Project Proposal
COMS4998: Deep Learning

Teammates: Natania Wolansky (nrw2107),
Guy Aridor (ga2449), Jisha Jacob (jsj2130)

Games of Incomplete Information – Modeling Responses to Irrational Behaviour using Deep Learning

Problem Statement and Goal of Project

Modeling strategies in Heads-Up No-Limits Poker (HUNL) is difficult because it is a game of incomplete information, in which the full state (the cards of the opponent) is not known. There has been recent breakthrough work (see [4], [1]) in searching for Nash-equilibrium strategies in this space. Yet this ignores the fact that average players often deviate from rationality in the context of games [3], exhibiting behaviors such as risk-aversion [2].

The goal of the project is therefore to train two Poker-playing bots – one trained by self-learning and one trained by using the historical examples of poker hands and subsequent actions. We wish to see whether learning how people really behave, rather than trying to learn a Nash-equilibrium, will allow the latter bot to exploit non-rational behavior, improving play against regular-level players.

Previous Work

To overcome the challenge of states lacking unique values in imperfect-information games, two notable approaches have been explored. Prior to the approach of multi-valued states taken by Noam et. al [1], AI DeepStack used the technique of modifying the definition of state into a joint belief state across all players [4], which proved inferior to the top two HUNL players. Recursive reasoning approaches to solve imperfect-information games have been explored since 2008, with counterfactual regret minimization being the most notable technique exploiting the degree of incomplete information in an extensive game and employing self-play to do recursive reasoning.

Data/Method and Evaluation

Our project primarily builds off the work of [1]. We follow their approach and consider multi-valued states in determining optimal strategies in imperfect information games. In determining the multiple-values associated with each state they utilize both a "bias" approach and a "self-generative" approach that attempt to approximate the Nash equilibrium strategy. However, we will consider a third approach and learn these values based on observed behavior of human players in those states. We will utilize the data on observed behavior from http://poker.cs.ualberta.ca/irc_poker_database.html and will follow the implementation details detailed in the [1] paper.

To evaluate our bots against average players (rather than other, rational poker bots) We will match our rational AI and our historical AI against a holdout set of real poker games found in the online resource. Then, we will compare performances of the bots in different scenarios – against bad, good, and great players – to try to identify any domains in which the historical training gives an edge.

References

[1] Noam Brown, Tuomas Sandholm, and Brandon Amos. Depth-limited solving for imperfect-information games. *arXiv preprint arXiv:1805.08195*, 2018.

[2] David Eil and Jaimie W. Lien. Staying ahead and getting even: Risk attitudes of experienced poker players. *Games and Economic Behavior*, 87, 2014.

[3] Jacob K Goeree and Charles A Holt. Ten little treasures of game theory and ten intuitive contradictions. *American Economic Review*, 91(5):1402–1422, 2001.

[4] Matej Moravčík, Martin Schmid, Neil Burch, Viliam Lisỳ, Dustin Morrill, Nolan Bard, Trevor Davis, Kevin Waugh, Michael Johanson, and Michael Bowling. Deepstack: Expert-level artificial intelligence in heads-up no-limit poker. *Science*, 356(6337):508–513, 2017.