האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

THE STRATEGIC VALUE OF RECALL

by

RON PERETZ

Discussion Paper # 470 November 2007

מרכז לחקר הרציונליות

CENTER FOR THE STUDY
OF RATIONALITY

Feldman Building, Givat-Ram, 91904 Jerusalem, Israel
PHONE: [972]-2-6584135 FAX: [972]-2-6513681
E-MAIL: ratio@math.huji.ac.il
URL: http://www.ratio.huji.ac.il/

The Strategic Value of Recall

Ron Peretz*

November 27, 2007

Abstract

This work studies the value of two-person zero-sum repeated games in which at least one of the players is restricted to (mixtures of) bounded recall strategies. A (pure) k-recall strategy is a strategy that relies only on the last k periods of history. This work improves previous results [4, 7] on repeated games with bounded recall. We provide an explicit formula for the asymptotic value of the repeated game as a function of the stage game, the duration of the repeated game, and the recall of the agents.

1 Introduction and Examples

Bounded recall is one of the alternatives proposed by Aumann [1] to model limited rationality in repeated games. Lehrer [4] studied infinitely repeated two-player zero-sum games where both players have bounded recall. Neyman and Okada [6, 7] study a setting in which one player is bounded while the other is fully rational. In [7] they examine specifically the case of bounded recall. The current work extends results of both [4] and [7].

Our main result is the following:[1]

Theorem 1.1. *Let $G =< I, J, g >$ be a two-player zero-sum game in strategic normal form. For every sequence of positive integers $\{T_k\}_{k=1}^{\infty}$ and every $h \neq$*

*Center of Rationality, the Hebrew University of Jerusalem. Research supported in part by Israel Science Foundation grant 263/03. This work is based on the author's Master's thesis (2005). The author would like to thank his advisor Prof. Abraham Neyman for his close guidance and endless patience.
Email: ronprtz@math.huji.ac.il.

[1]See theorem 1.1: the former result (regarding G^{T_k}) extends [7]. The later extends [4].

$\log_2 I$, *if* $\lim_{k\to\infty} \frac{\log_2 T_k}{k} = h$, *then*[2]

$$\begin{aligned}\lim_{k\to\infty} \mathrm{val} G^{T_k}[k,\infty] &= \\ \lim_{k\to\infty} \mathrm{val} G^{\infty}[k,T_k] &= \max_{\substack{\sigma\in\Delta(I):\\ H(\sigma)\geq h,\text{or}\\ H(\sigma)=0}} \min_{\tau\in J} G(\sigma,\tau).\end{aligned}$$

1.1 Examples

Let us denote

$$\begin{aligned} V_* &:= \max_{i\in I}\min_{j\in J} G(i,j) \\ \tilde{\nu}(h) &:= \max_{\substack{\sigma\in\Delta(I):\\ H(\sigma)\geq h}} \min_{\tau\in J} G(\sigma,\tau) \quad = \min_{\tau\in\Delta(J)} \max_{\substack{\sigma\in\Delta(I):\\ H(\sigma)\geq h}} G(\sigma,\tau) \\ \nu(h) &:= \max_{\substack{\sigma\in\Delta(I):\\ H(\sigma)\geq h,\text{or}\\ H(\sigma)=0}} \min_{\tau\in J} G(\sigma,\tau) = \tilde{\nu}(h)\vee V_*. \end{aligned}$$

Consider the "matching pennies" game described in Figure 1. Since the optimal strategy in this game is $(\frac{1}{2},\frac{1}{2})$, which is also the one with maximal entropy, the theorem says, roughly, that if $\frac{\log_2 T_k}{k} < 1$, then the value of the repeated game (in either one of the settings), ν, is "equal" to the value of the one-stage game.

The function $\nu(h)$ is continuous at every point except maybe $h = \log_2(I)$. In the above example, ν is *not* continuous at that suspicious point $h = 1$. It can be shown that $\lim_{k\to\infty} \frac{\log_2 T_k}{k} = h$ implies the convergence of the value of the repeated games if and only if ν is continuous at h.

[2] $G^{\alpha}[k,m]$ is the α-fold repeated version of G where player one is restricted to k-recall strategies and player two is restricted to m-recall strategies.

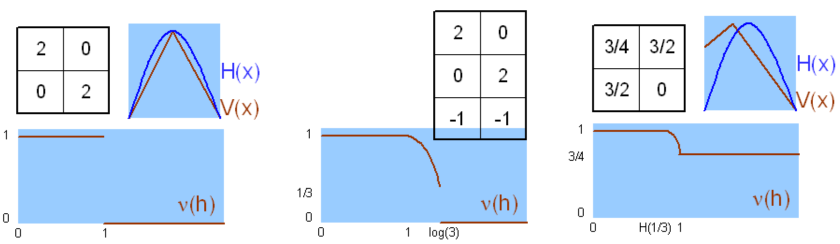

Figure 1: Examples (left to right): "matching pennies," "matching pennies+," and a game with a continuous ν.

The third example in Figure 1 is a game where ν *is* continuous at the suspicious point $h = 1$. This is because one of the pure strategies ensures the same payoff as the strategy with maximal entropy $(\frac{1}{2}, \frac{1}{2})$.

Finally, let us look at the "matching pennies+" game. The third alternative of player one is strongly dominated in the one-step game. Nevertheless, in the repeated game, player one can gain from playing the myopic inferior third alternative occasionally. An intuitive explanation is that by playing the third alternative rather than the first or second, player one can encode information about the history beyond her recall, and it turns out that memory is valuable in repeated games with bounded complexity.

