Research Article

Bluffing as a Rational Strategy in a Simple Poker-Like Game Model

Andrea Guazzini[1] and Daniele Vilone[2]

[1] *Department of Education and Psychology, University of Florence, Via San Salvi 12, Building 26, 50135 Florence, Italy*
[2] *Statistical Materials Modeling Laboratory (CNR-IENI), Via R. Cozzi 53, 20125 Milano, Italy*

Correspondence should be addressed to Andrea Guazzini; andrea.guazzini@gmail.com

We present a simple adaptive learning model of a poker-like game, by means of which we show how a bluffing strategy emerges very naturally and can also be rational and evolutionarily stable. Despite their very simple learning algorithms, agents learn to bluff, and the most bluffing player is usually the winner.

1. Introduction

Among the concepts formulated and elaborated by the cognitive psychology in the last century, the study of the process of development of problem-solving strategies has enabled a meaningful improvement in the comprehension of the mental activity and its relations with the cerebral circuits.

At this level of description, the concept of cognitive process is defined as the interconnected performances of some elementary cognitive activities that operate on and affect mental contents, representing the fundamental issue to bridge at a theoretical level the superior cognitive functions and the human behaviour [1, 2]. Such a concept is used in a wider sense to mean the act of knowing and may be interpreted in a social or cultural sense to describe the emergent development of knowledge, concepts, or strategies.

Cognitive psychologists argue that the mind can be understood in terms of information processing, especially when processes as abstraction, categorization, knowledge, expertise, or learning are involved [3–5].

The concept of cognitive process is defined both in terms of result of the parallel elaboration of several well-defined and functionally independent neural moduli and in terms of a sort of "software" able to optimize the integration among different cognitive functions by the adaptation to the different environmental/informational circumstances [6–8].

The target of the present paper is to formulate a cognitive model of a poker-like game in which the players are able to develop strategies by learning from experience. Such a task, better known as *problem solving*, allows to investigate effectively the relation and coupling between the dynamics of cognitive process and the *environment*.

Moreover, the study of poker is of great and general interest in complex systems, because it is strictly related with sociophysics, decision theory, and behaviour evolution. Actually, the study of human behaviour and, more in general, of social phenomena has been faced in the last years utilizing the tools of complex systems physics [9]. Poker-like games provide a very good instance of strategic dilemma where agents must optimize their income in conditions of imperfect information (see Section 2) or where it is not clear which is the true optimal strategy, a case which is very common in real human interactions [10, 11]. Understanding the mechanisms which underlie poker is then very useful for a comprehension of human psychology and to get hints of how we have to approach to more complicated models of human society.

2. Poker-Like Games

Poker is an interesting testsbed for artificial intelligence research [12–15]. It is a game of imperfect information, where multiple competing agents must deal with probabilistic knowledge, risk assessment, and possible deception, not unlike decisions made in the real world. The so-called "opponent modelling" is another difficult problem in decision-making applications, and it is essential to achieve

high performances in poker. Moreover, poker has a rich history of study in several academic fields. Economists and mathematicians have applied a variety of analytical techniques to poker-related problems [16–18]. For example, the earliest investigations in game theory, by luminaries such as John von Neumann and John Nash, used simplified poker to illustrate the fundamental principles.

There is an important difference between board games and popular card games like bridge and poker. In board games, players have complete knowledge of the entire game state since everything is visible to all participants. In contrast, bridge and poker involve imperfect information since the other players' cards are not known.

From a computational point of view, it is important to distinguish the lack of information from the possibility of chance moves. The former involves uncertainty about the current state of the world, in particular situations where different players have access to different information. The latter involves only uncertainty about the future, uncertainty which is resolved as soon as the future materializes. Perfect and imperfect information games may involve an element of chance; examples of games from all four categories are shown in Table 1.

The presence of chance elements does not need major changes to the computational techniques used to solve a game. In fact, the cost of solving a perfect information game with chance moves is not substantially greater than solving a game with no chance moves. By contrast, the introduction of imperfect information increases the complexity of the problem.

