Successful Nash Equilibrium Agent for a 3-Player Imperfect-Information Game

Sam Ganzfried

www.ganzfriedresearch.com

The main mathematical result is the proof of the existence in any game of at least one equilibrium point. Other results concern the geometrical structure of the set of equilibrium points of a game with a solution, the geometry of sub-solutions, and the existence of a symmetrical equilibrium point in a symmetrical game.

As an illustration of the possibilities for application a treatment of a simple three-man poker model is included.

Table of Contents

A Three-Man Poker Game

As an example of the application of our theory to a more or less realistic case we include the simplified poker game given below. The rules are as follows:

(1) The deck is large, with equally many high and low cards, and a hand consists of one card.

(2) Two chips are used to ante, open, or call.

(3) The players play in rotation and the game ends after all have passed or after one player has opened and the others have had a chance to call.

(4) If no one bets the antes are retrieved.

(5) Otherwise the pot is divided equally among the highest hands which have bet.

We find it more satisfactory to treat the game in terms of quantities we call "behavior parameters" than in the normal form of "Theory of Games and Economic Behavior." In the normal form representation two mixed strategies of a player may be equivalent in the sense that each makes the individual choose each available course of action in each particular situation requiring action on his part with the same frequency. That is, they represent the same behavior pattern on the part of the in-

Scope and applicability of game theory

- Strategic multiagent interactions occur in all fields
  - Economics and business: bidding in auctions, offers in negotiations
  - Political science/law: fair division of resources, e.g., divorce settlements
  - Biology/medicine: robust diabetes management (robustness against "adversarial" selection of parameters in MDP)
  - Computer science: theory, AI, PL, systems; national security (e.g., deploying officers to protect ports), cybersecurity (e.g., determining optimal thresholds against phishing attacks), internet phenomena (e.g., ad auctions)

Game theory background

	rock	paper	scissors
Rock	0,0	-1, 1	1, -1
Paper	1,-1	0, 0	-1,1
Scissors	-1,1	1,-1	0,0

- Players
- Actions (aka pure strategies)
- Strategy profile: e.g., (R,p)
- Utility function: e.g., $u_1(R,p) = -1$, $u_2(R,p) = 1$

Zero-sum game

	rock	paper	scissors
Rock	0,0	-1, 1	1, -1
Paper	1,-1	0, 0	-1,1
Scissors	-1,1	1,-1	0,0

- Sum of payoffs is zero at each strategy profile: e.g., $u_1(R,p) + u_2(R,p) = 0$
- Models purely adversarial settings

Mixed strategies

- Probability distributions over pure strategies
- E.g., R with prob. 0.6, P with prob. 0.3, S with prob. 0.1

Best response (aka nemesis)

- Any strategy that maximizes payoff against opponent's strategy
- If P2 plays (0.6, 0.3, 0.1) for r,p,s, then a best response for P1 is to play P with probability 1

Nash equilibrium

- Strategy profile where all players simultaneously play a best response
- Standard solution concept in game theory
  - Guaranteed to always exist in finite games [Nash 1950]
- In Rock-Paper-Scissors, the unique equilibrium is for both players to select each pure strategy with probability 1/3

- Theorem [Nash 1950]: Every game in strategic form G, with a finite number of players and in which every player has a finite number of pure strategies, has an equilibrium in mixed strategies.

Minimax Theorem

- Minimax theorem: For every two-player zero-sum game, there exists a value v* and a mixed strategy profile σ* such that:
  a. P1 guarantees a payoff of at least v* in the worst case by playing σ^*_1
  b. P2 guarantees a payoff of at least -v* in the worst case by playing σ^*_2
- v* (= v_1) is the *value* of the game
- All equilibrium strategies for player i guarantee at least v_i in the worst case
- For RPS, v* = 0

- “That’s just a fixed point theorem.”

- Theorem [von Neumann’s Minmax Theorem 1928]: Every two-player zero-sum game in which every player has a finite number of pure strategies has a value in mixed strategies.

