Universidad de San Andrés
Departamento de Economía
Licenciatura en Economía

Perfect Bayesian Equilibrium in Kuhn Poker

MARTIN IÑAKI LORIENTE Y JUAN CRUZ DIEZ
Legajos: 31075 y 31033

Mentores: Lucía Quesada y Christian A. Ruzzier

Buenos Aires, Argentina

10/10/2023

Abstract: In 1950, Harold W. Kuhn introduced a simplified version of poker referred to as Kuhn Poker and solved it using the notion of Nash Equilibrium. His pioneering work inspired subsequent scholars who applied similar methodologies to other poker versions. In contrast, we adopt a different procedure by employing Harsanyi's approach to reach a Perfect Bayesian Equilibrium (PBE), a concept that emerged two decades after Kuhn's original solutions. While computational techniques have greatly advanced the analysis of various poker variations, achieving a PBE remains elusive. Some studies suffer from methodological flaws, as they overlook the importance of incorporating beliefs into their analysis. In our research, we also conducted a rationality study and found that relaxing the sophistication of a player leads to a shift in optimal strategies towards more exploitative ones.

1 Introduction

Poker stands as an undeniable titan among card games, capturing the hearts and minds of more than 120 million online players worldwide (Offshore, 2021). Notably, a select few have amassed staggering fortunes, exceeding 10 million dollars in earnings from the prestigious World Series of Poker Main Event (Mansilla, 2020). Despite its ubiquitous appeal, poker remains an enigma, presenting an enduring challenge, even to computer science. Recent strides in artificial intelligence have gravitated toward Nash Equilibrium concepts and loss function minimization to nurture the growth of formidable algorithms. These breakthroughs have resulted in strategies of remarkable complexity, surpassing unaided achievements and even excelling in multiplayer poker games (Risk and Szafron, 2010; Brown and Sandholm, 2019). Noteworthy contributions in this realm can be found in Bowling *et al.* (2017), where the authors approximate an equilibrium akin to Perfect Bayesian Equilibrium (PBE) by harnessing computational methods to minimize a penalty function. Additionally, in their 2017 study, Moravčík *et al.* harnessed the power of information asymmetry reasoning and deep learning techniques to enhance poker strategies. While these contributions are undeniably significant, pursuing a refined equilibrium provides valuable insights. Artificial intelligence struggles to explain why some actions lead to equilibrium while others do not, and these findings lack robust verification. We can delve into these concepts through simplified poker scenarios.

Simple games often serve as gateways to gaining insight into the complexities of more profound real-world scenarios. Game theory, since its inception, has ardently sought to unveil the intricacies of bluffing within poker (Borel, 1938; von Neumann and Morgenstern, 1947; Bellman and Blackwell, 1949; Nash and Shapley, 1950), with bluffing remaining a pivotal subject of poker research (Norman, 2012; Cassidy, 2015). Remarkably, Harold William Kuhn, in "Contributions to the Theory of Games Vol. I" introduced "A Simplified two-person-poker" (Kuhn, 1950). In this rendition, Kuhn simplified poker to a duel between two players, with a deck compounded by three cards numbered 1, 2, and 3. The hierarchy is clear: 3 trumps 2, and both beat 1. Players must compulsorily bet 1 monetary unit, known as the ante, to initiate the game, with a singular opportunity to increase the pot by 1. No re-bets complicate matters. Player A, named Ana, commences. Ana's card influences her decision to *bet* or *check*. If she *bets*, player B, christened Bill, must choose to *call* or *fold*, culminating in a showdown. If Ana *checks*, Bill faces a similar decision: to either *bet* or *check*. Should he *bet*, Ana confronts the possibility of calling or folding, and the game concludes. If both *check*, they must show their cards directly, and the one who wins keeps the pot, composed of the antes. Like poker, this simplified version is a zero-sum game. Interested readers can partake in this game by clicking **here**, configuring parameters $r = m = 1$ and $n = 2$ (Prisner, 2014). Nevertheless, how should we solve this game?

The advent of Nash Equilibrium in 1951 revolutionized economic studies. It provided a solution concept where strategies preclude any unilateral deviation to increase expected utility (Nash, 1951). However, a refinement was needed to address sequentiality and incomplete information, which ultimately materialized as PBE many years after Kuhn's pioneering work (Fudenberg and Tirole, 1991). In PBE, strategies form a Bayesian Equilibrium within each continuation game, hinging on specified beliefs updated via Bayes' rule when feasible. Kuhn Poker operates in the realm of incomplete information, wherein players possess only partial knowledge of opponents' cards and strategies. In poker, players navigate decisions grounded in their beliefs about their opponents' cards.

By applying Harsanyi's (1967) approach, we recast the game of incomplete information, reinterpreting it as one of imperfect information. Nature begins by selecting realizations of random variables determining player types: their cards. These random variables are concealed from the players, with each player only privy to their respective card. This paradigm shift into imperfect information means information sets encompass multiple decision nodes. Kuhn (1950) did not recognize this facet in his solutions, as more sophisticated methods emerged after his work. Prisner's (2014) solutions to many versions of this game, like Kuhn's original work, grapple with methodological flaws due to a similar oversight. This problem holds for Borel (1938) and von Neumann and Morgenstern (1947), who tackled simplified poker variants without incorporating beliefs, leading to flawed solutions (Norman, 2012). This issue persists in modern attempts to solve intricate iterations of Kuhn Poker employing artificial intelligence (Szafron *et al.*, 2013; Billingham, 2018; van der Werf, 2022) and in mathematical solutions to other similar games like VNM Poker (Prisner, 2014; Valenzuela, 2021).

