Playing off-line games with bounded rationality

Jérôme Renault
CEREMADE
Université Paris Dauphine
Pl. du Marechal de Lattre de Tassigny
F–75775 Paris Cedex 16, France
renault@ceremade.dauphine.fr

Marco Scarsini
Dipartimento di Scienze Economiche e Aziendali
LUISS
Viale Pola 12
I–00198 Roma, Italy
and HEC, Paris
mscarsini@luiss.it

Tristan Tomala
CEREMADE
Université Paris Dauphine
Pl. du Marechal de Lattre de Tassigny
F–75775 Paris Cedex 16, France
tomala@ceremade.dauphine.fr

March 14, 2007

Abstract

We study a two-person zero-sum game where each player chooses simultaneously a sequence of actions, and the payoff is the average of a one-shot payoff over the joint sequence. We consider the maxmin value of the game where players are restricted to strategies implemented by finite automata. We study the asymptotics of this value and a complete characterization in the matching pennies case. We extend the analysis of this game to the case of strategies with bounded recall.

Key words: Zero-sum games, bounded recall, automata, periodic sequences, de Bruijn graphs.

MSC 2000 Subject Classification: Primary 91A05. Secondary 91A26.

OR/MS subject classification: Primary: Games/group decisions, noncooperative.

1 Introduction

The common assumption of perfectly rational agents has been questioned by several papers in game theory and a whole literature was born, where players are subject to some constraint in their ability to compute or to remember. Therefore only strategies that are not computationally too demanding are available to them.

Bounds on players' rationality can be expressed in different forms. For instance one stream of literature considers games played by finite automata (see, e.g., Neyman (1985, 1998); Rubinstein (1986); Abreu and Rubinstein (1988); Kalai and Stanford (1988); Ben-Porath (1990, 1993); Neyman and Okada (2000b); Gossner and Hernández (2003, 2006); Gossner et al. (2003); Bavly and Neyman (2005)).

A partially different stream of literature deals with players who can remember only the most recent actions taken (see, e.g., Lehrer (1988, 1994); Sabourian (1998); Gossner et al. (2003); Bavly and Neyman (2005); Renault et al. (2006)).

Other bounds on the complexity of the players have been considered, e.g., by Neyman and Okada (1999, 2000a). This list is by no means exhaustive.

Many of the existing papers consider zero-sum games and study the effect of different restrictions in the players' rationality on the outcome of the game. Our paper goes in this direction and deals with repeated two-person zero-sum games with imperfect monitoring where the signal is trivial, that is, each player observes only her own actions. This corresponds to playing a normal-form game where the two players choose simultaneously an infinite sequence of actions. We will consider only pure strategies for the repeated game.

First we consider the above game played by automata, and we focus on the maxmin of the game when the automata have different size. The main result in this section is that if player 1 is an automaton with size $2m$ and player 2 is an automaton of size m, then player 1 can almost guarantee the value *in mixed strategies* of the stage game, up to an error of order $1/m$.

When the stage game is "matching pennies," then an exact result can be obtained for the above maxmin, for every possible size of the two automata.

Then we consider players with bounded recall. This model is much more complicated to analyze than the one with automata. Only some results about matching pennies are given. They show a counterintuitive nonmonotone behavior of the maxmin of the repeated game, when the recall of the first player is equal to the recall of the second player plus one.

The proofs are based on some arithmetic arguments about periodic sequences. For the model with bounded recall we use in addition some results on de Bruijn graphs and sequences. These sequences have already appeared in some bounded rationality models (see, e.g., Challet and Marsili (2000); Piccione and Rubinstein (2003); Liaw and Liu (2005); Gossner and Hernández (2006); Renault et al. (2006))

The paper is organized as follows. Section 2 describes the model. Section 3 deals with games played by automata. Section 4 studies games with bounded recall.

2 Off-line games

We start with a finite zero-sum game $G = (\mathcal{A}, \mathcal{B}, g)$ where $\mathcal{A}, \mathcal{B}$ are nonempty finite sets and $g : \mathcal{A} \times \mathcal{B} \to \mathbb{R}$. Player 1 chooses $a \in \mathcal{A}$, player 2 chooses $b \in \mathcal{B}$ and the payoff $g(a, b)$ is paid by player 2 to player 1.

In the associated off-line game Γ, player 1 chooses an $\mathcal{A}$-valued infinite sequence $x = (x_i)_{i\geq 1}$, player 2 chooses a $\mathcal{B}$-valued infinite sequence $y = (y_i)_{i\geq 1}$, and the associated payoff is

$$\gamma(x, y) = \lim \frac{1}{t} \sum_{i=1}^{t} g(x_i, y_i), \tag{2.1}$$

where lim denotes a Banach limit, i.e. a linear mapping on the set of bounded sequences such that $\liminf \leq \lim \leq \limsup$. The use of a Banach limit (usual in repeated games) will be immaterial in most of the paper since we shall deal mostly with converging sequences.

We shall use the following notations throughout the paper. For a finite set $\mathcal{A}$, we let $\Delta(\mathcal{A})$ be the set of probability distributions on $\mathcal{A}$. We use the same symbol for the multilinear extension of g, i.e., given a two finite sets $\mathcal{A}, \mathcal{B}$, a function $g : \mathcal{A} \times \mathcal{B} \to \mathbb{R}$ and two distributions $\mu \in \Delta(\mathcal{A})$, $\nu \in \Delta(\mathcal{B})$, we define

$$g(\mu, \nu) = \sum_a \sum_b \mu(a)\nu(b)g(a, b), \tag{2.2}$$

and we shall identify the degenerate distribution at a point x with the point x itself. We denote by $\mathrm{val}(G)$ the value of the game $G = (\mathcal{A}, \mathcal{B}, g)$ in mixed strategies,

$$\mathrm{val}(G) = \max_{\mu\in\Delta(\mathcal{A})} \min_{b\in\mathcal{B}} g(\mu, b) = \min_{\nu\in\Delta(\mathcal{B})} \max_{a\in\mathcal{A}} g(a, \nu),$$