- He listened carefully, with his head cocked slightly to one side and his fingers tapping. Nash started to describe the proof he had in mind… But before he had gotten out more than a few disjointed sentences, von Neumann interrupted, jumped ahead to the as yet unstated conclusion of Nash’s argument, and said abruptly, “That’s trivial, you know. That’s just a fixed point theorem.”

Exploitability

- Exploitability of a strategy is difference between value of the game and performance against a best response
  - Every equilibrium has zero exploitability
- Always playing rock has exploitability 1
  - Best response is to play paper with probability 1

Nash equilibria in two-player zero-sum games

- Zero exploitability – “unbeatable”
- Exchangeable
  - If (a,b) and (c,d) are NE, then (a,d) and (c,b) are too
- Can be computed in polynomial time by a linear programming (LP) formulation
- Top poker AI programs such as Libratus and DeepStack attempted to approximate Nash equilibrium strategies in heads-up no-limit Texas hold ‘em (which is two-player zero sum)

Nash equilibria in multiplayer and non-zero-sum games

- None of the two-player zero-sum results hold
- There can exist multiple equilibria, each with different payoffs to the players
- If one player follows one equilibrium while other players follow a different equilibrium, overall profile is not guaranteed to be an equilibrium
- If one player plays an equilibrium, he could do worse if the opponents deviate from that equilibrium
- Computing an equilibrium is PPAD-hard

Most common “criticisms” of my and others’ research in computer poker

1. The approaches for two-player zero-sum games are not applicable to games with more than two players!
2. Even if the approaches were applicable, Nash equilibrium has no guarantees in games with more than two players, so approximating one would be useless!

