Endgame Solving: The Surprising Breakthrough that Enabled Superhuman Two-Player No-Limit Texas Hold ’em Play

Sam Ganzfried
School of Computing and Information Sciences
Florida International University
sganzfri@cis.fiu.edu

Abstract

While computer programs for no-limit Texas hold ’em have become steadily stronger over the past decade, the last several years have seen a sharp improvement that has enabled agents to defeat the best humans in the world. While several advances have fueled this final spike, the one that has been identified as the most significant had previously been written off due to the fact that it has no theoretical guarantee, highlighted by a simple example.

Sequential games of perfect information can be solved in linear time by a straightforward backward induction procedure in which solutions to endgames are propagated up the tree. However, this fails in imperfect-information games because different endgames can contain nodes that belong to the same information set and cannot be treated independently.

Definition 1. *E is an* endgame[1] *of game G if:*

1. *The set of E's nodes is a subset of the set of G's nodes.*
2. *If s' is a child of s in G and s is a node in E, then s' is also a node in E.*
3. *If s is in the same information set as s' in G and s is a node in E, then s' is also a node in E.*

For example, we can consider endgames in poker where several rounds of betting have taken place and several public cards have already been dealt. We can assume players have a joint distribution of private information from nodes prior to the endgame (i.e., the *trunk*) that are induced from precomputed base approximate-equilibrium strategies using Bayes’ rule. Given this distribution as input, we can then solve individual endgames in real time more accurately.

Unfortunately, this approach has fundamental flaws. It turns out that even if we computed an exact equilibrium in the trunk (which is an unrealistically optimistic assumption in large games) and in the endgame, the combined strategies for the trunk and endgame may fail to be an equilibrium in the full game. One obvious reason for this is that the game may contain many equilibria, and we might choose one for the trunk that does not match up correctly with the one for the endgame; or we may compute different equilibria in different endgames that do not balance appropriately. However, Proposition 1 shows that it is possible for this procedure to output a non-equilibrium strategy profile in the full game even if the full game has a unique equilibrium and a single endgame.

Proposition 1. *There exist games—even with a unique equilibrium and a single endgame—for which endgame solving can produce a non-equilibrium strategy profile.*

Proof. Consider a sequential version of Rock-Paper-Scissors where player 1 acts, then player 2 acts without observing player 1’s action. This game has a single endgame—when it is player 2’s turn—and a unique equilibrium—where each player plays each action with probability $\frac{1}{3}$. Now suppose we restrict player 1 to follow the equilibrium in the trunk. Any strategy for player 2 is an equilibrium in the endgame, because each one yields expected payoff 0. In particular, suppose our equilibrium solver outputs the pure strategy Rock. This is clearly not an equilibrium of the full game. □

Some early research used endgame solving for limit Texas hold ’em agents out of necessity due to limited scalability of existing approaches [Gilpin and Sandholm, 2006; 2007]. But for the next several years it was abandoned in favor of offline approaches that solve abstracted versions of the entire game due to the shortcomings described above. However, despite the adequacy of these holistic approaches, endgame solving was reinvestigated and applied to no-limit Texas hold ’em in 2013 due to the potential benefits of focused computation on the portion of the game tree that has been reached [Ganzfried and Sandholm, 2013; 2015]. This approach was used by the agent Claudico that competed in the inaugural Brains vs. AI competition against the strongest human two-player no-limit Texas hold ’em specialists in the world [Ganzfried, 2015]. In fact, the best human player in the world, Doug Polk, has relayed to me in personal communication that the final round strategy of Claudico computed by the endgame solver was the strongest component. Fueled by this promise, there has been a flurry of subsequent research further exploring this new paradigm [Burch *et al.*, 2014; Moravcik *et al.*, 2016], culminating in two agents DeepStack and Libratus that were recently able to successfully defeat human professional players [Brown and Sandholm, 2017; Moravcík *et al.*, 2017]. These agents both apply endgame solving in very different ways: Libratus used a supercomputer to solve both the turn and river rounds in real time while DeepStack viewed all rounds as independent endgames with leaf payoff values estimated using deep learning.

[1] Note that endgame differs from the traditional notion of *subgame*, which requires a root node that is alone in its information set.

References

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

[Burch *et al.*, 2014] Neil Burch, Michael Johanson, and Michael Bowling. Solving imperfect information games using decomposition. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2014.

[Ganzfried and Sandholm, 2013] Sam Ganzfried and Tuomas Sandholm. Improving performance in imperfect-information games with large state and action spaces by solving endgames. In *AAAI Workshop on Computer Poker and Incomplete Information*, 2013.

[Ganzfried and Sandholm, 2015] Sam Ganzfried and Tuomas Sandholm. Endgame solving in large imperfect-information games. In *Proceedings of the International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2015.

[Ganzfried, 2015] Sam Ganzfried. Reflections on the first man vs. machine no-limit Texas hold 'em competition. *SIGecom Exchanges*, 4.2, 2015.

[Gilpin and Sandholm, 2006] Andrew Gilpin and Tuomas Sandholm. A competitive Texas Hold'em poker player via automated abstraction and real-time equilibrium computation. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, 2006.

[Gilpin and Sandholm, 2007] Andrew Gilpin and Tuomas Sandholm. Better automated abstraction techniques for imperfect information games, with application to Texas Hold'em poker. In *Proceedings of the International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2007.

[Moravcik *et al.*, 2016] Matej Moravcik, Martin Schmid, Karel Ha, Milan Hladik, and Stephen J. Gaukrodger. Solving imperfect information games using decomposition. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2016.

[Moravcík *et al.*, 2017] Matej Moravcík, Martin Schmid, Neil Burch, Viliam Lisý, Dustin Morrill, Nolan Bard, Trevor Davis, Kevin Waugh, Michael Johanson, and Michael H. Bowling. Deepstack: Expert-level artificial intelligence in no-limit poker. *CoRR*, abs/1701.01724, 2017.