

The Dynamics of Human Behaviour in Poker

Marc Ponsen ^a Karl Tuyls ^b Steven de Jong ^a
Jan Ramon ^c Tom Croonenborghs ^d Kurt Driessens ^c

^a *Universiteit Maastricht, Netherlands*

^b *Technische Universiteit Eindhoven, Netherlands*

^c *Katholieke Universiteit Leuven, Belgium*

^d *Biosciences and Technology Department, KH Kempen University College, Belgium*

Abstract

In this paper we investigate the evolutionary dynamics of strategic behaviour in the game of poker by means of data gathered from a large number of real-world poker games. We perform this study from an evolutionary game theoretic perspective using the Replicator Dynamics model. We investigate the dynamic properties by studying how players switch between different strategies under different circumstances, what the basins of attraction of the equilibria look like, and what the stability properties of the attractors are. We illustrate the dynamics using a simplex analysis. Our experimental results confirm existing domain knowledge of the game, namely that certain strategies are clearly inferior while others can be successful given certain game conditions.

1 Introduction

Although the rules of the game of poker are simple, it is a challenging game to master. There exist many books written by domain experts on how to play the game (see, e.g., [2, 4, 9]). A general consensus is that a winning poker strategy should be adaptive: a player should change the style of play to prevent becoming too predictable, but moreover, the player should adapt the game strategy based on the opponents. In the latter case, players may want to vary their actions during a specific game, but they can also consider changing their overall game strategy over a series of games (e.g., play a more aggressive or defensive style of poker).

Although some studies exist on modeling poker players and providing a best-response given the opponent model (see, e.g., [1, 8, 10]), not much research focuses on overall strategy selection. In this paper we address this issue by investigating the evolutionary dynamics of strategic player behaviour in the game of poker. We perform this study from an evolutionary game-theoretic perspective using the Replicator Dynamics (RD) [5, 6, 11, 12]. More precisely, we investigate the dynamic properties by studying how players switch between different strategies (based on the principle of selection of the fittest), under different circumstances, what the basins of attraction of the equilibria look like, and what the stability properties of the attractors are.

A complicating factor is that the RD can only be applied straightforwardly to simple normal form games as for instance the Prisoner's Dilemma game [3]. Applying the RD to poker by assembling the different actions in the different phases of the game for each player will not work, because this leads to an overly complex table with too many dimensions. To address this problem, overall strategies (i.e., behaviour over a series of games, henceforth referred to as *meta strategies*) of players may be considered. Using these meta strategies, a heuristic payoff table can then be created that enables us to apply different RD models and perform our analysis. This approach has been used before in the analysis of behaviour of buyers and sellers in automated auctions [7, 13, 14]. Conveniently, for the game of poker several meta strategies are already defined in literature. This allows us to apply RD to the game of poker. An important difference with previous work, is that we use real-world poker games from which the heuristic payoff table is derived, as opposed to the artificial data used in the auction studies. We observed poker games played on a poker website, in which human players competed for real money at various stakes.

Therefore, the contributions of this paper are twofold. First, we provide new insights in the dynamics of strategic behaviour in the complex game of poker using RD models. These insights may prove useful for strategy selection by human players but can also aid in creating strong artificial poker players. Second, unlike other studies, we apply RD models to real-world human data.

The remainder of this paper is structured as follows. We start by explaining the poker variant we focus on in our research, namely No-Limit Texas Hold'em poker, and describe some well-known meta strategies for this game. Next we elaborate on the Replicator Dynamics and continue with a description of our methodology. We end with experiments and a conclusion.

2 Background

In this section we will first briefly explain the rules of the game of poker. Then we will discuss meta strategies as defined by domain experts.

2.1 Poker

Poker is a card game played between at least two players. In a nutshell, the object of the game is to win games (and consequently win money) by either having the best card combination at the end of the game, or by being the only active player. The game includes several betting rounds wherein players are allowed to invest money. Players can remain active by at least matching the largest investment made by any of the players, or they can choose to fold (i.e., stop investing money and forfeit the game). In the case that only one active player remains, i.e., all other players chose to fold, the active player automatically wins the game. The winner receives the money invested by all the players.