While less refined equilibrium notions may suffice in some simultaneous games or contexts with perfect information, they buckle under the weight of private information in sequential games. To thrive, we must scrutinize each information set, extracting optimal actions in harmony with our beliefs about the unfolding situation aligning with our chosen strategies. Perfect

Bayesian Equilibrium coalesces strategy profiles with descriptions of players' beliefs across their information sets (Watson, 2013). These beliefs reflect players' assessments regarding each other's types, adapting as new information surfaces. PBE is invaluable for dissecting games where players continually gain insights and adjust their strategies and beliefs accordingly.

Some authors have correctly solved simplified poker games (Reiley *et al.*, 2008). Nonetheless, our work aspires to a loftier pinnacle within PBE solutions in poker due to the intricate interplay of beliefs. Keen observers will notice that this game encompasses no subgames apart from itself due to these elaborate interactions.

It is necessary to acknowledge that all Nash Equilibriums serendipitously qualify as PBE in this specific case, though this may not hold for all games. However, PBE might still encompass equilibria that appear counterintuitive. This challenge arises because PBE imposes no restrictions on beliefs in scenarios of zero probability, often described as an off-equilibrium path. Hence, a more nuanced refinement, such as the concept in Kreps and Wilson (1982) of consistent beliefs, could narrow the spectrum of permissible beliefs in off-equilibrium path situations, potentially reshaping Kuhn's solutions.

2 Solving the Game

We begin solving the game by studying the dominated strategies. In $Figure$ 1 we can find the extensive form of the Kuhn Poker. In $Table$ 1 we show the definitions of all the actions and beliefs used on the extensive form. In this case, the order in which we exclude strategies from the game does not affect the equilibrium.

First, there are some evident flaws with folding with a 3 after a bet. This is because players know it would generate a higher payoff to call, as there is no better card. This situation holds for both players. The same logic appears when a player calls or folds against a bet with a 1. Calling with a 1 always leads to losing more money than just folding. However, in both cases, we are talking of weakly dominated strategies because if the other player never bets these decisions will not affect the expected payoff. So, strategies that include these actions will be weakly dominated by those with the same actions except for the flawed ones.

Then, there is also a more sophisticated dominated strategy. When a player has a 2, he should not *bet*. Why? Because he can only enlarge his losses but never his income. The other player can only have a 1 or a 3. If we are on the first case, the second player will never call; if we are on the second case, it is strictly worse for the first player to bet. When we join both cases, we can see that strategies that include betting with a 2 are dominated by the same strategies that play *check* with a 2. In the case of A, this is strictly dominated after we remove the previous weakly dominated actions for B, as it reduces the expected payoff. For B, betting with a 2 is weakly, and not strictly, dominated because if A plays a strategy that will always bet at the start with a 3, then the payoff will not be reduced by betting with a 2 later, as A will not hold a 3. Player B has another group of weakly dominated strategies. These are the ones that play *check* with a 3. If B has a positive belief of c_2^A (A calling a bet when she has a 2), he should always play *bet* with a 3.

We will exclude strictly and weakly dominated strategies to determine a solution in this game. These strategies are highlighted in red within the extensive form for clarity. This procedure seems natural for a poker game where experienced players aim to maximize their expected earnings and minimize potential losses. They avoid actions that might lead to greater losses without increasing their expected earnings and refrain from playing strategies that limit their income without reducing their losses. This is also consistent with the literature on random mistakes while employing a strategy. Even if an equilibrium path is reached where player A will never *call* a bet with a 2, if player B believes there is a little possibility of A making a mistake, he will always want to *bet* with a 3. We will not apply this mistake procedure, but it is an insight that may be useful to understand why we should not validate an equilibrium with weakly dominated strategies.

As we look for an equilibrium where players have consistently updated beliefs, we need Ana and Bill to determine them with Bayes' rule whenever possible. Then, we will calculate it for every belief, and we will use this information with caution. We must apply the rule only when the node is reached with a positive probability. If this does not happen, we are facing less restrictions. The definition of PBE requires the receiver's belief to be consistent with the sender's strategy and Bayes' rule. However, it does not require anything about signals the sender never sends since Bayes' rule is not applicable for events with zero probability.

Behavioral strategies provide a simplified approach to tackling this complex game. They differ from mixed strategies. While both share similarities, behavioral strategies are notably more handy in our context. In mixed strategies, one must consider numerous potential strategies for the game and then decide which probabilities to employ. In contrast, behavioral strategies involve crafting explicit probabilities as instructions that cover every conceivable situation. Rather than randomizing these instructions, we enumerate all available options and assign probabilities to each. Like mixed strategies, other players remain unaware of the specific actions employed, as the player implements a random way to determine which option to choose. Consequently, the behavior remains unpredictable. However, unlike mixed strategies, the player has to repeatedly randomize while playing rather than making a single initial randomization.