2 Preliminaries

2.1 The games $G^{\alpha}[k, m]$

Let $G =< I, J, g >$ be a two-person zero-sum game, and let k and T be positive integers.

Definition 2.1. A *history* in the repeated version of G is a finite sequence of action profiles. We define the set of all histories, $\mathfrak{H}$, by

$$\mathfrak{H} = \bigcup_{n=0}^{\infty} (I \times J)^n$$

Definition 2.2. A pure strategy σ of player one (resp. two) in the repeated version of G is a function $\sigma : \mathfrak{H} \to I$ (resp. J).

Definition 2.3. A k-recall strategy is a strategy that relies on the last k elements of a history. That is, σ is k-recall iff for every $n > k$ and every $h \in (I \times J)^n$

$$\sigma(h) = \sigma(h_{n-k+1}, \ldots, h_n).$$

Note that the set of k-recall strategies is finite.

One may think of a repeated game as a game in extended (tree) form. A strategy in the repeated game induces strategies at every sub-game corresponding to a history. Since the sub-games of a repeated game are isomorphic to it, we may state the following:

Definition 2.4. Let τ be a (pure or behavior) strategy in a repeated game. Let h be a history of actions in that game. The induced strategy $\tau_{|h}$ – "τ given h" is defined by

$$\tau_{|h}(g) = \tau(hg)$$

Note that if τ is k-recall, so is $\tau_{|h}$.

Definition 2.5. The *play* induced by a strategy σ for player one and a strategy τ for player two is an infinite sequence of action profiles $a_1, a_2, \ldots$, defined recursively by

- $a_1 = (\sigma(\emptyset), \tau(\emptyset))$;
- $a_{n+1} = (\sigma(a_1, \ldots, a_n), \tau(a_1, \ldots, a_n))$.

Definition 2.6. The game $G^T(k, \infty)$ is a two-person zero-sum game in which

- the set of strategies for player one is the set of k-recall strategies;
- the set of strategies for player two is the set of all strategies in the repeated version of G;
- the payoff function is a function of the play induced by the strategies of the players,

$$\frac{1}{T} \sum_{n=1}^{T} g(a_n)$$

Definition 2.7. The game $G^\infty(k, T)$ is a two-person zero-sum game in which

- the set of strategies for player one is the set of k-recall strategies;
- the set of strategies for player two is the set of T-recall strategies;
- the payoff function is a function of the play induced by the strategies of the players,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} g(a_n)$$

Note that the above limit exists since the play in this setting enters a loop at some stage.

2.2 Oblivious strategies and sequences

A (pure) oblivious strategy in a repeated game is a strategy that ignores the history of actions of the other players. It can, therefore, be viewed as a sequence of actions. In this work we focus on k-recall obilvious strategies.

Definition. Given a finite alphabet A and a positive integer k, we define the set of k-recall sequences $B(k, A) \subset A^{\mathbb{Z}}$ by

$$B(k, A) = \left\{(a_t) \in A^{\mathbb{Z}}{:}\forall t, t' \ (a_{t-k}, \ldots, a_{t-1}) = (a_{t'-k}, \ldots, a_{t'-1}) \to a_t = a_{t'}\right\}$$

Since the set A^k is finite, every k-recall sequence must have a period $\leq |A|^k$. A special case is the sequences with period exactly $|A|^k$.

Definition. We define the set of de Bruijn sequences of order k over the alphabet A, $DB(k, A)$ by

$$\begin{aligned} DB(k,A) = \{(a_t) \in B(k,A) &: \\ &\forall (b_1,\ldots,b_k) \in A^k \ \exists t \in \mathbb{Z} \text{ s.t. } (a_{t+1},\ldots,a_{t+k}) = (b_1,\ldots,b_k)\} \end{aligned}$$

Note that we define de Bruijn sequences to be periodic ($\mathbb{Z}$-indexed). Not only do de Bruijn sequences exist, but there are lots of them. The following result is due to N. G. de Bruijn [3]:

Theorem 2.8. $|DB(k,A)| = |A|!^{|A|^{k-1}}$

2.3 Entropy

For completeness we provide a few standard notions in information theory. The reader may refer to [2] for further study.

Definition. Let $p = (p_1, \ldots, p_n) \in \Delta(n)$. The entropy of p, $H(p)$ is defined by

$$H(p) = -\sum_{i=1}^{n} p_i \log_2(p_i)$$

Definition. Let $p = (p_1, \ldots, p_n), q = (q_1, \ldots, q_n) \in \Delta(n)$. The Kullback-Leibler divergence of q from p, $D(p||q)$ is defined by

$$D(p||q) = \sum_{i=1}^{n} p_i \log_2 \frac{p_i}{q_i}$$

Definition. Let X be a random variable that assumes values in a finite set I. We define the entropy of X, $H(X)$ to be the entropy of its distribution mass. That is,

$$H(X) = -\sum_{i \in I} \mathbf{P}(X = i) \log_2(\mathbf{P}(X = i))$$

Definition. Let X and Y be random variables that assume values in finite sets I and J respectively. (X, Y) is a $I \times J$-valued random variable. We define the entropy of X conditioned on Y, $H(X|Y)$ by:

$$H(X|Y) = H(X,Y) - H(Y)$$

2.4 The method of types

We look at finite sequences over a finite alphabet A. We think of the alphabet A as the vertices of the simplex $\Delta(A)$ and so we can refer to the average of the elements of the sequence.

