COMP423 - Reinforcement Learning and Dynamic Optimization
Poker Project Part 1 Report

Pantourakis Michail AM 2015030185
School of Electrical and Computer Engineering
Technical University of Crete

Date submitted: 30 June 2023

1 Introduction

A significant step towards understanding today's advancements in Reinforcement Learning (RL) is by first understanding the fundamental theory of Markov Decision Processes for modelling decision making problems and the "exact" model-based algorithms that optimally solve these problems. However, these approaches are limited by their requirement for complete knowledge of the environment (model). A major breakthrough in RL was thus the development of the Q Learning algorithm, the first model-free learning algorithm with guaranteed convergence to the optimal policy.

In this report I present my course project implementing these algorithms and analyzing their results on a simplified, "toy" version of Poker. I will start by first establishing the rules of this simplified game, the opponents designed as part of the environment, my model representation and how I implemented all these in Python. With the problem formulation being established, I will then continue with how I applied Policy Iteration (model-based) and Q Learning (model-free) algorithms to this problem, reporting key observations, and most importantly, analyzing their behavior compared to the nature of this game and what theory predicts.

2 Environment

- Simplified game and rules
- what was adapted from rlcard
- 2 opponent models
- State-action representation
- State space creation

3 Model-based solution: Policy Iteration algorithm

- Implementation
- Average Results
- Analysis of Optimal Policy per opponent
- Demonstration of optimality

4 Model-free solution: Q Learning algorithm

- Implementation
- Analysis of hyperparameter tuning
- Average Results
- Convergence analysis

5 Conclusion

- Main point for game size & performance of algorithms
- Suggested next steps/improvements