Cooperation, Repetition, and Automata

Abraham Neyman[1]

November, 1995

[1]Neyman: Institute of Mathematics, The Hebrew University of Jerusalem, Givat Ram, Jerusalem 91904, ISRAEL *e-mail: aneyman @ math.huji.ac.il* ⟩

Abstract

This talk studies the implications of bounding the complexity of players' strategies in long term interactions. The complexity of a strategy is measured by the size of the minimal automaton that can implement it.

A finite automaton has a finite number of states and an initial state. It prescribes the action to be taken as a function of the current state and its next state is a function of its current state and the actions of the other players. The size of an automaton is its number of states.

The results study the equilibrium payoffs per stage of the repeated games when players' strategies are restricted to those implementable by automata of bounded size.

The first talk will concentrate maily on the 0-sum case and address the following topics/questions.

1 What is the relation between the bounds of the automata sizes and the quantitative advantage of the player with the larger bound. (Theorems 1 and 3 of the enclosed paper)

2 What is the duration (number of repetition) needed for an unrestricted player to exploit fully his advantage over a player with bound automata (Conjecture 2 of the enclosed paper including a positive solution of its second part).

3 The existence of a deterministic periodic sequence (with period n) which is asymptotically random for every automata of size $o(n/\log n)$ (Proposition 2 of the enclosed paper).

1 Introduction

The simplest strategic game quickly gives rise to a game of formidable complexity when one considers a finitely-repeated version of it.This is because the number of pure strategies in the repeated game grows as a double exponential of the number of repetitions. To just write down in decimal form the number of pure strategies available to a player in the hundred-times repeated Prisoner's dilemma would require more digits than the number of all the letters in all of the books in the world. This chapter examines the implications of restricting the set of strategies to those that are implementable by finite automata of bounded size. Such restrictions place a bound on the complexity of strategies and they can (dramatically) alter the equilibrium play of a repeated game.

When we try to argue that an outcome is or is not an equilibrium in a game there are direct references to all possible strategies in that game. In the case of the hundred-times repeated Prisoner's dilemma it would obviously be an impossible task to merely write out all of the strategies, let alone construct the huge matrix which would constitute the explicit representation of this game in normal (strategic) form. Moreover, many of the strategies in the finitely or infinitely repeated game are extremely complicated. They may involve actions contingent on so many possible past events that it would be nearly impossible to describe them; even writing them down or writing a program to execute some of these strategies is practically impossible. We consider here a theory that limits the strategies available to a player in a repeated game. The restriction is to those strategies that are implementable by bounded size automata- the simplest theoretical model of a computer. It turns out that the equilibria of the resulting strategic game can be dramatically different from those of the original game.

In the finitely repeated Prisoner's dilemma, it is well known that all equilibria, and all correlated equilibria or communication equilibria, result in the repeated play of (defect, defect). This is in striking contrast to the experimental observation that real players do not choose always the dominant action of defecting, but in fact achieve some mode of cooperation.

The present approach justifies cooperation in the finitely repeated Prisoner's dilemma, as well as in other finitely repeated games, without departing from the hypothesis of strict utility maximization, but under the added assumption that there are bounds (possibly very large) on the complexity of the strategies that players may use.

There are other methods of restricting strategies. I am not going to advocate here that the avenue we are taking is superior. Each one of the possible approaches has its pros and cons.

The paper surveys some results about the equilibrium payoffs of repeated games when players' strategies in the repeated game are resricted. It contains also several new results, e.g., Propositions 2, 3, 4, 5, 6, 7 and 8. There is no attempt here to survey all results related to the title, and therefore several important and related papers are not covered in this survey.