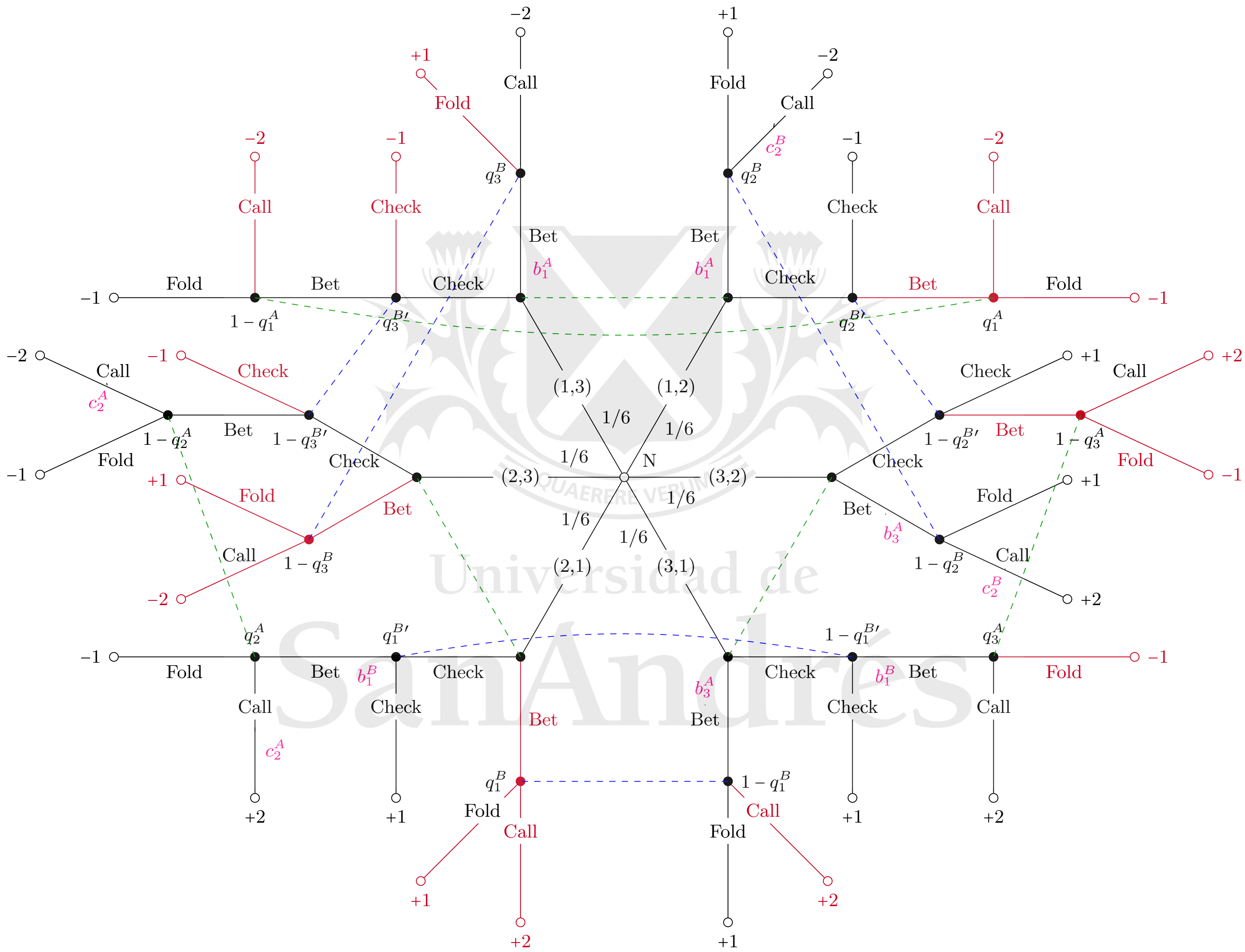

Figure 1: Kuhn Poker Player A Player B Dominated Strategies

Variable	Definition
Actions	
b_1^A	Probability that A plays *Bet* on the start when she has a 1
c_2^A	Probability that A plays *Call* at the end when she has a 2
b_3^A	Probability that A plays *Bet* on the start when she has a 3
b_1^B	Probability that B plays *Bet* after A checks when he has a 1
c_2^B	Probability that B plays *Call* after A checks when he has a 2
Beliefs	
q_1^A	Probability that A, when she gets a 1, assigns to B having a 2, conditional on B betting
$1-q_1^A$	Probability that A, when she gets a 1, assigns to B having a 3, conditional on B betting
q_1^B	Probability that B, when he gets a 1, assigns to A having a 2, conditional on A betting
$1-q_1^B$	Probability that B, when he gets a 1, assigns to A having a 3, conditional on A betting
$q_1^{B'}$	Probability that B, when he gets a 1, assigns to A having a 2, conditional on A checking
$1-q_1^{B'}$	Probability that B, when he gets a 1, assigns to A having a 3, conditional on A checking
q_2^A	Probability that A, when she gets a 2, assigns to B having a 1, conditional on B betting
$1-q_2^A$	Probability that A, when she gets a 2, assigns to B having a 3, conditional on B betting
q_2^B	Probability that B, when he gets a 2, assigns to A having a 1, conditional on A betting
$1-q_2^B$	Probability that B, when he gets a 2, assigns to A having a 3, conditional on A betting
$q_2^{B'}$	Probability that B, when he gets a 2, assigns to A having a 1, conditional on A checking
$1-q_2^{B'}$	Probability that B, when he gets a 2, assigns to A having a 3, conditional on A checking
q_3^A	Probability that A, when she gets a 3, assigns to B having a 1, conditional on B betting
$1-q_3^A$	Probability that A, when she gets a 3, assigns to B having a 2, conditional on B betting
q_3^B	Probability that B, when he gets a 3, assigns to A having a 1, conditional on A betting
$1-q_3^B$	Probability that B, when he gets a 3, assigns to A having a 2, conditional on A betting
$q_3^{B'}$	Probability that B, when he gets a 3, assigns to A having a 1, conditional on A checking
$1-q_3^{B'}$	Probability that B, when he gets a 3, assigns to A having a 2, conditional on A checking

Table 1: Actions and beliefs

We begin using Bayes' rule, which calculates the conditional probability of reaching each node given our prior information. We give the Bayes' rule for the belief $q_1^{B'}$:

$$q_1^{B'} = \frac{\frac{1}{6}}{\frac{1}{6} + \frac{1}{6}(1 - b_3^A)} = \frac{1}{2 - b_3^A}$$

In this example, player B has drawn a 1 and has to develop beliefs of what card player A got, either a 2 or a 3. For the top node to be reached, player A will always play *check* as she has a 2, so it is reached with probability $\frac{1}{6}$. The probability of reaching the information set will be the probability of reaching the top node, which is $\frac{1}{6}$, plus the probability that the bottom node is reached, which is $\frac{1}{6}$ times the probability player A checks with a 3, $(1 - b_3^A)$.

