Acta Univ. Sapientiae, Informatica, **6,** 1 (2014) 71–88

A computational model of outguessing in two-player non-cooperative games

Tamás László BALOGH
University of Debrecen
email: tamas.balogh@econ.unideb.hu

János KORMOS
University of Debrecen
email: kormos.janos@inf.unideb.hu

Abstract. Several behavioral game theory models aim at explaining why “smarter“ people win more frequently in simultaneous zero-sum games, a phanomenon, which is not explained by the Nash equilibrium concept. We use a computational model and a numerical simulation based on Markov chains to describe player behavior and predict payoffs.

1 Introduction

Since the birth of experimental economics, thousands of experiments have been conducted to observe the behavior of decision makers in different situations (see e.g. [4]).

However, the most famous equilibrium concept—the Nash equilibrium [13]—has proved to be unable to explain the outcome of several game theoretical experiments, predicting that human thinking is more complicated than pure rationality.

Game theory has also proved to be a useful modelling tool for network situations, e.g. telecommunication problems. For a detailed survey in this field, we refer the reader to [15]. An application can be found in [1].

Computing Classification System 1998: G.3, I.6, J.4
Mathematics Subject Classification 2010: 60J20, 62P20, 68U20, 91A10
Key words and phrases: computational game theory, non-cooperative games, outguessing, Markov chains, reasoning, communication networks
DOI:10.2478/ausi-2014-0019

The phanomenon that human behavior is not purely rational in certain interactive network situations led researchers to construct behavioral game theory models. Recently, several models have been built to explain experimental results (e.g. [2, 7]).

A popular model class aiming at explaining how players outguess each other is the group of iterative reasoning models. Iterative reasoning has been applied in many settings to the Rock-Paper-Scissors game or the Beauty contest-type games ([3, 5, 8, 11, 12, 16]).

The concept of iterative reasoning and the corresponding main results are presented in [4, pages 205–236]. A simplified concept for non-cooperative, two-person, simultaneous games can be defined as follows. If Player A plays a certain action, while Player B plays the best response to this action, then we say that Player B outguessed Player A and played according to 1-reasoning. If now Player A outguesses Player B, then Player A plays according to 2-reasoning. Following this rule, the level of reasoning can be any k positive integer, where the concept is defined as k-reasoning.

In this paper we investigate simultaneous, two-person, zero-sum, repeated games that do not have a pure strategy Nash equilibrium and the players' decisions depend only on their actions in the previous round of the game. Here, the stochastic processes of the players' decisions and their expected payoffs can be described by Markov chains.

Our main goal is to point out why "smarter" people win more frequently in some well-known zero-sum games. There are several ways to define smartness. Our definition of smartness is connected to the concept of iterative reasoning and is introduced later on in Section 3.

We focus on modelling players' optimal strategy choices and expected payoffs. These are both stochastic processes given a certain bimatrix game and the level of iterative reasoning according to which players make their decisions.

We constructed a Matlab script that carries out the requested numerical analysis for any simultaneous, two-person bimatrix game. In our paper we present the relating analytical results, describe our concept and recall some numerical results and visualizations. Our Matlab script is also attached for testing and experimental purposes.

The rest of the paper is organized as follows. Section 2 recalls some important results in the field of Markov chains that are related to our topic. Section 3 describes our concept and provides numerical evidence. Section 4 describes the Matlab script. Finally, Section 5 concludes.

2 Markov chains—definitions and some important results

It is necessary to recall some basic results from the field of Markov chains that we use in the upcoming sections. For a more detailed analysis, we refer the reader to [6, 9, 10, 14]. This section is a brief summary of Chapter 4 in [6, pages 119–155], that is related to our concept. The proofs are always omitted.

Definition 1 *Let S be a countable (finite or countably infinite) set. An S-valued random variable X is a function from a sample space ω into S for which $\{X = x\}$ is an event for every $x \in S$.*

Here S need not be a subset of $\mathbb{R}$, so this extends the notion of a discrete random variable (or vector). The concepts of distribution, jointly distributed random variables, and so on, extend in the obvious way. The expectation of X, however, is not meaningful unless $S \subset R$. On the other hand, the conditioning random variables in a conditional expectation may be S-valued, and all of the results about conditional expectation generalize without difficulty.

Definition 2 *A matrix $P = (P(i,j))_{i,j\in S}$ with rows and columns indexed by S is called a one-step transition matrix if $P(i,j) \geq 0$ for all $i,j \in S$ and $\sum_{j\in S} P(i,j) = 1$ for all $i \in S$.*

In particular, the row sums of a one-step transition matrix are equal to 1. We call $P(i,j)$, the entry in row i and column j of the matrix P, a one-step transition probability.