Definition. Let $s = s_1, \ldots, s_l$ be a sequence of elements of an alphabet A. The *empirical distribution* of s is the following quantity:

$$\text{emp}(s) = \frac{1}{l} \sum_{i=1}^{l} s_i$$

Conversely, given a distribution $q \in \Delta(A)$ we can look at the set of sequences with empirical distribution q.

Definition. Given an alphabet A, a positive integer l, and a probability measure $q \in \Delta(A)$, we define $T^q(l)$ by

$$T^q(l) = \left\{ s \in A^l : \text{emp}(s) = q \right\}$$

Proposition 2.9. *If $T^q(l) \neq \emptyset$ then*

$$\frac{2^{H(q)l}}{l^{|A|}} \leq |T^q(l)| \leq 2^{H(q)l}$$

Proposition 2.10 (large deviation). *Let $x = x_1, \ldots, x_l$ be a sequence of i.i.d. random variables with common distribution $q \in \Delta(A)$. Let $p \in \Delta(A)$, $T^p(l) \neq \emptyset$; then*

$$\frac{2^{-D(p||q)l}}{l^{|A|}} \leq \mathbf{P}(\text{emp}(x) = p) \leq 2^{-D(p||q)l}$$

Proofs. See [2] □

2.5 Approximated iids

The next lemma and the following corollary provide a simple criterion to determine whether a given oblivious strategy secures the value of its empirical distribution.

Lemma 2.11 (Neyman-Okada [6]). *Let $p \in \Delta(I)$. For every $\epsilon > 0$ there exists a $\delta > 0$ such that if (x, y) is a $(I \times J)$-valued random variable satisfying (i) $H(x|y) > H(p) - \delta$, and (ii) $\|x - p\| < \delta$, then $\mathbf{E}[g(x, y)] \geq \min_{j \in J} G(p, j) - \epsilon$.*

Corollary 2.12. *Let $p \in \Delta(I)$. Let $\{T_k\}_{k=1}^{\infty}$ be a sequence of positive integers, and let x^k be I^{T_k}-valued random variables satisfying (i)* $\liminf_k \frac{1}{T_k} H(x^k) \geq H(p)$*, and (ii)* $\mathbf{E}[\text{emp}(x^k)] \to_{k\to\infty} p$*. Then, for every sequence of strategies τ^k of player two in the games G^{T_k}, we have* $\liminf G^{T_k}(x^k, \tau^k) \geq \min_{j\in J} G(p, j)$*.*

Proof. Given a pure strategy τ^k of player two we denote by (x, y) the $(I \times J)$-valued random variable (x_t^k, τ_t^k), where t is a random variable that assumes values in $\{1, \ldots, T_k\}$ uniformly and independently of x^k. $G^{T_k}(x^k, \tau^k) = G(x, y)$. (x, y) surely satisfy condition (ii) of Lemma 2.11 and so it remains to verify condition (i) of the lemma. Since conditional entropy is a concave function, and y_n^k is a function of $x_1^k, \ldots, x_{n-1}^k$, we have

$$H(x|y) \geq \frac{1}{T_k} \sum_{n=1}^{T_k} H(x_n^k | y_n^k) \geq \frac{1}{T_k} \sum_{n=1}^{T_k} H(x_n^k | x_1^k, \ldots, x_{n-1}^k) = \frac{1}{T_k} H(x^k)$$

and hence

$$\liminf_k H(x|y) \geq H(p)$$

□

3 A Proof for Theorem 1.1 (Part I)

In this section we prove most of Theorem 1.1. Namely, the part that refers to the game $G^{T_k}[k, \infty]$, i.e., the finitely repeated game in which player one is restricted to (mixtures of) k-recall strategies and player two is fully rational. The second part, which refers to $G^{\infty}[k, m_k]$, will be proved in Section 4.

Henceforth let $G =< I, J, g >$ be a two-person zero-sum game.

3.1 The function ν

Define a set function $K : [0, log_2(I)] \Rightarrow \Delta(I)$ by

$$K(h) = \{p \in \Delta(I) : h \leq H(p)\}$$

The entropy function is concave and continuous; therefore $K(h)$ is convex and compact. By von Neumann's theorem the game in which player one is restricted to strategies in $K(h)$ has a value

$$\tilde{\nu}(h) = \max_{\sigma\in K(h)} \min_{j\in J} \sum_{i\in I} \sigma(i) g(i, j) = \min_{\tau\in\Delta(J)} \max_{\sigma\in K(h)} \sum_{i\in I, j\in J} \sigma(i)\tau(j)g(i, j)$$

Recall that

$$V_* = \max_{i \in I} \min_{j \in J} g(i,j)$$

$$\nu(h) = \begin{cases} \tilde{\nu}(h) \vee V_* & 0 \le h \le \log_2 I \\ V_* & h > \log_2 I \end{cases}$$

Proposition 3.1. *ν is non-increasing and continuous at any point other than $h = \log_2(I)$, where it is continuous from the left.*

Proof. The entropy function is continuous; therefore the graph of K, which is the area below the graph of the entropy function, is closed. Hence K is upper semi-continuous. The entropy function is concave; therefore the graph of K is convex. Hence K is lower semi-continuous. $\Delta(I)$ is compact; therefore K is continuous with respect to the Hausdorff metric.