Due to the complexity (both conceptual and algorithmic) of dealing with imperfect information games, this problem has been largely ignored at the computational level until the introduction of randomized strategies concept.

Once randomized strategies are allowed, the existence of "optimal strategies" in imperfect information games can be proved. In particular, this means that there exists an optimal randomized strategy for poker in the same way as there exists an optimal deterministic strategy for chess. Indeed, Kuhn showed for a simplified poker game that the optimal strategy does use randomization [19].

The optimal strategy has several advantages: the player cannot do better than this strategy if playing against a good opponent; furthermore, the player does not do worse even if his strategy is revealed to his opponent; that is, the opponent gains no advantage from figuring out the first player's strategy.

Another interesting result of such researches is the existence of an optimal strategy for the gambler in poker game. As first observed in a simple poker-like game by Kuhn [19], behaviors such as bluffing, that seem to arise from the psychological makeup of human players, are actually game theoretically optimal.

One of the earliest and most thorough investigations of poker appears in the classical treatise on game theory "Games and Economic Behavior" by von Neumann and Morgenstern [20], where a large section was devoted to the formal analysis of "bluffing" in several simplified variants of a two-person poker game with either symmetric or asymmetric information.

Table 1

	Perfect information	Imperfect information
No chance	Chess	Inspection game
Chance	Monopoly	Poker

Indeed, the general considerations concerning poker and the mathematical discussions of the different versions of the game were carried out by von Neumann as early as 1926. Recognizing that "bluffing" in poker "is unquestionably practiced by all experienced players," von Neumann and Morgenstern identified two reasons for bluffing. "The first is the desire to give a "false" impression of strength in "real" weakness; the second is the desire to give a "false" impression of weakness in "real" strength" [20].

Solutions to these simplified poker-like games as well as a large class of both zero-sum and nonzero-sum games were unified by the concept of mixed strategy, a probability distribution over the player's set of actions. The importance of mixed strategies to the theory of games and its applications to the social and behavioral sciences stems from the fact that for many interactive decision processes there can be no Nash equilibria in pure strategies.

Using randomization and adaptive learning as key concepts to modelize into a computational scaffolding of the cognitive processes, we believe that this area of research is more likely to produce insights about superior cognitive strategies because of their intrinsically structures. Finally, comparing them with the real human strategy, it is possible both to investigate the role of environmental factors on cognitive strategies development and to validate some theoretical psychological assumptions.

2.1. Summary. The target of this paper is to show how bluffing strategies can arise naturally as a mathematical property of a very simple model, without any references to psychological assessments. For this reason, we present a simple model of poker-like game with only two players and only two possible strategies: folding and calling, which each agent assumes simultaneously. Such oversimplified game, as we will see, allows to catch the fundamental mechanisms underlying the phenomenon of bluffing. The fact that bluffing naturally emerges already in a very simple version of the game seems to suggest that such a strategy is perfectly rational and can be mathematically characterized.

3. The Model

The model we are going to define and analyse is probably one of the greatest simplifications possible of a game of chance with imperfect information.

Here we have two players: at the beginning of each hand of the game, they put one coin as the entry pot. Then, they pick a "card" from a pack: each card has an integer value between 0 and $N - 1$ (i.e., there are N cards overall). At this point, according to the value of their card, the players decide to call or instead to fold. If both players call, they put another coin in the pot, and who holds the highest card

wins: the winner gets the entire pot (four coins). If one of the players folds, the "caller" wins and gets the entry pot (two coins). Finally, if nobody calls, both players take back the coin they had put as entry pot. Mathematically, when the player i holds the card of value n, he decides to call according to the probability distribution $P_i(n)$, with of course $i = 1, 2$ and $n \in \{0, 1, \dots, N-1\}$. After every hand, both players update their strategy. More precisely, if one folds, nothing happens; if the agent i, thanks to the card n, calls and wins (because he has the highest card or because the opponent folds), the probability that he calls holding the card n will change in this way

$$P_i(n) \longrightarrow P_i(n) + \mu_i\left[1 - P_i(n)\right]. \tag{1}$$

Analogously, if the agent i loses (i.e., if he calls but the opponent has a higher card than him), the probability $P_i(n)$ will change instead in the following way:

$$P_i(n) \longrightarrow \mu_i P_i(n). \tag{2}$$

In (1) and (2), the coefficient μ_i is the learning factor (LF), which can also be seen as a sort of risk propensity of the player i. The LFs of the players are set at the beginning of the game and will never change. Moreover, it can assume a value between 0 (no risk propensity at all) and 1 (maximum risk propensity possible): actually, a player with $\mu_i = 0$ and the card n does not increase $P_i(n)$, even though he wins and sets $P_i(n) = 0$ as soon as he loses; instead, with $\mu_i = 1$, he sets $P_i(n) = 1$ when he wins but does not decrease $P_i(n)$ if he loses. Finally, it is easy to notice that (1) and (2) ensure that $P_i(n)$ will always stay in the interval $[0, 1]$.

3.1. Numerical Results. In this section, we will present the most remarkable results of the simulations of the simple model defined previously.

First of all, for simplicity, we set the LF μ_1 of the "player 1" equal to 0.5, and then we checked the dynamics by varying μ_2: actually it is the difference between the LFs which essentially determines the main features of the dynamics, as we saw in several simulations. In particular, we can distinguish three cases: $\mu_2 < \mu_1$, $\mu_2 = \mu_1$, and $\mu_2 > \mu_1$.

3.1.1. Case $\mu_2 < \mu_1$. In Figures 1 and 2, the behaviour of the money of both players is shown as a function of time, where the time unit is a single hand of the game: we set $N = 25$, $\mu_2 = 0.3$, and $\mu_2 = 0.48$, respectively, and the results are averaged over 10^4 and 10^5 iterations, respectively; in both cases, we let the agents play 10^4 hands. For simplicity, we considered players with an infinite amount of money available, and we gauged to zero the initial amount. Additionally, the initial calling distributions $P_i(n)$ are picked randomly for each n.

As it can be seen, the first player, with higher LF, wins over his opponent, with smaller LF, and the money gained by player 1 increases with time. Moreover, the smaller μ_2, the faster and bigger the winnings of player 1. Previous figures show that on average, the player with bigger risk propensity finally overwhelms the other one, and this means that in a single match the most risk-inclined player has a bigger probability to win, and such probability increases as the difference $\mu_1 - \mu_2$ increases in its turn. The fact that risking is
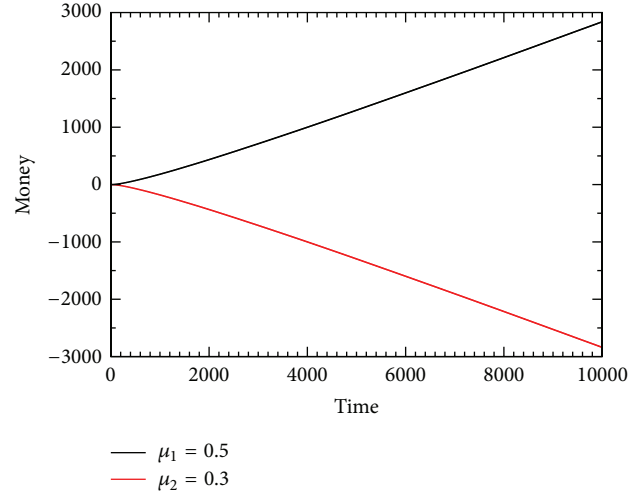

FIGURE 1: Behaviour of the money won (or lost) by the two players versus time, that is, versus the hands played, averaged over 10^4 iterated matches. The LFs of the players are here $\mu_1 = 0.5$ and $\mu_2 = 0.3$.
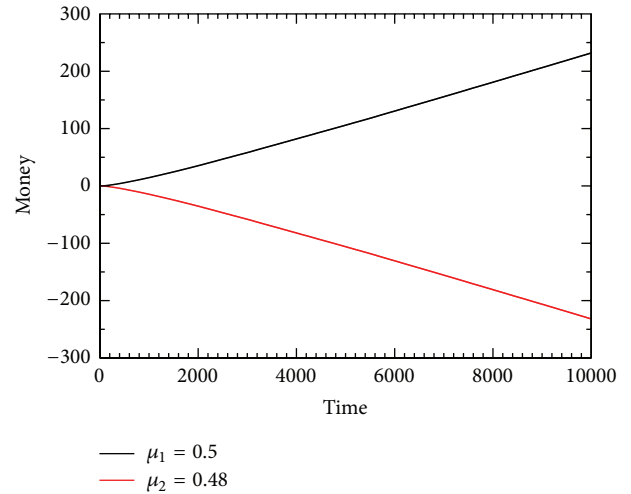