In this paper we focus on the most popular poker variant, namely No-Limit Texas Hold'em. This game includes 4 betting rounds (or phases), respectively called the pre-flop, flop, turn and river phase. During the first betting round, all players are dealt two private cards (what we will now refer to as a player's *hand*) that are only known to that specific player. To encourage betting, two players are obliged to invest a small amount the first round (the so-called small- and big-blind). One by one, the players can decide whether or not they want to participate in this game. If they indeed want to participate, they have to invest at least the current bet. This is known as *calling*. Players may also decide to *raise* the bet. If they do not wish to participate, players *fold*, resulting in possible loss of money they bet thus far. During the remaining three betting phases, the same procedure is followed. In every phase, community cards appear on the table (respectively 3 in the flop phase, and 1 in the other phases). These cards apply to all the players and are used to determine the card combinations (e.g., a pair or three-of-a-kind may be formed from the player's private cards and the community cards).

2.2 Meta strategies

There exists a lot of literature on winning poker strategies, mostly written by domain experts (see, e.g., [2, 4, 9]). These poker strategies may describe how to best react in detailed situations in a poker game, but also how to behave over large numbers of games. Typically, experts describe these so-called meta strategies based on only a few features. For example, an important feature in describing a player's meta strategy is the percentage of times this player voluntarily sees the flop (henceforth abbreviated as *VSF*), since this may give insight in the player's hand selection. If a particular player chooses to play more than, let's say, 40% of the games, he or she may play with less quality hands (see [9] for hand categorization) compared to players that only see the flop rarely. The standard terminology used for respectively the first approach is a *loose* and for the latter a *tight* strategy. Another important feature is the so-called *aggression-factor* of a player (henceforth abbreviated as *AGR*). The aggression-factor illustrates whether a player plays offensively (i.e., bets and raises often), or defensively (i.e., calls often). This aggression factor is calculated as:

$$\frac{\%bet + \%raise}{\%calls}$$

A player with a low aggression-factor is called *passive*, while a player with a high aggression-factor is simply called *aggressive*. The thresholds for these features can vary depending on the game context. Taking into

account these two features, we can construct four meta strategies, namely: 1) loose-passive (LP), 2) loose-aggressive (LA), 3) tight-passive (TP), and 4) tight-aggressive (TA). Again note that these meta-strategies are derived from poker literature.

Experts argue that the TA strategy is the most profitable strategy, since it combines patience (waiting for quality hands) with aggression after the flop. One could already claim that any aggressive strategy dominates all passive strategies, simply by looking at the rules of the poker game. Note that games can be won by having the best card combination, but also by betting all opponents out of the pot. However, most poker literature will argue that adapting a playing style is the most important feature of any winning poker strategy. This applies to detailed poker situations, i.e., varying actions based on current opponent(s), but also varying playing style on a broader scale (e.g., switching from meta strategy). We will next investigate how players (should) switch between meta strategies in the game of No-Limit Texas Hold'em poker.

3 Methodology

In this section we concisely explain the methodology we will follow to perform our analysis. We start by explaining Replicator Dynamics (RD) and the heuristic payoff table that is used to derive average payoffs for the various meta strategies. Then we explain how we approximate the Nash equilibria of interactions between the various meta strategies. Finally, we elucidate our algorithm for visualizing and analyzing the dynamics of the different meta strategies in a simplex plot.

3.1 Replicator Dynamics

The RD [11, 16] are a system of differential equations describing how a population of strategies evolves through time. The RD presumes a number of agents (i.e., individuals) in a population, where each agent is programmed to play a pure strategy. Hence, we obtain a certain mixed population state x , where x_i denotes the population share of agents playing strategy i . Each time step, the population shares for all strategies are changed based on the population state and the rewards in a payoff table. Note that single actions are typically considered in this context, but in our study we look at meta strategies.

An abstraction of an evolutionary process usually combines two basic elements, i.e., selection and mutation. Selection favors some population strategies over others, while mutation provides variety in the population. In this research, we will limit our analysis to the basic RD model based solely on selection of the most fit strategies in a population. Equation 1 represents this form of RD.

$$\frac{dx_i}{dt} = [(Ax)_i - x \cdot Ax] x_i \quad (1)$$

In Equation 1, the state x of the population can be described as a probability vector $x = (x_1, x_2, \dots, x_n)$ which expresses the different densities of all the different types of replicators (i.e., strategies) in the population, with x_i representing the density of replicator i . A is the payoff matrix that describes the different payoff values that each individual replicator receives when interacting with other replicators in the population. Hence $(Ax)_i$ is the payoff that replicator i receives in a population with state x , whereas $x \cdot Ax$ describes the average payoff in the population. The growth rate $\frac{dx_i}{dt} / x_i$ of the proportion of replicator i in the population equals the difference between the replicator's current payoff and the average payoff in the population. For more information, we refer to [3, 5, 15].