We showcase the Bayes' rules for all the beliefs in the game. Notice that $q_1^A = q_1^B = 0$ because the players with a 2 will never bet as it is a dominated strategy, so beliefs that the other player did play *bet* with a 2 are equal to 0. The same occurs with $q_3^A = q_3^B = 1$. In this case, the player has a 3, so when the other player plays *bet*, he is confident that he is betting with a 1, as betting with a 2 is dominated. Our reader can visually find the null beliefs at the red nodes defined in the extensive form.

The Bayes' rules of the game result:

$$
\begin{array}{lll}
q_1^A = 0 & q_1^B = 0 & q_1^{B'} = \dfrac{1}{2 - b_3^A} \\
q_2^A = \dfrac{b_1^B}{1 + b_1^B} & q_2^B = \dfrac{b_1^A}{b_1^A + b_3^A} & q_2^{B'} = \dfrac{1 - b_1^A}{2 - b_1^A - b_3^A} \\
q_3^A = 1 & q_3^B = 1 & q_3^{B'} = \dfrac{1 - b_1^A}{2 - b_1^A}
\end{array}
$$

You should notice that these probabilities are in the interval $[0, 1]$ and that we can find the remaining ones simply by subtracting each belief from 1; for example, $1 - q_1^A = 1 - 0 = 1$.

As we now understand the application of Bayes' rules, we develop the process for calculating the Perfect Bayesian Equilibrium. To do it, we will first exhibit each player's conditions to play each possible action willingly. We can call them the incentive compatibility constraints. Remember, we analyze each action separately under the logic of behavioral strategies. We are assessing the remaining actions, having excluded those that were dominated. So, we only have five: b_1^A, c_2^A, b_3^A, b_1^B and c_2^B (Definitions in *Table* 1).

A will weakly want $b_1^A = 1$ if $c_2^B \leq \frac{1}{3}$	B will weakly want $b_1^B = 1$ if $q_1^{B'} \geq \frac{1}{3-3c_2^A}$
A will weakly want $b_1^A = 0$ if $c_2^B \geq \frac{1}{3}$	B will weakly want $b_1^B = 0$ if $q_1^{B'} \leq \frac{1}{3-3c_2^A}$
A will be indifferent with b_1^A if $c_2^B = \frac{1}{3}$	B will be indifferent with b_1^B if $q_1^{B'} = \frac{1}{3-3c_2^A}$
A will weakly want $c_2^A = 1$ if $q_2^A \geq \frac{1}{4}$	B will weakly want $c_2^B = 1$ if $q_2^B \geq \frac{1}{4}$
A will weakly want $c_2^A = 0$ if $q_2^A \leq \frac{1}{4}$	B will weakly want $c_2^B = 0$ if $q_2^B \leq \frac{1}{4}$
A will be indifferent with c_2^A if $q_2^A = \frac{1}{4}$	B will be indifferent with c_2^B if $q_2^B = \frac{1}{4}$
A will weakly want $b_3^A = 1$ if $c_2^B \geq b_1^B$	
A will weakly want $b_3^A = 0$ if $c_2^B \leq b_1^B$	
A will be indifferent with b_3^A if $c_2^B = b_1^B$	

We have three possible decisions for each of the five actions, leaving us with $3^5 = 243$ options. PBE usually starts by searching for separating or pooling strategies (Levin, 2002). In that case, different types of player A use different actions so information is perfectly informed, or different types of A use the same actions so no information is transmitted in A's actions. All these pure strategies do not conform an equilibrium in Kuhn's game. This also applies to semi-separating situations, where multiple types of the player select the same specific actions, while a single type chooses other actions. In such cases, some degree of learning takes place. As none of these ideas helped, we must search for mixed strategies in equilibrium.

We will check for PBE to get all possible equilibrium by individually analyzing the 243 possible behavioral strategies. We divided the strategies by groups and subgroups using the compatibility constraints. For the first group, we define the first action of Ana. We know that with a 2 she will always *check*, but with a 1 or a 3 she can either *bet*, *check*, or play a probability between zero and one. So, with that in mind, we get the group of 9 possible first actions in which A can play a set of pure or mixed strategies.

Then, we will have a subgroup from each possible group that will contain the actions of Bill. This subgroup will follow the first group as it is the second step in the game. In this case, player B can *call* or *fold* with a 2 after a bet from player A, and *check* or *bet* with a 1; he will always *bet* with a 3. This can also be either pure, or he can mix the probabilities of his actions and is composed of 9 possible plans. Note that even though player A might never *bet*, we include player B's probability of calling with a 2 as a strategy must be a complete contingent action plan.

Finally, we have a final subgroup as we reach the third step in the game. Ana has to decide between calling or folding against Bill's bet at this stage. We get a set of 3 possible actions: player A calls, folds, or randomizes. In this case, we also need to define player A strategy even though the information set may not be reached as the game may always end.