Definition 3 *We say that $\{X_n\} n \geq 0$ is a Markov chain in a countable state space S with one-step transition matrix P if $X_0, X_1, \ldots$ is a sequence of jointly distributed S-valued random variables with the property that*

$$P(X_{n+1} = j | X_0, \ldots, X_n) = P(X_{n+1} = j | X_n) = P(X_n, j) \tag{1}$$

for all $n \geq 0$ and $j \in S$.

We assume that the sequence $X_0, X_1, \ldots$ is indexed by time, and if we regard time n as the present, the first equation in (1), known as the Markov property, says that the conditional distribution of the state of the process one time step into the future, given its present state as well as its past history, depends only on its present state. The second equation in (1) tells us that $P(X_{n+1} = j | X_n = i) = P(i,j)$ does not depend on n. This property is called time homogeneity.

The distribution of X_0 is called the initial distribution and is given by $(i) := P(X_0 = i), i \in S$.

A Markov chain can be described by specifying its state space, its initial distribution, and its one-step transition matrix.

Given a Markov chain $\{X_n\} n \geq 0$ in the state space S with one-step transition matrix P, it can be shown that, for every $m \geq 1$, $i_0, i_1, \ldots, i_m \in S$, and $n \geq 0$, $P(X_{n+1} = i_1, \ldots, X_{n+m} = i_m | X_n = i_0) = P(i_0, i_1)P(i_1, i_2) \ldots P(i_{m-1}, i_m)$.

Definition 4 *We define the m-step transition matrix* P^m *of the Markov chain by*

$$P^m(i,j) = \sum_{i_1,\ldots,i_{m-1}} \cdots \sum_{\in S} P(i, i_1)P(i_1, i_2)P(i_{m-1}, j) \tag{2}$$

Notice that the superscript m can be interpreted as an exponent, this is, the m-step transition matrix is the mth power of the one-step transition matrix. This is valid both when S is finite and when S is countably infinite. It is easy to check that this allows us to generalize (2), obtaining $P(X_{n+m} = j | X_0, \ldots, X_n) = P(X_{n+m} = j | X_n) = P_m(X_n, j)$ for all $n \geq 0$, $m \geq 1$, and $j \in S$.

Given $i \in S$, let us introduce the notation $P_i(\cdot) = P(\cdot | X_0 = i)$, with the understanding that the initial distribution is such that $P(X_0 = i) > 0$. It can be shown that

$$P_i(X_1 = i_1, \ldots, X_m = i_m) = P(i, i_1)P(i_1, i_2) \cdots P(i_{m-1}, i_m) \tag{3}$$

for all $i_1, \ldots, i_m \in S$.

Given $j \in S$, let us introduce the notation T_j for the first hitting time of state j (or first return time if starting in state j) and N_j for the number of visits to state j (excluding visits at time 0). More precisely, $T_j = \min\{n \geq 1 : X_n = j\}$ and $N_j = \sum_{n=1}^{\infty} 1_{\{X_n = j\}}$, where $\min \emptyset = \infty$. If also $i \in S$, we define $f_{ij} = P_i(T_j < \infty) = P_i(N_j \geq 1)$. This is the probability that the Markov chain, starting in state i, ever visits state j (or ever returns to state i if $j = i$). We can now define transient and recurrent states.

Definition 5 *We define state* j *to be transient if* $f_{jj} < 1$ *and to be recurrent if* $f_{jj} = 1$.

Some important features are pointed out in the next propositions.

Theorem 6 *Letting* $m \to \infty$ *it can be shown that*

$$P_i(N_j = \infty) = \begin{cases} 0 & \text{if } j \text{ is transient} \\ f_{ij} & \text{if } j \text{ is recurrent.} \end{cases} \tag{4}$$

Theorem 7 *For a Markov chain in* S *with one-step transition matrix* P*, state* $j \in S$ *is*

$$\text{transient if } \sum_{n=1}^{\infty} P_{j,j}^{n} < \infty, \tag{5}$$

$$\text{recurrent if } \sum_{n=1}^{\infty} P_{j,j}^{n} = \infty. \tag{6}$$

What is more, given that $i, j \in S$ is distinct, if state i is recurrent and $f_{ij} > 0$, then state j is also recurrent and $f_{ji} = 1$.

Let us define irreducible Markov chains.

Definition 8 *A Markov chain in* S *with one-step transition matrix* P *to be irreducible if* $f_{ij} > 0$ *for all* $i, j \in S$.