by $\underline{v}(G)$ the maxmin in pure strategies,

$$\underline{v}(G) = \max_{a\in\mathcal{A}} \min_{b\in\mathcal{B}} g(a, b),$$

and by $\overline{v}(G)$ the minmax in pure strategies,

$$\overline{v}(G) = \min_{b\in\mathcal{B}} \max_{a\in\mathcal{A}} g(a, b).$$

For a nonempty finite set $\mathcal{A}$, the set of $\mathcal{A}$-valued sequences is denoted $\mathcal{A}^\omega$. A sequence $x = (x_t)_{t\geq 1}$ is n-periodic $x_{t+n} = x_t$ for each t. A sequence is periodic if it is n-periodic for some $n \geq 1$. The set of all periodic sequences is denoted by $S(\mathcal{A})$. For each x in $S(\mathcal{A})$, we let $\mathbf{per}(x)$ be the smallest n such that x is n-periodic. For each $n \geq 1$, we let $S_n(\mathcal{A})$ be the set of periodic sequences x such that $\mathbf{per}(x) \leq n$, and we let $S'_n(\mathcal{A})$ be the set of n-periodic sequences, i.e. all sequences x such that $\mathbf{per}(x)$ divides n.

Let $\Delta_n(\mathcal{A})$ be the set of probability distributions which are fractional in n, that is $\mu \in \Delta_n(\mathcal{A})$ if for every $a \in \mathcal{A}$, the value $n\mu(a)$ is an integer (with an abuse of notation we write $\mu(a)$ instead of $\mu(\{a\})$). An n-periodic sequence x induces an *empirical distribution* $\mu_x \in \Delta_n(\mathcal{A})$, where:

$$\mu_x(a) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{x_i=a\}}.$$

3 Off-line games played by automata

The main goal of the paper is to study the maxmin in pure strategies of the off-line game, when players are restricted to boundedly complex strategies. A common way to model bounded complexity is to consider strategies implemented by finite automata.

An automaton (say for player 1) is a tuple (Q, q_1, f, h) where Q is a finite nonempty set of states, $q_1 \in Q$ is an initial state, $f : Q \to \mathcal{A}$ is the action function and $h : Q \to Q$ is the transition function. An automaton generates a sequence of actions $(x_t)_{t\geq 1}$ as follows:

$$x_1 = f(q_1),\ q_2 = h(q_1), \dots, q_{t+1} = h(q_t),\ x_{t+1} = f(q_{t+1}),$$

and so on. Since the set of states is finite, the sequences of states and of actions generated by the automaton are eventually periodic (periodic from some stage on) with period of length no more than $|Q|$ (the cardinality of Q). In the off-line game with the payoff as in (2.1), the transient phase of the automaton is irrelevant: given an eventually periodic sequence of actions for player 1, modifying finitely many terms to make it periodic does not change the payoff. So, without loss of generality, we may view the automaton as a periodic sequence. Conversely, note that any sequence of actions with period n can be played by an automaton with n states. We thus identify the set $S_n(\mathcal{A})$ with the set of strategies induced by automata with no more than n states. For $n, m \geq 1$ consider the quantity,

$$V_{n,m}(G) = \max_{x\in S_n(\mathcal{A})} \min_{y\in S_m(\mathcal{B})} \gamma(x, y),$$

which is the best payoff that player 1 can guarantee with an automaton with at most n states against player 2, whose automaton has at most m states.

Clearly, if $n \leq m$, then $V_{n,m}(G) = \underline{v}(G)$. Furthermore $V_{n,m}(G)$ is non-decreasing in n and non-increasing in m.

3.1 Properties

The off-line game Γ has generally no value in pure strategies and its value in mixed strategies is $\mathrm{val}(G)$:

$$\sup_{x\in\mathcal{A}^\omega} \inf_{y\in\mathcal{B}^\omega} \gamma(x, y) = \underline{v}(G).$$

In fact, if we fix a sequence x of player 1, player 2 may choose a sequence y such that for each stage t, y_t minimizes $g(x_t, b)$ over $b \in \mathcal{B}$. Thus, $\sup_{x\in\mathcal{A}^\omega} \inf_{y\in\mathcal{B}^\omega} \gamma(x,y) \leq \underline{v}(G)$. On another hand if player 1 plays constantly an action a that maximizes $\min_{b\in\mathcal{B}} g(a,b)$ over $\mathcal{A}$, one has $\gamma(x,y) \geq \underline{v}(G)$ for each y.

Analogously

$$\inf_{y\in\mathcal{B}^\omega} \sup_{x\in\mathcal{A}^\omega} \gamma(x,y) = \overline{v}(G).$$

Moreover, by playing finite support mixed strategies in Γ, both players guarantee $\mathrm{val}(G)$. Player 1 (resp. player 2) guarantees $\mathrm{val}(G)$ by drawing an action at random according to an optimal mixed strategy in G and playing constantly the selected action.

The main objects of our study are maxmin values in pure strategies. Player 2 can *defend* the value of the game with constant strategies. Formally, for each $x \in \mathcal{A}^\omega$, there exists $y \in S_1(\mathcal{B})$ such that $\gamma(x,y) \leq \mathrm{val}(G)$. To see this, let $x \in \mathcal{A}^\omega$. For each $t \geq 1$ and $a \in \mathcal{A}$ define

$$\mu_{x,t}(a) = \frac{1}{t}\sum_{i=1}^{t} \mathbf{1}_{\{x_i=a\}}.$$

For each $y \in S_1(\mathcal{B})$ constantly equal to b we have

$$\frac{1}{t}\sum_{i=1}^{t} g(x_i, y_i) = g(\mu_{x,t}, b)$$

Let $\mu_x(a) = \lim_t \mu_{x,t}(a)$, where lim is a Banach limit (the usual limit may not always exist). Since lim is linear, this defines $\mu \in \Delta(\mathcal{A})$ such that $\gamma(x,y) = g(\mu, b)$. If player 2 chooses b that minimizes $g(\mu,\cdot)$, then $\gamma(x,y) \leq \mathrm{val}(G)$.

Therefore for each n, m, $\underline{v}(G) \leq V_{n,m}(G) \leq \mathrm{val}(G)$.