Given this approach, we solved the game by checking if the players' beliefs were consistent with the strategies. We have the following example of the procedure:

- Player A *bets* with some probability with a 1 but leaves some probability to *check*, $b_1^A \in (0,1)$, and always *checks* with a 3, $b_3^A = 0$.
- By the incentive compatibility constraints, $c_2^B = \frac{1}{3}$ and $c_2^B \leq b_1^B$.
- Given this, Bayes' rules require that the following beliefs are $q_1^{B'} = \frac{1}{2}$, $q_2^B = 1$.
- But as $q_2^B = 1$ this requires by the constraints that $c_2^B = 1$. This is an absurd, as Ana would need $c_2^B = \frac{1}{3}$ to want to hold her strategy.

The group described before shows a contradiction as the beliefs are inconsistent with what players would like to do, so we do not need to check the subgroups in the other steps, as there cannot be any equilibrium with that set of actions. We continue this process for all the possible combinations.

After rejecting 240 options, we only have three left: $b_1^A = \frac{b_3^A}{3}$, $c_2^A = \frac{1}{3} + \frac{b_3^A}{3}$, $b_3^A \in (0,1)$, $b_1^B = \frac{1}{3}$, $c_2^B = \frac{1}{3}$ and the two corner solutions with an identical composition but $b_3^A = 0$ or $b_3^A = 1$. Understanding them as three different equilibria under the same structure is relevant because the incentive compatibility constraint changes in the two exceptional cases. We can write these behavioral strategies for Ana and Bill as $A : (b_1^A = \frac{b_3^A}{3}, b_2^A = 0, b_3^A \in [0,1], c_1^A = 0, c_2^A = \frac{1}{3} + \frac{b_3^A}{3}, c_3^A = 1), B : (b_1^B = \frac{1}{3}, b_2^B = 0, b_3^B = 0, c_1^B = 0, c_2^B = \frac{1}{3}, c_3^B = 1)$, where we have the probabilities for each player to *bet* and to *call* with every card in their corresponding nodes.

We have found infinite equilibria under a structure where Ana bluffs. Specifically, all equilibria with $b_3^A > 0$ imply $b_1^A > 0$. In those cases, she decides to *bet* even with the worst card. This structure is logical when we think about it. Ana would like to always *bet* with a 3 and never with a 1, but as she can only play *bet* on the first round, this would give Bill enough information to know that he should never *call* when he sees a *bet*. In order to mitigate this issue and optimize her expected payoff, Ana opts to reduce her bets when she has a 3 and increases them when she has a 1. Conversely, Bill can perfectly predict her actions in equilibrium, but as he does not know Ana's card, he cannot take great advantage. In the end, playing *bet* with a 1 one-third of the time and a *call* with a 2 one-third of the time is part of a best response of B, so A can be indifferent to betting with a 1. On her side, the probability of Ana calling this *bet* will increase proportionally with her *bet* with a 3. We can find in the equilibrium the insights of bluffing in a context where both players do not know the other player's type. As von Neumann claimed: "Real life consists of bluffing, of little tactics of deception, of asking yourself what is the other man going to think I mean to do" (Bronowski, 1973).

Behavioral strategies have proven efficient in solving the game. Let us go from them to a mixed strategy. Equilibrium behavioral strategies yield three numbers between 0 and 1 for A and two for B. The process to transform these numbers is natural: we take all the pure strategies and multiply them for the probabilities corresponding to each information set. We show the procedure for one example of A and then state the results. Let us define B_1 as betting with a 1, C_2 calling with a 2, and C_3 calling with a 3. Then, the probability that A plays the pure equilibrium that combines these decisions in her mixed equilibrium strategy is $P_A(B_1, C_2, B_3) = \frac{b_3^A}{3}(\frac{1}{3} + \frac{b_3^A}{3})b_3^A = \frac{(b_3^A)^2}{9} + \frac{(b_3^A)^3}{9}$. We take $\frac{b_3^A}{3}$ corresponding to B_1 because it is the probability that the behavioral strategy assigns to A betting with a 1. In the same way, we take $\frac{1}{3} + \frac{b_3^A}{3}$ corresponding to C_2 and b_3^A to B_3. When we have the opposite option in a probability, we do the complement by subtracting the probability the behavioral strategy assigns to one.

$$P_A(B_1, C_2, B_3) = \frac{(b_3^A)^2}{9} + \frac{(b_3^A)^3}{9} \qquad P_B(B_1, C_2) = \frac{1}{9}$$
$$P_A(B_1, C_2, C_3) = \frac{b_3^A}{9} - \frac{(b_3^A)^3}{9} \qquad P_B(C_1, C_2) = \frac{2}{9}$$
$$P_A(B_1, F_2, B_3) = \frac{2(b_3^A)^2}{9} - \frac{(b_3^A)^3}{9} \qquad P_B(B_1, F_2) = \frac{2}{9}$$
$$P_A(B_1, F_2, C_3) = \frac{2b_3^A}{9} - \frac{(b_3^A)^2}{3} + \frac{(b_3^A)^3}{9} \qquad P_B(C_1, F_2) = \frac{4}{9}$$
$$P_A(C_1, C_2, B_3) = \frac{b_3^A}{3} + \frac{2(b_3^A)^2}{9} - \frac{(b_3^A)^3}{9}$$
$$P_A(C_1, C_2, C_3) = \frac{1}{3} - \frac{b_3^A}{9} - \frac{(b_3^A)^2}{3} + \frac{(b_3^A)^3}{9}$$
$$P_A(C_1, F_2, B_3) = \frac{2b_3^A}{3} - \frac{5(b_3^A)^2}{9} + \frac{(b_3^A)^3}{9}$$
$$P_A(C_1, F_2, C_3) = \frac{2}{3} - \frac{11b_3^A}{9} + \frac{2(b_3^A)^2}{3} - \frac{(b_3^A)^3}{9}$$

