A Demonstration of the Polaris Poker System

Michael Bowling, Nicholas Abou Risk, Nolan Bard, Darse Billings, Neil Burch, Joshua Davidson, John Hawkin, Robert Holte, Michael Johanson, Morgan Kan, Bryce Paradis, Jonathan Schaeffer, David Schnizlein, Duane Szafron, Kevin Waugh, Martin Zinkevich

Department of Computing Science
University of Alberta
Edmonton, AB, Canada T6G 2E8
poker.cs.ualberta.ca

ABSTRACT

Poker games provide a useful testbed for modern Artificial Intelligence techniques. Unlike many classical game domains such as chess and checkers, poker includes elements of imperfect information, stochastic events, and one or more adversarial agents to interact with. Furthermore, in poker it is possible to win or lose by varying degrees. Therefore, it can be advantageous to adapt ones' strategy to exploit a weak opponent. A poker agent must address these challenges, acting in uncertain environments and exploiting other agents, in order to be highly successful. Arguably, poker games more closely resemble many real world problems than games with perfect information. In this brief paper, we outline Polaris, a Texas Hold'em poker program. Polaris recently defeated top human professionals at the Man vs. Machine Poker Championship and it is currently the reigning AAAI Computer Poker Competition winner in the limit equilibrium and no-limit events.

1. INTRODUCTION

Extensive games are useful for modelling multi-agent environments with imperfect information. They encapsulate how multiple agents can interact with a stochastic environment where each agent is privileged to its own private information and has its own definition of utility. These games can be used to model many real world problems and hence their study is of high practical value. A popular extensive game used as a testbed for Artificial Intelligence research is poker. Poker games include imperfect information, stochastic events, and adversaries to interact with. Though many types of extensive games can be directly solved, it is easy to devise poker games that are far too large for these techniques to be practical. In fact, many poker games played by enthusiasts have far too many game states to even record a single strategy. For these reasons, there are many practical challenges in the domain of poker that make it a compelling test domain.

Of all poker games, variants of Texas Hold'em are currently the most popular among human players. Texas Hold'em games are structured as follows. Prior to the hand, one player is designated the *small blind* and another the *big blind*. Each of these players make forced bets prior to play. All players are then dealt two private cards face down. A round of betting commences starting with the player to the left of the big blind. The structure of the betting round depends on the type of game being played. Two common types of betting rounds are *limit* and *no-limit*. After the betting round is finished, three community cards are placed face up for everyone to use to make a hand. This round is called the *flop*. The flop betting round starts from the small blind. Two more rounds follow, the *turn* and the *river*, with one community card each. After the river betting round finishes, the player with the best five card poker hand wins the pot. Texas Hold'em refers to the described game when limit betting is used.

Over the last three years, the AAAI Computer Poker Competition[5] has been used as the measuring stick for progress in the domain of poker (in particular, variants of Texas Hold'em poker). Programs made by the University of Alberta's Computer Poker Research Group have consistently performed very well in this competition. In the 2008 competition, Polaris, one of the group's programs, finished first in the limit equilibrium and no-limit events. Both of these events feature two player games, which are referred to as heads-up games. Another entry from the group, Poki, finished first in the limit ring event, which was a six player game. Just prior to the AAAI Competition, Polaris defeated a team of six professional poker players at the second Man vs. Machine Poker Championship at heads-up Texas Hold'em. This is the first time that a poker program has statistically defeated a group of human professionals.

In this brief paper, we will overview the underlying design of Polaris. The system itself consists of two main parts. First, we compute a variety of strong poker strategies. This computation can take weeks and is done prior to play. Second, an online component intelligently selects which strategy to play against a particular opponent. This component can quickly converge onto which precomputed strategy performs best against an opponent. The combination of these systems allows Polaris to rapidly select a strong poker strategy that can effectively exploit an opponent. The demonstration accompanying this paper will allow observers to test their luck against Polaris, just as the professional players did at the Man vs. Machine Competition.

2. COMPUTING STRATEGIES

Computing a good poker strategy is a complicated process. For instance, in a game with imperfect information, it is mathematically correct to bluff some percentage of the time. That is, by making a suboptimal, but deceptive, action with some probability, an opponent should adapt by proceeding to showdowns with more mediocre hands. If a strategy never bluffs, an observant opponent can more easily identify situations where it is likely have a strong hand. Bluffing is only one aspect required of a strong strategy.

Cite as: A Demonstration of the Polaris Poker System, Michael Bowling, Nicholas Abou Risk, Nolan Bard, Darse Billings, Neil Burch, Joshua Davidson, John Hawkin, Robert Holte, Michael Johanson, Morgan Kan, Bryce Paradis, Jonathan Schaeffer, David hnizlein, Duane Szafron, Kevin Waugh, Martin Zinkevich, *Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009),* Decker, Sichman, Sierra and Castelfranchi (eds.), May, 10–15, 2009, Budapest, Hungary, pp. 1391–1392

Unfortunately, directly writing and tuning a program with all this knowledge is difficult. As a result, it is hard to create a successful program based on a strategy from an expert.