FIGURE 2: Behaviour of the money won (or lost) by the two players versus time, that is, versus the hands played, averaged over 10^5 iterated matches. The LFs of the players are here $\mu_1 = 0.5$ and $\mu_2 = 0.48$.

convenient gets confirmed in the next figures, where the final calling distributions for both players are depicted.

As we can see, in both cases, the winner is characterized by higher calling probabilities than his opponent's ones for every value of n. Moreover, we have $P_i(n = 0) > 0$, which is the most explicit evidence of the emergence of bluffing.

3.1.2. Case $\mu_2 = \mu_1$. In this case, both players have the same winning probability, as it is well shown in **Figure 5**: indeed, having the same LF, they have also exactly the same behaviours so that in a single match nobody is able to overwhelm definitively the opponent.
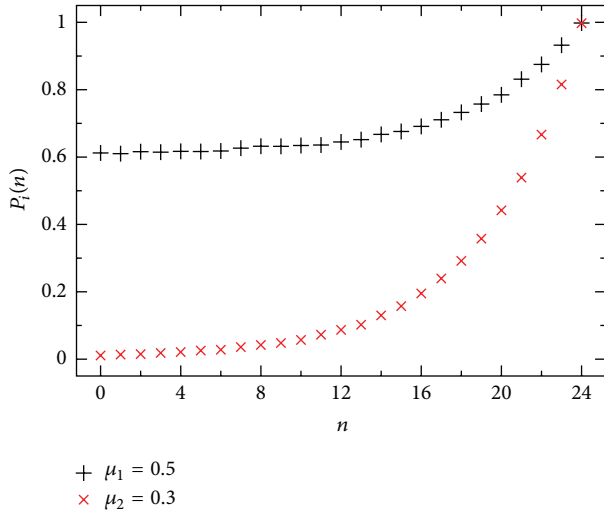
+ $\mu_1 = 0.5$
× $\mu_2 = 0.3$

FIGURE 3: Calling probabilities as functions of the card value n for both players. Black symbols: player 1 ($\mu_1 = 0.5$); red symbols: player 2 ($\mu_2 = 0.3$). Data took after 10^4 hands of the game and averaged over 10^4 iterations. Random initial distribution for every $P_i(n)$.