Here, we can see the mixed strategies of equilibrium for both players. These are the positive probabilities assigned to pure strategies. A simple way to check that this procedure was proper is that the sum of all these probabilities for each player equals one independent of the value of b_3^A. We also need to define the beliefs in equilibrium. We have three cases. For $b_3^A = 0$ we have $q_1^A = 0$, $q_1^B \in [0,1]$, $q_1^{\mathrm{B}'} = \frac{1}{2}$, $q_2^A = \frac{1}{4}$, $q_2^B = \frac{1}{4}$, $q_2^{\mathrm{B}'} = \frac{1}{2}$, $q_3^A = 1$, $q_3^B \in [0,1]$, $q_3^{\mathrm{B}'} = \frac{1}{2}$. For $b_3^A = 1$ we have $q_1^A = 0$, $q_1^B = 0$, $q_1^{\mathrm{B}'} = 1$, $q_2^A = \frac{1}{4}$, $q_2^B = \frac{1}{4}$, $q_2^{\mathrm{B}'} = 1$, $q_3^A \in [0,1]$, $q_3^B = 1$, $q_3^{\mathrm{B}'} = \frac{2}{5}$. For $b_3^A \in (0,1)$ we have $q_1^A = 0$, $q_1^B = 0$, $q_1^{\mathrm{B}'} = \frac{1}{2-b_3^A}$, $q_2^A = \frac{1}{4}$, $q_2^B = \frac{1}{4}$, $q_2^{\mathrm{B}'} = \frac{3-b_3^A}{6-4b_3^A}$, $q_3^A = 1$, $q_3^B = 1$, $q_3^{\mathrm{B}'} = \frac{3-b_3^A}{6-b_3^A}$. Some beliefs are free because they are never reached in equilibrium. This is an example of the lack of PBE conditions under them that we discussed earlier.

Now, what are the expected payoffs of both players? We can find this number for any infinite equilibrium, which will be all equal. There are many ways to get to this number; for example, we can extract the payoff for any pure strategy involved inside a mixed strategy against the mixed strategy of the rival. If we do this, we will get the same payoffs Kuhn found by achieving a NE:

$$U_A = \frac{1}{6}\left[b_1^A(1 - 3c_2^B) + c_2^A(3b_1^B - 1) + b_3^A(c_2^B - b_1^B) - b_1^B\right]$$
$$U_A = \frac{1}{6}\left[b_1^B(3c_2^A - b_3^A - 1) + c_2^B(b_3^A - 3b_1^A) + (b_1^A - c_2^A)\right]$$

We can see both equilibrium payoffs with these conditions as it is a zero-sum game: $U_A = -\frac{1}{18}$, $U_B = \frac{1}{18}$. This result is consistent with the dealer advantage found in poker by Bowling *et al.* (2017) and of common knowledge in competitive poker. The intuition behind this is that the last player observes the previous actions and gets an information rent. It means that being capable of receiving signals of the other player type, their cards, is useful even on equilibrium.

3 Rationality

Kuhn Poker simplifies real poker, so we want to translate these intuitions into more popular versions. No Limit Texas Hold'em Poker is the most played type of poker. It is a table game with incomplete information that game theory still needs to solve due to its complexity (Sandholm, 2010). This is why progress has been made with some hands of cards to determine specific ranges of best plays or to solve simplified versions (Bowling *et al.*, 2017). Thus, given hands and positions, considering what the opponents did, we have a range of best responses called Game Theory Optimal Strategy (GTO poker). This way of playing poker is trendy as it is often used as a synonym for Nash Equilibrium, allowing players not to be exploited (Acevedo, 2019). Are these GTO poker strategies always the best choice, or are there specific situations where employing an exploitative strategy, such as targeting an opponent's weaknesses like frequent bluffing or under-betting strong hands, would be more effective? We may find some intuitions here.

Rationality assumptions imply that two sophisticated players are facing, that they are cold so that no emotions may affect their decisions, and that they care about the game and analyze it to take the best possible actions. When rationality assumptions apply, which could be the case in poker tournaments, where much money is disputed and players are sophisticated, we could translate intuitions from our Kuhn's Poker equilibrium. First, players should avoid letting their rivals exploit the information they reveal. The best example is Ana's actions in her first node. She decides to bet more with the best cards, but this does not mean she will always bet with highs and never with lows. She has to take a position in the trade-off between showing information about her type and placing money in the game. We believe this principle would hold, although we cannot solve a complete poker game. Another idea we can sustain is the seconds' mover advantage. We also suspect that seeing information about other players' types with their actions could be better than being first in poker games.