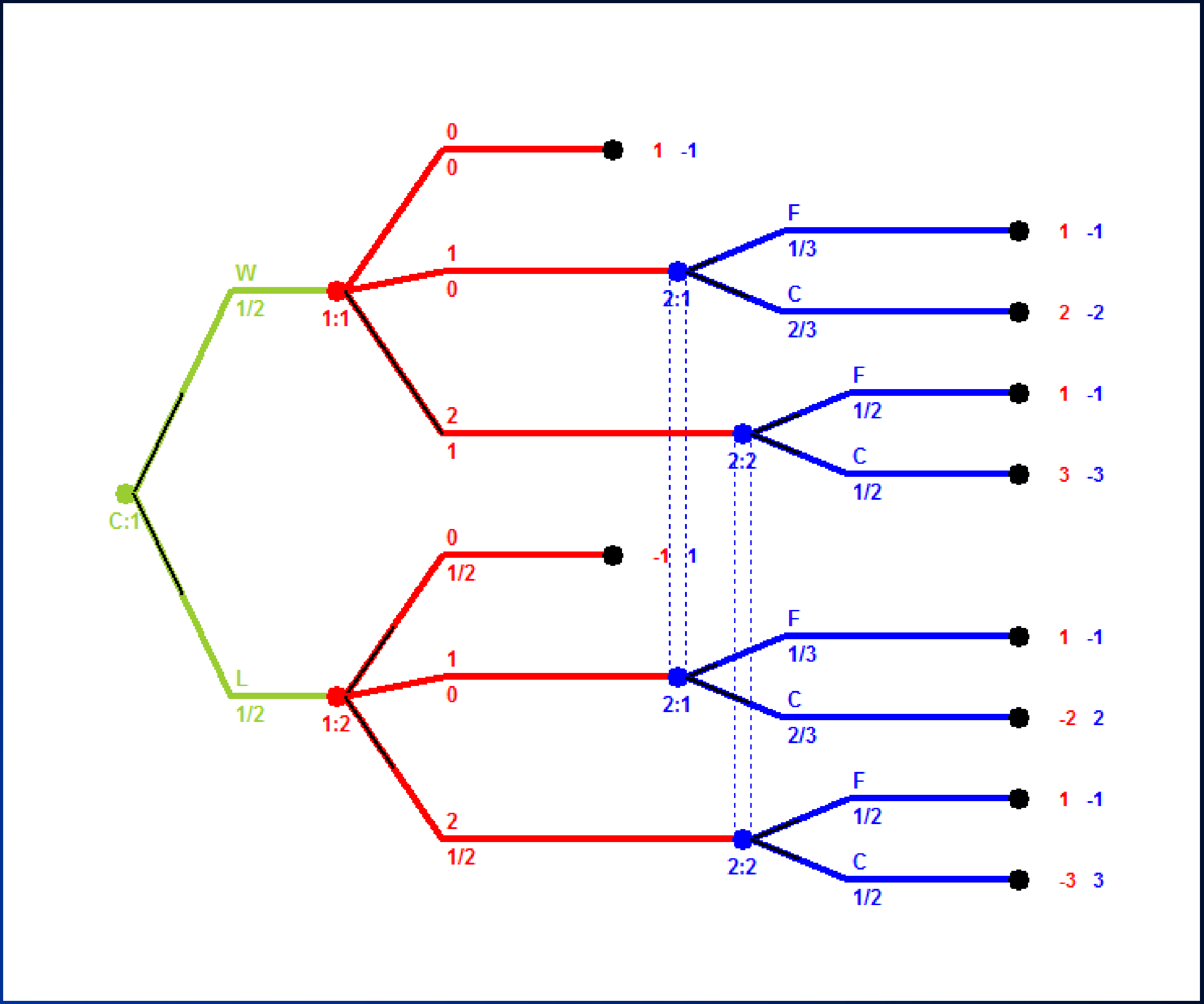
Frameworks and directions

- Standard paradigm
  - Abstraction, equilibrium-finding, reverse mapping (action translation and post-processing)
- New paradigms
  - Incorporating qualitative models (can be used to generate human-understandable knowledge)
  - Real-time endgame solving
- Domain-independent approaches
- Approaches are applicable to games with more than two players
  - Direct: abstraction, translation, post-processing, endgame solving, qualitative models, exploitation algorithm
  - Equilibrium algorithms also, but lose guarantees
  - Safe exploitation, but guarantees maximin instead of value

Imperfect information

- In many important games, there is information that is private to only some agents and not available to other agents
  - In auctions, each bidder may know his own valuation and only know the distribution from which other agents' valuations are drawn
  - In poker, players may not know private cards held by other players

Extensive-form representation



Extensive-form games

- Two-player zero-sum EFGs can be solved in polynomial time by linear programming
  - Scales to games with up to 10^8 states
- Iterative algorithms (CFR and EGT) have been developed for computing an ε-equilibrium that scale to games with 10^{17} states
  - CFR also applies to multiplayer and general sum games, though no significant guarantees in those classes
  - (MC)CFR is self-play algorithm that samples actions down tree and updates regrets and average strategies stored at every information set

Battle of the sexes

- Imagine a couple that agreed to meet this evening, but cannot recall if they will be attending the opera or a football match (and the fact that they forgot is common knowledge). The husband would prefer to go to the football game. The wife would rather go to the opera. Both would prefer to go to the same place rather than different ones. If they cannot communicate, where should they go?

	Opera	Football
Opera	(3,2)	(0,0)
Football	(0,0)	(2,3)

Fig. 1. Battle of the sexes game.

Equilibria of Battle of the Sexes

- Two pure equilibria (O,O) and (F,F)
- One mixed with prob 3/5 on preferred outcome

- Clearly in this game the success of playing a Nash equilibrium depends heavily on the strategy chosen by the other player. For example, if the wife follows her strategy from the first equilibrium and plays Opera, but the husband follows his strategy from the second Nash equilibrium and plays Football, the wife will receive the worst possible payoff of 0 despite following a Nash equilibrium. While this example is just for a two-player game, the same phenomenon can occur in games with more than two players (though as described above it cannot occur in two-player zero-sum games).