The next theorem states that the distance between $V_{2m,m}(G)$ and $\mathrm{val}(G)$ is of order $1/m$. This shows that, to guarantee the fully rational solution of the game, here the value, player 1 needs only to be twice more complex than player 2.

Theorem 3.1. *For each $m \geq 2$,*

$$\mathrm{val}(G) - \frac{\|G\|}{m} \leq V_{2m,m}(G) \leq \mathrm{val}(G),$$

where $\|G\| := \max_b \sum_a |g(a,b)|$.

The key to this theorem is to prove that when player 1 chooses a sequence with period n, player 2 has a best reply whose period divides n. In particular, when player 1 chooses a sequence with a prime period, the best that player 2 can do is to respond by a constant sequence. Hence, when player 2 has complexity m, player 1 with complexity $2m$ may choose a prime period p such that $m < p < 2m$, and guarantee the value of the stage game up to $1/m$.

We turn now to the formal proof of Theorem 3.1. We start by studying the problem of computing a best reply of player 2 within $S_m(\mathcal{B})$ against a periodic sequence $x \in S(\mathcal{A})$. Fix thus such a sequence x with $\mathbf{per}(x)$, and an integer m. In the sequel we let $p = \gcd(n, m)$ (the greatest common divisor of n and m), and $q = \operatorname{lcm}(n, m)$ (the least common multiple of n and m). We consider the problem of finding an m-periodic, $\mathcal{B}$-valued sequence y that minimizes the average of g over a joint period of (x, y), that is

$$\min_{y \in S'_m(\mathcal{B})} \lim_t \frac{1}{t} \sum_{i=1}^{t} g(x_i, y_i) = \min_{y \in S'_m(\mathcal{B})} \frac{1}{q} \sum_{i=1}^{q} g(x_i, y_i).$$

The n-periodic sequence x is the repetition of a word of length n with letters in $\mathcal{A}$, which is denoted $\tilde{x} = (x_1, \ldots, x_n)$. Likewise, we write $\tilde{y} = (y_1, \ldots, y_m)$. There are two integers u, v such that $q = un$ and $q = vm$, so that within a period of the bivariate sequence (x, y), $\tilde{x}$ is repeated u times and $\tilde{y}$ is repeated v times. For each $j \in \{1, \ldots, m\}$, we consider the x_i's that y_j meets in the sequence (x, y), that is we consider $\{x_{j+tm} : t \in \mathbb{N}\}$, and we look at the indices of these x_i's *within a period of* x.

For each integer $t \in \mathbb{Z}$, let $[t]_n$ be the smallest positive element of the class of t modulo n, that is $[t]_n \in \{1, \ldots, n\}$ and $t - [t]_n$ is a multiple of n. We let for each $j \in \{1, \ldots, m\}$,

$$T_{n,m,j} = \{[j + tm]_n \,:\, t \in \mathbb{Z}\}.$$

We also denote $\mu_{x,m,j}$ the empirical distribution induced by the set of x_i's met by y_j, that is,

$$\mu_{x,m,j}(a) = \frac{1}{|T_{n,m,j}|} \sum_{i \in T_{n,m,j}} \mathbf{1}_{\{x_i = a\}}. \tag{3.1}$$

Hence, using notation (2.2),

$$\frac{1}{q} \sum_{i=1}^{q} g(x_i, y_i) = \frac{1}{m} \sum_{j=1}^{m} g(\mu_{x,m,j}, y_j).$$

Therefore, to solve

$$\min_{y \in S'_m(\mathcal{B})} \frac{1}{q} \sum_{i=1}^{q} g(x_i, y_i),$$

one just has to choose y_i that minimizes $g(\mu_{x,m,j}, b)$ over $b \in \mathcal{B}$.

Lemma 3.2. *For every sequence $y \in S'_m(\mathcal{B})$, and every pair of indices $j, j' \in \{1, \ldots, m\}$, we have:*

$$[j]_p = [j']_p \Longrightarrow T_{n,m,j} = T_{n,m,j'},$$

with $p = \gcd(n, m)$.

Proof. Assume $[j]_p = [j']_p$, i.e. $j' = j + kp$ for some integer k. Let $i \in T_{n,m,j'}$. Then there exists two integers s, t such that

$$i = j' + tm + sn = j + kp + tm + sn.$$

From Bezout's identity (see e.g. Jones and Jones (1998)) there exist two integers c, d such that $p = cn + dm$. It follows that

$$i = j + (kc + s)n + (kd + t)m,$$

and thus $i \in T_{n,m,j}$. The conclusion is obtained by symmetry. $\square$

Lemma 3.2 shows that if two letters in $\tilde{y}$ have the same rank modulo $p = \gcd(n, m)$, then they meet the same set of letters of the sequence x. At optimum, these two letters can be chosen to be the same and thus y can be chosen p-periodic.

Corollary 3.3. (a) *For every $x \in S(\mathcal{A})$, the problem*

$$\min_{y \in S'_m(\mathcal{B})} \lim_t \frac{1}{t} \sum_{i=1}^{t} g(x_i, y_i)$$

has a solution y such that $\mathbf{per}(y)$ divides $\mathbf{per}(x)$.

(b) *For every $x \in S(\mathcal{A})$, the problem*

$$\min_{y \in S_m(\mathcal{B})} \lim_t \frac{1}{t} \sum_{i=1}^{t} g(x_i, y_i)$$

has a solution y such that $\mathbf{per}(y)$ divides $\mathbf{per}(x)$.

We may now prove Theorem 3.1.

Proof of Theorem 3.1. Given Proposition 3.4 (**??**), it is enough to prove that there exists $x \in S_{2m}(\mathcal{A})$ such that, for each $y \in S_m(\mathcal{B})$,

$$\gamma(x, y) \geq \mathrm{val}(G) - \frac{\|G\|}{m}.$$

Bertrand's postulate, first proved by Chebyshev, states that for every integer $m \geq 2$, there exists a prime number p such that $m < p < 2m$ (see e.g., Nagell (1964)).