But what happens in a Casino game where some players may not be entirely rational? This could be the case if Bill is a professional player, but Ana is just trying to have fun while caring about her money. In real life, players are somewhat sophisticated and have different reasoning levels (Camerer *et al.*, 2004; Feng *et al.*, 2021). This seems to happen with complex games like chess or poker, where complexity leads to errors in the perception of alternatives' values. People underestimate how bad very complex wrong actions are and how good very complex right actions are (Salant and Spenkuch, 2023). Also, cognitive skills have a causal impact on sophistication, and individuals who are not probabilistically sophisticated tend to have low cognitive ability (Carpenter *et al.*, 2013; Almeida and Rangel, 2014). It has been proven that the heterogeneity of cognitive skills of groups of players affects the strategic interaction, and the less intelligent player could also learn a sophisticated strategy from the other player if we were on a repeated game (Proto *et al.*, 2021). In this situation, cognitive capacity could prove to change our solution (Fé *et al.*, 2022). As B is a more sophisticated player, he could manipulate A's beliefs to make her think that he is not (Bhatt *et al.*, 2010). These are problems that regular Game Theory or computational approaches cannot solve. This is because these methods do not calculate exploitative strategies against weak players; they only provide its core strategy in equilibrium. So, we will need to shock the assumptions made for rationality and force the equilibrium concept to make it more realistic in poker under new assumptions.

Let us face the same game, but now we will relax rationality for player A. Ana will be a recreational player, and she will believe she can play like a sophisticated player. We will model this as if both players can see the equilibrium strategies, but Ana deviates from it because of an external shock called $\sigma \in (0, \frac{1}{3}]$. This shock may be defined as noise, and it may be related to a distraction, a confused desire to bet, or just a low sophistication. We assume that Ana plays this strategy: $A : (b_1^A = \frac{1}{3} + \sigma, b_2^A = 0, b_3^A = 1, c_1^A = 0, c_2^A = \frac{2}{3} + \sigma, c_3^A = 1)$. As Bill is a poker expert, he sees that Ana will not play as in equilibrium. Now, he has to decide which strategy maximizes his payoff against Ana. If he plays the previous optimal strategy, he would expect a payoff of $\frac{1}{18}$, which is exactly the one expected in equilibrium before. But we will calculate his best response strategy.

With the incentive compatibility constraints and the corresponding Bayes' updates for B, we will find that the best response for B is to play $b_1^B = 0$ and $c_2^B = 1$, followed by the already defined actions. Bill would play a pure strategy where he continuously checks with a 1 and bets with a 2. He wants to play a pure strategy because Ana lacks the sophistication to adapt to his decision. He no longer needs to choose a mixed strategy that implements a random way to determine which option to choose. Bill will turn to this pure strategy even when the shock is little. This discontinuous jump in decisions is natural: Ana's strategy was the one that allowed Bill to mix his actions, so she randomized it perfectly to give the same expected payoff to both decisions. Now that Ana is deviating, Bill will adopt a pure strategy even with a shock close to 0. This new strategy would give him an expected payoff of $\frac{1}{18} + \frac{\sigma}{2}$, increasing his utility.

We can compare the payoffs and see that when player A deviates following her external shock, her expected utility stays at $-\frac{1}{18}$, meaning that a deviation from the equilibrium would make her expect the same payoff. Furthermore, if player B updates his strategy, given the external shock in Ana's strategy, her expected payoff lowers from $-\frac{1}{18}$ to $-\frac{1}{18} - \frac{\sigma}{2}$. This means that when we change our assumptions of the capability of the players to reach equilibrium, the attitude a sophisticated player should take may vary. This change takes a structure natural for poker players. There is an existing debate between the GTO strategies and the exploitative ones. This analysis should shed some light on this debate, showing that, under different assumptions, sometimes much more realistic, an exploitative strategy is better than a GTO one.

To test the robustness of our result, we can check if this result is something particular to a shock on the first player. Then, we return to the basic game and add a shock to Bill's play on equilibrium with $\sigma \in (0, \frac{1}{3}]$. So, we assume Bill will play $B : (b_1^B = \frac{1}{3} + \sigma, b_2^B = 0, b_3^B = 0, c_1^B = 0, c_2^B = \frac{1}{3} + \sigma, c_3^B = 1)$. Ana, a poker expert, sees that Bill will not play as in equilibrium. She would expect a payoff of $-\frac{1}{18}$ if she maintains the previous optimal strategy. However, her best response is to play $b_1^A = 0, c_2^A = 1$ with $b_3^A \in [0, 1]$. Under this best response, her new payoff is $-\frac{1}{18} + \frac{\sigma}{3}$, greater than before.

We can conclude that the debate of GTO vs. Exploitative strategies makes sense when our assumptions about rationality tremble. However, we also show that the equilibrium of the previous section is fragile to little shocks. This result is mainly driven by breaking the indifference principle that allows players to take mixed strategies. Both cases of non-pure rationality show that softly relaxing the rationality concept leads to new best responses for a sophisticated player against a casual one. We found that little changes around the actions of the previous equilibrium turn the best possible sophisticated strategy into an exploitative one.

4 Conclusions

Nash, Kuhn, and von Neumann stand as pioneers in the field of Game Theory, revolutionizing our understanding of strategic interactions. Their work has spurred an extensive body of literature, ushering in significant advancements that have equipped us with tools to refine and build upon earlier research.

These advancements have become increasingly relevant in the academic sphere, as many scholars continue to place unwavering faith in the accuracy of Nash Equilibrium-based solutions. However, our analysis reveals that blind adherence to this equilibrium concept is not always warranted. Even in cases where previous solutions appear sound, as seen in the context of Kuhn Poker, a more nuanced approach is essential. This game involves incomplete information, requiring the incorporation of beliefs into the equation. While this approach may introduce complexities in the calculations, it bestows invaluable insights into the game's dynamics and the motivations underlying each equilibrium strategy.

Our examination has reaffirmed the "last player advantage", a phenomenon previously identified in poker literature. Furthermore, through a rationality analysis, we have discerned that sophisticated players should transition from equilibrium to exploitative strategies under adjusted assumptions, more in line with common scenarios in casual poker. This shift in strategy opens up exciting avenues for future research. We also show that equilibrium is fragile to little shocks. Subsequent studies could explore diverse models for representing unsophisticated players, thereby testing the robustness of our findings. Additionally, these investigations could extend their scope to solve historical games correctly, shedding new light on classic problems in Game Theory.