2 The Model

2.1 Strategic games

Let G be an n-person game, $G = (N, A, r)$, where $N = \{1, 2, \ldots, n\}$ is the set of players, $A = \times_{i \in N} A_i$,A_i is a finite set of actions for player i, $i = 1, \ldots, n$, and $r = (r^i)_{i \in N}$ where $r^i : A \to I\!R$ is the payoff function of player i. The set A_i is called also the set of pure strategies of player i. We denote by $r : A \to I\!R^N$ the vector valued function whose ith component is r^i, i.e., $r(a) = (r^1(a), \ldots, r^n(a))$. We use also the more detailed description of G, $G = (N; (A_i)_{i \in N}; (r^i)_{i \in N})$, or $G = ((A_i)_{i \in N}; (r^i)_{i \in N})$ for short. For any finite set (or measurable space) B we denote by $\Delta(B)$ the set of all probability distributions on B. For any player i and any n-person game G, we denote by $v^i(G)$ his individual rational payoff in the mixed extension of the game G, i.e., $v^i(G) = \min \max r^i(a^i, \sigma^{-i})$ where the max ranges over all pure strategies of player i, and the min ranges over all $N \backslash \{i\}$-tuples of mixed strategies of the other players, and r^i denotes also the payoff to player i in the mixed extension of the game. We denote by $u^i(G)$ the individual rational payoff of player i in pure strategies, i.e., $u^i(G) = \min \max r^i(a^i, a^{-i})$ where the max ranges over all pure strategies of player i, and the min ranges over all $N \setminus \{i\}$ -tuples of pure strategies of the other players. Obviously $u^i(G) \geq v^i(G)$. We denote by $w^i(G)$ the max min of player i where he maximizes over his mixed strategies and the min is over the pure strategies of the other players, i.e., $w^i(G) = \max_{x \in \Delta(A_i)} \min_{a^{-i} \in A_{-i}} r^i(x, a^{-i})$ where $A_{-i} = \times_{j \neq i} A_j$. Recall that the minimax theorem asserts that for a two person game G, $v^i(G) = w^i(G)$. For any game G in strategic form we denote by $E(G)$ the set of all equilibrium payoffs in the game G, and by $F(G)$ the convex hull of all payoff vectors in the one shot game, i.e., $F(G) = co(r(A))$. Given a 2-person 0-sum game G we denote by $\mathrm{Val}(G)$ the minimax value of G, i.e., $\mathrm{Val}(G) = v^1(G)$.

2.2 The repeated games G^T and G^*.

Given an n-person game, $G = ((A_i)_{i\in N}; (r^i)_{i\in N})$, we define a new game in strategic form $G^T = ((\Sigma^i(T))_{i\in N}; (r^i_T)_{i\in N})$ which models a sequence of T plays of G, called *stages*. After each stage, each player is informed of what the others did at the previous stage, and he remembers what he himself did and what he knew at previous stages. Thus, the information available to each player before choosing his action at stage t is all past actions of the players in previous stages of the game. Formally, let H_t, $t = 1, \ldots, T$, be the cartesian product of A by itself $t-1$ times, i.e., $H_t = A^{t-1}$, with the common set theoretic identification $A^0 = \{\emptyset\}$, and let $H = \cup_{t=1}^T H_t$. A pure strategy σ^i of player i in G^T is a function $\sigma^i : H \to A_i$. Obviously, H is a disjoint union of H_t, $t = 1, \ldots, T$ and therefore one often defines $\sigma^i_t : H_t \to A_i$ as the restriction of σ to H_t. We denote the set of all pure strategies of player i in G^T by $\Sigma^i(T)$. The set of pure strategies of player i in the infinitely repeated game G^* is denoted by Σ^i, i.e., $\Sigma^i = \{\sigma^i : \cup_{t=1}^\infty H_t \to A_i\}$.

Any N-tuple $\sigma = (\sigma^1, \ldots, \sigma^n) \in \times_{i\in N}\Sigma^i(T)$ (Σ^i) of pure strategies in G^T (in G^*) induces a play $\omega(\sigma) = (\omega_1(\sigma), \ldots, \omega_T(\sigma))$ $(\omega(\sigma) = (\omega_1(\sigma), \omega_2(\sigma), \ldots))$ defined by induction: $\omega_1(\sigma) = (\sigma^1(\emptyset), \ldots, \sigma^n(\emptyset)) = \sigma(\emptyset)$ and $\omega_t(\sigma) = \sigma(\omega_1(\sigma), \ldots, \omega_{t-1}(\sigma))$ or in other words $\omega^i_1(\sigma) = \sigma^i(\emptyset)$ and $\omega^i_t(\sigma) = \sigma^i(\omega_1(\sigma), \ldots, \omega_{t-1}(\sigma)) = \sigma^i_t(\omega_1(\sigma), \ldots, \omega_{t-1}(\sigma))$.

