

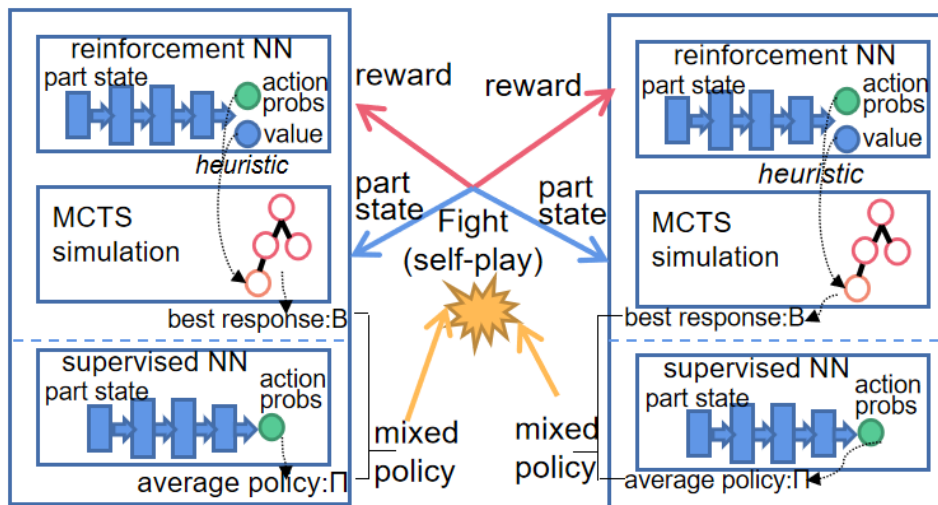
A Monte Carlo Neural Fictitious self-
play approach to approximate Nash
Equilibrium in imperfect-information
dynamic games

Li ZHANG, Yuxuan CHEN, Wei WANG, Ziliang HAN,
Shijian Li, Zhijie PAN, Gang PAN

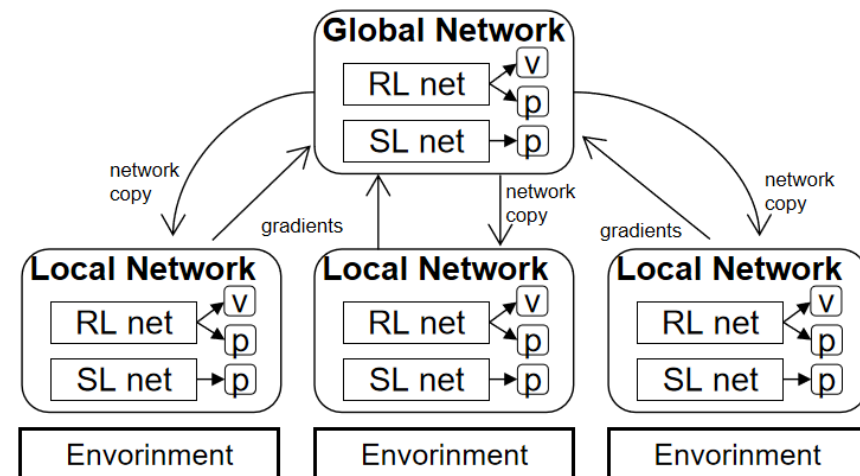
Frontiers of Computer Science, DOI: 10.1007/s11704-020-9307-6

Problems & Ideas

- Problems of Reinforcement Learning Algorithms on Imperfect-Information Dynamic Games
 - Many algorithms don't consider the Game Theory and are not guaranteed to approach optimized results (Nash Equilibrium, NE)
 - Neural Fictitious Self-Play (NFSP) can achieve approximate Nash Equilibrium, but may not converge in games of large state-action space
- Ideas: Combine Monte Carlo Tree Search (MCTS) with NFSP
 - MCTS controls the policy exploration to focus on higher-valued actions
 - MCTS can reduce optimality deficit, so it has lower exploitability to NE