— $\mu_1 = 0.5$
— $\mu_2 = 0.5$

FIGURE 5: Behaviour of the money won (or lost) by the two players versus time, that is, versus the hands played, throughout a single match. The LFs of the players are here $\mu_1 = \mu_2 = 0.5$.
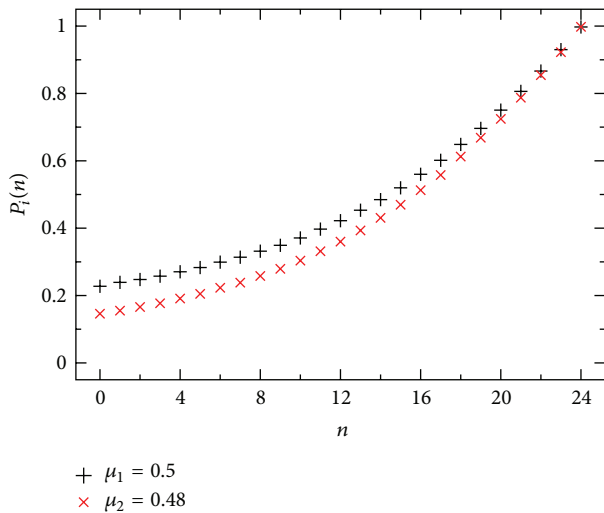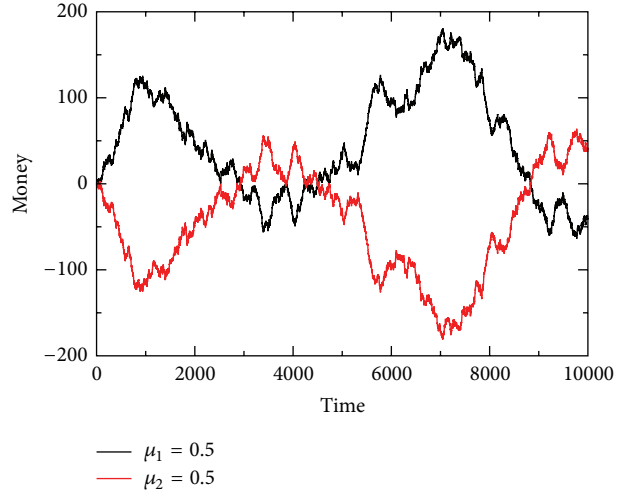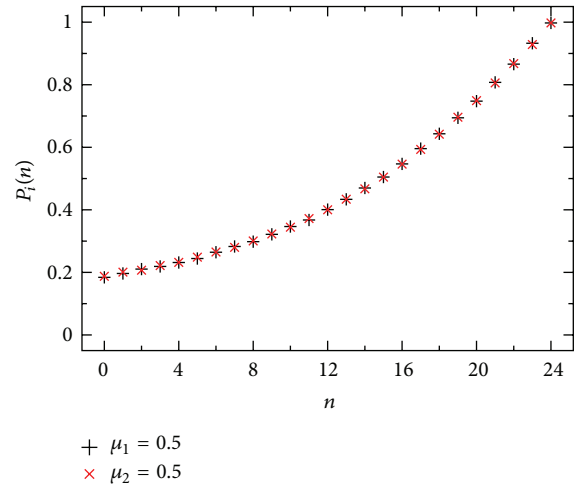

+ $\mu_1 = 0.5$
× $\mu_2 = 0.48$

FIGURE 4: Calling probabilities as functions of the card value n for both players. Black symbols: player 1 ($\mu_1 = 0.5$); red symbols: player 2 ($\mu_2 = 0.48$). Data took after 10^4 hands of the game and averaged over 10^5 iterations. Random initial distribution for every $P_i(n)$.

+ $\mu_1 = 0.5$
× $\mu_2 = 0.5$

FIGURE 6: Calling probabilities as functions of the card value n for both players. Black symbols: player 1 ($\mu_1 = 0.5$); red symbols: player 2 ($\mu_2 = 0.5 = \mu_1$). Data took after 10^4 hands of the game and averaged over $2.5 \cdot 10^4$ iterations. Random initial distribution for every $P_i(n)$.

Naturally, it is also easy to forecast the behaviour of the final calling probabilities of the agents: they will be equal, with $P_i(n = 0) > 0$ for both $i = 1$ and $i = 2$.

3.1.3. Case $\mu_2 > \mu_1$. In this case, the results are qualitatively equal to the ones of case $\mu_2 > \mu_1$; only now it is player 2 which defeats player 1, as shown in Figures 7 and 8.

3.2. Discussion. The first conclusion we can obtain just from the numerical results is that in this game bluffing emerges necessarily as rational strategy. Moreover, the player who bluffs more finally wins. This can be easily understood watching Figures 3, 4, 6, and 8. Indeed, while in general the calling

probabilities of both players tend to have the same value for $n \to N - 1$, for small n the player with higher LF, that is, with higher risk propensity, bluffs much more than his opponent: this means that even holding a poor card, the "risk-lover" will call and unless his opponent has a very good point will get the entry pot.

