Super-Human AI for Strategic Reasoning: Beating Top Pros in Heads-Up No-Limit Texas Hold'em

Tuomas Sandholm
Carnegie Mellon University, Computer Science Department, Pittsburgh, PA, USA, and
Strategic Machine, Inc., Pittsburgh, PA, USA
sandholm@cs.cmu.edu

Abstract

Poker has been a challenge problem in AI and game theory for decades. As a game of imperfect information it involves obstacles not present in games like chess and Go, and requires totally different techniques. No program had been able to beat top players in large poker games. Until now! In January 2017, our AI, *Libratus*, beat a team of four top specialist professionals in heads-up no-limit Texas hold'em, which has 10^{161} decision points. This game is the main benchmark challenge for imperfect-information game solving. *Libratus* is the only AI that has beat top humans at this game. *Libratus* is powered by new algorithms in each of its three main modules:

1. computing blueprint (approximate Nash equilibrium) strategies before the event,
2. novel nested endgame solving during play, and
3. fixing its own strategy to play even closer to equilibrium based on what holes the opponents have been able to identify and exploit.

These domain-independent algorithms have potential applicability to a variety of real-world imperfect-information games such as negotiation, business strategy, cybersecurity, physical security, military applications, strategic pricing, product portfolio planning, certain areas of finance, auctions, political campaigns, and steering biological adaptation and evolution, for example, for medical treatment planning.

1 Introduction

Poker has been a challenge problem in game theory, mathematics, operations research, and AI for decades. For example, it was the only application in John Nash's seminal PhD dissertation that introduced the Nash equilibrium solution concept. As a game of imperfect information, poker involves obstacles not present in games like checkers, chess and Go. Therefore, it requires totally different techniques.

No program had been able to beat top humans in large poker games. Until now! In January 2017, our AI, *Libratus*, beat a team of four top specialist professionals in heads-up no-limit Texas hold'em, which has 10^{161} decision points. This game is the main benchmark challenge for imperfect-information game solving. *Libratus* is the only AI that has beat top humans at this game.

2 *Libratus's* main modules

Libratus is powered by new algorithms in each of its three main modules, as discussed in the following subsections.

2.1 Computing a blueprint strategy

Libratus has an algorithm for computing an approximate Nash equilibrium strategy (which is also an approximate minmax strategy here) to an abstraction of the game. It provides a high-level blueprint for the strategy of the AI. The new algorithm is an improved variant of the Monte Carlo Counterfactual Regret Minimization algorithm. The main new aspect involves sampling an agent's very-negative-regret actions with decreasing probabilities. This enhances the algorithm's speed significantly, thereby enabling finer-grained abstraction to be solved.

State-of-the-art game abstractions are imperfect-recall abstractions where the agent forgets some aspects of the past in order to computationally afford a more refined model of the present. Because of this, the game model has multiple paths to the same abstract state. This causes the different paths to "fight" over what should be done at that state, and this compromises solution quality. The new equilibrium-finding algorithm also mitigates this problem by, in effect, reducing the in-degree of the abstract states by discounting some of the paths.

2.2 Subgame solving

Libratus has a new subgame-solving algorithm which repeatedly calculates a more detailed strategy for the subgames reached as play progresses, using the above-mentioned blueprint strategy as guidance. The new aspects of this algorithm include the following.

- Safe subgame solving taking into account the opponent's mistakes in the hand so far. The subgame solver can afford to give back to the opponent the amount the opponent has given us in the game so far via mistakes—while still being totally safe (i.e., no worse than the precomputed Nash equilibrium approximation blueprint).

We use this observation to expand the strategy space that the subgame solver can safely optimize over, thereby giving it flexibility to play better than prior subgame solvers against other (non-mistake) hands the opponent might hold.

- Typically, subgame solving in imperfect-information games has been conducted only once when a subgame is reached. In contrast, Libratus resolves each *remaining* subgame after each opponent move in the subgame. This enables finer-grained abstractions and also avoids the downsides of reverse mapping the opponent's out-of-abstraction action to an in-abstraction action—because the opponent's exact action is added into the remaining subgame.
- Subgame solving starts earlier in the game (on any betting round where the pot is large enough, but no later than the beginning of the third betting round).
- There is no card abstraction in the subgame solving.
- Noise is added to the action abstraction before subgame solving. This makes *Libratus* difficult to play because it changes its bet sizing after every hand.

2.3 Self-improver

Libratus has a self-improving module that augments the precomputed blueprint over time to play even closer to Nash equilibrium based on what holes (out-of-abstraction actions, i.e., bet sizes in poker) the opponents have been able to identify and exploit. This is in stark contrast to prior approaches to learning in games, where the goal has typically been opponent modeling and exploitation—an approach that tends to open the agent up to counter-exploitation and causes the strategy to be opponent specific. In contrast, *Libratus's* self-improvements are universal.

3 Real-world applications

The algorithms are domain independent. They have potential applicability to a variety of real-world imperfect-information games such as negotiation, business strategy, cybersecurity, physical security, military applications, strategic pricing, product portfolio planning, certain areas of finance, auctions, political campaigns, and steering biological adaptation and evolution, for example, for medical treatment planning.

Acknowledgments

This invited talk is based on joint work with my PhD student Noam Brown. We thank MS student Ben Clayman for computing statistics of the play of our AIs against humans.

This material is based on research supported by the National Science Foundation under grant IIS-1617590 and the ARO under award W911NF-17-1-0082, as well as XSEDE computing resources provided by the Pittsburgh Supercomputing Center. The *Brains vs. AI* competition was sponsored by Carnegie Mellon University, Rivers Casino, GreatPoint Ventures, Avenue4Analytics, TNG Technology Consulting, Artificial Intelligence, Intel, and Optimized Markets, Inc.