Let μ be an optimal mixed strategy for player 1 in G. There exists $\mu^p \in \Delta_p(\mathcal{A})$ such that

$$\|\mu - \mu^p\|_\infty := \max_a |\mu(a) - \mu^p(a)| \leq \frac{1}{p}.$$

The mapping $\mu \mapsto \min_b g(\mu, b)$ is $\|G\|$-Lipschitz, so that

$$\min_b g(\mu^p, b) \geq \mathrm{val}(G) - \frac{\|G\|}{p} \geq \mathrm{val}(G) - \frac{\|G\|}{m}.$$

There exists a sequence x such that $\mathbf{per}(x) = p$ and $\mu_x = \mu^p$. To show this, it is enough to order the action set $\mathcal{A} = \{a_1, \ldots, a_K\}$ and play in sequence a_1, $p\mu^p(a_1)$ times, …, a_K, $p\mu^p(a_K)$ times. From Corollary 3.3, a best reply of player 2 , i.e. a sequence that achieves $\min_{y \in S_m(\mathcal{B})} \gamma(x, y)$, can be chosen such that $\mathbf{per}(y)$ divides p. As p is prime, then $\mathbf{per}(y) = 1$, that is y is constant (say equal to b) and $\gamma(x, y) = g(\mu^p, b)$. Thus, $\min_{y \in S_m(\mathcal{B})} \gamma(x, y) = \min_b g(\mu^p, b)$ which completes the proof. □

We use now the previous construction to prove that there exists a sequence of player 1 which guarantees $\mathrm{val}(G)$ against any periodic sequence of player 2.

Proposition 3.4. *There exists $x \in \mathcal{A}^\omega$ such that for each $y \in S(\mathcal{B})$, $\gamma(x, y) \geq \mathrm{val}(G)$.*

Proof of Proposition 3.4. We construct $x^* \in \mathcal{A}^\omega$ such that for each $y \in S(\mathcal{B})$, $\gamma(x^*, y) \geq \mathrm{val}(G)$. Let $(p_n)_n$ denote the sequence of prime numbers. For each n, denote by $x_n \in \mathcal{A}^{p_n}$ a word generating a sequence x with smallest period p_n and $\mu_x = \mu^{p_n}$ where μ^{p_n} is, as above, a $(1/p_n)$-approximation of an optimal mixed strategy of player 1. Such a word was just constructed in the proof of Theorem 3.1.

The sequence x^* is constructed by concatenating those words. For all $n \geq 1$, call *superword* and denote $\tilde{x}_n$ the repetition of x_n, $(p_n - 1)!$ times. Then $\tilde{x}_n$ has length $N_n := (p_n)!$. Choose then a sequence of integers k_n such that

$$\frac{N_{n+1}}{k_n N_n} \to 0, \quad \text{as } n \to \infty. \tag{3.2}$$

The sequence x^* is such that $\tilde{x}_1$ is repeated k_1 times, $\tilde{x}_2$ is repeated k_2 times, …, $\tilde{x}_n$ is repeated k_n times, and so on.

Let $y \in S(\mathcal{B})$, set $u = \mathbf{per}(y)$. For k large enough, $p_k > u$, hence u divides $p_k!$. Since $\tilde{x}_k$ has length $p_k!$, we have that y and $\tilde{x}_k$ have $p_k!$ as common period. The average payoff over this period is thus the one yielded by y and the periodic repetition of $\tilde{x}_k$. Denote this payoff $\gamma(\tilde{x}_k, y)$. As p_k is prime and $u > p_k$, the best payoff that y can achieve against $\tilde{x}_k$ is $\min_b g(\mu^{p_k}, b)$. Thus,

$$\gamma(\tilde{x}_k, y) \geq \mathrm{val}(G) - \frac{\|G\|}{p_k}.$$

Therefore, if we let γ_k be the average payoff yielded by the k-th superword against y, we have $\liminf \gamma_k \geq \mathrm{val}(G)$. Condition (3.2) ensures that the length of a superword is negligible with respect to the sum of the lengths of all the preceding superwords. Thus, the limit of the average payoff,

$$\lim \frac{1}{t} \sum_{i=1}^{t} g(x_i, y_i)$$

is the limit of the Césaro average of the $(\gamma_k)_k$, which yields, $\gamma(x, y) \geq \mathrm{val}(G)$. □

3.2 Matching pennies played by automata

We give now a sharper result for the following matching pennies game, denoted G^* in the sequel:

$$\begin{array}{c|c|c|} & L & R \\ \cline{2-3} T & 1 & 0 \\ \cline{2-3} B & 0 & 1 \\ \cline{2-3} \end{array}. \tag{3.3}$$

The value of this game is 1/2 and each player has a unique optimal mixed strategy which is $(1/2, 1/2)$. We get here a characterization of $V_{n,m}(G^*)$. For every pair of integers n, m, let

$$P(n,m) = \inf \left\{ p \text{ odd} : p \text{ divides } n, \text{ and } \frac{n}{p} \leq m \right\}$$

and $P(n,m) = +\infty$ if there is no such p.

For instance, if $n \leq m$, then $P(n,m) = 1$. If $m < n = 2^k$, then $P(n,m) = +\infty$. If n is prime and $n > m$, then $P(n,m) = +\infty$.

Theorem 3.5. *For every pair of integers n, m,*

$$V_{n,m}(G^*) - \mathrm{val}(G^*) = \frac{-1}{2 \max_{k \leq n} P(k,m)}.$$

Again we use the fact that when choosing his best reply, player 2 may choose a period that divides the period n of player 1. The key argument is as follows. Assume that $n = pq$ with p odd, and that player 2 chooses a sequence with period q. Then each letter of player 2 faces a distribution of actions of player 1 which is fractional in p. Such a distribution must depart from the optimal strategy $(1/2, 1/2)$ by at least $1/2p$. We turn now to the formal proof.

Proof of Theorem 3.5. Here $\mathcal{A} = \{T, B\}$ and $\mathcal{B} = \{L, R\}$.