By Proposition 2, if a Markov chain in S with one-step transition matrix P is irreducible, then either all states in S are transient or all are recurrent. This allows us to refer to an irreducible Markov chain as either transient or recurrent.

Now we turn to the analysis of the asymptotic behavior of Markov chains.

Let π be a probability distribution on S satisfying

$$\pi_j = \sum_{i \in S} \pi_i P(i, j), \quad j \in S. \tag{7}$$

Regarding π as a row vector, this condition is equivalent to $\pi = \pi P$. Iterating, we have

$$\pi = \pi P = \pi P^2 = \ldots = \pi P^n, \; n \geq 1. \tag{8}$$

In particular, if $\{X_n\} n \geq 0$ is a Markov chain in S with one-step transition matrix P and if X_0 has distribution π_0, then X_n has distribution π_n for each $n \geq 1$. For this reason, a distribution π satisfying (7) is called a stationary distribution for the Markov chain.

We need one more definition to state an important result.

Definition 9 *The period* $d(i)$ *of state* $i \in S$ *is defined to be* $d(i) = g.c.d. D(i)$, $D(i) = \{n \in N : P^n(i, i) > 0\}$*, where g.c.d. stands for greatest common divisor.*

We first notice that every state of an irreducible Markov chain has the same period.

Note that if $i, j \in S$ are such that $f_{ij} > 0$ and $f_{ji} > 0$, then $d(i) = d(j)$. This allows us to speak of the period of an irreducible Markov chain.

Definition 10 *If the period is 1, we call the chain aperiodic.*

We can now describe the asymptotic behavior of the n-step transition probabilities of an irreducible aperiodic Markov chain.

Theorem 11 *If an irreducible aperiodic Markov chain in* S *with one-step transition matrix* P *has a stationary distribution* π*, then it is recurrent and*

$$\lim_{n\to\infty} P^n(i,j) = \pi(j)\ i,j \in S. \tag{9}$$

Furthermore, $\pi(i) > 0$ *for all* $i \in S$*.*

It follows from the previous statement that if an irreducible aperiodic Markov chain in S with one-step transition matrix P has no stationary distribution, then

$$\lim_{n\to\infty} P^n(i,j) = 0\ i,j \in S. \tag{10}$$

Thus, an irreducible aperiodic Markov chain in a finite state space S has a stationary distribution.

Theorem 12 *Let* $\{X_n\}_{n\geq 0}$ *be an irreducible aperiodic recurrent Markov chain in* S *with one-step transition matrix* P*. Then one of the following conclusions holds:*

(a) $E_i[T_i] < \infty$ *for all* $i \in S$*, and* P *has a unique stationary distribution* π *given by*

$$\pi(i) = \frac{1}{E_i[T_i]},\ i \in S. \tag{11}$$

(b) $E_i[T_i] = \infty$ *for all* $i \in S$*, and* P *has no stationary distribution.*

If (a) holds, then the chain is said to be positive recurrent, and Equation (9) holds. If (b) holds, then the chain is said to be null recurrent, and Equation (10) holds.

In the following section we define our concept and point out its relationship with Markov chains.

3 The outguessing equilibrium

This section introduces our notion of outguessing equilibrium. We first define the family of games we analyze.

Definition 13 *In a two person simultaneous normal form game we denote the players by* $i = 1, 2$. *We denote by* S_i *the pure strategy set of player* i, *where* $s_i \in S_i$ *and* $S = \times_{i=1}^{2} S_i$ *The utility (or payoff) of any player* i *is given by* $u_i(s_i, s_{-i}) \in \mathbb{R}$, *where* s_{-i} *denotes the strategy chosen by the other player.*

Our main assumptions are as follows.

Assumption 1 *We restrict attention to generic games, i.e. where the best response correspondance is a function. That means that there exists only one best response for any action of any of the two players.*

Assumption 2 *The game does not have a pure strategy Nash equilibrium.*

Notice that if the game had a pure strategy Nash equilibrium, mixed strategies and probability distributions would not have to be dealt with.

Assumption 3 *The game is repeated, the rounds are denoted by 1,2,...,*n*,...*

Assumption 4 *The players are assumed to keep in mind the strategy profile of the previous round (i.e. their own previous choice and their opponent's previous choice) and nothing else.*

Assumption 5 *Players are assumed to play according to 0-reasoning, 1-reasoning, 2-reasoning, ...,* k*-reasoning, or a according to a probability distribution of the different reasoning levels. The distributions are exogenously given and do not change among different rounds of the game.*

The definition of the different reasoning levels are discussed in Section 1. Besides, we define 0-reasoning by playing the same strategy as in the previous round.

