Modeling for Social Sciences – MATH231 – FLAME University

Game Theory Optimality & Exploitative Strategies in Poker

Arjun Grover

Abstract. When first learning to play poker, players are told to avoid betting outside the range of half pot to full pot, to consider the pot odds, implied odds, fold equity from bluffing, and the key concept of balance. Any play outside of what is seen as standard can quickly give away a novice player. But where did these standards come from and what happens when a player strays from standard play? This paper will explore the key considerations of making game theory optimal (GTO) plays in heads-up (two player) no limit Texas hold'em. To those new to the game, it involves dealing two cards that are revealed only to the player they are dealt to (hole cards), and five community cards that are revealed with rounds of betting in between. Hands are compared by looking at the highest five card poker hand that can be made with a player's hole cards combined with the community cards. This paper will focus on signalling, exploitative strategies and game theory optimal (GTO) play in heads-up poker based on examples of game scenarios from my own real life experience playing poker.

PS. *Gambling is an addiction, and Poker can be construed as gambling due to the presence of luck despite the variance in play style and presence of strategic thinking. Stay safe! Be addicted to Pull Ups instead!*

1. Introduction

Poker is a game that has been extensively studied from a mathematical stand-point, as it is interesting from a game theory standpoint and highlights considerations that must be made when making decisions under uncertainty and deals with expected value of strategies over time. It is a game with strategies that are not immediately intuitive and the value of those strategies are only seen over a large number of hands. To reduce complexity, this paper will focus on heads-up (two player) poker. To those new to the game, the game begins with each player being dealt two cards which are hidden from the other player. A round of betting takes place, where there are four actions available to the players: check, bet, call, raise. A player can check or bet if no amount has yet been made in the current round of betting and a player can call (match the amount bet by the opponent) or raise (bet an additional amount on top of opponent's bet) if the opponent bets. After the initial round of betting (pre-flop), the first three community cards (visible to both players) come out (flop). Another round of betting proceeds before the fourth card comes out and likewise before the fifth and final card. After all cards are out, there is one last round of betting before the players' hands are compared (showdown). The complexity of poker arises from inferring probabilities through the many rounds of betting and making decisions that consider events in the future. To understand the mathematics behind playing optimally, we dissect the game into constrained sub-problems, but the concepts derived through these examples are relevant in real play.

2. Literature Review

Poker is a game that has long fascinated mathematicians and game theorists. In recent years, there has been increasing interest in using game theory to model poker and develop optimal strategies for play.

One approach to modeling poker is to use game theory to develop game theory optimal (GTO) strategies. These strategies are designed to be unexploitable, meaning that they cannot be consistently beaten by an opponent who is capable of adjusting their own strategy. GTO strategies involve balancing your range of hands and actions, so that your opponent cannot predict what you will do based on your past behaviour (Les, 2019).

While GTO strategies can be effective in many situations, they are not always the most profitable approach to poker. Exploitative strategies, which involve taking advantage of an opponent's weaknesses or tendencies, can sometimes be more profitable. For example, if you notice that your opponent is consistently folding to aggression, you might use an exploitative strategy by betting aggressively and bluffing more frequently (Krieger, 2006).

Signaling models can also be used to develop successful strategies in poker. Discrete signaling, such as using chips or verbal cues to indicate your actions, can be used to communicate your intentions to other players. Continuous signaling, such as facial expressions or body language, can be used to read other players and gain insight into their hands and strategies (Vink et al., 2020).

Overall, the use of game theory, GTO strategies, exploitative strategies, and signaling models can all be effective approaches to modeling and playing poker. However, successful play requires a combination of skill, experience, and adaptability to the specific situation at hand.

3. Pot Odds

Definition 3.1. We refer to a ***made hand*** as a poker hand that is already guaranteed given a player's hole cards and currently revealed community cards.

Definition 3.2. We refer to a ***draw*** as a hand that can be made given certain community cards come out.

Example 3.1. Suppose **Shubhangi (A)** has A♦A♠ and **Santy (B)** has 5♥6♥. The community cards on the ***turn*** (stage of game where 4 community cards have been revealed) are K♦4♥3♣Q♥. Shubhangi has a made hand of a pair of aces and Santy has a draw to a straight. Now if both players knew each other's cards, they would agree that if the last card is a 2 or 7 of any suit, Santy wins, otherwise Shubhangi wins. In this world of perfect information, neither Shubhangi nor Santy would bet on the river (when the last card comes out), because the winner would be clear.