Set
$$r_T(\sigma) = \frac{r(\omega_1(\sigma)) + \ldots + r(\omega_T(\sigma))}{T}.$$

We define R_T or R for short to be the function from plays of G^T to the associated payoffs, i.e., $R_T : A^T \to \mathbf{R}^N$ is given by

$$R_T(a_1, \ldots, a_T) = \frac{r(a_1) + r(a_2) + \ldots + r(a_T)}{T}.$$

Two pure strategies σ^i and τ^i of player i in G^T (in G^*) are *equivalent* if for every $N\backslash\{i\}$ tuple of pure strategies $\sigma^{-i} = (\sigma^j)_{j\in N\backslash\{i\}}$, $\omega_t(\sigma^i, \sigma^{-i}) = \omega_t(\tau^i, \sigma^{-i})$ for every $1 \le t \le T$ ($1 \le t$). The equivalence classes of pure strategies are called *reduced strategies*. For $i \in N$ let $A_{-i} = \times_{j\ne i} A_j$. Then an equivalence class of pure strategies is naturally identified with a function $\bar{\sigma}^i : \cup_{t=0}^{\infty}(A_{-i})^t \to A_i$.

2.3 Finite automata

We will consider strategies of the repeated games which are described by means of automata, (which are also sometimes referred to as Moore machines or exact automata). An *automaton* for player i consists of a finite state space M; an initial state $q_1 \in M$; a function f that describes the action to be taken as a function of the different states of the machine, $f : M \to A_i$, where A_i denotes the set of actions of player i; and a transition function g that determines the next state of the machine as a function of its present state and the action of the other players, i.e., $g : M \times A_{-i} \to M$. Thus, an automaton of player i is represented by a 4-tuple $\langle M, q_1, f, g\rangle$. The *size* of an automaton is the number of states.

This machine, the automaton, will change its state in the course of playing a repeated game. At every state $q \in M$, f determines what action it will take. The next state of the automaton is determined by the current state and the action taken by the other players. We can think of such an automaton as playing a repeated game. It starts in its initial state q_1, and plays at the first stage of the game the action assigned by the action function f, $f(q_1)$. Thus, $f(q_1) = a_1^i$ is the action of the player at stage 1. The other players' action at this stage is $b_1 = a_1^{-i} \in A_{-i}$. Thus the history of play before the start of stage 2 is the n-tuple of actions, (a_1^i, a_1^{-i}), played at the first stage of the game . As a function of the present state, and the other players' actions, the machine is transformed into a new state which is given by the transition function g. The new state of the machine is $q_2 = g(q_1, b_1)$. The action that player i takes at stage 2, a_2^i, is described by the function $f : f(q_2) = f(g(q_1, b_1))$, and denoting by a_2^{-i} the action of the other players in stage 2, $(f(q_2), a_2^{-i})$ is the pair of actions played in the second stage of the repeated game, and so on.

What is the state of the machine at stage t of the game? The machine moves to a new state which is a function of the state of the machine in the previous stage and the action played by the other players. Thus $q_t = g(q_{t-1}, a_{t-1}^{-i})$, is the new state of the automaton at stage t, and player i takes at stage t the action $f(q_t) = f(g(q_{t-1}, a_{t-1}^{-i}))$ and so on.

Define inductively,

$$g(q, b_1, \ldots, b_t) = g(g(q, b_1, \ldots, b_{t-1}), b_t),$$

where $b_j \in A_{-i}$. The action prescribed by the automaton for player i at stage t is $f(g(q_1, a_1^{-1}, \ldots, a_{t-1}^{-i}))$ where a_j^{-i}, $1 \leq j < t$, is the $N\backslash\{i\}$ tuple of actions at stage j. Therefore, any automaton $\alpha = \langle M, q_1, f, g \rangle$ of player i induces a strategy σ_α^i in G^T that is given by $\sigma_\alpha^i(\emptyset) = f(q_1)$ and

$$\sigma_\alpha^i(a_1, \ldots, a_{t-1}) = f(g(q_1, a_1^{-i}, \ldots, a_{t-1}^{-i})).$$

Note also that an automaton α of player i induces also a strategy σ_α^i of player i in the infinitely repeated game G^*. A strategy σ^i of player i in G^* (in G^T) is *implemented* by the automaton α of player i if σ^i is equivalent to σ_α^i, i.e., if for every $\sigma^{-i} \in \times_{j \neq i} \Sigma^j$ $(\Sigma^j(T))$, $\omega(\sigma^i, \sigma^{-i}) = \omega(\sigma_\alpha^i, \sigma^{-i})$.