Fortunately, with recent computational advances in solving for equilibria in large extensive games, we can rely on game theory to learn strong poker strategies that perform at, or above, the level of top human players. Texas Hold'em itself is an extensive game, but it is far to large to solve directly. We rely on abstraction techniques to create smaller extensive games that we hope model the original game reasonably well. These smaller extensive games can be solved using standard techniques to obtain a strategy that can be used to play the original game.

Polaris uses counterfactual regret minimization[4] to solve abstract poker games. Using this technique, we can solve games with as many as 10^{12} states with a few days of computation. Unfortunately, since the program learns its own strategy, we cannot easily adjust its style of play. By using *tilts* of the game's payoffs, *i.e.* modified payoffs, we are able to shift the strategy produced and, as a side effect, alter its style. This allows us to compute a variety of individually strong poker strategies, each of which likely performs better against a certain type of opponent. For more information on the abstraction techniques used and tilts, please see *Robust Strategies and Counter-Strategies: Building a Champion Level Computer Poker Program*[2].

3. EXPLOITING AN OPPONENT

If the objective is merely to win in expectation, then playing an equilibrium strategy is sufficient regardless of ones' opponent. To win with the highest expectation is an entirely different challenge. Many previous approaches to solving this problem attempted to use past play to model an opponent's actions. Using this model, one could then devise a strategy tailored to defeating that specific opponent. Unfortunately, this approach has many practical downsides when applied to Texas Hold'em. First, it often requires an extremely large amount of training data to come up with an accurate model of an opponent. Most matches played by humans consist of hundreds of hands, whereas a good modelling technique can require millions of hands to be accurate. Second, a model built from past play is likely to be inaccurate in uncommon situations. In other words, if you were to watch two humans play, most of the hands are biased towards a small subset of the game states. Most techniques to build counter-strategies from a model require at least some knowledge of these uncommon situations. Third, these counter-strategies can be extremely brittle. That is, if the model is inaccurate, used against a different opponent, or against an adapting opponent, then the strategy is likely to do poorly. Many of these detriments are being addressed by recent work by the Computer Poker Research Group[3], but currently Polaris takes a different approach to opponent modelling.

Instead of attempting to model an opponent, Polaris consists of a number of strong poker strategies. Each strategy on its own is the solution to a perturbed (*i.e.*, tilted) abstract poker game and therefore each has the advantages of a strong Game Theory-based strategy. These strategies play differently and these different styles can perform better or worse against a particular opponent. When playing against an unknown opponent, any of the strategies should perform well, but in an online setting we can use techniques to evaluate which particular strategy we believe will do the best.

This problem of selecting a strategy is akin to a bandit problem, but we can perform off policy estimates of how a strategy that did not even play would have performed. More specifically, after one single hand of play by a single strategy, we can evaluate the performance of all of our strategies. This estimate is done using importance sampling[1]. Importance sampling allows us to, in essence, generate an observation for each of our strategies from a single hand. These extra observations allow us to more rapidly determine which strategy is likely the best against a particular opponent. Empirically, we have found that this technique can quickly converge (50 hands) to the best strategy available, which makes it practical to use against human and computer opponents. Using hindsight analysis, we can verify that strategy selected was indeed the correct choice.

4. CONCLUSION

Polaris is a world class computer Texas Hold'em program. It makes use of modern abstraction and equilibria solving techniques to create strong poker strategies. These strategies are then selected during play to attempt to exploit weaker opponents. This adaptive system has proven itself as a world class Texas Hold'em player and as the individual components are further enhanced, Polaris can only improve.

Acknowledgments

Polaris is a product of the ongoing work of the Computer Poker Research Group at the University of Alberta. The hard work of the group's members, both past and present, is responsible for Polaris' achievements.

5. REFERENCES

[1] M. Bowling, M. Johanson, N. Burch, and D. Szafron. Strategy evaluation in extensive games with importance sampling. In *Proceedings of the Twenty-Fifth International Conference on Machine Learning (ICML)*, pages 72–79, 2008.

[2] M. Johanson. Robust strategies and counter-strategies: Building a champion level computer poker player. Master's thesis, University of Alberta, 2007.

[3] M. Johanson, M. Zinkevich, and M. Bowling. Computing robust counter-strategies. In *Advances in Neural Information Processing Systems 20 (NIPS)*, 2008. To appear (8 pages). A longer version is available as a University of Alberta Technical Report, TR07-15.

[4] M. Zinkevich, M. Johanson, M. Bowling, and C. Piccione. Regret minimization in games with incomplete information. In *Advances in Neural Information Processing Systems 20 (NIPS)*, 2008.

[5] M. Zinkevich and M. Littman. The AAAI computer poker competition. *Journal of the International Computer Games Association*, 29, 2006. News item.