The function $p \mapsto \min_{j \in J} \sum_{i \in I} p(i)g(i,j)$ is continuous in $\Delta(I)$; therefore taking its maximum is continuous with respect to the Hausdorff metric. Composing the two functions gives $\tilde{\nu}$. This shows that ν is continuous in $[0, \log_2(I)]$. At $h > \log_2(I)$, $\nu(h)$ is constant.

The fact that ν is non-increasing follows from the fact that $\tilde{\nu}$ is non-increasing, which follows from the fact that K is monotone. □

Given Proposition 3.1, we can conclude this part of the proof with the following propositions:

Proposition 3.2. *For every $p \in \Delta(I)$, if $limsup_{k \to \infty} \frac{\log_2 T_k}{k} < H(p)$, then* $\liminf \mathrm{val} G^{T_k}[k, \infty] \ge \min_{j \in J} G(p, j)$.

Proposition 3.3. *For every $h \in \mathbb{R}_+$, if $liminf_{k \to \infty} \frac{\log_2 T_k}{k} > h$, then* $\limsup \mathrm{val} G^{T_k}[k, \infty] \le \nu(h)$.

In the next two sections we will provide proofs of Propositions 3.2 and 3.3.

3.2 $\liminf \mathrm{val} G^{T_k}[k, \infty]$ is at least $\nu(h)$

The proof of Proposition 3.2 relies on the richness of k-recall sequences among all sequences of length T_k. For every $\epsilon > 0$ consider the set of all k-recall sequences of length T_k with empirical distribution within a distance of ϵ from p. Let us denote this set $C_k = C_k(\epsilon)$.

Proposition 3.4. *Under the conditions of Proposition 3.2 $\forall \epsilon > 0$ $\exists K \in \mathbb{N}$ such that $\forall k > K$*

$$|C_k(\epsilon)| > 2^{(H(p)-\epsilon)T_k}.$$

Proof. The richness of k-recall sequences is treated in Section A of the Appendix. See Proposition A.2. □

Proof of Proposition 3.2. For every k let the mixed strategy of player one σ_k be the uniform distribution over the set C_k provided by Proposition 3.4. The sequences in C_k are implementable by oblivious k-recall strategies, and, by Neyman-Okada's Criterion (Lemma 2.12), the expected payoff is asymptotically large enough. □

Alternatively, one can consider a slightly different (perhaps more natural) distribution as a strategy for player one: the probability mass of T_k independent p trials restricted to sequences that are implementable by oblivious k-recall strategies. Verifying that this strategy "works" can be done by either Neyman-Okada's Criterion or large deviation considerations (e.g., Theorem 1 in [5]).

3.3 $\limsup \mathrm{val} G^{T_k}[k,\infty]$ is at most $\nu(h)$

Given $0 \leq h$ and an integer k we will construct a mixed strategy $\tau = \tau^k$ for player two in the games $G^\infty[k,\infty]$. Then we will prove that for any sequence of positive integers $\{T_k\}_{k=1}^\infty$ satisfying $\liminf_{k\to\infty} \frac{\log_2 T_k}{k} > h$, τ^k ensures that the $\limsup$ of the mean payoff in the first T_k steps is at most $\nu(h)$.

The strategy τ will be a mixture of pure strategies of the form τ_x, where $x = x_1, x_2, \ldots$ is an infinite sequence of actions of player two. We will first describe τ_x as a function of x and discuss its properties (for a special class of sequences). Then we will define τ by providing a probability measure over the possible values for x. The last part of the proof is the analysis of the payoff that τ secures.

We begin by defining a mapping s from the set of all finite histories to the set $\{0,\ldots,k\}$ (Figure 2). The set $\{0,\ldots,k\}$ can be interpreted in this context as the "states of mind" of the strategy τ. Note that the mapping s

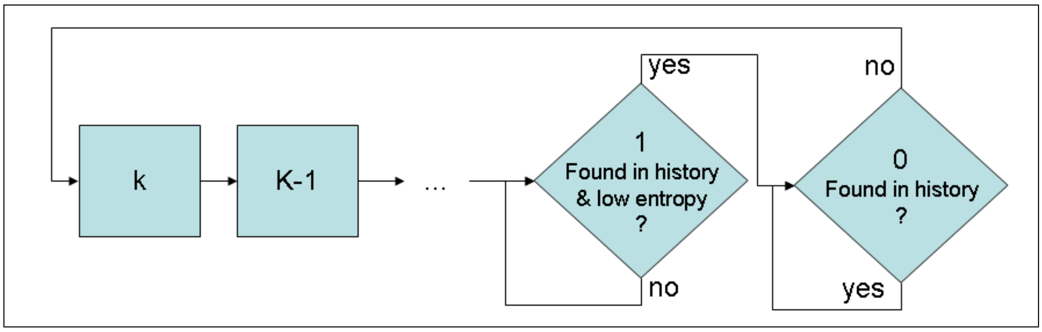

Figure 2: An illustration of the function s as a flow chart.

does not depend on the sequence x (it does, however, depend on k).