A finite sequence of actions $a_1, \ldots, a_t$ is *compatible* with the pure strategy σ^i of player i in G^*, if for every $1 \leq s \leq t$, $\sigma^i(a_1, \ldots, a_{s-1}) = a_s^i$. Given a strategy σ^i of player i in G^*, any sequence of actions $a_1, \ldots, a_t$, induces a strategy $(\sigma^i | a_1, \ldots, a_t)$ in G^*, by

$$(\sigma^i | a_1, \ldots, a_t)(b_1, \ldots, b_s) = \sigma^i(a_1, \ldots, a_t, b_1, \ldots, b_s).$$

Proposition 1 *The number of different reduced strategies that are induced by a given pure strategy σ^i of player i in G^* and all σ^i-compatible sequences of actions equals the size of the smallest automaton that implements σ.*

2.4 Repeated games with finite automata

Given a strategic game G and positive integers $m_1, \ldots, m_n$, we define $\Sigma^i(T, m_i)$ ($\Sigma^i(m_i)$) to be all pure strategies in $\Sigma^i(T)$ (in Σ^i) that are induced by an automaton of size m_i. Note that if a strategy is induced by an automaton of size m_i and $m_i' \geq m_i$ then it is also induced by an automaton of size m_i'. The game $G^T(m_1, \ldots, m_n)$ is the strategic game $(N; (\Sigma^i(T, m_i))_{i \in N}; r_T)$ where r_T here is the restriction of our earlier payoff function r_T to $\times_{i \in N} \Sigma^i(T, m_i)$.

The play in the supergame G^* which is induced by an n-tuple of strategies $\sigma = (\sigma^i)_{i \in N}$ with $\sigma^i \in \Sigma^i(m_i)$ enters a cycle of length $d \leq \prod_{i \in N} m_i$ after a finite number of stages. Indeed, if at stages t and s the n-tuple of states of the automata coincide, then for every nonnegative integer r, $\omega_{t+r}(\sigma) =$

$\omega_{s+r}(\sigma)$. As the number of different n-tuples of automata states is bounded by $\prod_{i\in N} m_i$ the periodicity follows. Therefore, the limiting average payoff per stage is well defined whenever all players are restricted to strategies which are implemented by finite automata. The game $G^*_\infty(m_1,\ldots,m_n)$ or $G(m_1,\ldots,m_n)$ for short, is the strategic game $(N;(\Sigma^i(m_i))_{i\in N};r_\infty)$ where r_∞ is defined as the limit of our earlier payoff function r_T as $T\to\infty$.

3 Zero-Sum games with finite automata

In this section we present results of the value of 2-person 0-sum repeated games with finite automata. Results concerning zero-sum games are important for the study of the non-zero sum case by specifying the individual rational payoffs and thus the effective "punishments."

Consider the two-person zero-sum game of matching pennies:

1	-1
-1	1

Assume that player 1, the row player, and player 2, the column player, are restricted to play strategies that are implemented by automata of size m_1 and m_2 respectively. Recall that we are considering the mixed extension of the game in which the pure strategies of player i are those implemented by an automaton of size m_i. An easy observation is that for every m_1 there exists a sufficiently large m_2 , a pure strategy $\tau\in\Sigma^2(m_2)$ and a positive integer T such that for any $t\geq T$ and $\sigma\in\Sigma^1(m_1)$, $r^1(\omega_t(\sigma,\tau)) = -1$. Therefore, we conclude in particular that for the above matching pennies game $G=(A,B,h)$, for every m_1 there exists m_2 such that $\text{Val}(G(m_1,m_2)) = \max_A \min_B h(a,b)$. Moreover this statement is valid for any two-person zero-sum game $H=(A,B,h)$. Theorem 1 of Ben-Porath (1993) asserts that if $m_2\geq m_1|\Sigma^1(m_1)|$, where for a set X, $|X|$ denotes the number of elements in X, then

$$\text{Val }(H(m_1,m_2)) = \max_{a\in A}\min_{b\in B} h(a,b).$$

Note that $|\Sigma^1(m)|$ is of the order of an exponential function of $m\log m$. However, it turns out that if the larger bound m_2 is subexponential in m_1, player 2 is unable to use effectively in the long run his larger bound. Indeed,