- Even three-player zero-sum games are not special, as any two-player general-sum game can be converted into a three-player zero-sum game by adding a “dummy” third player whose payoff equals negative the sum of the other two players' payoff. Furthermore, even if we wanted to compute a Nash equilibrium, it has been proven to be PPAD-complete and is widely conjectured that no efficient algorithm exists, though several heuristic approaches have been developed for strategic-form games with varying degrees of success in different settings.
- There have also been techniques developed that approximate Nash equilibrium to a provably very small degree of approximation error in a 3-player imperfect-information game.

Computing Nash equilibria in games with more than two players

- Developed new algorithms for computing ε-equilibrium strategies in multiplayer imperfect-information stochastic games
  - Models multiplayer poker tournament endgames
- Most successful algorithm, called PI-FP, used a two-level iterative procedure
  - Outer loop is variant of policy iteration
  - Inner loop is an extension of fictitious play
- Proposition: If the sequence of strategies determined by iterations of PI-FP converges, then the final strategy profile is an equilibrium.
- We verified that our algorithms did in fact converge to ε-equilibrium strategies for very small ε in a three-player poker tournament with high blinds restricted to all-in or fold strategies.

- The problem of how to create strong agents for non-zero-sum and multiplayer games, and in particular the question of whether Nash equilibrium strategies are successful, remains an open problem -- perhaps the most important one at the intersection of artificial intelligence and game theory. Of course, the most successful approach would not just simply follow a solution concept and would also attempt to learn and exploit weaknesses.
- Note that this would be potentially very helpful for two-player zero-sum games as well, as Nash equilibrium may not fully exploit mistakes of suboptimal opponents as much as successful exploitative agents even for that setting.

- However, successfully performing opponent exploitation is very difficult, particularly in very large games where the number of game iterations and observations of the opponents' play is small compared to the number of game states. And furthermore, such approaches are susceptible to being deceived and counterexploited by sophisticated opponents. It is clear that pure exploitation approaches are insufficient to perform well against a mix of opponents of unknown skill level, and that a strong strategy rooted in game-theoretic foundations is required.

Strongest existing agents for large multiplayer games

- The strongest existing agents for large multiplayer games have been based on approaches that attempt to approximate Nash equilibrium strategies. In particular, they apply the counterfactual regret minimization algorithm, which has also been used for two-player zero-sum games and has resulted in super-human level play for both limit Texas hold ‘em and no-limit Texas hold ‘em. These agents have performed well in the 3-player limit Texas hold 'em division of the Annual Computer Poker Competition which is held annually at the AI conferences AAAI or IJCAI.

- Counterfactual regret minimization is an iterative self-play algorithm that is proven to converge to Nash equilibrium in the limit for two-player zero-sum games. It can be integrated with various forms of Monte Carlo sampling in order to improve performance both theoretically and in practice. For multiplayer and non-zero-sum games the algorithm can also be run, though the strategies computed are not guaranteed to form a Nash equilibrium. It was demonstrated that it does in fact converge to an ε-Nash equilibrium (strategy profile in which no agent can gain more than ε by deviating) in the small game of 3-player Kuhn poker, while it does not converge to equilibrium in Leduc hold 'em. It was subsequently proven that it guarantees converging to a strategy that is not dominated and does not put any weight on iteratively weakly-dominated actions. While for some small games this guarantee can be very useful (e.g., for two-player Kuhn poker a high fraction of the actions are iteratively-weakly-dominated), in many large games (such as full Texas hold 'em) only a very small fraction of actions are dominated and the guarantee is not useful.

- Other approaches based on integrating the fictitious play algorithm with MDP-solving algorithms such as policy iteration have been demonstrated experimentally to converge to ε-equilibrium for very small ε in a no-limit Texas hold 'em poker tournament endgame. It has been proven that if these algorithms converge, then the resulting strategy profile constitutes a Nash equilibrium (while CFR does not have such a guarantee); however, the algorithms are not proven to converge in general, despite the fact that they did for the game that was experimented on.

