

Ivey: A Reinforcement Learning Poker Engine

Aaron Jin, Ryan Cheng

Department of Computer Science, Stanford University

aaronjin@stanford.edu
rcheng07@stanford.edu

Stanford
Computer Science

Overview

This study develops poker agents using reinforcement learning (RL), focusing on preflop decision-making in heads-up No-Limit Texas Hold'em. Beginning with a push-fold framework, we expanded to include sophisticated betting options and implemented a Q-learning based system that generates strategies through self-play. Our single-agent learning paradigm, where one agent learns both positions, outperforms traditional dual-agent approaches and develops sophisticated preflop strategies that approach theoretical optimality.

Data Architecture & Features

- Training data generated through **iterative self-play** with **Q-table** structure.
- State space includes **169 unique preflop hands**: 13 pairs, 78 suited, and 78 offsuit combinations.
- Binary position encoding (first/second to act).
- Action history tracks complete sequence of betting decisions in current hand.
- Training batch consists of **100K hands per epoch**, multiple decision points per hand.
- Pre-computed equity lookup tables stored for quick hand strength evaluation.
- Stack and pot sizes normalized to big blinds.
- Legal action masking to prevent invalid moves.
- Position-specific exploration rates optimize asymmetric learning patterns.

References

- Noam Brown, Adam Lerer, Sam Gross, and Tuomas Sandholm. 2019. Deep counterfactual regret minimization. In *Proceedings of the 36th International Conference on Machine Learning*.
- Sergiu Hart and Andreu Mas-Colell. 2000. A simple adaptive procedure leading to correlated equilibrium.
- Michael Bowling, Martin Zinkevich, Michael Johanson and Carmelo Piccione. 2007. Regret minimization in games with incomplete information. In *Advances in Neural Information Processing Systems (NeurIPS)*.
- Michael Johanson, Michael Bowling, Neil Burch and Oskari Tammelin. 2007. Heads-up limit hold'em poker is solved. In *Communications of the ACM, Volume 60, Issue 11*.
- Michael Bowling, Viliam Lisý, Marc Lanctot. 2015. Online monte carlo counterfactual regret minimization for search in imperfect information games. In *AAMAS*.

Results

Decision Charts

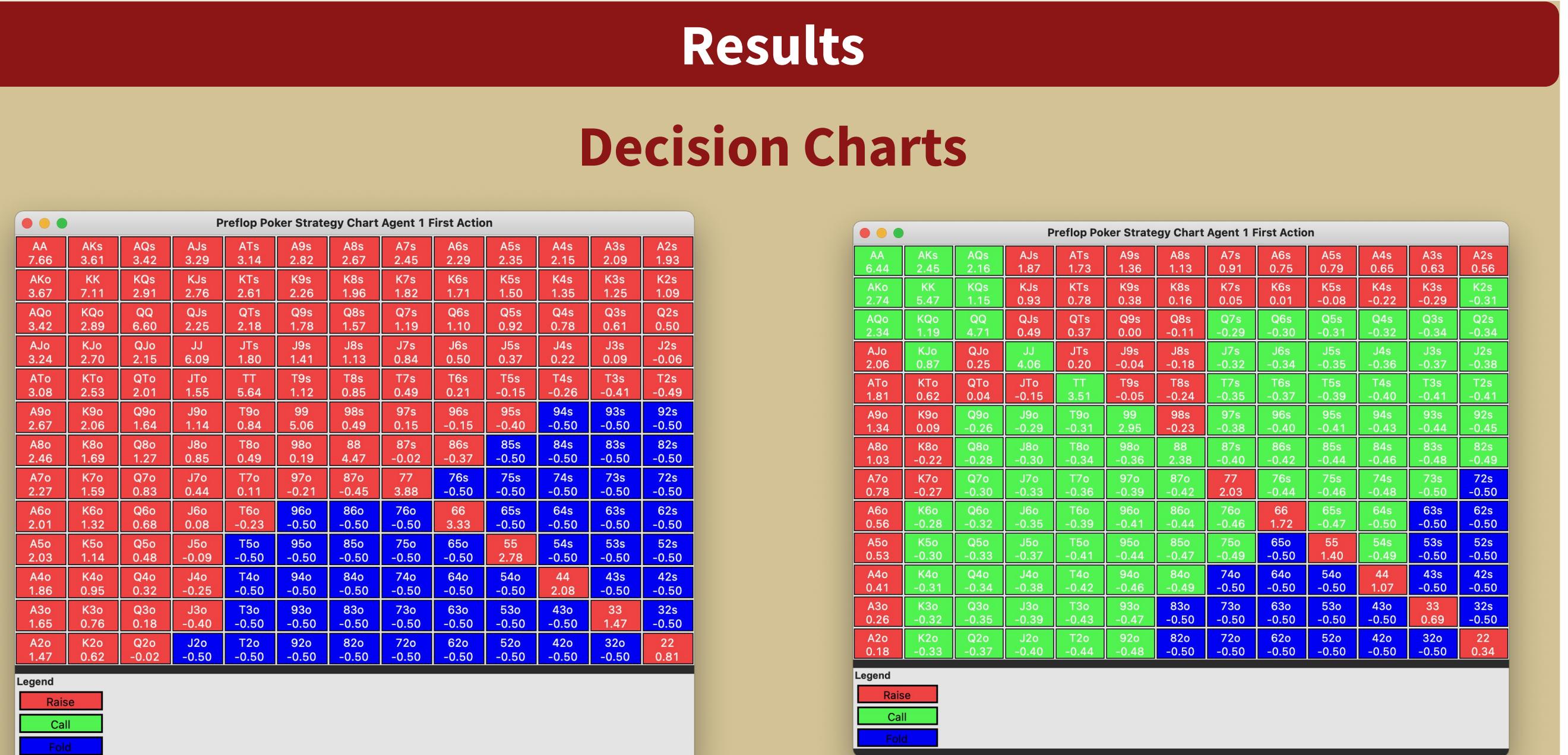


Figure 1: Agent 1 vs. Random

