Optimal Poker Strategies Are Shaped Like Florida*

Sam Ganzfried
sam@ganzfriedresearch.com
Miami Beach, Florida

Abstract

A common situation in poker is that a player has made a bet and the opposing player must decide whether to call the bet or to fold. Two important theoretical concepts are often utilized in making this decision. First is the *minimum defense frequency* needed to call vs. the betsize in order to ensure that the opponent cannot profitably bluff us with his weakest hands. And second is the *range advantage* of a player: the probability he has a superior hand to the opponent given how the situation has developed so far. In general the player with the range advantage will bet more aggressively, since he has a stronger distribution of hands than the opponent. When a player has a strong range advantage, the opposing player will be forced to call with significantly lower frequency than the minimum defense frequency, because the betting player will have a strong hand very often. Thus, it is natural to investigate what the optimal tradeoff should be between the degree of range advantage and the correct calling frequency. We generate and solve a large sample of games with randomly generated hand distributions, and create a graph of the optimal defense frequency vs. range advantage. As it turns out, this graph looks very similar to the map of the state of Florida. Based on this similarity, we single out several poker situations that exemplify notable cases: Miami, Pensacola, Tallahassee, Jacksonville, St. Petersburg, Orlando, and Key West.

1 Introduction

Poker is a very popular game played by humans, and creating strong computer agents for poker has been a major AI challenge problem. Recently programs for two-player no-limit Texas hold 'em have competed against the strongest human players in a sequence of high-profile "Brains vs. Artificial Intelligence" competitions [2, 1] culminating in the agent Libratus that won decisively in 2017. Independently a second agent DeepStack was also created that defeated human players in 2017, though the humans were not experts in the specific game format [4]. Despite these successes, there have been relatively few takeaways from the research that aspiring human players can readily apply to improve their poker game; while several of the "hand histories" from these competitions are publicly available, it is difficult to construe strategic insights that will generalize from just observing several isolated situations. In a separate new line of research, a new approach has been developed for computing strategic rules that are easily understood by humans [3]. This has led to deduction of two new fundamental rules of poker strategy, concerning when a player should make an extremely small bet (*80-20 rule*) and an extremely large bet (*all-in rule*). We continue this avenue of research by exploring for the first time the connection between two key poker strategy concepts. The first concept called *range advantage* concerns the degree to which one player has an advantage over the other in a given situation, and the second called *optimal defense frequency* concerns the amount of time the opposing player must call a bet in order to prevent himself from being exploited.

Consider the following simplified poker game. First player 1 is dealt a card from a n-card deck (cards are the numbers 1–n), according to a probability distribution p_1 (e.g., $p_1(2)$ is the probability that player 1 is dealt a 2). We assume similarly that player 2 is dealt a card from 1–10 according to p_2. We assume the cards are independent, and that it is possible for both players to be dealt the same card (so that our experiments will not be confounded by the additional issue of *card removal*). Players initially start with a stack S and there is a pot of size P, with a deck size equal to $n = 10$. Then player 1 can either *check* or bet a size in $\{b_{1i}\}$. Facing a bet from player 1, player 2 can then call or fold. Facing a check from player 1, player 2 can check or bet a size in $\{b_{2i}\}$. Then player 1 can choose to call or fold facing a bet.

For our experiments we will assume the pot and stack sizes P and S both equal 1 (if they differed we could divide by the size of the pot to normalize). For both players we only allow a bet size equal to pot (1) for b_{1i}, b_{2i}. Of course this setting can allow arbitrary values for these parameters and multiple bet sizes, but for concreteness we decided to start with these values.

The *range advantage* for player 1 under the distributions p_1, p_2 is player 1's *equity* in the hand given the distributions. This is equal to the probability that player 1 will have a better hand under p plus one-half times the probability that they have the same hand. So we define the Range Advantage as:

$$\text{Range Advantage}(p_1, p_2) = \left(\sum_{j=1}^{n} \sum_{i=j+1}^{n} p_1(i) p_2(j) \right) + \left(\frac{1}{2} \sum_{i=1}^{n} p_1(i) p_2(i) \right) \tag{1}$$

For a bet of size b and a pot of size P, the *minimum defense frequency* (MDF) is defined to be the quantity $\gamma = \frac{P}{b+P}$. Suppose player 2 calls a bet of size b into a pot of size P with probability γ, and suppose that player 1 has an extremely weak hand that he is contemplating betting as a bluff. With probability γ the bluff would be called and player 1 would lose b, while with probability $1 - \gamma$ player 2 would fold and player 1 would win P. Thus the expected profit for player 1 from betting is:

$$\gamma \cdot (-b) + (1 - \gamma) P = P - \gamma(b + P) = P - \frac{P}{b + P} \cdot (P + b) = 0.$$