Theorem 1 *(Ben-Porath, 1986, 1993). Let $H = (A, B, h)$ be a two person 0-sum game in strategic form, and let $(m(n))_{n=1}^{\infty}$ be a sequence of positive integers with*

$$\lim_{n\to\infty} \frac{\log m(n)}{n} = 0.$$

Then,

$$\liminf_{n\to\infty} \mathrm{Val}\,(H(n, m(n))) \geq \mathrm{Val}\,(H).$$

Proof. W.l.o.g. we assume that $n \leq m(n)$. For every sequence $a = (a_1^1, a_2^1, \ldots)$ of actions of player 1 we denote by σ^a the pure strategy of player 1 with $\sigma_t^a(*) = a_t^1$. Note that if a is k-periodic then $\sigma^a \in \Sigma^1(k)$. For every k, $\sigma^1(k)$ denotes the mixed strategy σ^X of player 1 where $X = (X_1, X_2, \ldots)$ is a random k-periodic sequence of actions of player 1, with $X_1, X_2, \ldots, X_k$ i.i.d and the distribution of X_t is an optimal strategy of player 1 in the one shot game. It follows that for every pure strategy τ of player 2 and every $t \leq k$, $E_{\sigma^1(k),\tau}(h(a_t, b_t)|\mathcal{H}_t) \geq \mathrm{Val}(H)$, where $\mathcal{H}_t$ denotes the algebra generated by the actions $a_1, b_1, \ldots, a_{t-1}, b_{t-1}$ in stages $1, \ldots, t-1$. Therefore $Prob_{\sigma^1(k),\tau}(\sum_{t=1}^k h(a_t, b_t)/k < \mathrm{Val}(H) - \varepsilon) \leq e^{-C(\varepsilon)k}$ with $C(\varepsilon) > 0$. Therefore for every finite set $\mathcal{T} \subset \Sigma^2$,

$$Prob_{\sigma^1(n)}(\min_{\tau\in\mathcal{T}} h_n(\sigma^a, \tau) \leq \mathrm{Val}(H) - \varepsilon) \leq |\mathcal{T}| \exp(-C(\varepsilon)n). \qquad (1)$$

Let τ be a pure strategy of player 2 which is implemented by an automaton of size $m(n)$, and set $\mathcal{T} = \{(\tau \mid b_1, \ldots, b_t)\}$. Then for every n periodic sequence a and every positive integer s,

$$\sum_{t=s+1}^{s+n} h(\omega_t(\sigma^a, \tau)) \geq \min_{\tau\in\mathcal{T}} \sum_{t=s+1}^{s+n} h(\omega_t(\sigma^a, \tau)).$$

As $|\mathcal{T}| \leq m(n)$, and $\sigma^1(n)$ is a mixture of (at most $|A|^n$) pure strategies of the form $\sigma^a \in \Sigma^1(n)$, the result follows from (1). ∎

It is worth mentioning that the proof implies a stronger result. Setting $\Sigma_g^i(m)$ to be all strategies σ^i such that for each t $|\{(\sigma^i \mid b_1, \ldots, b_t) : b_j \in B\}| \leq m$, and $\sigma^1(n)$ as constructed in the proof, we conclude that under the same condition as in the theorem,

$$\liminf_{n\to\infty} Val\ H(\{\sigma^1(n)\}, \Sigma_g^2(m(n))) \geq \ \mathrm{Val}\,(H).$$

This stronger result implies that whenever $\lim_{n\to\infty} \log m(n)/n = 0$, for every n there exists a *random* n-periodic sequence of actions of player 1, (σ^X), which guarantees approximately the value $\mathrm{Val}(H)$ against any strategy in $\Sigma_g^2(m(n))$. Note that for every pure strategy σ of player 1, there exists a strategy $\tau \in \Sigma_g^2(1)$ with $h_t(\sigma,\tau) \leq \max_{a\in A}\min_{b\in B} h(a,b)$. The next result asserts that when $m(n)\log m(n) = o(n)$ as $n \to \infty$, then there is a *deterministic* n-periodic sequence of actions of player 1, a, such that σ^a guarantees approximately $\mathrm{Val}(H)$ when player 2 is restricted to strategies in $\Sigma^2(m(n))$.