$$s(\emptyset) = k$$

$$s(a_1, \ldots, a_{t+1}) =$$

$$\begin{cases} s(a_1,\ldots,a_t)-1 & \text{if } s(a_1,\ldots,a_t) > 1, \\ 1 & \text{if } s(a_1,\ldots,a_t) = 1, \text{ and} \\ & \forall s < t \; ((a_{s-k+2},\ldots,a_{s+1}) \neq (a_{t-k+2},\ldots,a_{t+1}), \\ & \quad \text{or } H(\text{emp}(a^1_{t-k+2},\ldots,a^1_{t+1})) \geq h) \\ 0 & \text{if } s(a_1,\ldots,a_t) = 1, \text{ and} \\ & \exists s < t \text{ s.t. } ((a_{s-k+2},\ldots,a_{s+1}) = (a_{t-k+2},\ldots,a_{t+1}), \\ & \quad \text{and } H(\text{emp}(a^1_{t-k+2},\ldots,a^1_{t+1})) < h) \\ & \text{or} \\ & \text{if } s(a_1,\ldots,a_t) = 0, \text{ and} \\ & \exists s < t \text{ s.t. } (a_{s-k+2},\ldots,a_{s+1}) = (a_{t-k+2},\ldots,a_{t+1}) \\ k & \text{otherwise.} \end{cases}$$

In words: the initial state is k. States 2 to k simply count down to 1. State 1 looks in the history for an occurrence of the past k actions as long as the actions of player one have *low* entropy. If found, it goes to state 0; otherwise it stays in state 1. State 0 looks in the history for an occurrence of the past k actions. If found, it stays in state 0; otherwise it jumps back to state k.

Given a play $a_1, a_2, \ldots$ we denote $s_t = s(a_1, \ldots, a_{t-1})$. Note that

- $s_{l+1}, \ldots, s_{l+k} \neq 0$, iff $s_{l+k} = 1$;
- if player one is limited to k-recall, then whenever $s_t = 0$ player two can predict the next action of player one, a^1_t.

Now we are ready to define τ_x. For every finite history $h = a_1, \ldots, a_t$ ($t = 0$ indicating the empty history) define

$$\tau_x^k(h) = \begin{cases} \gamma((a_{t+1}^1)) & \text{if } s(h) = 0, \\ x_{t+1} & \text{otherwise,} \end{cases}$$

where $\gamma : I \to J$ is such that $\forall i \in I\ g(i, \gamma(i)) \leq V_*$.

We shall now define τ by specifying a random choice of x. Choose a strategy in the one-shot game $q \in \Delta(J)$ such that for every $p \in K(h)$ $\sum_{i \in I, j \in J} p(i)q(j)g(i,j) \leq \tilde{\nu}(h)$. Define $x = x_1, \ldots, x_{k^2}, x_1, \ldots$ to be a k^2-periodic sequence of i.i.d. random variables with common distribution q.

For every strategy τ_x in the support of τ (defined by a k^2-periodic sequence x) and every k-recall strategy for player one σ, consider the induced play $a_1, a_2, \ldots$ and the sets

$$\begin{aligned} L^{(i)} = \{l|\ &0 \leq l \leq T_k,\ H(\mathrm{emp}(a_{l+1}^1, \ldots, a_{l+k}^1)) < h, \\ &s_{l+k} = 1, s_{l+k+1} = \ldots = s_{l+k+i} = 0, s_{l+k+i+1} \neq 0\} \\ L = \cup_{i=0}^{\infty}&L^{(i)} \\ A = \{(&a_{l+1}, \ldots, a_{l+k})|l \in L\} \end{aligned}$$

Form the definition of s we have for every $i > 0$

$$L^{(i)} = \{l|s_{l+k} \neq 0, s_{l+k+1} = \ldots = s_{l+k+i} = 0, s_{l+k+i+1} \neq 0\}. \tag{3.1}$$

Analogously we define

$$\begin{aligned} N^{(i)} &= \{n|s_{n-i-1} \neq 0, s_{n-i} = \ldots = s_{n-1} = 0, s_n \neq 0\} \\ &= \{l + k + i + 1|\ l \in L^{(i)}\}, \\ N &= \cup_{i=0}^{\infty} N^{(i)}. \end{aligned}$$

Every time a k-tuple in A recurs along the play it is followed by a sequence in which $s_i = 0$ (the predictive state) that is longer than the previous occurrence; therefore the map $l \mapsto (a_{l+1}, \ldots, a_{l+k})$ is injective when restricted to each $L^{(i)}$. Hence

$$\left|L^{(i)}\right| = \left|N^{(i)}\right| \leq |A| \leq k^2 2^{hk}$$

The latter inequality follows by bounding by 2^{hk} the number of sequences of length k of actions of player one with empirical distribution $< h$, and by k^2 the number of possible reactions of player two. Hence $\left|\cup_{i<k^3} L^{(i)}\right| \leq k^5 2^{hk}$. By (3.1), $\left|\cup_{i \geq k^3} L^{(i)}\right| \leq T_k/k^3$. Consequently, for large k

$$|L \cup N| < T_k/k^{2.99}$$

Hence, for almost every $t \in \{0, \ldots, T_k\}$, none of the numbers $t-k, t-k+1, \ldots, t+k^2$ is in $L \cup N$; therefore, for almost every t, either

(i) $s_{t+1} = \ldots = s_{t+k^2} = 0$, or

(ii) $s_{t+1}, \ldots, s_{t+k^2} \neq 0$, and $\forall 0 \leq l < k^2$ $H(\mathrm{emp}(\mathrm{a_{t+l+1}}, \ldots, \mathrm{a_{t+l+k}})) \geq h$.