- The empirical success of the 3-player limit Texas hold 'em agents in the Annual Computer Poker Competition suggests that CFR-based approaches which are attempting to approximate Nash equilibrium are promising for multiplayer games. However, the takeaway is not very clear.
- First, the algorithms are not guaranteed to converge to equilibrium for this game, and there is no guarantee on whether the strategies used by the agents constitute a Nash equilibrium or are even remotely close to one.
- Furthermore, there were only a small number of opposing agents submitted to the competition who may have questionable skill level, so it is not clear whether the CFR-based approaches actually produce high-quality strategies or whether they just produced strategies that happened to outperform mediocre opponents and would have done very poorly against strong ones.
- While these CFR-based approaches are clearly the best so far and seem to be promising, they do not conclusively address the question of whether Nash equilibrium strategies can be successful in practice in interesting multiplayer games against realistic opponents.

Three-player Kuhn poker

- Three-player Kuhn poker is a simplified form of limit poker that has been used as a testbed game in the AAAI Annual Computer Poker Competition for several years. There is a single round of betting. Each player first antes a single chip and is dealt a card from a four-card deck that contains one Jack (J), one Queen (Q), one King (K), and one Ace (A). The first player has the option to bet a fixed amount of one additional chip or to check (remain in the hand but not bet an additional chip). When facing a bet, a player can call (i.e., match the bet) or fold (forfeit the hand). No additional bets or raises beyond the additional bet are allowed (while they are allowed in other common poker variants such as Texas hold 'em, both for the limit and no-limit variants). If all players but one have folded, then the player who has not folded wins the pot, which consists of all chips in the middle. If more than one player have not folded by the end there is a showdown, at which the players reveal their private card and the player with the highest card wins the entire pot (which consists of the initial antes plus all additional bets and calls).

- As one example of a play of the game, suppose the players are dealt Queen, King, Ace respectively, and player 1 checks, then player 2 checks, then player 3 bets, then player 1 folds, then player 2 calls; then player 3 would win a pot of 5, for a profit of 3 from the amount he started the hand with.

- Note that despite the fact that 3-player Kuhn poker is only a synthetic simplified form of poker and is not actually played competitively, it is still far from trivial to analyze, and contains many of the interesting complexities of popular forms of poker such as Texas hold 'em. First, it is a game of imperfect information, as players are dealt a private card that the other agents do not have access to, which makes the game more complex than a game with perfect information that has the same number of nodes. Despite the size, it is not trivial to compute Nash equilibrium analytically, though recently an infinite family of Nash equilibria has been computed [Szafron et al. AAMAS '13].

- The equilibrium strategies exhibit the phenomena of bluffing (i.e., sometimes betting with weak hands such as a Jack or Queen), and slow-playing (aka trapping) (i.e., sometimes checking with strong hands such as a King or Ace in order to induce a bet from a weaker hand).

- To see why, suppose an agent X played a simple strategy that only bet with an Ace or sometimes a King. Then the other agents would only call the bet if they had an Ace, since otherwise they would know they are beat (since there is only one King in the deck, if they held a King they would know that player X held an Ace). But now if the other agents are only calling with an Ace, it is unprofitable for player X to bet with a King, since he will lose an additional chip whenever another player holds an Ace, and will not get a call from a worse hand; it would be better to check and then potentially call with hopes that the other player is bluffing (or to fold if you think the player is bluffing too infrequently). A better strategy may be to bet with an Ace and to sometimes bet with a Jack as a bluff, to put the other players in a challenging situation when holding a Queen or King. However, player X may also want to sometimes check with an Ace as well so that he can still have some strong hands after he checks and the players are more wary of betting into him after a check.

