center p-1.5o

Playing Matches:

Match 1:	Student: score_center p-1.5o vs	Random	Result: 16 to 4
Match 2:	Student: score_center p-1.5o vs	MM_Null	Result: 17 to 3
Match 3:	Student: score_center p-1.5o vs	MM_Open	Result: 16 to 4
Match 4:	Student: score_center p-1.5o vs	MM_Improved	Result: 12 to 8
Match 5:	Student: score_center p-1.5o vs	AB_Null	Result: 15 to 5
Match 6:	Student: score_center p-1.5o vs	AB_Open	Result: 15 to 5
Match 7:	Student: score_center p-1.5o vs	AB_Improved	Result: 10 to 10

Results:

Student: score_center p-1.5o 72.14%

Evaluating: Student: score_center p-2o

Playing Matches:

Match 1:	Student: score_center p-2o vs	Random	Result: 17 to 3
Match 2:	Student: score_center p-2o vs	MM_Null	Result: 19 to 1
Match 3:	Student: score_center p-2o vs	MM_Open	Result: 16 to 4
Match 4:	Student: score_center p-2o vs	MM_Improved	Result: 12 to 8
Match 5:	Student: score_center p-2o vs	AB_Null	Result: 15 to 5
Match 6:	Student: score_center p-2o vs	AB_Open	Result: 11 to 9
Match 7:	Student: score_center p-2o vs	AB_Improved	Result: 15 to 5

Results:

Student: score_center p-2o 75.00%

Evaluating: Student: score_freecells

Playing Matches:

Match 1:	Student: score_freecells vs	Random	Result: 19 to 1
Match 2:	Student: score_freecells vs	MM_Null	Result: 19 to 1
Match 3:	Student: score_freecells vs	MM_Open	Result: 14 to 6
Match 4:	Student: score_freecells vs	MM_Improved	Result: 14 to 6
Match 5:	Student: score_freecells vs	AB_Null	Result: 15 to 5
Match 6:	Student: score_freecells vs	AB_Open	Result: 16 to 4
Match 7:	Student: score_freecells vs	AB_Improved	Result: 12 to 8

Results:

Student: score_freecells 77.86%

Evaluating: Student: score_freecells p-1.5o

Playing Matches:

Match 1:	Student: score_freecells p-1.5o vs	Random	Result: 19 to 1
Match 2:	Student: score_freecells p-1.5o vs	MM_Null	Result: 20 to 0
Match 3:	Student: score_freecells p-1.5o vs	MM_Open	Result: 13 to 7
Match 4:	Student: score_freecells p-1.5o vs	MM_Improved	Result: 14 to 6
Match 5:	Student: score_freecells p-1.5o vs	AB_Null	Result: 17 to 3

Match 6: Student: score_freecells p-1.5o vs AB_Open Result: 10 to 10
Match 7: Student: score_freecells p-1.5o vs AB_Improved Result: 13 to 7

Results:

Student: score_freecells p-1.5o 75.71%

Evaluating: Student: score_freecells p-2o

Playing Matches:

Match 1:	Student: score_freecells p-2o vs	Random	Result: 19 to 1
Match 2:	Student: score_freecells p-2o vs	MM_Null	Result: 19 to 1
Match 3:	Student: score_freecells p-2o vs	MM_Open	Result: 14 to 6
Match 4:	Student: score_freecells p-2o vs	MM_Improved	Result: 12 to 8
Match 5:	Student: score_freecells p-2o vs	AB_Null	Result: 18 to 2
Match 6:	Student: score_freecells p-2o vs	AB_Open	Result: 13 to 7
Match 7:	Student: score_freecells p-2o vs	AB_Improved	Result: 11 to 9

Results:

Student: score_freecells p-2o 75.71%

Evaluating: Student: score_moves_center

Playing Matches:

Match 1:	Student: score_moves_center vs	Random	Result: 19 to 1
Match 2:	Student: score_moves_center vs	MM_Null	Result: 17 to 3
Match 3:	Student: score_moves_center vs	MM_Open	Result: 12 to 8
Match 4:	Student: score_moves_center vs	MM_Improved	Result: 14 to 6
Match 5:	Student: score_moves_center vs	AB_Null	Result: 14 to 6
Match 6:	Student: score_moves_center vs	AB_Open	Result: 10 to 10
Match 7:	Student: score_moves_center vs	AB_Improved	Result: 12 to 8

Results:

Student: score_moves_center 70.00%

Evaluating: Student: score_moves_center p-1.5o

Playing Matches:

Match 1:	Student:	score_moves_center	p-1.5o	vs	Random	Result:	18	to	2
Match 2:	Student:	score_moves_center	p-1.5o	vs	MM_Null	Result:	17	to	3
Match 3:	Student:	score_moves_center	p-1.5o	vs	MM_Open	Result:	14	to	6
Match 4:	Student:	score_moves_center	p-1.5o	vs	MM_Improved	Result:	13	to	7
Match 5:	Student:	score_moves_center	p-1.5o	vs	AB_Null	Result:	18	to	2
Match 6:	Student:	score_moves_center	p-1.5o	vs	AB_Open	Result:	14	to	6
Match 7:	Student:	score_moves_center	p-1.5o	vs	AB_Improved	Result:	11	to	9

Results:

Student: score_moves_center p-1.5o 75.00%

Evaluating: Student: score_moves_center p-2o

Playing Matches:

Match 1:	Student:	score_moves_center	p-2o	vs	Random	Result:	20	to	0
Match 2:	Student:	score_moves_center	p-2o	vs	MM_Null	Result:	16	to	4
Match 3:	Student:	score_moves_center	p-2o	vs	MM_Open	Result:	17	to	3
Match 4:	Student:	score_moves_center	p-2o	vs	MM_Improved	Result:	14	to	6
Match 5:	Student:	score_moves_center	p-2o	vs	AB_Null	Result:	18	to	2
Match 6:	Student:	score_moves_center	p-2o	vs	AB_Open	Result:	16	to	4
Match 7:	Student:	score_moves_center	p-2o	vs	AB_Improved	Result:	12	to	8

Results:

Student: score_moves_center p-2o 80.71%

Evaluating: Student: score_moves_freecells

Playing Matches:

Match 1:	Student:	score_moves_freecells	vs	Random	Result:	19	to	1
Match 2:	Student:	score_moves_freecells	vs	MM_Null	Result:	16	to	4
Match 3:	Student:	score_moves_freecells	vs	MM_Open	Result:	13	to	7
Match 4:	Student:	score_moves_freecells	vs	MM_Improved	Result:	17	to	3
Match 5:	Student:	score_moves_freecells	vs	AB_Null	Result:	15	to	5
Match 6:	Student:	score_moves_freecells	vs	AB_Open	Result:	12	to	8
Match 7:	Student:	score_moves_freecells	vs	AB_Improved	Result:	10	to	10

Results:

Student: score_moves_freecells 72.86%

Evaluating: Student: score_moves_freecells p-1.5o

Playing Matches:

Match 1:	Student: score_moves_freecells p-1.5o vs	Random	Result: 17 to 3
Match 2:	Student: score_moves_freecells p-1.5o vs	MM_Null	Result: 16 to 4
Match 3:	Student: score_moves_freecells p-1.5o vs	MM_Open	Result: 17 to 3
Match 4:	Student: score_moves_freecells p-1.5o vs	MM_Improved	Result: 14 to 6
Match 5:	Student: score_moves_freecells p-1.5o vs	AB_Null	Result: 14 to 6
Match 6:	Student: score_moves_freecells p-1.5o vs	AB_Open	Result: 13 to 7
Match 7:	Student: score_moves_freecells p-1.5o vs	AB_Improved	Result: 12 to 8

Results:

Student: score_moves_freecells p-1.5o 73.57%

Evaluating: Student: score_moves_freecells p-2o

Playing Matches:

Match 1:	Student: score_moves_freecells p-2o vs	Random	Result: 18 to 2
Match 2:	Student: score_moves_freecells p-2o vs	MM_Null	Result: 16 to 4
Match 3:	Student: score_moves_freecells p-2o vs	MM_Open	Result: 13 to 7
Match 4:	Student: score_moves_freecells p-2o vs	MM_Improved	Result: 13 to 7
Match 5:	Student: score_moves_freecells p-2o vs	AB_Null	Result: 16 to 4
Match 6:	Student: score_moves_freecells p-2o vs	AB_Open	Result: 13 to 7
Match 7:	Student: score_moves_freecells p-2o vs	AB_Improved	Result: 12 to 8

Results:

Student: score_moves_freecells p-2o 72.14%

Evaluating: Student: score_moves_center_freecells

Playing Matches:

Match 1:	Student: score_moves_center_freecells vs	Random	Result: 19 to 1
----------	--	--------	-----------------

Match 2: Student: score_moves_center_freecells vs MM_Null Result: 20 to 0
Match 3: Student: score_moves_center_freecells vs MM_Open Result: 12 to 8
Match 4: Student: score_moves_center_freecells vs MM_Improved Result: 15 to 5
Match 5: Student: score_moves_center_freecells vs AB_Null Result: 18 to 2
Match 6: Student: score_moves_center_freecells vs AB_Open Result: 15 to 5
Match 7: Student: score_moves_center_freecells vs AB_Improved Result: 13 to 7

Results:

Student: score_moves_center_freecells 80.00%

Evaluating: Student: score_moves_center_freecells p-1.5o

Playing Matches:

Match 1: Student: score_moves_center_freecells p-1.5o vs Random Result: 19 to 1
Match 2: Student: score_moves_center_freecells p-1.5o vs MM_Null Result: 18 to 2
Match 3: Student: score_moves_center_freecells p-1.5o vs MM_Open Result: 17 to 3
Match 4: Student: score_moves_center_freecells p-1.5o vs MM_Improved Result: 13 to 7
Match 5: Student: score_moves_center_freecells p-1.5o vs AB_Null Result: 16 to 4
Match 6: Student: score_moves_center_freecells p-1.5o vs AB_Open Result: 17 to 3
Match 7: Student: score_moves_center_freecells p-1.5o vs AB_Improved Result: 15 to 5

Results:

Student: score_moves_center_freecells p-1.5o 82.14%

Evaluating: Student: score_moves_center_freecells p-2o

Playing Matches:

Match 1: Student: score_moves_center_freecells p-2o vs Random Result: 18 to 2

Match 2: Student: score_moves_center_freecells p-2o vs MM_Null Result: 1
8 to 2
Match 3: Student: score_moves_center_freecells p-2o vs MM_Open Result: 1
5 to 5
Match 4: Student: score_moves_center_freecells p-2o vs MM_Improved Result: 1
3 to 7
Match 5: Student: score_moves_center_freecells p-2o vs AB_Null Result: 1
9 to 1
Match 6: Student: score_moves_center_freecells p-2o vs AB_Open Result: 1
5 to 5
Match 7: Student: score_moves_center_freecells p-2o vs AB_Improved Result: 1
5 to 5

Results:

Student: score_moves_center_freecells p-2o 80.71%

Evaluating: Student: score_combined

Playing Matches:

Match 1:	Student	vs	Random	Result: 19 to 1
Match 2:	Student	vs	MM_Null	Result: 17 to 3
Match 3:	Student	vs	MM_Open	Result: 15 to 5
Match 4:	Student	vs	MM_Improved	Result: 12 to 8
Match 5:	Student	vs	AB_Null	Result: 12 to 8
Match 6:	Student	vs	AB_Open	Result: 15 to 5
Match 7:	Student	vs	AB_Improved	Result: 12 to 8

Results:

Student: score_combined 72.86%