3.2 The Heuristic Payoff Table

The heuristic payoff table represents the payoff table of the poker game for the different meta strategies the different agents can employ. In essence it replaces the Normal Form Game (NFG) payoff table for the atomic actions. For a complex game such as poker it is impossible to use the atomic NFG, simply because the table has too many dimensions to be able to represent it. Therefore, we look at heuristic strategies as outlined in Section 2.2.

Let's assume we have A agents and S strategies. This would require S^A entries in our NFG table. We now make a few simplifications, i.e., we do not consider different types of agents, we assume all agents can choose from the same strategy set and all agents receive the same payoff for being in the same situation. This setting corresponds to the setting of a symmetric game. This means we consider a game where the payoffs for playing a particular strategy depend only on the strategies employed by the other agents, but not on who

is playing them. Under this assumption we can seriously reduce the number of entries in the heuristic payoff table. More precisely, we need to consider the different ways of dividing our A agents over all possible S strategies. This boils down to:

$$\binom{A + S - 1}{A}$$

Suppose we consider 3 heuristic strategies and 6 agents, this leads to a payoff table of 28 entries, which is a serious reduction from $3^6 = 729$ entries in the general case. As an example the next table illustrates what the heuristic payoff table looks like for three strategies S_1, S_2 and S_3 .

$$P = \left(\begin{array}{ccc|ccc} S_1 & S_2 & S_3 & U_1 & U_2 & U_3 \\ s_1 & s_2 & s_3 & u_1 & u_2 & u_3 \\ \dots & & & & & \dots \end{array} \right)$$

Consider for instance the first row of this table: in this row there are s_1 agents that play strategy S_1 , s_2 agents that play strategy S_2 and s_3 agents play strategy S_3 . Furthermore, u_i is the respective expected payoff for playing strategy S_i . We call a tuple $(s_1, s_2, s_3, u_1, u_2, u_3)$ a profile of the game. To determine the payoffs u_i in the table, we compute expected payoffs for each profile from real-world poker data we assembled. More precisely, we look in the data for the appearance of each profile and compute from these data points the expected payoff for the used strategies. However, because payoff in the game of poker is non-deterministic, we need a significant number of independent games to be able to compute representative values for our table entries. In Section 4 we provide more details on the data we used and on the process of computing the payoff table.

3.3 Approximating Nash Equilibria

In this section we describe how we can determine which of our restpoints of the RD are effectively Nash equilibria (so note that a restpoint of the RD is not necessarily Nash). The approach we describe is based on work of Walsh et al. and Vytelyngum et al. [13, 14]. A Nash equilibria occurs when no player can increase its payoff by changing strategy unilaterally. For the sake of clarity we follow the notation of [14].

The expected payoff of an agent playing a strategy $j \in S^1$, given a mixed-strategy p (the population state), is denoted as $u(e_j, p)$. This corresponds to $(Ax)_i$ in Equation 1. The value of $u(e_j, p)$ can be computed by considering the results from a large number of poker games with a player playing strategy j and the other agents selected from the population, with a mixed-strategy p . For each game and every strategy, the individual payoffs of agents using strategy j are averaged. The Nash equilibrium is then approximated as the argument to the minimisation problem given in Equations 2 and 3.

$$v(p) = \sum_{j=1}^S (\max[u(e_j, p) - u(p, p), 0])^2 \quad (2)$$

$$p_{nash} = \operatorname{argmin}_p [v(p)] \quad (3)$$

Here, $u(p, p)$ is the average payoff of the entire population and corresponds with term $x \cdot Ax$ of Equation 1. Specifically, p_{nash} is a Nash equilibrium if and only if it is a global minimum of $v(p)$, and p is a global minimum if $v(p) = 0$. We solve this non-linear minimisation problem using the Amoeba non-linear optimiser [14].