The exogenously given distribution over the set of reasoning levels is defined as follows.

Definition 14 *For any player* i *and any reasoning level* k *let* P_{ik} *denote the probability of acting according to* k*-reasoning.*

A player is considered smarter than its opponent if his expected reasoning level is higher than that of his opponent. This is how we grab the difference in the complexity of human thinking and try to point out why smarter people may win more frequently in several strategic interactions.

We begin the analysis with the description of the equilibrium concept for the simplest case, where both players have two strategies each.

3.1 The 2-by-2 model

Initially, we restrict attention to two-player 2x2 games with the following general payoff matrix:

			Player 2	
			q	$1-q$
			Left	Right
Player 1	p	Top	$u_{TL}; v_{TL}$	$u_{TR}; v_{TR}$
	$1-p$	Bottom	$u_{BL}; v_{BL}$	$u_{BR}; v_{BR}$

Table 1: The 2-by-2 game

According to Table 1, Player 1's strategies are Top and Bottom, while Player 2 can choose between Left and Right. $p, 1-p, q, 1-q$ are the respective strategy choice probabilities. Finally, u_{ij}, v_{ij} (where $i \in \{T, B\}$ and $j \in \{L, R\}$) are the two players' payoff levels given a certain strategy pair.

According to Assumption 2, we assume that the game does not have a pure strategy Nash-equilibrium. A necessary and sufficient condition for this is

$$u_{TL} > u_{BL}, u_{BR} > u_{TR}, v_{TR} > v_{TL}, v_{BL} > v_{BR}. \tag{12}$$

This means that the best responses of both players are given for any action of their opponent. E.g. if Player 1 chooses Top, then Player 2's best response is Right, as $v_{TR} > v_{TL}$.

For games that do not have a pure strategy Nash equilibrium, the classical solution is the mixed strategy Nash equilibrium. As a reference point, we provide the formulas for calculating the Nash-equlibrium mixing probabilities of the two players for the game using the notations of Table 1:

$$p_{nash} = \frac{v_{BL} - v_{BR}}{v_{BL} - v_{BR} + v_{TR} - v_{TL}}, \tag{13}$$

$$q_{nash} = \frac{u_{BR} - u_{TR}}{u_{BR} - u_{TR} + u_{TL} - u_{TR}}. \tag{14}$$

However, the mixed strategy Nash equilibrium has been criticized, as several experiments pointed out that it does not describe player behavior properly (e.g. [2, 7]). As described in the introduction, these findings led researchers to construct behavioral game theory models that may explain the way of strategic thinking more precisely.

Our model tries to provide a mathematical framework for player behavior.

We introduce our concept of play history in the next definition.

Definition 15 *We use the notion history for the strategy profile of the previous round of the game.*

The history of the game described by Table 1 can be the following: (Top;Left), (Top;Right), (Bottom;Left) and (Bottom;Right).

Depending on the history, we can define four different games, where the strategies and the payoffs are the same. The only difference is that both players keep the history in mind and this has an influence on their decisions, i.e. their strategy mixing probabilities.

The payoff and probability matrices with the four different histories are as follows.

			Player 2	
			q_{TL}	$1 - q_{TL}$
			Left	Right
Player 1	p_{TL}	Top	$u_{TL}; v_{TL}$	$u_{TR}; v_{TR}$
	$1 - p_{TL}$	Bottom	$u_{BL}; v_{BL}$	$u_{BR}; v_{BR}$

Table 2: The game with (Top, Left) history

			Player 2	
			q_{TR}	$1 - q_{TR}$
			Left	Right
Player 1	p_{TR}	Top	$u_{TL}; v_{TL}$	$u_{TR}; v_{TR}$
	$1 - p_{TR}$	Bottom	$u_{BL}; v_{BL}$	$u_{BR}; v_{BR}$

Table 3: The game with (Top, Right) history

			Player 2	
			q_{BL}	$1 - q_{BL}$
			Left	Right
Player 1	p_{BL}	Top	$u_{TL}; v_{TL}$	$u_{TR}; v_{TR}$
	$1 - p_{BL}$	Bottom	$u_{BL}; v_{BL}$	$u_{BR}; v_{BR}$

Table 4: The game with (Bottom, Left) history

			Player 2	
			q_{BR}	$1-q_{BR}$
			Left	Right
Player 1	p_{BR}	Top	$u_{TL}; v_{TL}$	$u_{TR}; v_{TR}$
	$1-p_{BR}$	Bottom	$u_{BL}; v_{BL}$	$u_{BR}; v_{BR}$

Table 5: The game with (Bottom, Right) history

An example for the game in Table 1 can be as follows.