Proposition 3.5. $\limsup_{k \to \infty} \max_{\sigma \text{ with } k\text{-recall}} G^{T_k}(\sigma, \tau^k) \leq \nu(h)$

Proof. Fix $k > 0$, and an arbitrary k-recall strategy for player one σ. σ and τ^k induce a play $a = a_1, a_2, \ldots$ and a sequence of states $s_t = s(a_1, \ldots, a_{t-1})$. Clearly, a and s are random variables (functions of $x = x_1, \ldots, x_{k^2}$). Let t be another r.v. independent of x, assuming values in $\{0, \ldots, T_k\}$ with uniform distribution. Define a real valued r.v. W by

$$W = \frac{1}{k^2} \sum_{i=1}^{k^2} G(a_{t+i}) \tag{3.2}$$

The expectation of W approximates the actual payoff – $1/T_k \sum_{t=1}^{T_k} G(a_t)$. $|\mathbf{E}[W - G(a_t)]| = \mathrm{O}(\frac{k^2}{T_k})$; therefore the objective is to bound $\mathbf{E}W$ from above.

In case (i) W is at most V_*. If the probability of case (ii) does not converge to zero (as $k \to \infty$), then for every sub-sequence bounded away from zero the actions of player two $x_{t+1}, \ldots, x_{t+k^2}$ conditioned on (ii) form an approximated i.i.d. sequence and therefore, by lemma 2.11, secure a payoff of $\tilde{\nu}(h)$.

In the case of $h > \log_2 I$, the condition in (ii) cannot be satisfied and hence (i) occurs with probability close to one and V_* is secured. □

4 Refinements, Extensions, and Remarks

4.1 A proof for Theorem 1.1 (Part II)

In this section we prove the remaining part of Theorem 1.1 – the part that refers to the game $G^\infty[k, T_k]$. Our proof for this part follows the track of the proof for the other part with minor adjustments. We begin by strengthening Theorem A.1.

Theorem 4.1 (Theorem A.1, strengthened). *Let $\{T_k\}_{k=1}^\infty$ be a sequence of positive integers, $T_k \to \infty$. Let $x_1, x_2, \ldots$ be a sequence of i.i.d. random variables with common distribution p (over a finite set of values). If $\limsup \frac{\log_2 T_k}{k} < H(p)$, then there exists a sequence $\{T'_k\}_{k=1}^\infty$ such that*

1. *$\forall k$ $T_k \leq T'_k$,*

2. $\mathbf{P}\big(\forall t \neq s \in \mathbb{Z}_{T'_k} \quad (x_{t+1}, \ldots, x_{t+k}) \neq (x_{s+1}, \ldots, x_{s+k})\big) = \exp(-\mathrm{o}(T'_k))$*;*

where the elements of $\mathbb{Z}_T$ *are identified with their representatives* $1, \ldots, T$ *in* $\mathbb{Z}$ *in the obvious way.*

Proof. A close look at the proof of Theorem A.1 (and the construction of the sets C_k in particular) shows that it actually proves this stronger statement. □

The next two propositions are analogous to Propositions 3.2 and 3.3.

Proposition 4.2. *For every* $p \in \Delta(I)$*, if* $\limsup_{k\to\infty} \frac{\log_2 T_k}{k} < H(p)$*, then* $\liminf \mathrm{val} G^\infty[k, T_k] \geq \min_{j\in J} G(p, j)$*.*

Proposition 4.3. *For every* $h \in \mathbb{R}_+$*, if* $\liminf_{k\to\infty} \frac{\log_2 T_k}{k} > h$*, then* $\limsup \mathrm{val} G^\infty[k, T_k] \leq \nu(h)$*.*

Proof of Proposition 4.2 (sketch). Take the integers $T'_k \geq kT_k$ provided by Theorem 4.1. Let the oblivious strategy of player one – $\sigma = \sigma_1, \sigma_2, \ldots$ – consist of a mixture of all the T'_k-periodic sequences in $B(k, I)$ with the distribution induced from the T'_k-fold Cartesian product of p. Theorem 4.1 implies that

$$\liminf \frac{1}{T_k} H(\sigma_1, \ldots, \sigma_{T_k}) \geq H(p), \tag{4.1}$$

$$\lim \mathbf{E}[\mathrm{emp}(\sigma)] = p, \tag{4.2}$$

$$T'_k \gg T_k \tag{4.3}$$

Let τ be a (pure or behavior) T_k-recall strategy for player two. Denote by $i_1, j_1, i_2, j_2, \ldots$ the play induced by σ and τ. For every $t = 1, 2, \ldots$, consider the mean empirical distribution of the joint actions of players one and two during the period $t+1$ till $t + T'_k$:

$$Q = Q_t = \mathbf{E}[\mathrm{emp}(i_{t+l}, j_{t+l})_{l=1}^{T'_k}]$$

The required mean payoff during this period is obtained by Lemma 2.11 and[3]

$$H^Q(i|j) \geq \frac{1}{T'_k}\sum_{l=1}^{T'_k} H(i_{t+l}|j_{t+l}) \quad (4.4)$$

$$\geq \frac{1}{T'_k}\sum_{l=1}^{T'_k} H(i_{t+l}|i_{t+1},\ldots,i_{t+l-1},\tau_{|h_t}) \quad (4.5)$$

$$= \frac{1}{T'_k} H(i_{t+1},\ldots,i_{t+T'_k}|\tau_{|h_t}) \quad (4.6)$$

$$\geq \frac{1}{T'_k}[H(i_{t+1},\ldots,i_{t+T'_k}) - H(\tau_{|h_t})] \geq H(p) - \mathrm{o}(1) \quad (4.7)$$

where h_t denotes the history up to step t, $h_t = i_1, j_1, \ldots i_t, j_t$. □

Proof of Proposition 4.3 (sketch). Let τ^* be an optimal strategy for player two in the game $G^{T_k}(k,\infty)$. To ensure the same payoff in the game $G^{\infty}(k,\infty)$ repeat τ^* every T_k stages. That is, play the strategy τ defined by

$$\tau(a_1,\ldots,a_t) = \tau^*(a_{sT_k+1},\ldots,a_{sT_k+r})$$

where $sT_k + r = t$ and $r < t$.

