TwoPlusTwo Magazine

NoGTO, the Theory of Live Play, Neural Networks Prevail, Part 1

by Brian Space
Two Plus Two Magazine, Vol. 17, No. 7

Top Live Players Would Out Perform Bots

I have a lot of skill and experience that I bring to live poker. In my journey to improve, a long-standing question lingered. I was more successful when I did what I wanted to do, rather than when I made a decision based on poker theory or informed by a particular model. To be clear, my plays almost always made technical sense retrospectively -- to a degree that was beyond my cognitive abilities in the moment. Further, the intuitive plays I was driven toward can be thought of in terms of models. They could be framed as selecting an exploitive branch of the game tree or a particular kind of play that represents only a part of equilibrium strategy that was especially well suited to the complicated poker moment I found myself in.

I spent a lot of time in study, in therapy, and in poker playing examining this issue. I will share some of the personal and general insights that resulted. The take home message is the brain is a powerful neural network that can be trained to collate data and make superior model independent decisions. These choices can be interpreted across the frameworks used to make poker decisions. But, in my experience they are clearly better than a using a default model. They are nonetheless fuzzy as products of a complex neural network implicitly weighting factors from vastly different areas of live poker information. Brains are complicated.

To point, optimal live play – people playing the best live poker – do not derive their actions from a computational solution. I will call the analytic version mathematical poker and much effort goes into serious poker study along these lines. My assertion is that top live poker players play much better than a computer can in this complicated social environment.

The focus will be on no-limit Texas hold 'em (NLH) but the concepts are more general and apply aptly to big bet poker games where errors are especially costly. Consider, highly successful in-person poker players have a mixture of explicit and implicit skill in navigating the live arena. They play in a fashion that synthesizes the wide arena of information available to achieve what can be thought of as constrained optimal lie poker solutions. Such strategies need not be those that resemble equilibrium solutions to the mathematically determined game state. The best plays use knowledge of human behavior and emotion, social considerations, and accounting for the slow nature of live play.

Don't Think or Play Like a Robot in Live Poker

Thus, I will argue that while equilibrium solutions to poker situations will always form a solid basis for making poker decisions, they are best used from a mental framework that is vastly different. The mathematical derived style of poker has come to be known in the poker world, for worse more than better, as game theory optimal or GTO. Note, true GTO decision making is the optimal response given all the information available -- thus I am asserting the top-level live play eclipses the mathematically determinable solution by a wide margin. Live win rates reflect this reality as accomplished professionals routinely achieve win rates in excess of 10bb/hour at middle high stakes, which is far beyond what a GTO bot would achieve in the same circumstance (while facing substantial house rake). The GTO solution wins slowly but surely -- the life of a live player is short and this approach will not suffice.

I believe this is true at all levels of live poker and clearly the mixture of information available varies widely depending on the nature of one's opponents. Nonetheless, consistent big winners are "exploiting" relative to mathematical optimal strategies in a live, truly optimal approach informed by experience and both conscious and unconscious cues. Poker players can play "in the zone" in ways very similar to athletes playing at a high level where skill and intuition merge into what is can be called a flow state. Of course, such information can also be used rationally in constructing explicit exploits. Some combination of these attributes is often best.

One way to look at the complexity is to see how we naturally approach games with a model in mind when choosing an action. Again, this may be an explicit or implicit decision, or the total strategy may be a continuously updated ad hoc mixture of the two.

Rock Paper Scissors (1/3,1/3,1/3)

Let's consider a very simple game where there is no difficulty in understanding or implementing a true GTO strategy, Rock Paper Scissors (RPS). A first interesting insight is that there is no mystery to the solution, yet people still play with a variety of strategies few of which resemble an approximation to optimal. Let's learn from this simple game. The GTO solution is randomly throwing 1/3 of each choice — this is unexploitable and breaks even vs the same strategy as is the nature of a GTO solution. The strategy also has no advantage at equilibrium with itself. When the GTO strategy faces any other strategy, it wins while remaining unexploitable.

A limitation is that optimal doesn't maximally exploit known errors from an opponent. If we knew that, for example, our opponent was randomizing but with say 50% rock, and 25% paper and scissors, one can construct a maximally exploitive counter strategy to win the most vs the known error. This is easy to imagine. If our opponent only threw paper, we would always choose scissors and never lose -- the constrained optimization in the more complex case is a simple generalization of this to obtain optimal random frequencies. This is an important point. The GTO solution of throwing RPS with a 1/3 frequency would always win and be unexploitable but no sane person would use it against someone throwing 50% rock.

Such an optimal equilibrium strategy is always possible against a fixed, albeit possibly randomized, known opposing strategy. In this case, it is also constructed as maximally exploitive. Optimal is the best we can do given the information and constraints present. Optimal equilibrium means that, because our approach is the best possible counter, it cannot be improved so as long as our opponent's strategy or other constraints do not change. Therefore, the game will proceed forever with the two strategies producing a predictable outcome, one beating the other with some average win rate.