Example 16 *We assume that Player 1 chooses strategy Top, while Player B chooses strategy Left in the first round of the game. Thus, for the second round the history is (Top,Left). Let Player 1 play according to 0-reasoning with certainty and Player 2 according to 1-reasoning with certainty. Thus, Player 1 will remain at strategy Top, while Player 2 will choose his best response to Top with certainty, that is, Right. We arrived at the (Top, Right) strategy pair with certainty. Using the notations of Table 2, this means that* $p_{TL} = 1$*, while* $q_{TL} = 0$*. Clearly, even if the reasoning levels follow a more complicated distribution, then* $p_{TL} \in [0,1]$ *and* $q_{TL} \in [0,1]$*. As we arrived at (Top, Right) with certainty, (Top, Right) becomes the history for the third round of the game. Applying again that Player 1 plays according to 0-reasoning and Player 2 plays according to 1-reasoning with certainty, the (Top, Right) profile will occur in the third round of the game. With the same logic, the process can be continued till any* k*th round of the game.*

If we consider any 2-by-2 game that satisfies our assumptions and a first-round strategy profile and a distribution on the set of reasonng levels is given for both players, there emerge the following questions:

1. What is the ex ante strategy choice distribution of the two players if the number of rounds $n \to \infty$? Is there any limiting distribution?

2. What is the expected payoff of the players for each round if $n \to \infty$? Is the series of expected payoffs convergent?

To answer these questions, we will apply the theory of Markov chains.

3.2 The Markov chain model of the 2-by-2 game

The outguessing model can be interpreted as a Markov chain. According to Assumption 4, the players keep in mind only the actions of the previous round of the game. Let us define the Markov chain of the described 2-by-2 game.

Proposition 17 *The strategy profile sequence of the repeated game represents a Markov chain.*

Proof. The proof comes directly from Definition 3 and Assumption 4 that show that the strategy profile sequence $\{X_n\}$ $(n \geq 1)$ has the Markov property. □

In a 2-by-2 game we have 4 different strategy profiles, in our example these are (Top;Left), (Top;Right), (Bottom;Left) and (Bottom;Right), or in a shorter form: TL, TR, BL, BR. Thus, we can define the four states as follows:

State no.	Strategy profile
1	(Top;Left)
2	(Top;Right)
3	(Bottom;Left)
4	(Bottom;Right)

Table 6: States of the Markov chain

The transition matrix of the Markov chain can be obtained by using the data of the general payoff and probability matrices from the previous subsection.

Proposition 18 *The 4-by-4 transition matrix can be written as follows:*

$$T = \begin{pmatrix} p_{TL}q_{TL} & p_{TL}(1-q_{TL}) & (1-p_{TL})q_{TL} & (1-p_{TL})(1-q_{TL}) \\ p_{TR}q_{TR} & p_{TR}(1-q_{TR}) & (1-p_{TR})q_{TR} & (1-p_{TR})(1-q_{TR}) \\ p_{BL}q_{BL} & p_{BL}(1-q_{BL}) & (1-p_{BL})q_{BL} & (1-p_{BL})(1-q_{BL}) \\ p_{BR}q_{BR} & p_{BR}(1-q_{BR}) & (1-p_{BR})q_{BR} & (1-p_{BR})(1-q_{BR}) \end{pmatrix}.$$

Proof. The elements of the transition matrix are the probabilities of getting into a given state from a given previous state, i.e. the probabilities that a certain strategy profile will emerge given the strategy profile of the previous round. Using the previously defined p_{ij} and q_{ij} probabilities, and knowing that strategic decisions are independent from each other in a simultaneous game, we obtain the formula in the statement. □

Clearly, the transition matrix depends directly only on the players' probabilities of choosing a certain strategy with a given history. The transition matrix is independent from the construction of these probabilities. Thus, it remains the same for all models where the players' actions depend only on the previous round and a probability distribution is exogenously given for both players on the set of reasoning levels.

The following lemma indicates that only 0, 1, 2 and 3-reasoning levels are relevant for the given 2-by-2 game.

Lemma 19 *For 2-by-2 games and for all* $k \geq 4$, k*-reasoning is equivalent to* $(k-4)$*-reasoning.*

Proof. The proof comes directly from the inequalities in (12). □

We need one more definition to be able to state the key result of the paper.

Definition 20 *Let us denote the initial strategy distribution of the players by* π_0.

The distribution over the state space (the set of strategy pairs) in the nth round can obviously be calculated as follows:

$$\pi_n = T^n \pi_0. \tag{15}$$

The key result states that under certain conditions there exists a limiting distribution if $n \to \infty$.