Now suppose there is already ₹100 in the pot and Shubhangi can either check or bet before the river card comes out. If Shubhangi bets, Santy has the option to re-raise. There are 9 hearts remaining in the deck, which would give Santy a flush, beating Shubhangi. The remaining 35 cards would allow Shubhangi's aces to hold. Suppose Shubhangi is to act first. Since Shubhangi is favored to win the hand, Shubhangi has reason to bet here. The amount she should bet is derived from calculating expected value (EV).

The expected value is calculated as the probability of Shubhangi winning the pot times the new pot amount, deducted by the amount she bets. Note that this calculation emphasizes that as soon as Shubhangi places a bet, she should no longer consider that money to be her's to lose, but rather part of the pot that she can win (sunk cost).

$$E(A) = \frac{35}{44}(100+2x) - x$$

$$\approx 80 + 0.6x$$

Note that if the probability of winning here is less than $\frac{1}{2}$, it is not profitable to bet. This however is complicated when we consider a real game where both players do not have complete information and bluffing is a valid strategy.

Also note that Shubhangi's EV is strictly increasing as her bet increases if Santy always calls. Santy however, should only call if it is positive EV for him.

$$E(B) = \frac{9}{44}(100+2x) - x$$

$$\approx 20 - 0.6x$$

Santy should thus only call if Shubhangi's bet is below around ₹33 or $\frac{1}{3}$ of the pot pre-betting. This $\frac{1}{3}$ is what we refer to as pot odds. It is important to keep in mind that Santy can call larger bets or even re-raise because of something we refer to as implied odds, which take into consideration further betting on the river due to it being unknown who has the better hand.

4. Implied Odds

Implied odds refer to the potential to make more money when a draw hits. Remember that we previously assumed both players had complete information. This is not true in a real game, which means betting on the river can be profitable. In the case of our previous example, Shubhangi does not know what Santy has, so if Santy hits his flush, he can potentially make more off Shubhangi than was estimated by our EV calculations on the turn.

Example 4.1.

Let us continue with our previous example. If Santy hits a flush on the river, we will assume that he knows correctly that he has the better hand (for now we will ignore the possibility Shubhangi has a higher flush, because the probability is relatively low). Suppose Shubhangi bet ₹50 on the turn and Santy called. The final card comes 7♥. Now the pot is ₹200 and Shubhangi acts first. Recall that the board currently shows K♦4♥3♣Q♥7♥. Now Shubhangi doesn't know what Santy has and believes it's likely he has top pair (a king that pairs with the king showing on the board). Shubhangi thinks she can bet again to get value off of Santy. Here Santy can fairly safely call or re-raise Shubhangi's bet.

Let's look at what Shubhangi should do when the river card comes out. Suppose she's fairly certain Santy either hit his flush or just has the top pair on the board. Estimating the probabilities of these two cases is more complicated (has to take into account what kinds of hands Santy generally plays and the history of actions on the current hand), but it's fair to assume Santy has more hands involving kings in his range than two hearts.

This means that if Santy knows Shubhangi will bet on the river even if he hits his flush, he is willing to call larger bets from Shubhangi on the turn or even re-raise or bet if Shubhangi checks.

Definition 4.2. We refer to a player's *range* as the hands he plays in a given situation. In general, a player's range does not change from hand to hand. That is not to say that the player should be predictable (see Section 5.3. regarding balancing range).

5. Game Theory Optimal Strategies

5.1. Signaling Models

Game theory optimal strategies (GTO) involve finding a strategy that cannot be exploited by an opponent over the long term. In the context of poker, this means finding a strategy that is balanced and cannot be consistently beaten by an opponent who is capable of adjusting their own strategy. Both discrete and continuous signaling models can be used to develop GTO strategies in poker. Here are a few examples:

1. *Discrete signaling*: One way to use discrete signaling to develop a GTO strategy is to vary your actions randomly, rather than always using the same signal to indicate a particular action. For example, you might sometimes say "I raise" when you're bluffing, and sometimes say it when you have a strong hand. By varying your signals randomly, you make it harder for your opponent to detect a pattern and exploit your strategy.
2. *Continuous signaling*: Continuous signaling can also be used to develop a GTO strategy by incorporating "false tells" into your behavior. For example, you might intentionally display a certain facial expression or body language when you have a strong hand, and then use the same expression when you have a weak hand. This can make it difficult for your opponent to accurately read your tells and adjust their own strategy accordingly.

In both cases, the goal is to create a strategy that is balanced and cannot be consistently exploited by an opponent. By incorporating elements of randomness and deception into your signaling, you can make it more difficult for your opponent to accurately read your strategy and gain an advantage over you.