3.4 Simplex Analysis

The simplex analysis allows us to graphically and analytically study the dynamics of strategy changes. Before explaining this analysis, we first introduce a definition of a simplex. Given n elements which are randomly chosen with probabilities (x_1, x_2, \dots, x_n) , there holds $x_1, x_2, \dots, x_n \geq 0$ and $\sum_{i=1}^n x_i = 1$. We denote the set of all such probability distributions over n elements as Σ_n or simply Σ if there is no confusion possible. Σ_n is a $(n - 1)$ -dimensional structure and is called a *simplex*. One degree of freedom is lost due to the normality constraint. For example in Figure 1, Σ_2 and Σ_3 are shown. In the figures throughout the experiments we use Σ_3 , projected as an equilateral triangle as in Figure 1(b), but we drop the axes and

¹The use of S differs from that in Section 3.2. Here S represents the set of strategies, unlike the number of strategies in Section 3.2.

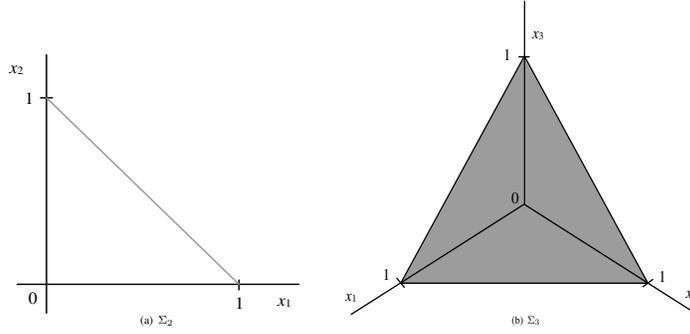


Figure 1: The unit simplices Σ_2 (a; left) and Σ_3 (b; right).

labels. Since we use four meta strategies and Σ_3 concerns only three, this implies that we need to show four simplexes Σ_3 , from each of which one strategy is missing.

Using the generated heuristic payoff table, we can now visualize the dynamics of the different agents in a simplex as follows. To calculate the RD at any point $s = (x_1, x_2, x_3)$ in our simplex, we consider N (i.e., many) runs with mixed-strategy s ; x_1 is the percentage of the population playing strategy S_1 , x_2 is the percentage playing strategy S_2 and x_3 is the percentage playing strategy S_3 . For each run, each poker agent selects their (pure) strategy based on this mixed-strategy. Given the number of players using the different strategies (S_1, S_2, S_3), we have a particular profile for each run. This profile can be looked up in our table, yielding a specific payoff for each player. The average of the payoffs of each of these N profiles gives the payoffs at $s = (x_1, x_2, x_3)$. Provided with these payoffs we can easily compute the RD by filling in the values of the different variables in Equation 1. This yields us a gradient at the point $s = (x_1, x_2, x_3)$.

Starting from a particular point within the simplex, we can now generate a smooth trajectory (consisting of a piecewise linear curve) by moving a small distance in the calculated direction, until the trajectory reaches an equilibrium. A trajectory does not necessarily settle at a fixed point. More precisely, an equilibrium to which trajectories converge and settle is known as an attractor, while a saddle point is an unstable equilibrium at which trajectories do not settle. Attractors and saddle points are very useful measures of how likely it is that a population converges to a specific equilibrium.

4 Experiments and results

We collected a total of 1599057 No-Limit Texas Hold'em games with 6 or more players starting. As a first step we needed to determine the strategy for a player at any given point. If a player played less than 50 games in total, we argue that we do not have sufficient data to establish a strategy, and therefore we ignore this player (and game). If the player played at least 50 games, we used an interval of 50 games to collect statistics for this specific player, and then determined the *VSF* and *AGR* values. We set the thresholds respectively to 0.35 and 2.0, i.e., if $VSF > 0.35$, then the player is considered loose (and tight otherwise), and if $AGR > 2$ then the player is considered aggressive (and passive otherwise). These are commonly used thresholds for a No-Limit Texas Hold'em game (see e.g., [2, 4, 9]). The resulting strategy was then associated with the specific player for all games in the interval of 50 games. Having estimated all players' strategies, it is now possible to determine the table configuration (i.e., the number of players playing any of the four meta strategies) for all games. Finally, we can compute the average payoffs for all strategies given a particular table configuration and produce a profile (see Section 3.2).

We plotted four simplexes that resulted from our RD analysis in Figure 2. Recall from Section 3.4 that these simplexes show the dynamic behavior of the participating players having a choice from three strategies. This means that the evolution of the strategies, employed in the population, is visualized for every possible initial condition of the game. The initial condition determines in which basin of attraction we end up, leading to some specific attractor or repeller. These restpoints (i.e. attractors or repellers) are potentially Nash equilibria.