Clearly $S_n(\mathcal{A}) = \cup_{k \leq n} S'_k(\mathcal{A})$ and

$$V_{n,m}(G^*) = \max_{k \leq n} \max_{x \in S'_k(\mathcal{A})} \min_{y \in S_m(\mathcal{B})} \gamma(x,y).$$

We will prove that

$$w_{k,m} := \max_{x \in S'_k(\mathcal{A})} \min_{y \in S_m(\mathcal{B})} \gamma(x,y) = \frac{1}{2} - \frac{1}{2P(k,m)}.$$

If $k \leq m$, this formula holds, as $P(k,m) = 1$ and $w_{k,m} = 0$, since we can choose $y = x$. From now on, assume $k > m$. We consider two cases.

Case 1: $P(k,m) < +\infty$.

Set $p = P(k, m)$ so that p is odd, divides k and

$$\ell := \frac{k}{p} \leq m.$$

We first show that for every x in $S'_k(\mathcal{A})$, there exists y in $S_m(\mathcal{B})$ such that

$$\gamma(x, y) \leq \frac{1}{2} - \frac{1}{2p}. \tag{3.4}$$

Let x in $S'_k(\mathcal{A})$ and let y be an ℓ-periodic sequence induced by the word $\tilde{y} = (y_1, \ldots, y_\ell)$. The joint period of (x, y) is then k, and the word $\tilde{y}$ is repeated p times. Each y_j meets thus p letters x_i, and for each j, $\mu_{x,\ell,j} \in \Delta_p(\mathcal{A})$, where $\mu_{x,\ell,j}$ is defined as in (3.1). One can thus choose $\tilde{y}$ such that

$$\gamma(x, y) = \frac{1}{\ell} \sum_{j=1}^{\ell} \min_{b \in \mathcal{B}} g(\mu_{x,\ell,j}, b).$$

A probability distribution $\mu \in \Delta_p(\mathcal{A})$ has the form

$$\mu = (\mu(T), \mu(B)) = \left(\frac{q}{p}, \frac{p-q}{p} \right),$$

with $q \in \{0, \ldots, p\}$, and

$$\min_{b \in \mathcal{B}} g(\mu, b) = \min(\mu(T), \mu(B)).$$

Since $p = 2d + 1$ for some integer d, then for each $q \in \{0, \ldots, p\}$,

$$\min \left(\frac{q}{p}, \frac{p-q}{p} \right) \leq \frac{d}{p} = \frac{1}{2} - \frac{1}{2p}.$$

Therefore, for the chosen y, (3.4) holds. We construct now $x \in S'_k(\mathcal{A})$ such that for each $y \in S_m(\mathcal{B})$,

$$\gamma(x, y) \geq \frac{1}{2} - \frac{1}{2p}.$$

By definition p is an odd divisor of k, and as $k > m$, $p > 1$. We let $k = Cp$ for some integer C, and $p = 2d + 1$ with d a positive integer. Let then x be the periodic sequence generated by the word

$$\tilde{x} = (x_1, \ldots, x_k) = \underbrace{T \ldots T}_{Cd \text{ times}} \underbrace{B \ldots B}_{C(d+1) \text{ times}}.$$

We claim that for this x, $\min_{y \in S_m(\mathcal{B})} \gamma(x, y)$ is achieved by the sequence which is constantly L and thus,

$$\min_{y \in S_m} \gamma(x, y) = \frac{Cd}{k} = \frac{d}{p}.$$

Let y that achieves this minimum. From Lemma 3.2, we may assume that $u := \mathbf{per}(y)$ divides k, so there is an integer D such that $k = Du$. Since $u \leq m < k$, we have $D \geq 2$. Let $(y_1, \ldots, y_u)$ be the word generating y, we claim that each letter y_j meets more B's than T's so at optimum, each y_j must be chosen equal to L.

For each $j \in \{1, \ldots, u\}$, y_j appears at stages $j + tu$, $t = 0, \ldots, D-1$. We just need to check that less than half of these dates are before the time of the last T, i.e. we check that,

$$|\{t = 0, \ldots, D-1 : j + tu \leq Cd\}| \leq \frac{D}{2}.$$

If $j + tu \leq Cd$, then $tu < Cd$. Thus,

$$t < \frac{Cd}{u} = \frac{Dd}{p} = D\frac{d}{2d+1} < \frac{D}{2},$$

which completes the proof in this case.

Case 2: $P(k,m) = \infty$.

In this case k has the form $k = p2^j$ with p odd and j nonnegative integer such that $k/p = 2^j > m$. We need to prove that $w_{k,m} = 1/2$. As $w_{k,m} \leq 1/2$, we prove that there exists $x \in S'_k(\mathcal{A})$ such that for each $y \in S_m(\mathcal{B})$, $\gamma(x,y) \geq 1/2$.

Consider then the sequence x with $\mathbf{per}(x) = 2^j$ induced by the word of length 2^j

$$\underbrace{T \ldots T}_{2^{j-1}\text{ times}}\ \underbrace{B \ldots B}_{2^{j-1}\text{ times}}.$$

Given this sequence x, $\min_{y \in S_m(\mathcal{B})} \gamma(x,y)$ is achieved by y such that $\mathbf{per}(y)$ divides 2^j, and since $m < 2^j$, $\mathbf{per}(y) = 2^\ell$ with $\ell \leq j-1$. Thus the period of y divides 2^{j-1}, and each letter in y meets as many T's and B's. Thus $\gamma(x,y) = 1/2$. □

The following is obtained directly from Theorem 3.5.

Corollary 3.6. (a) *If $n \leq m$, then $V_{n,m}(G^*) = 0$.*

(b) *For each N and $m < 2^N$, $V_{2^N,m}(G^*) = 1/2$.*

(c) *For each m, $V_{2m,m}(G^*) = 1/2$.*

Proof. (a) This is immediate.

(b) If $n = 2^N$ for some N, then n has no odd divisor other than 1. Hence, for $m < n$, $P(n,m) = +\infty$ and $V_{n,m}(G^*) = 1/2$.

(c) For each $m \geq 1$, there is a unique $N \geq 1$ such that $2^{N-1} \leq m < 2^N$ and thus $m < 2^N \leq 2m$. Thus in the formula for $V_{2m,m}(G^*)$, choose $k = 2^N \leq 2m$. □

4 Off-line games with bounded recall

Another commonly used measure of complexity of strategies is the recall, that is the number of past values of the sequence on which the next value depends.