5.2. Exploiting the Opponent

In actuality, the size of bets should not be proportional to how good your hand is, nor should you only bet when you have a good hand, as that is exploitable by the opponent over time. In the previous sections, we looked at examples constrained to a single hand, in which case we only care about maximizing EV on that hand. However, poker is all about beating the odds over time, so it's important to realize that a strategy optimized for a single hand may not be optimal or even profitable in the long run.

As a simple but realistic example, suppose your opponent only bets and raises hands that they think will win the pot, but still calls some of your bets with weaker hands (this is not an uncommon type of play from risk-adverse beginners). It's easy to exploit a player like this by simply using a strategy which folds to every bet or raise the opponent makes and still betting our good hands. Of course, eventually the opponent will catch on and counter-exploit by bluffing their hands if they know we will fold. On this end of the spectrum, suppose a player bluffs too many hands. To exploit this play style, we can afford to play a larger portion of hands and make large profits when we hit a top hand.

This leads us to the idea of balancing our range, or deciding the hands we play in a given situation such that an opponent cannot exploit our strategy.

5.3. Balance

To play non-exploitable game theory optimal (GTO) poker, ranges should be "balanced". Often, this means that we have a variety of possible hands in the eyes of the opponent in any situation. This means adding in a range of hands with which you bluff and not betting only when you have a good hand or betting a larger amount when you have the winning hand.

Definition 5.1. We define *defensive value* as the expected value of a strategy against the opponent's most exploitative strategy. Note the difference between this value and EV as we've previously looked at is that this assumes the opponent knows how we play and can exploit any patterns over time in our strategy.

A more rigorous definition of balanced strategy is minimizing the gap between defensive value (Definition 5.1) and expected value. In other words, the expected payoff of the strategy in a given hand should not change over time as your strategy is gradually exposed to your opponent: your opponent plays the same way regardless whether your strategy is known to them.

Definition 5.2. A *pure strategy* dictates a player's action in any situation i.e. the player will always make the same decision under given circumstances.

Definition 5.3. A *mixed strategy* is one in which the player assigns a probability distribution over all pure strategies (Definition 5.2).

Definition 5.4. *Nash equilibrium* is a strategy set in a multi-player game where neither player alone can increase their payoff. Because of this, it is a stable point where neither player wants to deviate from their current strategy.

Definition 5.5. A game in which the sum of all players' scores is equal to 0 is called a *zero-sum game*.

Definition 5.6. *Indifference* refers to a game state where a player gets the same expected payoff regardless what strategy is chosen.

Definition 5.7. An *indifference threshold* is a value for a parameter that a player can choose to force indifference (Definition 5.6) on the opponent.

It is a known fact of game theory that all multi-player games with finite payout matrices have at least one Nash equilibrium (Definition 5.4). Additionally, poker is a zero-sum game (Definition 5.5) and it is known that all zero-sum two-player games have an optimal strategy as long as mixed strategies (Definition 5.3) are allowed. This leads to the concept of indifference (Definition 5.6. By setting expected payoff equations equal to each other, we can obtain values for parameters that force a player to be indifferent to choosing among strategies). The value of the parameter found by solving these equations is an indifference threshold (Definition 5.7). Let us take a look at the following example.

Example 5.8.
Suppose Santy has three of a kind and on a particular board is only scared of Shubhangi having a flush. Let us assume that Shubhangi has a flush here 20% of the time. How often can Shubhangi bluff? For this example suppose there is ₹300 in the pot and Shubhangi can choose to bet a fixed amount of ₹100. To keep it simple, we will say Santy either calls or folds when Shubhangi bets.

How often should Shubhangi bluff here? If Shubhangi bets ₹100, Santy can pay ₹100 to potentially win ₹400. Suppose Shubhangi only bets when she has the flush. Santy can exploit this strategy by folding every time Santy bets, preventing him from getting any additional value from hitting his flush and taking the pot 80% of the time. Shubhangi has a defensive value of 0.2 x ₹300 = ₹60 with this strategy, where she only profits when she has the flush. Now suppose Shubhangi bets all her hands here. 20% of the time she has the flush and the other 80% of the time she has nothing. If Santy calls, his EV is 0.8 x ₹400 − ₹100 = ₹220 and if he folds, his EV is 0, so Santy will exploit Shubhangi's strategy here by always calling. The defensive value of Shubhangi's strategy is 0.2 x ₹400 − ₹100 = −₹20.

The two strategies mentioned so far (always checking a dead hand and always betting a dead hand) are what are known as pure strategies (Definition 5.2) and neither is optimal for Shubhangi in this situation. We know this, because both are exploitable – Santy alone can change his strategy and increase his payoff. This indicates we are not at a equilibrium point.

