Research Statement

Sam Ganzfried

"Big data" has revolutionized every industry. Massive datasets on the human genome have enabled significant advances in DNA sequencing, which will lead to breakthroughs in disease diagnosis and treatment; data on users' preferences have enabled recommendation systems that provide highly accurate predictions for purchases of new products; and traders capitalize on historical datasets to effectively predict future price trajectories of stocks.

The field of "data science" has produced valuable tools that have enabled these dramatic advances. These tools, which draw primarily from the fields of machine learning and statistics, extract as much *signal* as possible from the enormous, complex, and noisy datasets, and produce sophisticated *inferences* from historical data to make highly accurate *predictions* about the future.

We are now in the midst of an arms race that has fueled a massive bubble permeating industry and academia. Researchers and practitioners of "data science" are slugging it out to obtain larger and larger datasets and more efficient algorithms that can extract every last ounce of signal from them. It is inevitable that this bubble will burst, unless major measures are taken. Soon every bit of useful information will be sucked out of your purchase history on Amazon and Netflix to predict what you will enjoy next. Algorithmic traders have already begun to feel the pressure of this arms race; strategies that traded on a millisecond and even microsecond time scale based on pure speed advantages and statistical models have begun to suffer in favor of longer-term strategies that take into account fundamental *economic* principles. The marginal returns will soon be felt everywhere, and the bubble will burst dramatically, unless the approaches are combined with core theoretical principles that can withstand the test of time.

In many important situations that involve multiple self-interested agents, the goal is not simply to make a prediction about the future from the data: it is to make better *strategic decisions* based on these complex inferences and predictions. Data-driven predictions alone do not tell us what the optimal security threshold should be to prevent phishing attacks by hackers, how to optimally randomize security screens in airports, or whether to call a large all-in bet on the river with two pair; we must also reason about the *strategy* the opponents are employing, reason about the reasoning they are employing about our own strategy, and so on. In order to ultimately make better-informed *decisions* in complex multiagent environments, we must incorporate the tools of economic theory, and particularly, *game theory*.

Take no-limit Texas hold 'em poker for example. The game has 10^{165} states in its game tree, and making good strategic decisions is on par with reasoning in domains with extremely complex datasets. A "big data" approach would be to apply sophisticated machine learning algorithms to databases of historical data from prior games, to predict how a future opponent will play in similar situations; then we could exploit such an opponent by playing an appropriate response to this strategy. However, this approach has limitations. It assumes that the future opponents encountered are identical to the past ones depicted in the database; if the training and testing data are from completely different pools of the population, then this approach would have no performance guarantees. As agents are constantly improving and modifying their play, we would expect, for example, the current year's poker competition agents to be significantly stronger than prior agents: even the best approaches for learning from historical data could fare quite poorly against stronger opposition than what was trained on.

It would be preferable to employ a more *robust* approach that is not entirely dependent on a particular dataset of historical play. We would like to perform well against a wide variety of agents, including agents that can be significantly stronger than the ones in the dataset, who may be deceptive and adaptive.

In theory, there exists a single strategy for this game (and for any two-player *zero-sum* (i.e., competitive) game) that would guarantee being unbeatable against all opposing agents, regardless of their skill level, level of deception, or, generally, their strategy. This result is due to the Minimax Theorem, one of the fundamental results in game theory, and the “optimal” strategy is called a *Nash equilibrium*. If we were able to compute a Nash equilibrium for two-player no-limit Texas hold ’em, then we would guarantee that against *any* opponent we would either win or tie (in expectation). This would be true tomorrow, two weeks from now, or fifty years from now.

From a complexity-theoretic perspective, computing this strategy is easy; there exists a polynomial-time algorithm based on a linear programming formulation. However, this algorithm only scales to games with 10^8 states. More recently algorithms have been developed for approximating equilibrium strategies (they converge to equilibrium in the limit) that scale to 10^{15} states [1]. However, even this is a far cry from 10^{165}, the size of the version of two-player no-limit Texas hold ’em played in the AAAI Annual Computer Poker Competition. Approximating Nash equilibrium strategies in a domain of that magnitude involves, at a minimum, approaches for approximating the full 10^{165} game tree with a significantly smaller game of 10^{15} states that retains much of the strategic structure of the original game (i.e., *automated abstraction algorithms* [1, 11, 12]), approaches for interpreting actions for the opponent that have been removed from the abstraction [10], and additional approaches for extrapolating the equilibrium strategies from the abstract game to the full game [1, 14].