It is possible to formalize such considerations by writing the equations of the dynamics for the model at stake. Neglecting fluctuations, the "mean-field" equation ruling the money M_1 won (or lost) by the first player is

$$M_1(t+1) = M_1(t) + \sum_{n_1=0}^{N-1} \frac{1}{N} \left[\left(1 - \widehat{\Pi}_2^1\right) P_1(n_1; t) \right.$$

$$-\widehat{\Pi}_2^1\left(1 - P_1\left(n_1;t\right)\right)$$

$$+2P_1\left(n_1;t\right)\widehat{\Pi}_2^1\left(\mathscr{P}_2^1 - \mathscr{P}_1^2\right)\Big], \tag{3}$$

where $\mathscr{P}_2^1 = \Pr(n_1 > n_2)$ is the probability that the card n_1 held by player 1 is higher than the card n_2 held by player 2; analogously it is $\mathscr{P}_1^2 = \Pr(n_2 > n_1)$. On the other hand, $\widehat{\Pi}_2^1$ is an operator defined as follows

$$\widehat{\Pi}_2^1 \cdot X = \sum_{n_2=0}^{N-1}\left[\frac{P_2\left(n_2;t\right)}{N-1}\left(1 - \delta_{n_2,n_1}\right)X\right] \tag{4}$$

and represents the probability that player 2 calls from the point of view of player 1. Equation (3) can be rewritten as a differential equation in time, which can assume the form

$$\dot{M}_1\left(t\right) = \left[1 - 2\gamma_2\left(t\right)\right]\omega_1\left(t\right) - \left[1 - 2\gamma_1\left(t\right)\right]\omega_2\left(t\right) \tag{5}$$

with

$$\omega_i\left(t\right) = \frac{1}{N}\sum_{n=0}^{N-1}P_i\left(n;t\right) \quad i = 1, 2,$$

$$\tag{6}$$

$$\gamma_i\left(t\right) = \frac{1}{(N-1)^2}\sum_{n=0}^{N-1}nP_i\left(n;t\right) \quad i = 1, 2.$$

Since this is a zero-sum game, second player's money will be obtained by the relation $M_2(t) = -M_1(t)$.

Finally, the relation giving the temporal behaviour of the calling distributions $P_1(n;t)$ of player 1 (being the one of the opponent of analogous form) is

$$\dot{P}_1\left(n_1;t\right) = \frac{1}{N}P_1\left(n_1;t\right)\left[\left(1 - \widehat{\Pi}_2^1 + \widehat{\Pi}_2^1\mathscr{P}_2^1\right)\right.$$

$$\times\left[P_1\left(n_1;t\right) + \mu_1\left(1 - P_1\left(n_1;t\right)\right)\right]$$

$$\left.+\mu_1 P_1\left(m_1;t\right)\widehat{\Pi}_2^1\mathscr{P}_1^2\right]$$

$$+\frac{1 - P_1\left(n_1;t\right)}{N}P_1\left(n_1;t\right) - \frac{P_1\left(n_1;t\right)}{N}. \tag{7}$$

Now, (5), (6), and (7) are rather complicated, but some features of them can be determined without an explicit solution. Actually, it is straightforward to understand that we have

$$P_1\left(n;t\right) = P_2\left(n;t\right) \Longrightarrow \omega_1\left(t\right) = \omega_2\left(t\right),$$

$$\gamma_1\left(t\right) = \gamma_2\left(t\right) \Longrightarrow \dot{M}_1 = \dot{M}_2 = 0. \tag{8}$$

Now, since in our simulations we always started from the same initial $P_i(n)$ for all i, n, and from $M_1(0) = M_2(0) = 0$, this implies that for $\mu_1 = \mu_2$ both players must have on average the same calling distributions and then they should not gain nor lose money, apart from fluctuations: this is exactly what we found in Figures 5 and 6. It can also be shown that if $\mu_1 > \mu_2$, then we will have soon $P_1(n) \geq P_2(n)$ for all n (with the equality holding only for $n = N - 1$), so that (6) allow us to get $\dot{M}_1 > 0$, that is, the victory of player 1, as shown in Figures 1 to 4. Obviously, the opposite situation takes place for $\mu_1 < \mu_2$ (as shown in Figures 7 and 8).
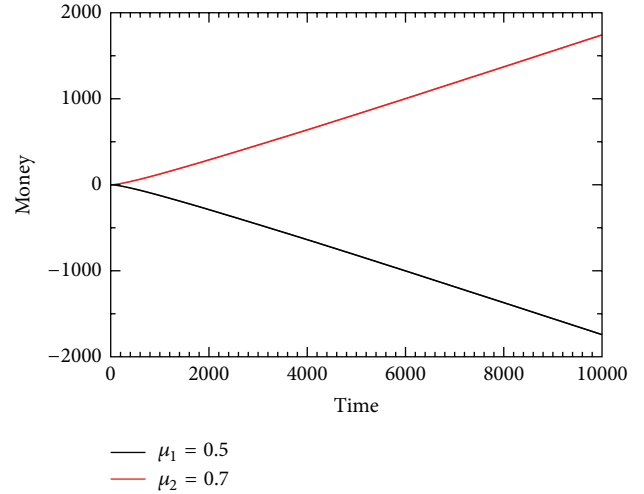