Now, we explore mixed strategies. Let P⟨A,bluff⟩ be faction of all hands Shubhangi has on the river that she bluffs with. Santy's EV for calling when Shubhangi bets can be computed as:

$$\mathrm{E(B)}\ \langle \mathrm{B, call}\rangle = \frac{P\langle \mathrm{A,bluff}\rangle}{0.2+P\langle \mathrm{A,bluff}\rangle}(₹400 - ₹100)$$

Alternatively, Santy can fold when Shubhangi bets.

$$\mathrm{E(B)}\ \langle \mathrm{B, fold}\rangle = ₹0$$

Shubhangi's EV can be computed as

$$\text{E(A)}\ \langle \text{B, call} \rangle = \frac{0.2}{0.2+P\langle \text{A,bluff} \rangle}\,(₹400 - ₹100)$$

$$\text{E(A)}\ \langle \text{B, fold} \rangle = ((0.2 + \text{P}\ \langle \text{A, bluff} \rangle)(₹300)$$

Shubhangi's strategy is least exploitable when Santy's EV for calling and folding are equal (i.e. Santy is not able to change his strategy to exploit Shubhangi even if over time he figures out how often Shubhangi bluffs). By setting E(B) ⟨B, call⟩ = E(B) ⟨B, fold⟩, we can solve for Shubhangi's optimal bluff frequency such that Santy is indifferent to calling versus folding. It turns out that it is optimal for Shubhangi to bluff around 6.7% of her hands.

Example 5.9.
Consider the general scenario where we only have one round of betting, the pot has b bets, Shubhangi can make a bet of size 1, and Santy can call or fold if Shubhangi bets. The payout matrix is as follows:

	Santy	Check-call	Check-fold
Shubhangi			
Winning hand	Bet	P + 1	P
	Check	P	P
Dead hand	Bet	-1	P
	Check	0	0

As we can see from the payout matrix, it is always in Shubhangi's favor to bet when she has a winning hand. It is less obvious what Shubhangi should do when she has a dead hand. Depending on Santy's calling versus folding frequency, it can be beneficial for Shubhangi to bluff a percentage of her dead hands.

According to the concept of indifference, Shubhangi wants to choose a bluffing frequency such that Santy's EV for calling is equal to his EV for folding. Let b represent $\frac{bluffs}{bluffs+valuebets}$.

$$\text{E(B)}\ \langle \text{call} \rangle = \text{b(P + 1)} - 1$$

$$\text{E(B)}\ \langle \text{fold} \rangle = 0$$

We have E(B) ⟨call = E(B)⟩fold when

$$\text{b} = \frac{1}{p+1}$$

Likewise, Santy should choose a calling frequency such that Shubhangi is indifferent to checking versus bluffing her dead hands. Let c be the frequency with which Santy calls.

$$\text{E(A)}\ \langle \text{check} \rangle = 0$$

$$E(A)\ \langle \text{bluff} \rangle = (1 - c)(P + 1) - 1$$

By setting these two EVs equal to each other, we find the value c with which Santy should call when Shubhangi bets.

$$c = \frac{P}{P+1}$$

It turns out these two quantities are quite useful, so we will give the quantity $\frac{1}{P+1}$ its own letter, α. Shubhangi's optimal bluff to bet ratio is equal to α and Santy's optimal calling frequency is equal to 1 − α.

This can be generalized to different bet sizes. Bets are generally thought about as a fraction of the pot (according to pot odds). Suppose Shubhangi can bet any fraction of the pot xP.

$$b = \frac{xP}{P+xP}$$

$$= \frac{x}{1+x}$$

6. Multi-Street Games

Thus, far we have mainly discussed single street (one round of action) scenarios, but in reality, action on a given street depends on everything that has happened before. In Example 5.8, we assumed that Shubhangi has a flush 20% of the time. In reality, this probability depends on everything that happened before the river.

Example 6.1.
Let's set up the following scenario:

- The board shows K♦4♥3♣Q♥.
- Shubhangi has a pair of aces.
- Santy has a hand from a distribution which contains $\frac{1}{10}$ hands with two hearts and $\frac{9}{10}$ dead hands.
- The pot contains ₹400, and players can either check or bet ₹100.
- Shubhangi is first to act.
- Santy is confident he has the winning hand if he hits his flush and a dead hand otherwise.

The flush comes around 20% of the time (in actuality, it's a little less but for simplicity's sake we will use 20%). From what we studied before, Shubhangi should bet and Santy should call if he has the odds. Note that implied odds should be considered here rather than just pot odds, because Santy can get more value on the river by hitting his flush.