Thus player 1 is precisely indifferent between betting as a bluff and not betting with his weak hand. If player 2 were to call with any probability below γ, then player 1 would be able to have a guaranteed profit by bluffing with all his weak hands; therefore, the value γ is the minimum amount of the time that player 2 must call (or raise if it is permitted) against a bet in order to prevent player 1 from profitably betting with all his weak hands. For our parameters of $b = 1, P = 1$ the MDF is $\alpha = \frac{1}{2}$. Note also that if player 2 calls with probability exceeding α, then player 1 would never want to bluff with his weak hands (in which case player 2 would never want to call with his medium-strength hands, etc.). So in general we would expect player 2 to call with frequency precisely equal to α; however this is not always the case. Consider an extreme example when player 1 is always dealt a 10 and player 2 is always dealt a 1. Then clearly player 2 would always prefer to fold to a bet because he will never have the best hand. Note that in this example player 1 has a range advantage equal to 1. In general, if player 1 has a significant range advantage, optimal play (i.e., according to Nash equilibrium strategies) may suggest that player 2 call with probability below α.

We generated 100,000 games with uniform random distributions for p_1 and p_2. We then solve these games for a Nash equilibrium. Let $c(i)$ be the optimal frequency for player 2 to call a bet from player 1 with card i according to the equilibrium strategy computed from p_1, p_2. Then the *optimal defense frequency* (ODF) c^* is the weighted sum of the call probabilities over all hands, i.e., $c^* = \sum_i (p_2(i) c(i))$. In our experiments we seek to understand the relationship between the range advantage and the optimal defense frequency which, as described above, may differ considerably from the minimum defense frequency value.

For each hand in the sample, we can add a data point where the x coordinate is the range advantage of the distributions p_1, p_2, and the y coordinate is the optimal defense frequency c^*. A scatterplot of points selected according to this process is shown in Figure 1. This figure happens to resemble the US state Florida: both have a flat "panhandle" extended to the left that connects to a curved peninsula below to the right.

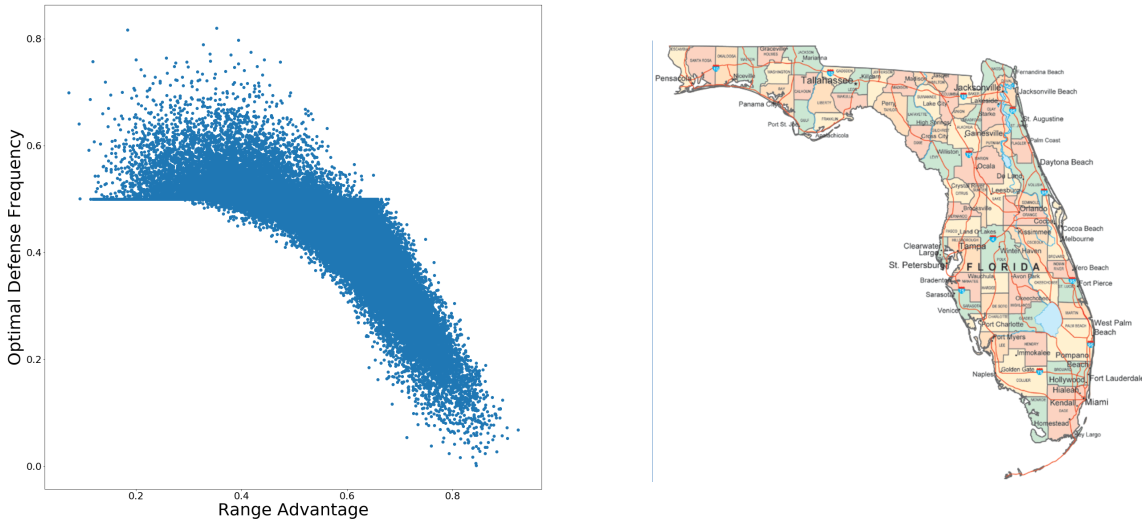

Figure 1: Scatter plot of range advantage vs. optimal defense frequency; Map of US state Florida

2 Optimal strategy for notable situations

The experiments depict a clear negative trend between the range advantage of player 1 and the optimal defense frequency for player 2. If we were to fit a straight line using linear regression over the dataset of 100,000 points (using 20-fold cross validation) this produces the line ODF $= -0.446 \cdot$ RA $+ 0.688$. For example, for a range advantage of 0.5 this would give optimal defense frequency of 0.465, and for range advantage of 0.9 this would give optimal defense frequency of 0.287. This produces a mean-squared error (MSE) of 0.002. However, we can see clearly from the figure that a curve (perhaps a quadratic one) would more appropriately model the relationship. But furthermore, we see that no single curve can produce an extremely accurate prediction; for example, for RA=0.7 there is a mass of points for values of ODF ranging from 0.2 up to 0.4, and predicting a single value is problematic. In order to obtain better predictive performance we would need to include additional features that can better predict ODF than just RA alone.