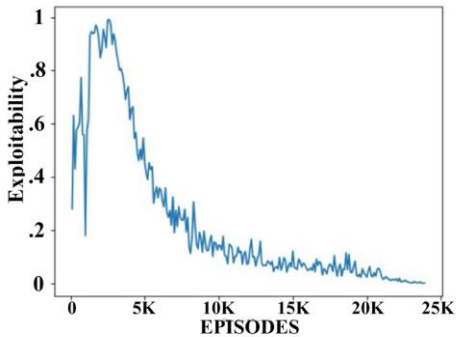
a) MC-NFSP Algorithm Overview



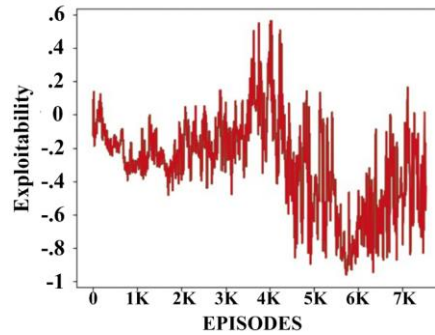
b) MC-NFSP can be further parallelized

Main Contributions

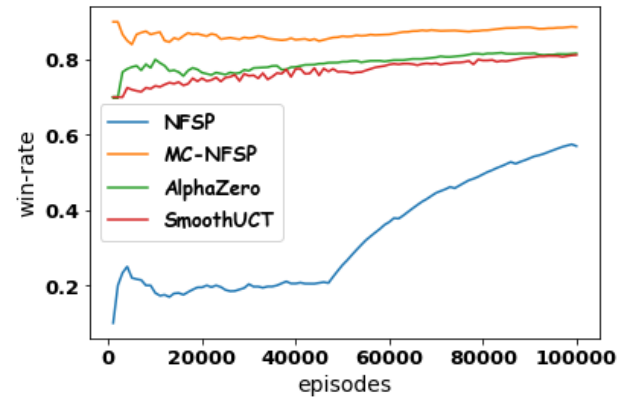
- Converge in games where NFSP can't
- Win recent algorithms in poker



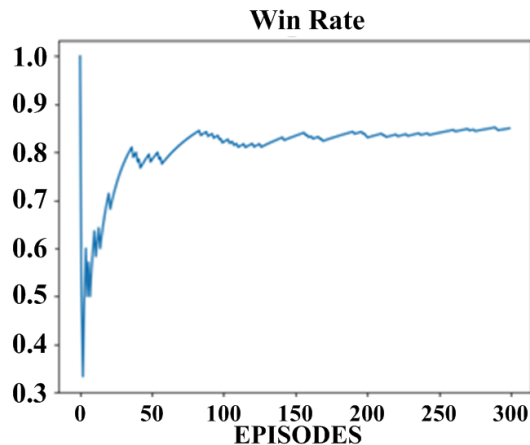
1. Exploitability of MC-NFSP in Othello



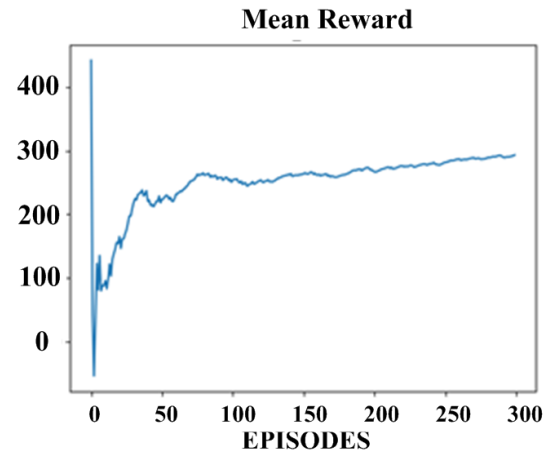
2. Exploitability of NFSP in Othello



- Win human experience in stochastic FPS games



1. Win Rate with Episodes



2. Mean Reward with Episodes