Assume Shubhangi and Santy make it to the river, and now there is ₹600 in the pot (Shubhangi bets on the turn and Santy calls). Now Shubhangi has no reason to bet here, because we have assumed Santy knows whether he has the winning hand at this point. Thus Shubhangi checks and Santy can choose to either bet or check. According to Example 5.9, Santy should bluff with $\alpha = \frac{1}{7}$ as many hands as she value bets with and Shubhangi should call with a frequency of $1 - \alpha = \frac{6}{7}$

However, this is actually incorrect, because our prior calculations relied on a single street game. Consider how this situation is different. Suppose Santy wants to bluff on the river. This means he had to have called Shubhangi's bet on the turn with a dead hand. Also note that Santy can only bluff on the river if a heart comes out. Thus in this multi-street game, Shubhangi should be considering indifference of Santy folding versus playing a dead hand through both streets. Let c be Shubhangi's optimal calling frequency on the river.

$$\mathrm{E(B)}\ \langle \text{dead hand, fold} \rangle = 0$$

$$\begin{aligned}\mathrm{E(B)}\ \langle \text{dead hand, play} \rangle &= \mathrm{P\ (flush)P\ (Shubhangi\ calls)}(-2) \\ &\quad + \mathrm{P\ (flush)P\ (Shubhangi\ folds)}(5) \\ &\quad + \mathrm{P\ (no\ flush)}(-1) \\ &= (0.2)(-2)\mathrm{c} + (0.2)(5)(1-\mathrm{c}) + (0.8)(-1)\end{aligned}$$

$$\mathrm{c} = \frac{1}{7}$$

So we see that it turns out Shubhangi's optimal calling frequency is actually $\frac{1}{7}$ rather than $\frac{6}{7}$ from analysis of a single street game. By analyzing a single street game, we are able to reason about strategies, but the determined frequencies cannot be blindly applied to multi-street games where there are added layers of complexity.

7. Conclusion

GTO strategy explores the concepts of balance and indifference which minimizes exploitability. When you have minimal knowledge of the opponent's play style, it is a good defensive strategy to play close to GTO, which aims to optimize for the worst case by minimizing your own exploitability. GTO strategy assumes an opponent who also plays optimally, or knows how to exploit weaknesses in any strategy. However, the assumption that the opponent is always perfectly rational and plays according to GTO strategy is rarely true and the discrepancy is what allows players who know how to take advantage profit. Even good players do not play a perfectly balanced game and open themselves up to exploitability in which case it is beneficial to play to their weaknesses whenever you have the information to do so.

To play exploitative poker, it is important to consider past information about the opponent's overall play style, ranges, and strategy as well as what actions took place on earlier streets of a given hand. However, keep in mind that strategies that stray too far from GTO can be counter-exploited. Suppose we have two perfectly rational players who start off with very different exploitable strategies. In theory, over time the two players would learn to exploit and counter-exploit each other's strategies and eventually their strategies would converge to near GTO.

In conclusion, depending on the opponent, playing GTO may not always be the most profitable, but it minimizes exploitability, making it a safe strategy to play against any opponent. Thus, we conclude a preliminary analysis of poker using game theory.

8. Works Cited

Bill Chen and Jerrod Ankenman, The mathematics of poker, ConJelCo, 2006.

Krieger, L. (2006). Hold'em excellence: From beginner to winner. Two Plus Two Publishing LLC.

Les, P. (2019). Game theory optimal poker: How to think like a pro. D&B Publishing.

Vink, M., Taghizadeh, M. R., & Heydenreich, B. (2020). To bluff or not to bluff? A game-theoretic analysis of poker signaling. Operations Research Letters, 48(3), 272-277.

Parsons, K.; Tingley, P.; Zajdela, E. (2021) When to Hold 'Em, Mathematics Magazine, 94:3, 201-212, DOI: 10.1080/0025570X.2021.1908785, https://doi.org/10.1080/0025570X.2021.1908785

David H. Reiley, Michael B. Urbancic & Mark Walker (2008) Stripped-Down Poker: A Classroom Game with Signaling and Bluffing, The Journal of Economic Education, 39:4, 323-341, DOI: 10.3200/JECE.39.4.323-341

Li, J. (2018). *Exploibility and Game Theory Optimality*. Retrieved April 19, 2023, from https://math.mit.edu/~apost/courses/18.204_2018/Jingyu_Li_paper.pdf

Easley, D., & Kleinberg, J. (2010). *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511761942

Pareto Optimality, Game Theory and Equilibria. (2008). Germany: Springer New York.