Proposition 21 *If* $P_{ik} > 0$ *(see Definition 14) for every* $i \in \{1, 2\}$ *and every* $k \in \{0, 1, 2, 3\}$ *and if* $n \to \infty$*, then there exists a limiting distribution* π *over the state space of the* $\{X_n\}$ $(n \geq 1)$ *Markov chain.*

Proof. If $P_{ik} > 0$ for every $i \in \{1, 2\}$ and every $k \in \{0, 1, 2, 3\}$, then it can easily be verified according to Definitions 8 and 13 that $\{X_n\}$ is an irreducible aperiodic Markov chain. Thus, according to Theorem 11 it is recurrent and has a limiting distribution. □

We arrived at our equilibrium concept. The outguessing equilibrium is defined as the limiting distribution.

Definition 22 *We call the limiting distribution* π *the outguessing equilibrium.*

According to the proof of Proposition 21, the key result is supported by Theorem 11: the strategy profile sequence of the repeated 2-by-2 game represents a Markov chain that has a limiting distribution.

As far as the players' expected payoffs are concerned, they can easily be determined by multiplying π' (π vector transposed) with the vector of the corresponding payoff levels.

By running our script, the limiting distribution π and the long-term expected payoffs can be calculated and visualized. It becomes clear that the player with the higher expected reasoning level has the higher expected payoff on the long run.

3.3 Numerical experiment—the matching pennies

In the matching pennies game, both players have to announce "heads" or "tails" at the same time. If the announcements are the same, Player 1 wins 1 from Player 2, otherwise Player 2 wins 1 from Player 1. The payoff matrix of the well-known zero-sum game is as follows:

			Player 2	
			q	$1-q$
			Left	Right
Player 1	p	Top	$1;-1$	$-1;1$
	$1-p$	Bottom	$-1;1$	$1;-1$

The Nash-equilibrium mixing probabilities are 50%-50% for both players.

Let us assume that in our model the initial strategy choice probabilities are 0.5 each (in the first round when there is no history). For the distributions over the set of reasoning levels (see Definition 14) let us assume that $P_{10} = 0.4$, $P_{11} = 0.2$, $P_{12} = 0.2$, $P_{13} = 0.2$, while $P_{10} = 0.2$, $P_{10} = 0.2$, $P_{10} = 0.4$, $P_{10} = 0.4$. Clearly, Player 2 is considered smarter due to his higher expected reasoning level.

We ran our script and the process of the expected payoffs calculated from $(\pi_1, \pi_2, \ldots, \pi_n, \ldots)$ and T is depicted in Figure 1.

It can easily be seen that these processes converge to certain limit values, a numerical evidence for Proposition 21. It is also verified for the Matching pennies that the smarter Player 2 (the above "+" sequence) has a higher expected payoff (0.232) than Player 1 (-0.232). According to the mixed strategy Nash equilibrium concept, both players would have zero expected payoff.

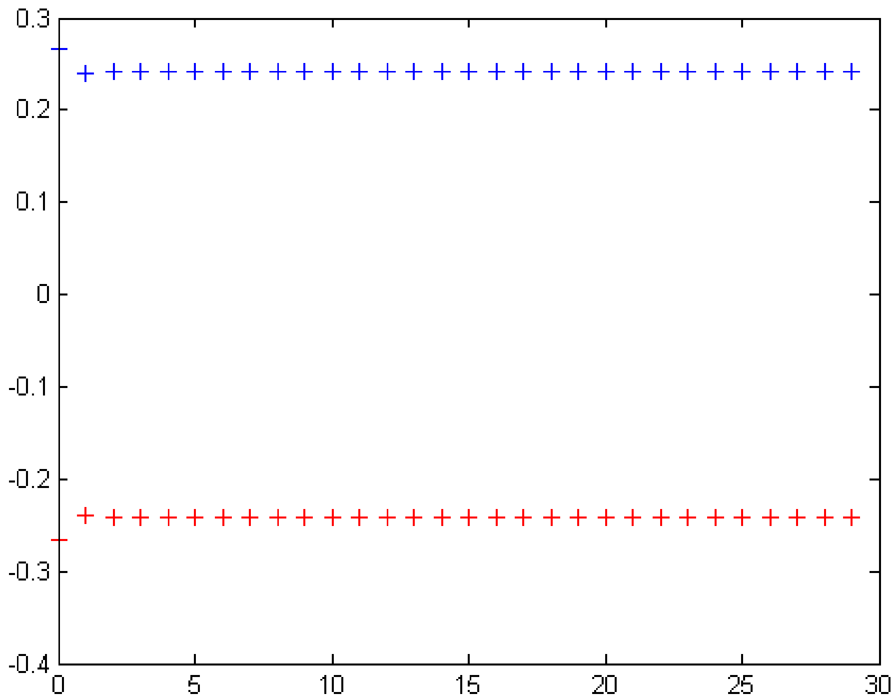

Figure 1: Long-term expected payoffs of the two players depending on the number of rounds; 2-by-2 case