The problem is that τ is not a T_k-recall strategy since it is a function of the stage number t. Player two can simulate τ by inserting a statistically significant yet short enough ($\ll T_k$) sequence of actions at the end of every period of T_k stages. The technical details are left to the reader. □

4.2 The special point $h = \log_2(I)$

In this section we give an example of a game G for which the fact that $\frac{\log_2 T_k}{k}$ converges does not imply that $\mathrm{val}G^{T_k}[k,\infty]$ converges.

Consider the game M, called the "matching pennies" game, which can be described in strategic normal form by the following matrix:

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

A close look at the proof of Theorem 1.1 shows that the argument that deals with $h > \log_2 I$ is valid also for $T_k \geq k^6 2^k$; therefore, the value of $M^{k^6\cdot 2^k}[k,\infty]$ converges to $V_*(M) = 0$.

Proposition 4.4. $\liminf_{k\to\infty} \mathrm{val}M^{2^k}[k,\infty] \geq \frac{1}{2}$

[3]Recall definition 2.4.

Proof. Let $x = x^k = \left\{x_n^k\right\}_{n\in\mathbb{Z}}$ be a random sequence that takes values in $DB(k,\{0,1\})$ with uniform distribution. Let y and t be random variables: $y \sim x$, $t \sim U\left\{1,\dots,2^k\right\}$, and x, y, and t are independent. Define a random variable $f = h(x_t|x_1 = y_1,\dots,x_{t-1} = y_{t-1})$. Then $\mathbf{E}f = \frac{1}{2^k}H(x) = \frac{1}{2^k}\log_2(|DB(k,\{0,1\})|) = \frac{1}{2}$. Define an event

$$X = \{\text{“}\exists k \le t' < t \text{ such that } (x_{t'-k+1},\dots,x_{t'-1}) = (x_{t-k+1},\dots,x_{t-1})\text{''}\}$$

X is the event that the $k-1$ actions right before time t occurred once more before time t, and hence the action at time t must be other than the previous one. Hence, $f\mid_X = 0$. Since every sequence of $k-1$ actions occurs exactly twice along a de Bruijn cycle, $\mathbf{P}(X) \to \frac{1}{2}$. $\mathbf{1}_{X^c} - f$ is non-negative and its expectation converges to zero; therefore $\forall\epsilon > 0$, $\mathbf{P}(f > 1-\epsilon) \to \frac{1}{2}$; and therefore $\liminf_{k\to\infty} M^{2^k}(x^k,\cdot) \ge \frac{1}{2}$. □

Appendix

A The richness of k-recall sequences

In this section we state and prove the following theorem:

Theorem A.1. *Let $\{T_k\}_{k=1}^{\infty}$ be a sequence of positive integers, $T_k \to \infty$. Let $x_1, x_2, \dots$ be a sequence of i.i.d. random variables with common distribution p (over a finite set of values). If $\limsup \frac{\log_2 T_k}{k} < H(p)$, then*

$$\mathbf{P}\big(\forall 0 \le t \ne s \le T_k \quad (x_{t+1},\dots,x_{t+k}) \ne (x_{s+1},\dots,x_{s+k})\big) = \exp(-\mathrm{o}(T_k)).$$

We begin by reducing Theorem A.1 to the following combinatorial statement:

Proposition A.2. *Let $\{T_k\}_{k=1}^{\infty}$ be a sequence of positive integers, $T_k \to_{k\to\infty} \infty$. Let A be a finite alphabet and $p \in \Delta(A)$. If $\limsup \frac{\log_2 T_k}{k} < H(p)$, then there exist sets $C_k \in A^{T_k+k}$ satisfying*

$$\begin{aligned} &\forall x \in C_k\ \forall 0 \le t < s \le T_k\ (x_{t+1},\dots,x_{t+k}) \ne (x_{s+1},\dots,x_{s+k}),\\ &\lim_k \max_{x\in C_k} \|\mathrm{emp}(x) - p\| = 0,\\ &\liminf_k \frac{\log_2|C_k|}{T_k} \ge H(p) \end{aligned} \tag{A.1}$$

Proposition. *Proposition A.2 implies Theorem A.1.*

Proof. First, since the number of non-empty types $|\{T^q(n) : q \in \Delta(A)\}|$ is polynomial in n, we may assume w.l.o.g. that the elements of C_k all have the same empirical distribution p_k. Take $x = x_1, \ldots, x_{T_k+k}$ i.i.d. random variables with distribution p.

$$\mathbf{P}(x \in C_k) = \mathbf{P}(\mathrm{emp}(x) = p_k) \cdot \frac{|C_k|}{|T^{p_k}(T_k + k)|} \tag{A.2}$$

By Proposition 2.10,

$$\mathbf{P}_{x\sim p^n}(\mathrm{emp}(x) = q) \geq \frac{2^{-D(q||p)n}}{n^{|A|}}; \tag{A.3}$$

and, by Proposition 2.9,

$$|T^q(n)| \leq 2^{H(q)n}. \tag{A.4}$$

Combining (A.1), (A.3), and (A.4) we get the required result. □

In the rest of this section we prove Proposition A.2.

