Libratus: Beating Top Humans in No-Limit Poker

Noam Brown
Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15217
noamb@cs.cmu.edu

Tuomas Sandholm
Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15217
sandholm@cs.cmu.edu

Abstract

Heads-up no-limit Texas Hold'em is a primary benchmark challenge for AI. Due to the hidden information in poker, techniques used for games like chess and Go are ineffective. We present Libratus, the first AI to defeat top human professionals in no-limit poker. Libratus's features three main components: pre-computing a solution to an abstraction of the game which provides a high-level blueprint for how the AI should play, a new nested subgame-solving algorithm which repeatedly calculates a more detailed strategy as play progresses, and a self-improving module which augments the pre-computed blueprint based on opponent behavior.

1 Introduction

Recreational games have long been benchmarks to evaluate the progress of the field of AI. In recent years AIs have successfully beaten top humans in chess Campbell *et al.* [2002] and Go Silver *et al.* [2016]. But these games both share the feature of being perfect information, in which both players see all the information that is available and know the exact state of the game. Poker is an imperfect-information game, in which players have access to hidden information. Most real-world strategic interactions, such as negotiations, security situations, and auctions, involve some amount of hidden information and can be modeled as imperfect-information games. Dealing with hidden information requires drastically different approaches for AI because the exact state is no longer known. No-limit Texas Hold'em has long been the primary benchmark challenge for imperfect-information games. Until Libratus, no AI had beaten top humans in no-limit Texas Hold'em.

Libratus features three main components. The first is an abstraction and equilibrium-finding algorithm. This reduces the size of the game from 10^{161} different decision points to a more manageable 10^{12} by grouping similar states together. This *abstract* game is solved to determine a detailed strategy for the first two out of four rounds in no-limit Texas Hold'em, but only a blueprint of a strategy in the third and fourth rounds. The second component of Libratus solves a finer-grained abstraction of the remaining game, taking into account the blueprint of the strategy for the entire game, when the third round is reached. The final component improves the first-component abstraction over time by observing which situations Libratus frequently encounters during play against an opponent that are not accurately represented in its abstraction. The abstraction is augmented to better estimate the value of those situations.

2 Abstraction and Equilibrium Finding

Heads up no-limit Texas hold'em involves 10^{161} decision points, making it infeasible to pre-compute a strategy for the entire game. However, many strategically similar situations can be treated identically at only a small cost. For example, the difference between a bet of $\$500$ and a bet of $\$501$ is tiny. In *Libratus* we first generated an *abstraction* of the full game that includes only a few of the $20,000$ possible actions, and buckets strategically similar poker hands together to reduce the game's complexity. The abstraction is very fine-grained on the first two rounds, but coarse on the final two rounds. This is acceptable because the agent will never play according to the abstraction strategy it computes in the final two rounds. Instead, it will generate an improved strategy in real time using subgame solving (explained next) when it reaches that point in the game.

The abstraction is solved via a distributed form of *Monte Carlo Counterfactual Regret Minimization* (MCCFR) Zinkevich *et al.* [2007]; Lanctot *et al.* [2009]. MCCFR is an algorithm that repeatedly traverses a sampled portion of the game tree and independently minimizes regret at every decision point encountered. Libratus improves upon the vanilla MCCFR algorithm by exploring actions in the game tree less frequently if those actions have performed poorly in the past.

3 Nested Subgame Solving

An imperfect-information subgame cannot be solved in isolation, because the Nash equilibrium strategy in other subgames affects the optimal strategy in the subgame that is reached during play. This is an important difference from perfect-information games. Nevertheless, we can approximate a good strategy in a subgame in real time if we have a good estimate of the value of reaching all subgames in an equilibrium. The first module estimated this value for every subgame. Using these subgame values as input, subgame solving creates and solves a finer-grained abstraction in the subgame that is reached.

This finer-grained abstraction does not use any card abstraction and uses a dense action abstraction. Rather than apply action translation, Libratus instead constructs and solves a new subgame every time an opponent chooses an action that is not in the finer-grained abstraction (in practice, it constructs a new subgame every time the opponent bets). This allows it to avoid the rounding error due to action translation and leads to much lower exploitability Brown and Sandholm [2017].

4 Self-Improving

Libratus uses a dense action abstraction on the first two rounds of no-limit Texas Hold'em, but if the opponent does not bet an amount that is in the abstraction then the bet is rounded to a nearby size that is in the abstraction. This leads to the strategy the AI uses to be slightly off. To improve upon this, every night during the competition the AI determined a small number of actions to add to the abstraction. The choice of actions was based on a combination of which actions the opponents were choosing most frequently, and how far those actions were from an existing action in the abstraction. Once an action was selected, a strategy was calculated for it in a similar manner to subgame solving, described in Section 3. From that point on, if that action (or a nearby one) were chosen by an opponent, then the newly solved subgame strategy would be used. This had the effect of reducing over time the rounding error due to action translation.

5 Demonstration

Attendees will have the opportunity to play against Libratus in heads up no-limit Texas Hold'em via the same web-based user interface that the top professional players used in their competition against the AI. One computer will be available for an attendee to use, while others can watch the game on a large panel monitor. A Youtube video of the precursor bot (played through the same user interface that will be available to NIPS attendees) can be seen at https://www.youtube.com/watch?v=phRAyF1rq0I.

References

Noam Brown and Tuomas Sandholm. Safe and nested endgame solving for imperfect-information games. In *AAAI-17 Workshop on Computer Poker and Imperfect Information Games*, 2017.

Murray Campbell, A Joseph Hoane, and Feng-Hsiung Hsu. Deep Blue. *Artificial intelligence*, 134(1-2):57–83, 2002.

Marc Lanctot, Kevin Waugh, Martin Zinkevich, and Michael Bowling. Monte Carlo sampling for regret minimization in extensive games. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS)*, pages 1078–1086, 2009.

David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587):484–489, 2016.

Martin Zinkevich, Michael Johanson, Michael H Bowling, and Carmelo Piccione. Regret minimization in games with incomplete information. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS)*, pages 1729–1736, 2007.