FIGURE 7: Behaviour of the money won (or lost) by the two players versus time, that is, versus the hands played, averaged over 10^4 iterated matches. The LFs of the players are here $\mu_1 = 0.5$ and $\mu_2 = 0.7$.
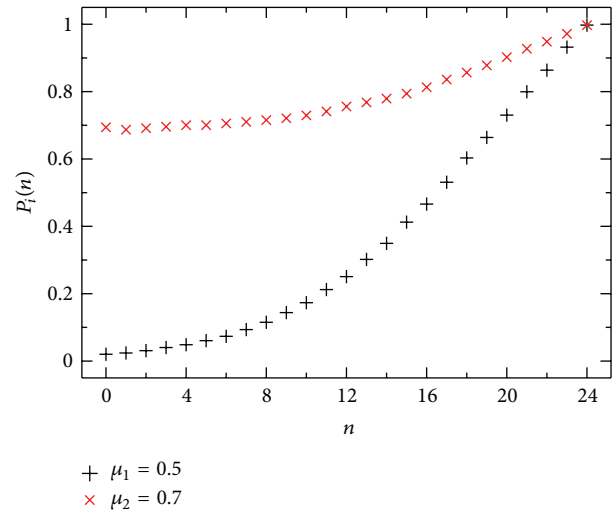

FIGURE 8: Calling probabilities as functions of the card value n for both players. Black symbols: player 1 ($\mu_1 = 0.5$); red symbols: player 2 ($\mu_2 = 0.7$). Data took after 10^4 hands of the game and averaged over 10^4 iterations. Random initial distribution for every $P_i(n)$.

4. Conclusions and Perspectives

In this work, we have developed a model to describe very roughly a cognitive processes dynamics. The model is among the most simple ones, but allows to capture the main behaviours of real dynamics, demonstrating how fundamental is bluffing as a rational strategy.

First of all this model confirms the role of LF (μ) as prominent. In a straightforward way, a direct connection between LF and bluffing tendency is here detectable. In fact even though both agents tend to develop bluff, the one with the greatest LF bluffs more than the other, ending up as

the winner. Finally this last evidence suggests that bluffing must be an evolutionary rational (stable) strategy.

Other, more realistic models have been developed [21], and for sure they can give us deeper information about the inner mechanisms driving to bluffing behaviours in poker-like games, but the emergence of bluff as a rational strategy already in a so simple toy model is with no doubt a very remarkable result. Indeed, more in-depth analyses of the real game can be found in the literature [13] which systematically survey every facet of poker (blind, flop, raise, etc.), but this is the first time that it is tried to catch the main features of it by means of a very "reductionist" approach, by means of the simplest version of the game. Shortly, if it has been thought until now that to catch an apparently complex behaviour, bluffing with effective outcomes, it is necessary to take into account (almost) all the rules of poker, we have shown here that such attitude is much more fundamental, and emerges naturally when few simple ingredients are present.

Moreover, for practical purposes, a simpler model which can be more easily understood turns out to be very useful because it will result easier also in the realization and the analysis of experimental tests with real human agents, and hence it will be possible to plan more accurately other kinds of experimental tools. In particular, it will be straightforward to utilize a tool as in reference [22] for a test with real players and to exploit data from poker-online web sites for a statistical analysis of the outcomes of real poker games.