Definition 4.1. Given an nonempty finite set $\mathcal{A}$, a sequence $x \in \mathcal{A}^\omega$ has recall $k \in \mathbb{N}$ if there exists a mapping $f : \mathcal{A}^k \to \mathcal{A}$ such that for each $t > k$, $x_t = f(x_{t-1}, \ldots, x_{t-k})$.

Such a sequence x is eventually periodic. As for automata, the transient phase is irrelevant for our purposes, so we let $M_k(\mathcal{A})$ be the set of periodic sequences with recall k. For a sequence $x \in M_k(\mathcal{A})$, we have $\mathbf{per}(x) \leq |\mathcal{A}|^k$. However, there are sequences with period $|\mathcal{A}|^k$ which are not of recall k. Take for example the sequence

$$\underbrace{T \ldots T}_{2^{k-1} \text{ times}} \underbrace{B \ldots B}_{2^{k-1} \text{ times}}.$$

Although of period 2^k, this sequence does not have recall k, otherwise the k last T's should be followed by a T (assuming $2^{k-1} > k$).

In this section we study the maxmin value of the off-line game where players are restricted by the size of the recall

$$W_{j,k}(G) = \max_{x \in M_j(\mathcal{A})} \min_{y \in M_k(\mathcal{B})} \gamma(x, y).$$

As in the previous section, if $j \leq k$, then $W_{j,k}(G) = \underline{v}(G)$. Moreover $W_{j,k}$ is non-decreasing in j and non-increasing in k.

4.1 Matching pennies with bounded recall

Since the analysis of off-line games with bounded recall is significantly more difficult, we concentrate on the matching pennies game G^* defined in (3.3). The proofs of our result shall use the tools of the previous section and the theory of de Bruijn graphs (see, e.g., de Bruijn (1946) and Yoeli (1962) for some properties of these graphs).

Definition 4.2. A directed graph D_k called a *de Bruijn graph* if

- the set of vertices of D_k is $\{T, B\}^k$,
- there is an edge from $x = (x_1, \ldots, x_k)$ to $y = (y_1, \ldots, y_k)$ if and only if $(x_2, \ldots, x_k) = (y_1, \ldots, y_{k-1})$.

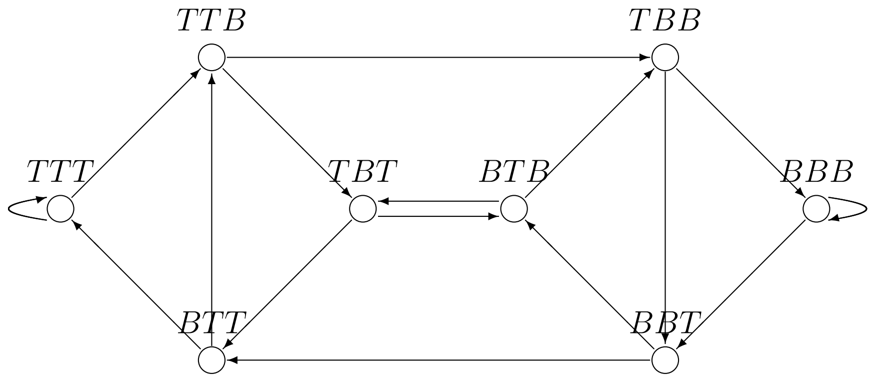

Figure 1. de Bruijn graph D_3

Consider player 1 with recall k. The set of possible recalls for player 1 is $\{T, B\}^k$. If the recall is the word $x \in \{T, B\}^k$ at some stage, the recall at the next stage is obtained by deleting the first letter of x and adding a new letter after x. If $x = (x_1, \ldots, x_k)$, the next recall is either $(x_2, \ldots, x_k, T)$ or $(x_2, \ldots, x_k, B)$. This defines a de Bruijn graph.

A sequence with recall k (for player 1) can the be viewed as a cycle in the de Bruijn graph D_k. Since D_k has 2^k vertices, the longest cycle has length 2^k. Since each vertex has as many outgoing as ingoing edges, such a cycle, called Hamiltonian cycle, exists (see, e.g., Bollobás (1998)). The associated sequence of Ts an Bs is called a *de Bruijn sequence.* A cycle of length 1 also exists (associated to the constant sequence $TTT\ldots$), but, more generally, the following proposition (Yoeli (1962)) shows that every length cycle is possible (see Lempel (1971) for a generalization to any finite alphabet).

Proposition 4.3. *For every p in $\{1, \ldots, 2^k\}$, there exists a cycle with length p in the de Bruijn graph D_k.*

The next lemma provides results similar to those obtained for automata.

Lemma 4.4. (a) *For every pair of integers (j, k), $0 \leq W_{j,k}(G^*) \leq 1/2$.*

(b) *If $j \leq k$, $W_{j,k}(G^*) = 0$.*

(c) *For every k, $W_{2^k,k}(G^*) = 1/2$.*

In Corollary 3.6 (c) player 1 can induce a period whose maximum length is twice the maximum length of the period induced by player 2. In Lemma 4.4 (c) player 1's maximum possible period is of length $2^{(2^k)}$, which is exponentially larger than the length of player 2's maximum possible period 2^k. Hence here, in accordance with known results on zero-sum games with bounded complexity (see Lehrer (1988), Ben-Porath (1990)), if player 1 is exponentially more complex than player 2, then she behaves like a fully rational player.

Proof of Lemma 4.4. (a) This follows from Proposition 3.4 as a constant sequence has recall 0.
(b) As before, if player 2 can use the same sequences as player 1, he can match at every stage.
(c) If $j = 2^k$, player 1 can choose the 2^{k+1}-periodic sequence x whose cycle is

$$\underbrace{T \dots T}_{2^k \text{ times}} \underbrace{B \dots B}_{2^k \text{ times}}.$$

Note that such x has indeed recall 2^k. Each sequence $y \in M_k$ has period $\mathbf{per}(y) \leq 2^k < \mathbf{per}(x)$. It follows then from the proof of Theorem 3.5, that for each such y, $\gamma(x, y) = 1/2$. □

The main concern of this section is the study of $W_{k+1,k}(G^*)$.