What we can immediately see from the plots is that both passive strategies LP and TP (except in plot a) are repellers. In particular the LP strategy is a strong repeller. This suggests that no matter what the game situation is, when playing the LP strategy, it is always rational to switch strategy to for example TA or LA.

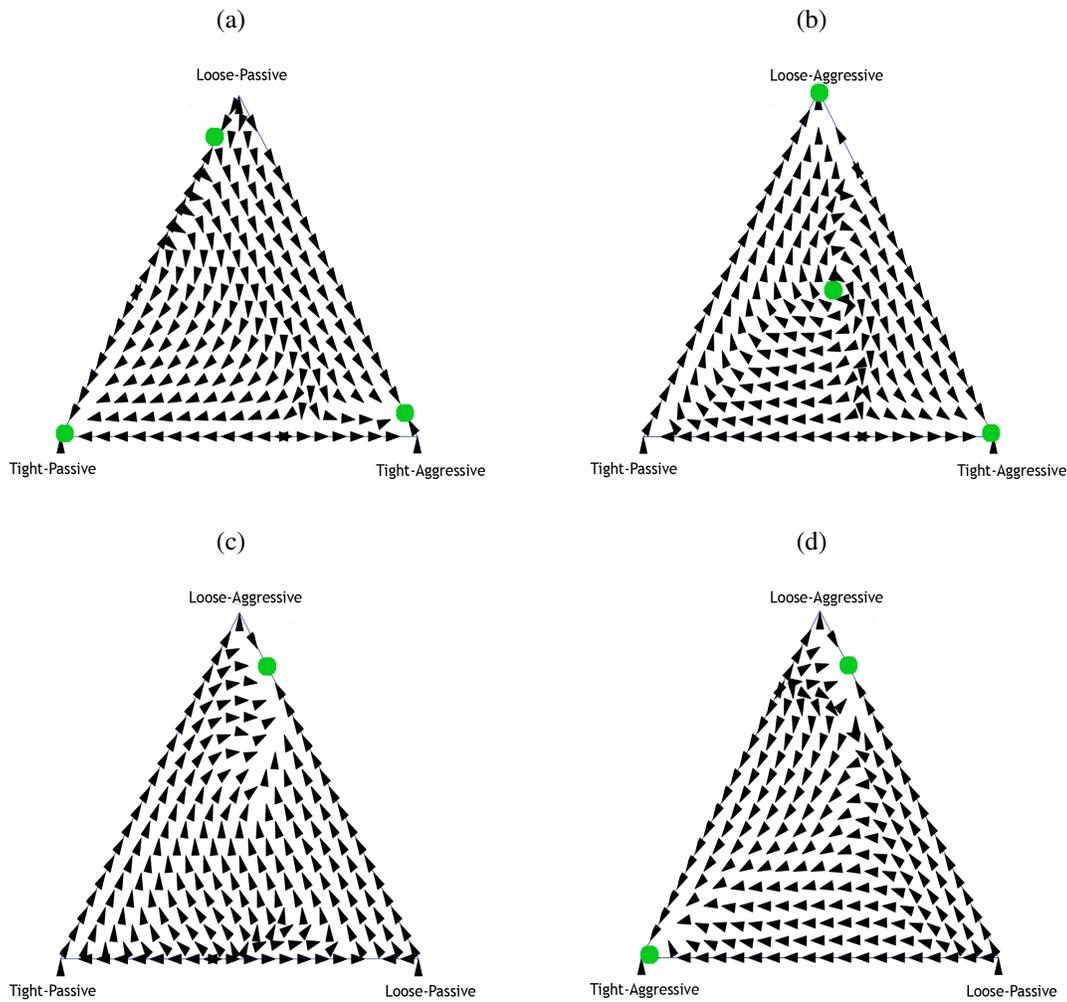


Figure 2: The direction field of the RD using the heuristic payoff table considering the four described meta-strategies. Dots represent the Nash equilibria.

This nicely confirms the claim made earlier (and in literature), namely that aggressive strategies dominate their passive counterparts.

The dots indicated on the plots represent the Nash equilibria of the respective games². Figure 2a contains three Nash equilibria of which two are mixed and one is pure. The mixed equilibrium at the axis TP-LP is evolutionarily unstable as a small deviation in a players' strategy might lead the dynamics away from this equilibrium to one of the others. The mixed equilibrium at the axis LP-TA is stable. As one can see this equilibrium lies close to the pure strategy TA. This means that TA is played with a higher probability than LP. Finally, there is also one stable pure equilibrium present, i.e., TP. Of the stable equilibria TP has the largest basin of attraction.