References

Acevedo, M. (2019). Modern poker theory: Building an unbeatable strategy based on gto principles.

Almeida, S., & Rangel, M. (2014). Probabilistic sophistication, sources of uncertainty, and cognitive ability: Experimental evidence.

Bellman, R., & Blackwell, D. (1949). Some two-person games involving bluffing. *Contributions to the Theory of Games*, (35), 600–605.

Bhatt, M. A., Lohrenz, T., Camerer, C. F., & Montague, R. (2010). Neural signatures of strategic types in a two-person bargaining game.

Billingham, J. (2018). Equilibrium solutions of three player kuhn poker with n ¿ 3 cards: A new numerical method using regularization and arc-length continuation.

Borel, E. (1938). Traité de calcul des probabilités et ses applications. *Applications aux jeux de hazard*, *4*.

Bowling, M., Burch, N., Johanson, M., & Tammelin, O. (2017). Heads-up limit hold'em póker is solved. *Contributions to the Theory of Games*, 97–104.

Bronowski, J. (1973). *The ascent of man* [Access date: 27th of september, 2023]. https://archive.org/details/the-ascent-of-man-ep13

Brown, N., & Sandholm, T. (2019). Superhuman ai for multiplayer poker. *Science*, *365*(6456), 885–890.

Camerer, C. F., Ho, T. H., & Chong, J.-K. (2004). A cognitive hierarchy model games. *The Quarterly Journal of Economics*.

Carpenter, J., Graham, M., & Wolf, J. (2013). Cognitive ability and strategic sophistication. *Department of Economics, Middlebury College, United States*.

Cassidy, C. J. (2015). Early round bluffing in poker. *The American Mathematical Monthly*, *122*(8), 726–744.

Fé, E., Gill, D., & Prowse, V. (2022). Cognitive skills, strategic sophistication, and life outcomes. *Journal of Political Economy, The University of Chicago Press*, *130*(10).

Feng, J., Qin, X., & Wang, X. (2021). A bayesian cognitive hierarchy model with fixed reasoning levels. *Journal of Economic Behavior and Organization*, *192*, 704–723.

Fudenberg, D., & Tirole, J. (1991). Perfect bayesian equilibrium and sequential equilibrium. *Journal of Economic Theory*, *53*(2), 236–260.

Harsanyi, J. C. (1967). Games with incomplete information played by "bayesian" players, i-iii. part i. the basic model. *Management Science, Theory Series*, *14*(3), 159–182.

Kreps, D. M., & Wilson, R. (1982). Sequential equilibria. *Econometrica*, *50*(4), 863–894.

Kuhn, H. W. (1950). A simplified two-person poker. *Contributions to the Theory of Games*, *1*(3), 97–104.

Levin, J. (2002). Dynamic games with incomplete information.

Mansilla, S. G. (2020). *Récords de la wsop: Los jugadores que más dinero ganaron* [Access date: 28th of september, 2023]. https://pokerlogia.com/records-poker/records-de-la-wsop-los-jugadores-que-mas-dinero-ganaron/

Moravčík, M., Schmid, M., Burch, N., Lisý, V., Morrill, D., Bard, N., Davis, T., Waugh, K., Johanson, M., & Bowling, M. (2017). Deepstack: Expert-level artificial intelligence in heads-up no-limit poker. *Science*, *356*(6337), 508–513.

Nash, J. F. (1951). Non-cooperative games. *Annals of Mathematics*, *54*(2), 105–116.

Nash, J. F., & Shapley, L. S. (1950). A simple three-person poker game. *Contributions to the Theory of Games*, *1*(3), 105–116.

Norman, T. W. L. (2012). Bluffing as a mixed strategy. *Magdalen College, Oxford.*

Offshore. (2021). *How many people play poker around the world?* [Access date: 28th of september, 2023]. https://fastoffshore.com/2021/10/online-poker-sector-overview-2021/

Prisner, E. (2014). Game theory through examples. *Mathematical Association of America.*

Proto, E., Rustichini, A., & Sofianos, A. (2021). Intelligence, errors and cooperation in repeated interactions.

Reiley, D. H., Urbanic, M. B., & Walker, M. (2008). Stripped-down póker: A classroom game with signaling and bluffing.

Risk, N. A., & Szafron, D. (2010). Using counterfactual regret minimization to create competitive multiplayer poker agents. *International Foundation for Autonomous Agents and Multiagent Systems*, *1*, 159–166.

Salant, Y., & Spenkuch, J. L. (2023). Complexity and satisficing: Theory with evidence from chess. *National Bureau of Economic Research.*

Sandholm, T. (2010). The state of solving large incomplete-information games, and application to póker.

Szafron, D., Gibson, R., & Sturtevant, N. (2013). A parameterized family of equilibrium profiles for three-player kuhn poker. *AAMAS.*

Valenzuela, A. T. (2021). *Vnm póker* [Access date: 28th of september, 2023]. https://torrejonvalenzuela.com/post/2021/03/16/vnm-p%C3%B3ker.html

van der Werf, L. (2022). Analysis of nash equilibria for kuhn poker and its extensions. *University of Twente.*

von Neumann, J., & Morgenstern, O. (1947). Theory of games and economic behavior. *Princeton University Press.*

Watson, J. (2013). Strategy : An introduction to game theory. *New York : W. W. Norton and Company.*