Theorem 4.5. (a) $W_{1,0}(G^*) = W_{2,1}(G^*) = 1/2$*,* $W_{3,2}(G^*) = 3/7$*,*

(b) $\lim_k W_{k+1,k}(G^*) = 1/2$*.*

Point (a) may suggest that the sequence $W_{k+1,k}(G^*)$ decreases away from 1/2, the intuition being that the advantage of having one extra slot of recall vanishes as k grows. Point (b) shows that it is not so.

Piccione and Rubinstein (2003, Section 5, Footnote 5) noticed that if player 1 plays a de Bruijn sequence of recall $k+1$, then player 2 with recall k must "have a frequency of mistakes of at least $1/(2(k+1))$." This statement implies that

$$W_{k+1,k}(G^*) \geq \frac{1}{2(k+1)}.$$

Our result shows that for large values of k, if player 1 plays a de Bruijn sequence of recall $k+1$, then player 2 with recall k has a frequency of mistakes close to 1/2.

Proof of Theorem 4.5. (a) Applying point c of Lemma 4.4 for $k = 0$ and $k = 1$ gives $W_{1,0}(G^*) = W_{2,1}(G^*) = 1/2$. We prove now that $W_{3,2}(G^*) = 3/7$. We first show that $W_{3,2}(G^*) \geq 3/7$.

Let x be the 3-recall strategy for player 1 that plays the 7-periodic sequence $TTTBBTB\ TTTBBTB \dots$. Any strategy y with recall 2 for player 2 has a period $\mathbf{per}(y) \leq 4 = 2^2$. From Theorem 3.5, the best sequence for player 2 can be chosen with a period that divides 7, thus with period 1. As the proportions of T's and B's are respectively 4/7 and 3/7, the best payoff that player 2 can get is 3/7.

We prove now $W_{3,2}(G^*) \leq 3/7$. Appling Theorem 3.5, we have

$$V_{7,4}(G^*) = \frac{1}{2} - \frac{1}{2 \max_{k \leq 7} P(k, 4)}.$$

We have $P(1,4) = P(2,4) = P(3,4) = P(4,4) = 1$, $P(5,4) = 5$, $P(6,4) = 3$ and $P(7,4) = 7$. Thus $V_{7,4}(G^*) = 3/7$. This means that within the set of sequences with recall 3, player 1 cannot do better than 3/7 with a sequence x such that $\mathbf{per}(x) \leq 7$. With recall 3, player 1 can play 8-periodic sequences but, up to circular permutations, there is only one such sequence which is the de Bruijn sequence,

$$BBBTBTTT\ BBBTBTTT\ \ldots.$$

But then, player 2 with recall 2 may play the 4-periodic sequence $LLRR\ LLRR\ldots$. The payoff is here $2/8 = 1/4 < 3/7$.

(b) By Bertrand's postulate, for each k there exists a prime number p such that $2^k < p < 2^{k+1}$. By Proposition 4.3, there exists a p-periodic sequence of T's and B's that corresponds to a strategy with recall $k+1$. This defines x in $M_{k+1}(\mathcal{A})$. As p is prime, the best sequence that player 2 may choose among $S_{2^k}(\mathcal{B})$ and thus among $M_k(\mathcal{B})$ is a constant sequence.

Let $T(x)$ and $B(x)$ be the respective numbers of T's and B's in a cycle of x, so that

$$\mu_x = \left(\frac{T(x)}{p}, \frac{B(x)}{p}\right).$$

The best payoff that player 2 can get is

$$\min\left(\frac{T(x)}{p}, \frac{B(x)}{p}\right),$$

and we just need to check that it is close to 1/2 when k is large. We assume w.l.o.g. $T(x) \geq B(x)$ and evaluate $B(x)$.

For each $i \geq k+2$, denote by $u_i = (x_{i-1}, \ldots, x_{i-(k+1)})$ the recall of player 1 before stage i, and denote by $B(u_i) = |\{j \in \{i-(k+1), \ldots, i-1\}, x_j = B\}|$ the number of B's appearing in u_i. The sequence $(u_i)_{i \geq k+2}$ is periodic with period p, and

$$\frac{1}{p}B(x) = \frac{1}{p}\left|\{i \in \{1, \ldots, p\}, x_i = B\}\right| = \frac{1}{p}\left(\sum_{i=k+2}^{k+1+p} \frac{1}{k+1} B(u_i)\right).$$

The point is that $u_{k+2}, u_{k+3}, \ldots, u_{p+k+1}$ are *distinct* elements of $\{T, B\}^{k+1}$, and $p > 2^k$, so more than half of the words in $\{T, B\}^{k+1}$ appear in this average.

• Assume first k even: $k = 2a$, with a in $\mathbb{N}$. Then half of the words in $\{T, B\}^{k+1}$ contain more T's than B's, and we get a lower bound by selecting the p elements with fewer B's. Better, we consider even less elements by taking average over the $2^k = 2^{2a}$ elements with less B's than T's. So

$$\frac{B(x)}{p} > \frac{1}{2^{2a}} \sum_{\ell=0}^{a} \frac{\ell}{2a+1} \binom{2a+1}{\ell} =: F(a).$$

Since,

$$\begin{aligned}\sum_{\ell=0}^{a} \ell \binom{2a+1}{\ell} &= (2a+1) \sum_{l=1}^{a} \frac{(2a)!}{(\ell-1)!\,(2a+1-\ell)!} \\ &= (2a+1) \sum_{\ell=1}^{a} \binom{2a}{\ell-1} \\ &= (2a+1) \sum_{\ell=0}^{a-1} \binom{2a}{\ell} \\ &= (2a+1) \left(2^{2a-1} - \frac{1}{2}\binom{2a}{a}\right),\end{aligned}$$

then

$$F(a) = 1/2 - \frac{1}{2^{2a+1}} \binom{2a}{a}.$$

So for k even,

$$\frac{B(x)}{p} > F(a) = 1/2 - \frac{1}{2^{2a+1}} \binom{2a}{a}.$$

• Assume now that $k = 2a+1$ is odd. Proceeding the same way,

$$\frac{B(x)}{p} > \frac{1}{2^{2a+1}(2a+2)} \left(\sum_{\ell=0}^{a} \ell \binom{2a+2}{\ell} + 1/2(a+1) \binom{2a+2}{a+1} \right),$$

and the right-hand side of this inequality is nothing but

$$\frac{1}{2} + \frac{1}{2^{2a+3}} \binom{2a+2}{a+1} - \frac{1}{2^{2a+1}} \binom{2a+1}{a+1}.$$