References

[1] L. A. Real, "Animal choice behavior and the evolution of cognitive architecture ," *Science*, vol. 253, no. 5023, pp. 980–986, 1991.

[2] L. A. Real, "Information processing and the evolutionary ecology of cognitive architecture," *American Naturalist*, vol. 140, pp. S108–S145, 1992.

[3] E. Brandstätter, G. Gigerenzer, and R. Hertwig, "The priority heuristic: making choices without trade-offs," *Psychological Review*, vol. 113, no. 2, pp. 409–432, 2006.

[4] J. R. Busemeyer and J. T. Townsend, "Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment," *Psychological Review*, vol. 100, no. 3, pp. 432–459, 1993.

[5] I. Erev and G. Barron, "On adaptation, maximization, and reinforcement learning among cognitive strategies," *Psychological Review*, vol. 112, no. 4, pp. 912–931, 2005.

[6] G. T. Fong, D. H. Krantz, and R. E. Nisbett, "The effects of statistical training on thinking about everyday problems," *Cognitive Psychology*, vol. 18, no. 3, pp. 253–292, 1986.

[7] C. R. Fox and L. Hadar, "Decisions from experience = sampling error + prospect theory: reconsidering Hertwig, Barron, Weber & Erev," *Judgment and Decision Making*, vol. 1, no. 2, pp. 159–161, 2006.

[8] R. Hau, T. J. Pleskac, and R. Hertwig, "Decisions from experience and statistical probabilities: why they trigger different choices than a priori probabilities," *Journal of Behavioral Decision Making*, vol. 23, no. 1, pp. 48–68, 2010.

[9] C. Castellano, S. Fortunato, and V. Loreto, "Statistical physics of social dynamics," *Reviews of Modern Physics*, vol. 81, no. 2, pp. 591–646, 2009.

[10] C. Camerer, *Behavioral Game Theory: Experiments on Strategic Interaction*, Princeton University Press, Princeton, NJ, USA, 2003.

[11] H. Gintis, *The Bounds of Reason*, Princeton University Press, Princeton, NJ, USA, 2009.

[12] I. Bárány and Z. Füredi, "Mental poker with three or more players," *Information and Control*, vol. 59, no. 1–3, pp. 84–93, 1983.

[13] D. Billings, A. Davidson, J. Schaeffer, and D. Szafron, "The challenge of poker," *Artificial Intelligence*, vol. 134, no. 1-2, pp. 201–240, 2002.

[14] H. Johansen-Berg and V. Walsh, "Cognitive neuroscience: who to play at poker," *Current Biology*, vol. 11, no. 7, pp. R261–R263, 2001.

[15] D. Koller and A. Pfeffer, "Representations and solutions for game-theoretic problems," *Artificial Intelligence*, vol. 94, no. 1-2, pp. 167–215, 1997.

[16] A. Rapoport, I. Erev, E. V. Abraham, and D. E. Olson, "Randomization and adaptive learning in a simplified poker game," *Organizational Behavior and Human Decision Processes*, vol. 69, no. 1, pp. 31–49, 1997.

[17] Z. Shen, "A study of a generalization of a card problem," *Applied Mathematics and Computation*, vol. 166, no. 2, pp. 385–410, 2005.

[18] J. Miekisz, "Evolutionary game theory and population dynamics," in *Multiscale Problems in the Life Sciences*, vol. 1940 of *Lecture Notes in Mathematics*, pp. 269–316, 2008.

[19] H. W. Kuhn, "A simplified two-person poker," in *Annals of Mathematics Studies*, vol. 24, pp. 97–103, 1950.

[20] J. Von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ, USA, 1947.

[21] A. Guazzini and D. Vilone, *The emergence of bluff in poker-like games [Ph.D. thesis]*, University of Florence, 2009.

[22] A. Guazzini, D. Vilone, F. Bagnoli, T. Carletti, and R. Lauro Grotto, "Cog-nitive network structure: an experimental study," *Advances in Complex Systems*, vol. 15, no. 6, Article ID 1250084, 2012.