Human-Like RL Chess Engine

Reid McIlroy-Young, Karthik Raja Kalaiselvi Bhaska

November 2018

Goal

To make the chess engine that acts more like a human, through supervised training and modification of it's risk sensitivity

Motivation

- Modern Chess Engines far exceed Human ability
- Humans are less likely to give up pieces
- Humans assume a pattern holds more than they should (Intuition)
- Humans will play to their opponent making a mistake over an optimal move
- In the end game, Humans tends to make more mistakes

Differences suggests – Mechanism of Learning in Humans and Machines are *different*

Human Like chess work

- Not much previous work on human like chess engines
- One useful paper: *A More Human Way to Play Computer Chess*
- Uses move table, not good for late game also not RL
- result was not particularly strong
- Automatic analysis of the positions is also possible
- Lots of work needs to be done

RL chess engines

- Started before *AlphaZero*, but were not very good, e.g. Giraffe
- You all know about *AlphaZero*, NN that evaluates on it's own, and uses this to do a tree search
- Based on Predictive + Upper Bound Tree Search, modification of UCB 1
- *AlphaZero* has been reimplemented in the open source Leela Chess project

ELO

A single number representing player skill, stronger players have higher ELOs

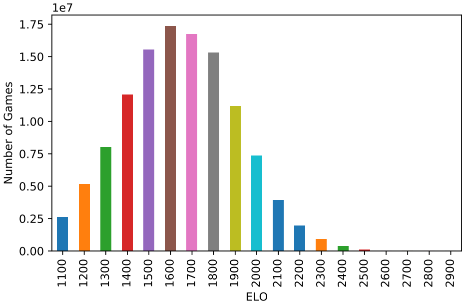


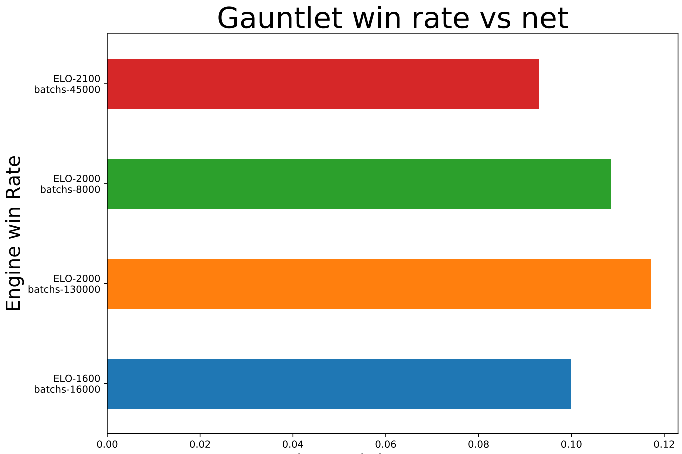
Figure: Distribution of games between similar ELO players

Supervised training

- All games from `database.lichess.org`
- Using 432,335,939 games as a training set extracted games between similar ELO players
- 22,971,939 as a holdout for later use
- used ranges of 100 to segregate the training data
- During training 10% was held out to get a training error
- Training error quickly converged to 30% on all runs, while MSE slowly decreased

Win rates

Results from an initial gauntlet of 2,500 games vs a variety of computer opponents, with increasing skill from 1,000 ELO to $3,000+$ ELO



KL divergence from humans

13,000 boards states from our holdout set with at least 50 human games had encountered and created a probability distribution over the human moves (P_{human}) and compared it to the output probabilities from the neural engines (P_{NN})

$$D_{KL}\left(P_{NN} || P_{human}\right)$$

Human like	kl Q mean	kl Q median	kl prob mean	kl prob median
1100-40000	2.24	2.03	1.73	1.52
1600-118000	2.16	1.92	1.65	1.43
2000-130000	2.11	1.85	1.59	1.35
2100-45000	1.90	1.65	1.36	1.13
Leela				
leela-1697.27	0.97	0.84	0.94	0.87
leela-2072.79	0.98	0.82	0.92	0.83

Conclusion

You can play our engine at: `lichess.org/@/haibrid_bot`
We are still looking for the best way to compare against human play

```
f3h4  (586 ) N:      65 (+ 0) (P:  1.82%) (Q:  0.04971) (U: 0.07317) (Q+U:  0.12288) (V:  0.9787)
d2d3  (288 ) N:      69 (+ 0) (P: 11.40%) (Q: -0.30271) (U: 0.43309) (Q+U:  0.13039) (V: -0.9104)
d1e1  (75  ) N:     156 (+ 0) (P:  7.60%) (Q: -0.00620) (U: 0.12872) (Q+U:  0.12251) (V: -0.9745)
h2h3  (400 ) N:     313 (+ 0) (P: 41.26%) (Q: -0.21944) (U: 0.34930) (Q+U:  0.12986) (V: -0.9979)
d2d4  (293 ) N:     557 (+ 0) (P:  7.91%) (Q:  0.08027) (U: 0.03768) (Q+U:  0.11795) (V: -0.8553)
c2c3  (259 ) N:    4765 (+37) (P: 10.18%) (Q:  0.12579) (U: 0.00563) (Q+U:  0.13143) (V: -0.9049)
```

Figure: Example of output for a board after 10000 rollbacks