These approaches produced two-player no-limit Texas hold ’em Tartanian7 that won the most recent AAAI Annual Computer Poker Competition, defeating each opposing agent with statistical significance (16 agents were submitted). An improved agent called Claudico competed against the strongest human players in the 2015 Brains vs. Artificial Intelligence competition. This was the first ever man vs. machine no-limit Texas hold ’em competition; the humans won by a margin that was statistically significant at the 90% confidence level, but not at the 95% level. An important feature of Claudico was a novel approach for computing strategies in real time in the portion of the game tree we have reached to a higher degree of accuracy [12] (the abstraction and equilibrium approaches described above are performed offline, in advance of gameplay). Doug Polk, a participant in the competition and widely regarded as the best two-player no-limit Texas hold ’em player in the world, commented that the “endgame solver” was the strongest part of the agent. I have recently written an article that puts the event into perspective within both the poker and academic communities, highlights the strengths and weaknesses of Claudico, and describes the key takeaways and future research directions, both in terms of computer poker and more broadly [5]. While the main goal was to produce the strongest possible poker agent, there are deeper theoretical questions related to each component of the agent. Endgame solving has been proven to have theoretical guarantees in certain games while it can lead to strategies with high exploitability in others (even if the full game has a single Nash equilibrium and just a single endgame is considered) [12]; it would be interesting to prove theoretical bounds on its performance on interesting game classes, perhaps classes that include variants of poker. Empirically the approach appears to be very successful on poker despite its lack of theoretical guarantees. The main abstraction algorithms that have been successful in practice are heuristic and have no theoretical guarantees (it is extremely difficult to prove meaningful theoretical guarantees when approximating a game with 10^{165} states by one with 10^{15} states). Recent work has presented an abstraction algorithm with bounds on the solution quality; however, it only scales to a tiny poker game with a five-card deck. It would be very interesting to bridge this gap between heuristics that work well in practice for large games with no theoretical guarantees, and the approaches with theoretical guarantees that have more modest scalability. There are also many exciting theoretical questions related to the action translation [10] and post-processing approaches [14].

Scalable algorithms for computing Nash equilibria have diverse applications, including cybersecurity (e.g., determining optimal thresholds to protect against phishing attacks), business (e.g., auctions and negotiations), national security (e.g., computing strategies for officers to protect airports), and medicine. For medicine, algorithms that were created in the course of research on poker have been applied to compute robust policies for diabetes management [2]; recently it has been proposed that equilibrium-finding algorithms are applicable to the problem of treating diseases such as the HIV virus that can mutate adversarially [15]. My research has been cited by and applied to research on jamming attacks and cyber security, trading agent design, security games for the protection of resources, dynamic resource allocation, sequential auctions, automated guidance for taxi service, robot planning, disease management, and emergency response.

This is not to say that the "game-theoretic" approach is wholly free of flaws. For one, the Minimax Theorem does not apply to games that are not zero sum or have more than two agents. These games can have many equilibria, each assigning different payoffs to the agents; if the opponents do not follow the equilibrium strategy that we have computed, then we can perform arbitrarily poorly. Furthermore, computing a Nash equilibrium in these game classes is challenging computationally (it is PPAD-complete and widely conjectured that no efficient algorithms exist). Despite this worst-case hardness result, I have developed approaches that provably computed an ϵ-equilibrium for very small ϵ in a three-player poker tournament [6, 7]. I have proven that some of these approaches can only converge to a Nash equilibrium (though they may not converge at all). However, even these very close approximations of Nash equilibrium have no performance guarantees against unknown opponents.

Even in two-player zero-sum games, the Nash equilibrium is not the end of the story (even if we are able to compute one exactly without requiring approximation). Against suboptimal opponents who are not playing an equilibrium strategy, we can often obtain a significantly higher payoff than the value of the game by learning to exploit their mistakes as opposed to following a static equilibrium strategy; for instance, if the opponent has played Rock in each of the first thousand iterations of rock-paper-scissors, it seems desirable to put additional probability mass on Paper beyond the equilibrium value of $\frac{1}{3}$.

