Poker Prime Team

Team members:

Adam Samulak, Richard Aguilar, Tony Chow

Problem Description

Poker is an imperfect information game. That means it is difficult to estimate the proper strategy for winning without knowing all the information.

Our goal is to develop agents that will focus more heavily on one strategy. We plan to explore how specialized agents will perform against each other and a general agent that focuses on all of strategies.

Related Work

- Moravčík et al. (2017). DeepStack: Expert-Level Artificial Intelligence in No-Limit Poker Science 356, 508–513. https://science.sciencemag.org/content/356/6337/508
- Vaswani et al. (2017). Attention Is All You Need. arXiv 1706.03762. https://arxiv.org/abs/1706.03762
- M. Lanctot, et al. (2009). Monte Carlo Sampling for Regret Minimization in Extensive Games. http://books.nips.cc/papers/files/nips22/NIPS2009_0363.pdf
- Brown, N., & Sandholm, T. (2019). Superhuman AI for multiplayer poker. Science 365, 885-890. https://science.sciencemag.org/content/365/6456/885
- Teófilo, L.F., Reis, L. P. 2011. Building a No Limit Texas Hold'em Poker Playing Agent based on Game Logs using Supervised Learning. Proceedings2ndInternational

Algorithms to be investigated

1) Attention transformer
   - Calculates the probability of a given hand
   - Low on storage (2 player game has 3,679,075,400 different sets of 9 cards)
2) Strategy algorithm candidates:
   - Counterfactual Regret Minimization (general agent)
   - Naive Base Classifier combined with eval. metrics

Poker Hand Probability

Poker Hand	Frequency	Probability
Nothing	1,302,540	0.501177
One Pair	1,098,240	0.422569
Two Pair	123,552	0.047539
Three of a Kind	54,912	0.021128
Straight	10,200	0.003925
Flush	5,108	0.001965
Full House	3,744	0.001441
Four of a Kind	624	0.000240
Straight Flush	36	0.000014
Royal Flush	4	0.000002
Total	2,598,960	1.000000

Roles

Stage 1:

All members: RL Card API, Tensorflow, Keras, Transformers

Stage 2:

Each member can either develop a custom algorithm for a strategy or use an existing one.

Evaluation

- Evaluation metrics total winnings, VPIP and PFR.
  - VPIP and PFR helps to classify agents into different categories that suggest the strategy an agent is using.

Timeline

Week 11/02:

- Understand how to use RL Card API and TF work
- Understand how to link RL Card API and the neural network

Week 11/10:

- Implement simple neural network examples (probability estimation) before transformer implementation
- Begin work on the transformer implementation

Week 11/17:

- Solve any issues that came up during the transformer implementation
- Test it with RL Card API

Week 11/24:

- Begin work on the strategy algorithms

Week 12/01:

- Continue the work on strategy algorithms
- Resolve any issues

Week 12/08:

- Resolve any issues with the algorithms and test them
- Report write up

Week 12/15:

- Presentation/demo