Hence,

$$\frac{B(x)}{p} \geq \frac{1}{2} - \frac{1}{2^{2a+1}} \binom{2a+1}{a+1}.$$

The proof is completed by noticing that both

$$\frac{1}{2^{2a}} \binom{2a}{a} \quad \text{and} \quad \frac{1}{2^{2a+1}} \binom{2a+1}{a+1}$$

go to zero as a goes to infinity. □

4.2 Same recall, more actions

To conclude the paper, we present an example showing that, in games with bounded recall, the complexity of the player may not be conveniently measured by the size of her recall. Consider the following game G^{**}. It is a variation of matching pennies where each action of player 1 is duplicated.

	L	R
T_1	1	0
T_2	1	0
B_1	0	1
B_2	0	1

Proposition 4.6. *$W_{k,k}(G^{**}) = 1/2$ for each $k \geq 1$.*

Proof of Proposition 4.6. With recall k, player 1 can play the following 2^{k+1}-periodic sequence: first play a de Bruijn sequence on the alphabet $\{T_1, T_2\}$ (of length 2^k) followed by a de Bruijn sequence on the alphabet $\{B_1, B_2\}$. With recall k, player 2 cannot produce a period greater than 2^k, and as the best reply has a period that divides 2^{k+1}, player 2 cannot get more than 1/2. □

Proposition 4.6 suggests that the actual power of player 1 with recall k depends on her number of actions.

Acknowledgment

The work of Marco Scarsini was partially supported by MIUR-COFIN. This paper was written while he was visiting CEREMADE, whose nice hospitality is gratefully acknowledged.

References

Abreu, D. and Rubinstein, A. (1988) The structure of Nash equilibrium in repeated games with finite automata. *Econometrica* **56**, 1259–1281.

Bavly, G. and Neyman, A. (2005) Online concealed correlation by boundedly rational players. Technical report, Center for the Study of Rationality, The Hebrew University of Jerusalem.

Ben-Porath, E. (1990) The complexity of computing a best response automaton in repeated games with mixed strategies. *Games Econom. Behav.* **2**, 1–12.

Ben-Porath, E. (1993) Repeated games with finite automata. *J. Econom. Theory* **59**, 17–32.

Bollobás, B. (1998) *Modern Graph Theory.* Springer-Verlag, New York.

de Bruijn, N. G. (1946) A combinatorial problem. *Nederl. Akad. Wetensch., Proc.* **49**, 758–764 = *Indagationes Math.* **8**, 461–467 (1946).

CHALLET, D. and MARSILI, M. (2000) Relevance of memory in minority games. *Phys. Rev. E* **62**, 1862–1868.

GOSSNER, O. and HERNÁNDEZ, P. (2003) On the complexity of coordination. *Math. Oper. Res.* **28**, 127–140.

GOSSNER, O. and HERNÁNDEZ, P. (2006) Coordination through De Bruijn sequences. *Oper. Res. Lett.* **34**, 17–21.

GOSSNER, O., HERNÁNDEZ, P., and NEYMAN, A. (2003) Online matching pennies. Technical report, Center for the Study of Rationality, The Hebrew University of Jerusalem.

JONES, G. A. and JONES, J. M. (1998) *Elementary Number Theory.* Springer-Verlag London Ltd., London.

KALAI, E. and STANFORD, W. (1988) Finite rationality and interpersonal complexity in repeated games. *Econometrica* **56**, 397–410.

LEHRER, E. (1988) Repeated games with stationary bounded recall strategies. *J. Econom. Theory* **46**, 130–144.

LEHRER, E. (1994) Finitely many players with bounded recall in infinitely repeated games. *Games Econom. Behav.* **7**, 390–405.

LEMPEL, A. (1971) m-ary closed sequences. *J. Combinatorial Theory Ser. A* **10**, 253–258.

LIAW, S.-S. and LIU, C. (2005) The quasi-periodic time sequence of the population in minority game. *Phys. A* **351**, 571–579.

NAGELL, T. (1964) *Introduction to Number Theory.* Second edition. Chelsea Publishing Co., New York.

NEYMAN, A. (1985) Bounded complexity justifies cooperation in the finitely repeated prisoners' dilemma. *Econom. Lett.* **19**, 227–229.

NEYMAN, A. (1998) Finitely repeated games with finite automata. *Math. Oper. Res.* **23**, 513–552.

NEYMAN, A. and OKADA, D. (1999) Strategic entropy and complexity in repeated games. *Games Econom. Behav.* **29**, 191–223. Learning in games: a symposium in honor of David Blackwell.

NEYMAN, A. and OKADA, D. (2000a) Repeated games with bounded entropy. *Games Econom. Behav.* **30**, 228–247.

Neyman, A. and Okada, D. (2000b) Two-person repeated games with finite automata. *Internat. J. Game Theory* **29**, 309–325.

Piccione, M. and Rubinstein, A. (2003) Modeling the economic interaction of agents with diverse abilities to recognize equilibrium patterns. *J. European Econom. Assoc.* **1**, 212–223.

Renault, J., Scarsini, M., and Tomala, T. (2006) A minority game with bounded recall. Technical report, CEREMADE, Université de Paris Dauphine.

Rubinstein, A. (1986) Finite automata play the repeated prisoner's dilemma. *J. Econom. Theory* **39**, 83–96.

Sabourian, H. (1998) Repeated games with M-period bounded memory (pure strategies). *J. Math. Econom.* **30**, 1–35.

Yoeli, M. (1962) Binary ring sequences. *Amer. Math. Monthly* **69**, 852–855.

March 14, 2007