- A full infinite family of Nash equilibria for this game has been computed and can be seen in the tables from a recent article by Szafron et al. The family of equilibria is based on several parameter values, which once selected determine the probabilities for the other portions of the strategies. One can see from the table that randomization and including some probability on trapping and bluffing are essential in order to have a strong and unpredictable strategy. Thus, while this game may appear quite simple at first glance, analysis is still very far from simple, and the game exhibits many of the complexities of far larger games that are played competitively by humans for large amounts of money.

Nash equilibrium-based agent

- One way wonder why it is worthwhile to create agents and experiment on three-player Kuhn poker, given that the game has been "solved," as described in the preceding section. First, as described there are infinitely many Nash equilibria in this game (and furthermore there may be others beyond those in the family computed in the prior work). So even if we wanted to create an agent that employed a Nash equilibrium "solution," it would not be clear which one to pick, and the performance would depend heavily on the strategies selected by the other agents (who may not even be playing a Nash equilibrium at all).

- This is similar to the phenomenon described for the Battle of the Sexes Game in the introduction, where even though the wife may be aware of all the equilibria, if she attends the Opera as part of the (O,O) equilibrium while the husband does football as part of the (F,F) equilibrium, both players obtain very low payoff despite both following equilibrium.

- A second reason is that, as also described in the introduction, Nash equilibrium has no theoretical benefits in three-player games, and it is possible that a non-equilibrium strategy (particularly one that integrates opponent modeling and exploitation) would perform better, even if we expected the opponents may be following a Nash equilibrium strategy, but particularly if we expect them to be playing predictably and/or making mistakes.

- So despite that the fact that exact Nash equilibrium strategies have been computed for this game, it is still very unclear what a good approach is for creating a strong agent against a pool of unknown opponents.

- For our agent we have decided to use a Nash equilibrium strategy that has been singled out as being more robust than the others in prior work and that obtains the best worst-case payoff assuming that the other agents are following one of the strategies given by the computed infinite equilibrium family. We depict this strategy in the following table. This table assigns values for the 21 free parameters in the infinite family of Nash equilibrium strategies.

- To define these parameters, a_{jk}, b_{jk}, and c_{jk} denote the action probabilities for players P1, P2, and P3 respectively when holding card j and taking an aggressive action (Bet (B) or Call (C)) in situation k, where the betting situations are defined in the table.

Situation	P1	P2	P3
1	–	K	KK
2	KKB	B	KB
3	KBF	KKBF	BF
4	KBC	KKBC	BC

TABLE III

BETTING SITUATIONS IN THREE-PLAYER KUHN POKER

P1	P2	P3
$a_{11} = 0$	$b_{11} = 0$	$c_{11} = 0$
$a_{21} = 0$	$b_{21} = 0$	$c_{21} = \frac{1}{2}$
$a_{22} = 0$	$b_{22} = 0$	$c_{22} = 0$
$a_{23} = 0$	$b_{23} = 0$	$c_{23} = 0$
$a_{31} = 0$	$b_{31} = 0$	$c_{31} = 0$
$a_{32} = 0$	$b_{32} = 0$	$c_{32} = 0$
$a_{33} = \frac{1}{2}$	$b_{33} = \frac{1}{2}$	$c_{33} = \frac{1}{2}$
$a_{34} = 0$	$b_{34} = 0$	$c_{34} = 0$
$a_{41} = 0$	$b_{41} = 0$	$c_{41} = 1$

TABLE I

PARAMETER VALUES USED FOR OUR NASH EQUILIBRIUM AGENT

- Prior work has actually singled out a range of strategies that receive the best worst-case payoff; above we have described the lower bound of this space, and we also experiment using the strategy that falls at the upper bound (Table 2).