Figure 2b contains 3 Nash equilibria of which one is mixed and two are pure. As one can see from the picture, the mixed Nash equilibrium is evolutionarily unstable, i.e., any small perturbation of this equilibrium immediately leads the dynamics away from it to one of the other pure Nash equilibria. This means that if one of the players would decide to slightly change its strategy at the equilibrium point, the dynamics of the entire population would drastically change. The mixed Nash equilibrium almost corresponds to the situation in which the three strategies are played with equal probability, i.e., a uniform distribution. The pure Nash equilibria LA and TA are both evolutionarily stable. LA has a larger basin of attraction than TA (similar to plot a), which does not completely correspond with the expectations of domain experts (it is assumed by

²Due to space constraints we only discuss the Nash equilibria of Figures 2a-2b and Figures 3a-3b. For completeness the equilibria of Figures 2c and 2d are also indicated.

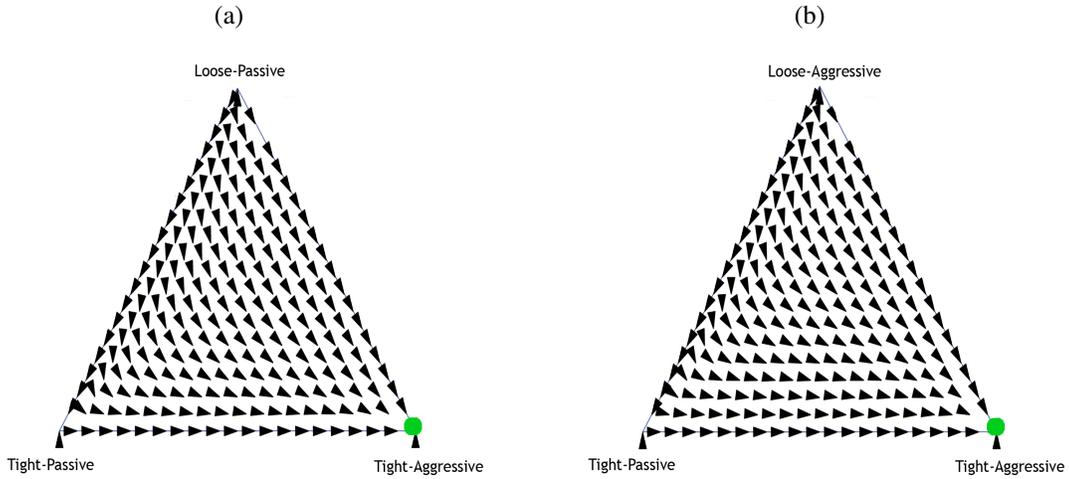


Figure 3: The direction field of the RD using the heuristic payoff table using data of games with active players at the flop.

domain experts that in general TA is the most profitable strategy).

One possible explanation is the following: we noticed that some strategies (depending on the used thresholds for *VSF* and *AGR*) are less played by humans compared to other strategies. Therefore, a table configuration with a large number of agents playing these scarcely played strategies, results in few instances and possibly a distorted average payoff due to the high variance of profits in the game of No-Limit Texas Hold'em. In particular, we observed that table configurations with many humans playing a tight strategy had only few instances (e.g., the payoffs used in plot a, with two tight strategies in the simplex, were calculated using 40% less instances compared to those in plot b). A severe constraint on the number of instances is currently our chosen representation for a profile. In the previous experiment, we used games with 6 or more starting players, and counted the number of occurrences of the four strategies. An alternative way of interpreting the data is only considering players active at the flop. Since most of the times only 4 or less players (and a maximum of 6 players in our data) are active at the flop, this results in fewer profiles. Basically, we generalize over the number of players starting at the beginning of the game and only focus on the interaction between strategies during the phases that most influence the average payoffs. The results from these experiments are illustrated in Figure 3.

In Figure 3a and 3b we have one pure Nash equilibrium being a dominant strategy, i.e., TA. These equilibria, and the evolution to them from any arbitrary initial condition, confirm the conclusions of domain experts.