From the graph, we are able to conclude several interesting takeaways based on several of the notable data points. Using the analogy to the map of Florida, we name these points by Florida's major cities.

Name	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	q_1	q_2	q_3	q_4	q_5	q_6	q_7	q_8	q_9	q_{10}	RA	ODF
Miami	.007	.002	.002	.118	.045	.108	.210	.075	.300	.132	.047	.451	.051	.020	.200	.032	.090	.072	.014	.023	.852	.129
Pensacola	.401	.012	.002	.115	.238	.068	.036	.022	.048	.012	.005	.008	.053	.002	.250	.114	.020	.004	.131	.413	.114	.500
Tallahassee	.160	.014	.049	.249	.192	.022	.065	.178	.046	.024	.058	.161	.063	.152	.091	.305	.054	.050	.033	.033	.500	.500
Jacksonville	.122	.006	.053	.095	.080	.075	.242	.021	.075	.231	.087	.366	.007	.165	.017	.017	.007	.214	.058	.062	.651	.500
St. Petersburg	.126	.009	.016	.000	.052	.077	.157	.021	.521	.021	.208	.107	.061	.040	.034	.016	.020	.301	.014	.197	.600	.325
Orlando	.054	.041	.234	.013	.057	.001	.013	.023	.483	.079	.049	.273	.032	.078	.011	.081	.180	.266	.025	.005	.700	.397
Key West	.119	.020	.024	.000	.036	.020	.108	.008	.525	.139	.045	.071	.053	.056	.005	.305	.204	.167	.053	.040	.726	.100
Atlanta	.077	.059	.052	.225	.017	.180	.223	.076	.083	.009	.027	.043	.019	.191	.074	.076	.016	.363	.150	.040	.353	.820

Table 1: Hand distributions (p_i for player 1 and q_i for player 2), range advantage, and optimal defense frequency for several notable situations.

- **Miami**
  In the Miami situation, which corresponds to the bottom right of the peninsula, player 1 has a large mass of strong hands 7–10 while player 2 has a very low mass of strong hands and high mass on weak hands, giving player 1 a range advantage of 0.852 and player 2 an optimal defense frequency of 0.129.

- **Pensacola**
  In the Pensacola situation, at the left of the "panhandle," player 1 has very high weight on weak hands while player 2 has most weight on strong hands, giving player 1 the very low range advantage 0.114 and player 2 an optimal defense frequency of 0.5.

- **Tallahassee**
  In the Tallahassee, at the center of the panhandle, both the range advantage and optimal defense frequency are 0.5.

- **Jacksonville**
  In the Jacksonville, at the right of the panhandle, player 1 has a sizable range advantage of 0.65 while the optimal defense frequency is 0.5.

- **St. Petersburg**
  In the St. Petersburg, player 1 has a range advantage of 0.65 due in large part to having over half his mass on a 9, while player 2, who has most mass on an 8 and also significant mass on 10, has an optimal defense frequency of only 0.325.

- **Orlando** In the Orlando, player 1 has almost half his mass on a 9 and also significant mass on a 3 while player 2 has significant mass on 2, 7, and 7, giving player 1 a range advantage of 0.7 and player 2 an optimal defense frequency of 0.397.

- **Key West** At Key West, which is an island city located below the southern tip of the peninsula, player 1, with over half the mass on 9, has a range advantage of 0.726, while player 2 only has an optimal defense frequency of 0.100 despite having sizable mass on 6–8.

- **Atlanta** Finally, to represent the northmost point, we have Atlanta, which is technically not in Florida but in the adjacent state of Georgia, with range advantage 0.353 and optimal defense frequency 0.820.

3 Conclusion

Any poker player looking for understandable theoretical takeaways for improving their strategy should study the 8 canonical solutions we have identified from the "Florida-shaped" relationship between the range advantage of the first player and the optimal defense frequency for the second player. For future work we would like to identify features for predicting the optimal defense frequency that can lead to an even crisper fit than range advantage, perhaps some that can lead to a strong low-variance linear prediction.

References

[1] Noam Brown and Tuomas Sandholm. Superhuman AI for heads-up no-limit poker: Libratus beats top professionals. *Science*, 2017.

[2] Sam Ganzfried. Reflections on the first man versus machine no-limit Texas hold 'em competition. *AI Magazine*, 38(2), 2017.

[3] Sam Ganzfried and Farzana Yusuf. Computing human-understandable strategies. *Games*, 8(4):1–13, November 2017.

[4] Matej Moravčík, Martin Schmid, Neil Burch, Viliam Lisý, Dustin Morrill, Nolan Bard, Trevor Davis, Kevin Waugh, Michael Johanson, and Michael Bowling. Deepstack: Expert-level artificial intelligence in heads-up no-limit poker. *Science*, 2017.