3.4 The 3-by-3 case

If we consider a 3-by-3 game and keep all our assumptions, then the propositions trivially remain valid. The only exception is Lemma 19. The modified version for 3-by-3 games is as follows.

Lemma 23 *For 3-by-3 games and for all* $\mathsf{k} \geq 6$*,* k*-reasoning is equivalent to* $(\mathsf{k}-6)$*-reasoning.*

Proof. The proof comes directly from the modified version of inequalities in (12) for 3-by-3 games. □

What is important is that Proposition 21 remains valid if $\mathsf{k} \in \{0, 1, 2, 3, 4, 5\}$. By running our script for 3-by-3 games, we can obtain numerical evidence for the results.

3.5 Numerical experiment—the Rock-paper-scissors game

The payoff matrix of the well-known Rock-paper-scissors game is as follows.

			Player 2		
			q_1 Rock	q_2 Paper	q_3 Scissors
	p_1	Rock	$0; 0$	$-1; 1$	$1; -1$
Player 1	p_2	Paper	$1; -1$	$0; 0$	$-1, 1$
	p_3	Scissors	$-1; 1$	$1; -1$	$0; 0$

We fixed the expected (average) reasoning level of Player 1 at 2.0 and that of Player 2 at 2.5 (not violating the conditions of Proposition 7). The expected payoffs are depicted in Figure 2 below.

Clearly, smarter Player 2 (crosses) ”beats” Player 1 (dotted crosses) on the long run. Player 2’s long term expected payoff lies at 0.038, while Player 1’s is -0.038.

4 Notes about the script

Our script was written in Matlab. Its inputs are the following values:

- the payoff matrix of the corresponding 2-by-2 or 3-by-3 game
- the players’ discrete probability distributions over the set of reasoning levels
- initial strategies (i.e. player behavior in the very first round–either a fixed strategy pair or an initial distribution)

The script works the following way. Firstly, from the given values the script calculates the transition matrix of the Markov chain. Then, the outguessing equilibrium (see Definition 13) and the long-term expected payoffs for both players are also calculated. Proposition 3 suggests that the “smarter” player (if there is one) beats its opponent on the long run.

Apart from the calculations, the power of the script is that the outguessing equilibrium concept can be tested for any 2-by-2 or 3-by-3 bimatrix game that does not have a Nash equilibrium on pure strategies.[1]

[1] Upon request, the authors provide the interested reader with the script with pleasure for testing purposes.

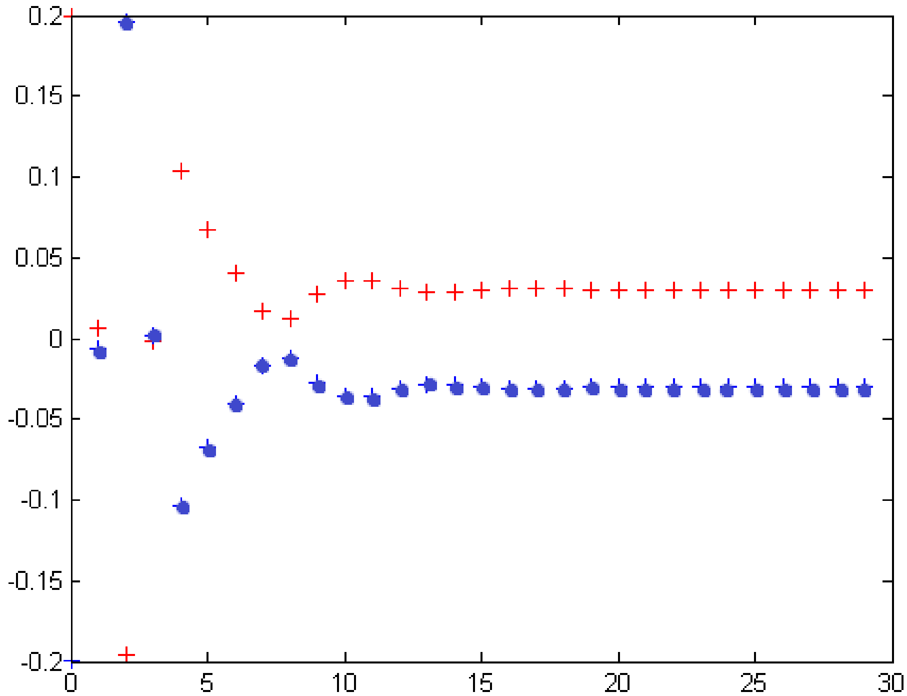

Figure 2: Long-term expected payoffs of the two players depending on the number of rounds; 3-by-3 case