5 Conclusion

In this paper we investigated the evolutionary dynamics of strategic behaviour of players in the game of No-Limit Texas Hold'em poker. We performed this study from an evolutionary game theoretic perspective using Replicator Dynamic models. We investigated the dynamic properties by studying how human players should switch between different strategies under different circumstances, and what the Nash equilibria look like. We observed poker games played at an online poker site and used this data for our analysis. Based on domain knowledge, we identified four distinct meta strategies in the game of poker. We then computed the heuristic payoff table to which we applied the Replicator Dynamic model. The resulting plots confirm that what is claimed by domain experts, namely that often aggressive strategies dominate their passive counterparts, and that the Loose-Passive strategy is an inferior one.

For future work, we will examine the interactions between the meta strategies among several other dimensions, namely, more detailed meta strategies (i.e., based on more features), a varying number of players, different parameter settings and different Replicator Dynamic models (e.g., including mutation). We are also interested in performing this study using simulated data (which we can generate much faster). Finally, since it is clear from our current experiments that the Loose-Passive strategy is an inferior one, we can focus

on the switching dynamics between the remaining strategies given the presence of a fixed number of players playing the Loose-Passive strategy. This way, we focus on the dynamics for the strategies that matter.

6 Acknowledgments

Marc Ponsen is sponsored by the Interactive Collaborative Information Systems (ICIS) project, supported by the Dutch Ministry of Economic Affairs, grant nr: BSIK03024. Jan Ramon and Kurt Driessens are post-doctoral fellow of the Research Foundation - Flanders (FWO). The authors wish to express their gratitude to P. Vytelingum for his insightful comments on the construction of the heuristics payoff table.

References

- [1] A. Davidson, D. Billings, J. Schaeffer, and D. Szafron. Improved opponent modeling in poker. In *Proceedings of The 2000 International Conference on Artificial Intelligence (ICAI'2000)*, pages 1467–1473, 2000.
- [2] D. Doyle Brunson. *Doyle Brunson's Super System: A Course in Power Poker*. Cardoza, 1979.
- [3] H. Gintis. *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction*. Princeton University Press, 2001.
- [4] D. Harrington. *Harrington on Hold'em Expert Strategy for No Limit Tournaments*. Two Plus Two Publisher, 2004.
- [5] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, 1998.
- [6] J. Maynard-Smith. *Evolution and the Theory of Games*. Cambridge University Press, 1982.
- [7] S. Phelps, S. Parsons, and P. McBurney. Automated trading agents versus virtual humans: an evolutionary game-theoretic comparison of two double-auction market designs. In *Proceedings of the 6th Workshop on Agent-Mediated Electronic Commerce*, New York, NY, 2004.
- [8] M. Ponsen, J. Ramon, T. Croonenborghs, K. Driessens, and K. Tuyls. Bayes-relational learning of opponent models from incomplete information in no-limit poker. In *Twenty-third Conference of the Association for the Advancement of Artificial Intelligence (AAAI-08)*, pages 1485–1487, Chicago, USA, 2008.
- [9] D. Slansky. *The Theory of Poker*. Two Plus Two Publisher, 1987.
- [10] F. Southey, M. Bowling, B. Larson, C. Piccione, N. Burch, D. Billings, and D. C. Rayner. Bayes' bluff: Opponent modelling in poker. In *Proceedings of the 21st Conference in Uncertainty in Artificial Intelligence (UAI '05)*, pages 550–558, 2005.
- [11] P. Taylor and L. Jonker. Evolutionary stable strategies and game dynamics. *Math. Biosci.*, 40:145–156, 1978.
- [12] K. Tuyls, P. 't Hoen, and B. Vanschoenwinkel. An evolutionary dynamical analysis of multi-agent learning in iterated games. *The Journal of Autonomous Agents and Multi-Agent Systems*, 12:115–153, 2006.
- [13] P. Vytelingum, D. Cliff, and N. R. Jennings. Analysing buyers and sellers strategic interactions in marketplaces: an evolutionary game theoretic approach. In *Proc. 9th Int. Workshop on Agent-Mediated Electronic Commerce*, Hawaii, USA, 2007.
- [14] W. E. Walsh, R. Das, G. Tesauro, and J. O. Kephart. Analyzing complex strategic interactions in multi-agent systems. In P. Gymtrasiwicz and S. Parsons, editors, *Proceedings of the 4th Workshop on Game Theoretic and Decision Theoretic Agents*, 2001.
- [15] J. W. Weibull. *Evolutionary Game Theory*. MIT Press, 1996.
- [16] E. Zeeman. Dynamics of the evolution of animal conflicts. *Journal of Theoretical Biology*, 89:249–270, 1981.