For $k = 1, 2, \ldots$, we shall construct sets C_k satisfying (A.1). We assume w.l.o.g. that $k < T_k$. Let $l = l(k)$ and $m = m(k)$ be integers with the following properties:

$$1 \ll l(k) \ll k \tag{A.5}$$

$$m(k)l(k) + 2l(k) < k \tag{A.6}$$

$$m(k)l(k) \sim k \tag{A.7}$$

Let A, p, and T_k be as given in Proposition A.2. Let $q = q(k) \in \Delta(A)$ such that

$$q(k) \to_{k\to\infty} p,$$
$$\forall k \; T^q(l) \neq \emptyset$$

Consider the alphabet $B = T^q(l)$. For every De Bruijn sequence $\hat{x} \in DB(B, m)$ we shall define a corresponding (infinite) sequence x in $A^{\mathbb{Z}}$. The set C $(= C_k)$ will consist of the elements 1 through $T_k + k$ of such sequences. Formally, $C := \{(x_1, \ldots, x_{T_k+k}) | x \text{ corresponds to some } \hat{x} \in DB(B, m)\}$.

Let α be a least probable element of A with respect to the probability mass q, and let $\beta \neq \alpha$ be another element of A.[4] Let $b = \alpha^l\beta$. That is, b is a word over the alphabet A that consists of l consecutive *alpha*s followed by one *beta*. The correspondence $\hat{x} \mapsto x$ is defined as follows:

$$x = \ldots b\hat{x}_1\hat{x}_2 \ldots \hat{x}_m b\hat{x}_{m+1}\hat{x}_{m+2} \ldots \hat{x}_{m+m}b \ldots$$

[4]We assume w.l.o.g. that A includes more than one element.

That is, x is the concatenation of the elements of $\hat{x}$ separated by a b after every mth element.

The first two lines of (A.1) hold trivially. It remains to verify that the last line of (A.1) holds. Since $DB(B,m)$ is invariant to shifts for every $T < |B|^m$ we have

$$\frac{\log |\{\hat{x}_1 \dots \hat{x}_T | \hat{x} \in DB(B,m)\}|}{T} \geq \frac{\log |DB(B,m)|}{|B|^m} = \frac{log_2(|B|!)}{|B|} \geq$$
$$\geq \log_2 |B| - \log_2 \log_2 |B| \quad \text{(A.8)}$$

The last inequality follows from the inequality of means $\frac{n}{\log_2 n} \leq \frac{n}{1+\frac{1}{2}+\ldots+\frac{1}{n}} \leq \sqrt[n]{n!}$. Let T be an integer, $\frac{T_k}{l(k)} \leq T \leq B^{m(k)}$. Substituting in (A.8) we obtain

$$\frac{\log_2 C_k}{T_k} \geq \frac{\log |\{\hat{x}_1 \dots \hat{x}_T | \hat{x} \in DB(B,m)\}|}{(l(k))T} \geq \frac{\log_2 |B| - \log_2 \log_2 |B|}{l(k)} \to H(p)$$

Finally, we have to verify that such a T exists. It is sufficient to show that $T_k \leq |B|^{m(k)}$, and indeed:

$$\frac{\log_2 |B|^{m(k)}}{k} = \frac{l(k)m(k)}{k} \frac{\log_2 |B|}{l(k)} \to H(p) > \limsup \frac{\log_2 T_k}{k} \quad \square$$

Remark. The assumption of Theorem A.1 that $\limsup \frac{\log_2 T_k}{k} < H(p)$ is necessary.[5] On the other hand, the theorem does not tell us how small $o(T_k)$ is. Finding an explicit expression for (the asymptotic of) that probability is of interest.

References

[1] Aumann R.J. (1981). "Survey of Repeated Games," in *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern*, 11-42. Mannheim: Bibliographisches Institut.

[2] Cover, T.M., and Thomas J.A. (1991). *Elements of Information Theory*, Wiley Interscience, New York.

[3] de Bruijn, N.G. (1946). "A Combinatorial Problem" *Nedel. Akad. Wetensch. Proc.* **49**, 758-764.

[5] By considering only s and t congruent to 1 modulo k, one obtains an instance of the well-studied "birthday" problem.

[4] Lehrer, E. (1994). "Finitely Many Players with Bounded Recall in Infinitely Repeated Games," *Games Econom. Behav.* **7**, 390-405.

[5] Neyman, A (1997). "Cooperation, Repetition and Automata," *Cooperation: Game-Theoretic Approaches*, NATO ASI Series F, 155. S. Hart, A. Mas-Colell, eds. Springer, Berlin, 233-255.

[6] Neyman, A., and Okada, D. (1997). "Repeated Games with Bounded Entropy," *Games Econom. Behav.* **30**, 228-247.

[7] Neyman, A., and Okada, D. (2005). "Growth of Strategy Sets, Entropy, and Nonstationary Bounded Recall," DP 411, Center for Rationality, Hebrew University.