Machine Learning Project: Report 2

Ignace Bleukx Quinten Bruynseraede

May 3, 2020

1 Introduction

1.1 Evaluation metrics

Evaluation of agents is traditionally done using **NashConv** and **exploitability**. We introduce these concepts here, using

Given a policy π,

1.2 Algorithm 1: Fictitious Self-Play

1.2.1 Extension: Neural Fictitious Self-Play

1.3 Algorithm 2: Counterfactual Regret Minimization

1.3.1 Extension: Regression Counterfactual Regret Minimization

1.3.2 Extension: Counterfactual Regret Minimization against best responder

1.3.3 Extension: Deep Counterfactual Regret Minimization

2 Kuhn Poker

- Which algorithm is most suitable to develop an agent to play Kuhn Poker, minimizing exploitability?
- Can we exploit properties of Kuhn Poker to optimize parameters?

3 Leduc Poker

- Which algorithm is most suitable to develop an agent to play Leduc Poker, minimizing exploitability?
- Can we exploit properties of Leduc Poker to optimize parameters?
- Can we combine agents into an ensemble that minimizes exploitability further than its parts?

[1]

References

[1] Karl Tuyls, Julien Perolat, Marc Lanctot, Georg Ostrovski, Rahul Savani, Joel Z Leibo, Toby Ord, Thore Graepel, and Shane Legg. Symmetric decomposition of asymmetric games. *Scientific reports*, 8(1):1–20, 2018.

Appendix

3.1 Time spent