5 Conclusions

Behavioral game theory has been dealing with the understanding of human behavior in strategic interactions. Among several different approaches, we have developed a behavioral model that aims at showing why "smarter" people outguess their opponents and win more frequently in some well-known zero-sum games.

Game theory is a useful modeling tool for network problems. We defined a behavioral model in a two-player non-cooperative network.

We used the concept of iterative reasoning to define smartness. The theory of Markov chains has proved to be a very useful technical tool to prove the main result of the paper. Namely, an outguessing equilibrium according to our definition exists and can even be calculated.

A Matlab script supports the calculations and provides numerical evidence for our concept.

The authors wish to emphasize that the introduced model can not only be applied for the games recalled in the examples, but for any conflict situation that can be modeled by bimatrix games.

Although the theoretical results are proved, and numerical evidence is also provided, there have remained some interesting questions which are out of the scope of this paper. One of these questions is rather technical: what types of Markov chains (e.g. periodic, absorbing etc…) can emerge given a specific bimatrix game and initial strategy profile? Another one deals with the game theoretic assumptions: if either the number of players, or the simultanity of decisions were altered, or we allowed for non-generic games, how would the equilibrium outcome change? These problems are left for future research.

Acknowledgements

This paper was supported in part by the TÁMOP-4.2.2C-11/1/KONV-2012-0001 project supported by the European Union, co-financed by the European Social Fund.

References

[1] T. Alpcan, T. Basar, A Globally Stable Adaptive Congestion Control Scheme for Internet-Style Networks with Delay, *IEEE/ACM Trans. On Networking* **13** (2005) 1261–1274. ⇒71

[2] T. L. Beard, R. Beil, Do people rely on the self-interested maximization of others? An experimental test, *Management Science* **40,** 2 (1994) 252–262. ⇒72, 78

[3] S.C. Cabrera, C. M. Capra, R. Gomez, *The effects of common advice on one-shot traveller's dilemma games: Explaining behavior through an introspective model with errors*, Washington and Lee University working paper, 2002. ⇒72

[4] C.F. Camerer, *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton University Press, 2003, pp. 199–264. ⇒71, 72

[5] C. M. Capra, J.K. Goeree, R. Gomez, C. Holt, Anomalous behavior in traveller's dilemma?, *Amer. Economic Review* **89,** 3 (1999) 678–690. ⇒72

[6] S. N. Ethier, *The Doctrine of Chances*, Springer, 2010, pp. 119–155. ⇒73

[7] J. K. Goeree, C. Holt, Stochastic game theory: For playing games, not just for doing theory, *Proc. National Academy of Sciences* **96** (1999) 10564–10567. ⇒ 72, 78

[8] T. Ho, C. F. Camerer, K. Weigelt, Iterated dominance and iterated best-response in experimental "p-beauty contests", *Amer. Economic Review* **88** (1998) 947–969. ⇒72

[9] S. P. Meyn, R. L. Tweedie, *Markov chains and stochastic stability*, Cambridge University Press, 2009. ⇒73

[10] S. P. Meyn, *Control Techniques for Complex Networks*, Cambridge University Press, 2007. ⇒73

[11] H. Moulin, *Game Theory for the Social Sciences*, New York University Press, 1986. ⇒72

[12] R. Nagel, Unravelling in guessing games: an experimental study, *Amer. Economic Review* **85,** 5 (1995) 1313–1326. ⇒72

[13] J. Nash, Non-cooperative games, *Annals of Mathematics* **54,** 2 (1951) 286-295. ⇒71

[14] J. R. Norris, *Markov chains*, Cambridge University Press, 1998. ⇒73

[15] I. A. Shah, S. Jan, I. Khan, S. Qamar, An overview of game theory and its applications in communication networks *Int. J. Multidisciplinary Sciences and Engineering* **3,** 4 (2012) 5–11. ⇒71

[16] Y. Sovik, *Impossible bets: an experimental study*, University of Oslo working paper, 1999. ⇒72

Received: February 16, 2014 • Revised: March 31, 2014