P1	P2	P3
$a_{11} = 0$	$b_{11} = \frac{1}{4}$	$c_{11} = 0$
$a_{21} = 0$	$b_{21} = \frac{1}{4}$	$c_{21} = \frac{1}{2}$
$a_{22} = 0$	$b_{22} = 0$	$c_{22} = 0$
$a_{23} = 0$	$b_{23} = 0$	$c_{23} = 0$
$a_{31} = 0$	$b_{31} = 0$	$c_{31} = 0$
$a_{32} = 0$	$b_{32} = 1$	$c_{32} = 0$
$a_{33} = \frac{1}{2}$	$b_{33} = \frac{7}{8}$	$c_{33} = 0$
$a_{34} = 0$	$b_{34} = 0$	$c_{34} = 1$
$a_{41} = 0$	$b_{41} = 1$	$c_{41} = 1$

TABLE II

PARAMETER VALUES USED FOR OUR SECOND NASH EQUILIBRIUM AGENT

Experiments against class project agents

- We experimented against 10 of 11 agents submitted recently for a class project (we ignored one agent that ran very slowly, which performed poorly).
  - www.ultimateaiclass.com
- These agents utilized a wide variety of approaches, ranging from neural networks to counterfactual regret minimization to opponent modeling to rule-based approaches.

- For each grouping of 3 agents we ran matches consisting of 3000 hands between each of the 6 permutations of the agents (with the same cards being dealt for the respective positions of the agents in each of the duplicated matches). The number of hands per match (3000) is the same value used in the Annual Computer Poker Competition, and the process of duplicating the matches with the same cards between the different agent permutations is a common approach that significantly reduces the variance. We ran 10 matches for each permutation of 3 agents. Table 4 shows the overall payoff (divided by 100,000) for each agent. The Nash agent received highest payoff. The results are very similar when using the upper and lower bound equilibrium strategies with the upper bound performing slightly better.

Nash	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
2.81 (UB)	2.25	1.18	2.54	-1.65	2.32	1.74	-1.34	-9.56	-3.48	1.42
2.81 (LB)	2.24	1.17	2.54	-1.66	2.32	1.74	-1.34	-9.54	-3.47	1.42

TABLE IV

EXPERIMENTS USING NASH AGENTS AGAINST CLASS PROJECT AGENTS

Conclusion

- Two main criticisms of prior work in computer poker and computational game theory:
  1. The approaches for two-player zero-sum games are not applicable to games with more than two players!
  2. Even if the approaches were applicable, Nash equilibrium has no guarantees in games with more than two players, so approximating one would be useless!

- I have argued previously that many of the state-of-the-art techniques utilized for the two-player poker agents are actually applicable to more than two players as well.
  - Direct: abstraction, translation, post-processing, endgame solving, qualitative models, exploitation algorithm
  - Equilibrium algorithms also, but lose guarantees
  - Safe exploitation, but guarantees maximin instead of value

- These new experiments demonstrate that an agent based on following an exact Nash equilibrium is able to outperform agents submitted for a recent class project that utilize a wide variety of approaches. This suggests that agents based on using Nash equilibrium strategies can in fact be successful in multiplayer games, despite the fact that they do not have a worst-case theoretical guarantee.

Conclusion

- Furthermore, for all game classes – two-player zero-sum games as well as non-zero-sum and multiplayer games – approaches that are able to (robustly) exploit opponents' mistakes, while also performing well against strong opposing agents, would be preferable to simply following a static Nash equilibrium strategy throughout. This has also been an active area within my research, both for two-player zero-sum and multiplayer games.
  - Game Theory-Based Opponent Modeling in Large Imperfect-Information Games AAMAS 2011
  - Safe Opponent Exploitation, EC 2012/TEAC 2015
  - Bayesian Opponent Exploitation in Imperfect-Information Games, CIG 2018

Future questions

- Obviously this analysis just holds for this one game, and remains to be seen whether Nash equilibrium strategies are effective in practice in other multiplayer games.
- Can it be shown that Nash equilibrium agent is successful in a general class of games?
- Can a theoretical performance guarantee be proven for a general class of games?
- Can a better solution concept for multiplayer games be developed?
  - Perhaps one that is a refinement of Nash equilibrium, or a different solution concept altogether.
  - Perhaps an independent concept or perhaps as stated before, in conjunction with opponent exploitation.