Humans Think in Models and Play like Humans

Another point concerns how complete solutions break down into models of the game. For example, in poker it is common to characterize bets as for value or bluffs. In some cases, these distinctions are clear. Other times, the actions required by an optimal solution may be irreducible – they are a part of the solution but not rationalizable in terms of a simpler model of the game. In the context of RPS, imagine asking why someone (or something?) just threw eleven papers in a row. If it's GTO play, the answer is it's part of the solution and no further insight is available beyond checking the random number generator used to execute the strategy. Such patterns will emerge from truly random play. (For a discussion of the complexities of randomness see Poker as War: reducible & Quantum Games, Randomness and Free Will?

The full mathematical solution space of poker has both rationalizable models and simplifications and non-reasoned elements that are just the ultimate function of the objective function (e.g. winning money)-- things you need to do to remain at equilibrium but do not easily align with any interpretation. These elements can't be extracted from simple models. Why we bet in poker is like this, call it for value, a bluff, a merge it can sometimes be best understood as part of the ultimate global optimization. So, in constructing strategies models are essential but also of very high but limited ultimate utility.

These kinds of considerations pass into poker strategy robustly. Nonetheless, complications ensue. Poker tends to be played multiplayer, and when one maximally exploits one strategy, the new exploitive strategy is itself exploitable by both an observant exploitee or another opponent in the game. Any net exploitation benefit is the difference between the gains and losses from being both exploiter and potentially exploitee yourself.

Still, let's go back to RPS and glean more insight. Why would anyone ever play a simply solved game? Consider: humans don't do randomness well. For example, the truly random solution will have us throw those eleven papers in a row at some frequency. People will not include these rare events at appropriate frequencies. In fact, in experiments where people attempt to imitate randomness, they tend to do so by making it look too perfectly balanced. Actual randomness includes streaks and ordered patterns that offend human sensibilities in that their conception of randomness implies a lack of local order.

Further, humans are emotional. Even perfectly random play will lead to winning and losing streaks. Imagine a recreational RPS player that over plays rock when losing and angry— they are easily exploitable by nudging up our paper frequency appropriately. One would then monitor the player's mood and adjust back to equilibrium as the player calms down or perhaps starts winning randomly. Perchance you are playing RPS for money against a very wealthy individual that just likes to win with paper and doesn't care that the resulting strategy loses overall. Thus, one can play RPS with an edge vs. real humans by exploiting whatever human frailties are manifesting themselves at any given time. Winning the game requires knowing what's going on.

Be a Model Citizen

Another point to consider, if someone is having fun losing at RPS, don't show up with a backpack and headphones and eat bean sprouts out of Tupperware in the gaming area without an obvious sense of humor. Those who do will soon find themselves playing with people that throw Rock, Paper, and Scissors with a fully randomized 1/3 frequency before you know it. People react to their environment and play better or worse, within their own skill set, depending on intrinsic and extrinsic motivation. In poker, bean sprouters always motivated me to play my best.

Casinos understand this concept – they even tell players optimal strategy at Black Jack or give away cards with the proper plays on them. Aside from the fact that they have a built in house advantage, they know people like to gamble and have fun so let them do so. If the atmosphere is fun, people don't try and play perfectly.

Further, because there is a lot of variance associated with randomized near even strategies, there will be winners and losers even in near optimal RPS. Human nature and experience suggest that high flying new RPS pros will be convinced that their mastery of the number 1/3 has led them to riches and domination vs. their unfortunate sprouted yet withering peers that keep having their rock covered in paper. Variance dominates when edges are small or nonexistent.

These kinds of considerations form a framework for playing live poker. The best live poker players construct optimal solutions on the fly with useful yet imperfect information, employing clues both implicitly or explicitly in constructing strategies and making plays. The strategies are exploitive yet informed by considerations of both the GTO and exploitive optimal equilibrium solutions. An analogy to mathematical poker is performing a simulation with constraints reflecting what one knows about the play of an opponent or group of opponents. This might involve inputting non-optimal opponent ranges, fixed nodes, or fixing non-optimal frequencies in constructing a counter-strategy.

The best live poker players are also fun to play with. They take their losses with grace and don't always try to get the best of every possible situation. They are grateful that less skilled players are willing to give them action and they give back in other ways.

The take home messages from Part 1 are:

1. People still play RPS and it is solved
2. There are winners at RPS (see Rock Paper Scissors Advanced Strategies).
3. There are only (good) games if people want to play with you.
4. One should use a variety of models to assess an action at any decision point. Some of these should incorporate human behaviors.
5. Interestingly, there are non-reducible strategic elements in games – things that are done optimally that do not appear in any simple model of the game. Put another way, actions are taken that are hard to easily classify in terms of obvious value in the simpler models of the game. Why do we make some bets can be a question without a simple answer?

Part 2 will continue this exploration with a focus on live poker games.

Brian Space is a scientist and professor seeking people to play Quantum Statistical Mechanics for money in Raleigh, NC. He plays poker but is no old man coffee. Remember the GTO strategy ignores that of the other players – you really want to play like that? His poker articles are available on his web site: http://drbrian.space/poker.html

Discuss the Magazine | Contribute to the Magazine

Poker Strategy and Other Topics – July 2021