Thus, for all game classes, there is a need to capitalize on the "big data" approach and learn from historical data. In order to make robust decisions in large-scale multiagent environments, we must integrate the "game-theoretic" and "big data" approaches, creating agents that obtain high payoff against weak opponents who make mistakes, yet also perform well against strong opponents who may be dynamic and adaptive. These agents will stand the test of time: they will take advantage of crucial data when it is available, though they will not be overfit to the latest trend due to their underlying foundations based on economic principles. This is the thesis statement of my dissertation [4]. I have already made significant progress in achieving this ambitious goal, and have developed approaches for robustly integrating learning with the game-theoretic approaches. I developed an algorithm that is successfully able to exploit weaknesses of opponents in extremely large imperfect-information games after only a small number of interactions [9]. It uses an approximate Nash equilibrium strategy as the prior, which is updated based on observations of opponents' play. This has led to a large performance improvement against a variety of opponents in two-player limit Texas hold 'em. I have also developed new approaches with theoretical guarantees even against strong deceptive opponents [13]. I have shown that in certain games it is actually possible to deviate from repeatedly playing a one-shot equilibrium strategy in order to exploit perceived weaknesses of an opponent, while still guaranteeing at least the value of the game in expectation against any opponent. Recently I have developed the first exact algorithm for opponent exploitation in imperfect-information games in a Bayesian setting, which utilizes the most natural prior distribution (Dirichlet) based on historical data [3].

At the end of the day we would like to enable humans to make important decisions, not produce massive binary strategy files that are only intelligible to computers. I have designed an algorithm that exploits qualitative information about equilibrium structure to improve the speed of equilibrium finding and produces strategies that are more human understandable [8]. I showed that for the final round of limit Texas hold 'em, equilibrium strategies for any input hand distribution will conform to one of three relatively simple qualitative action structures. Extracting visually-appealing representations from the massive files is an important step towards the goal of enabling humans to make robust decisions in complex multiagent environments.

References

[1] Noam Brown*, Sam Ganzfried*, and Tuomas Sandholm. Hierarchical abstraction, distributed equilibrium computation, and post-processing, with application to a champion no-limit Texas Hold'em agent. In *Proceedings of the International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2015. *Listed alphabetically.

[2] Katherine Chen and Michael Bowling. Tractable objectives for robust policy optimization. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS)*, 2012.

[3] Sam Ganzfried. Bayesian opponent exploitation in imperfect-information games. Manuscript, 2015.

[4] Sam Ganzfried. *Computing Strong Game-Theoretic Strategies and Exploiting Suboptimal Opponents in Large Games*. PhD thesis, Carnegie Mellon University, 2015. (Also published as Technical Report CMU-CS-15-104, Computer Science Department, Carnegie Mellon University).

[5] Sam Ganzfried. Reflections on the first man vs. machine no-limit Texas hold 'em competition. SIGecom Exchanges, Volume 14.2, 2015. Feature article.

[6] Sam Ganzfried and Tuomas Sandholm. Computing an approximate jam/fold equilibrium for 3-player no-limit Texas Hold'em tournaments. In *Proceedings of the International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2008.

[7] Sam Ganzfried and Tuomas Sandholm. Computing equilibria in multiplayer stochastic games of imperfect information. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*, 2009.

[8] Sam Ganzfried and Tuomas Sandholm. Computing equilibria by incorporating qualitative models. In *Proceedings of the International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2010.

[9] Sam Ganzfried and Tuomas Sandholm. Game theory-based opponent modeling in large imperfect-information games. In *Proceedings of the International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2011.

[10] Sam Ganzfried and Tuomas Sandholm. Action translation in extensive-form games with large action spaces: Axioms, paradoxes, and the pseudo-harmonic mapping. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2013.

[11] Sam Ganzfried and Tuomas Sandholm. Potential-aware imperfect-recall abstraction with earth mover's distance in imperfect-information games. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2014.

[12] 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.

[13] Sam Ganzfried and Tuomas Sandholm. Safe opponent exploitation. *ACM Transactions on Economics and Computation (TEAC)*, 2015. Special issue on selected papers from EC-12. Early version appeared in EC-12.

[14] Sam Ganzfried, Tuomas Sandholm, and Kevin Waugh. Strategy purification and thresholding: Effective non-equilibrium approaches for playing large games. In *Proceedings of the International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2012.

[15] Tuomas Sandholm. Steering evolution strategically: Computational game theory and opponent exploitation for treatment planning, drug design, and synthetic biology. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2015. Senior Member Track, Blue Skies Subtrack.