Proposition 2 *Let* $m : I\!N \to I\!N$ *with* $\lim_{n\to\infty} \frac{m(n)\log m(n)}{n} = 0$. *Then for every* n *there exists an* n*-periodic sequence of actions of player* 1, a, *such that*

$$\lim_{n\to\infty} (\inf\{h_t(\sigma^a,\tau) \mid \tau \in \Sigma^2(m(n)), t \geq n\}) = \mathrm{Val}(H).$$

Proof. Note that there is a positive constant K such that $|\Sigma^2(m(n))| \leq m(n)^{Km(n)}$. Let $k : I\!N \to I\!N$ be such that $\lim_{n\to\infty} \frac{m(n)\log m(n)}{k(n)} = 0$ and $\lim_{n\to\infty} k(n)/n = 0$. Let $X = (X_1,\ldots,X_{k(n)},\ldots)$ be a random n-periodic sequence of actions of player 1, where $X_1,\ldots,X_{k(n)}$ are i.i.d each distributed according to the distribution of an optimal mixed strategy of player 1 in the one shot game, and $(X_1,\ldots,X_n)$ is $k(n)$-periodic. As $\lim_{n\to\infty} \frac{m(n)\log m(n)}{k(n)} = 0$, it follows that for every positive constant $C > 0$,

$$\lim_{n\to\infty} |\Sigma^2(m(n))| \exp\left(-Ck(n)\right) = 0,$$

and therefore it follows from (1) that

$$\lim_{n\to\infty} \Pr(\min_{\tau\in\Sigma^2(m(n))} h_{k(n)}(\sigma^X,\tau) \leq \mathrm{Val}(H) - \varepsilon) = 0$$

and therefore there is an n-periodic sequence of actions a such that

$$\lim_{n\to\infty} (\inf\{h_t(\sigma^a,\tau) \mid \tau \in \Sigma^2(m(n)), t \geq n\}) \geq Val(H).$$

■

The next result follows from the proof of the result of Ben-Porath (1993), and is used in the proof of Theorems 5 and 6.

Theorem 2 *For every* $\varepsilon > 0$ *sufficiently small, if*

$$\exp(\varepsilon^2 m_1) \geq m_2 > 1,$$

then for every positive integer T,

$$\mathrm{Val}\,(H^T(m_1, m_2)) \geq \mathrm{Val}\,(H) - \varepsilon.$$

The next corollary is a restatement of Theorems 1 and 2 which provides a lower bound for equilibrium payoffs in nonzero sum repeated games with finite automata.

Corollary 1 *For every strategic game $G = (N, A, r)$, $i \in N$, and $\varepsilon > 0$ sufficiently small, if*

$$exp(\varepsilon^2 m_i) \geq m_j > 1 \qquad \textit{for every } j \neq i,$$

then for every $x \in E(G^T(m_1, \ldots, m_n))$, or $x \in E(G(m_1, \ldots, m_n))$,

$$x^i \geq w^i(G) - \varepsilon.$$

The next result asserts that if the bound on the sizes of the automata of player 2 is larger than an exponential of the sizes of the automata of player 1, then player 2 could hold player 1 down to his maxmin in pure strategies.

Theorem 3 *For every 2-person 0-sum game $H = (\{1, 2\}; (A, B); h)$, and every positive constant K with $K > \ln|A|$, if $m(n) \geq exp(Kn)$, then*

$$\mathrm{Val}\,(H(n, m(n)) \to \max_{a \in A} \min_{b \in B} h(a, b) \textit{ as } n \to \infty.$$

Proof. Let $K > \ln|A|$. It is sufficient to prove that for every $\varepsilon > 0$ there exists n_0 such that for every $n \geq n_0$ and every $m \geq \exp(Kn)$,

$$\mathrm{Val}\ (H(n, m)) \leq \max_{a \in A} \min_{b \in B} h(a, b) + \varepsilon.$$

Note that for every positive constant C there exists n_0 such that for every $n \geq n_0$, $\exp(Kn) \geq Cn^2|A|^n$. Therefore the theorem follows from the next lemma. ∎

Lemma 1 *For every $\varepsilon > 0$, there is a sufficiently large integer $K = K(\varepsilon)$, such that for every $m \geq Kn^2|A|^n$ there exists a strategy $\tau^* \in \Delta(\Sigma^2(m))$ such that for every $T \geq K^2 n^3 |A|^n$, and any strategy $\sigma \in \Sigma^1(n)$,*

$$h_T(\sigma, \tau^*) \leq \max_{a \in A} \min_{b \in B} h(a, b) + \varepsilon.$$

and therefore,

$$\mathrm{Val}\,(G^T(n,m)) \le \max_{a\in A}\min_{b\in B} h(a,b) + \varepsilon.$$

and

$$\mathrm{Val}\,(G(n,m)) \le \max_{a\in A}\min_{b\in B} h(a,b) + \varepsilon.$$

Proof. The idea of the proof is as follows: for every pure strategy $\sigma \in \Sigma^1(n)$ of player 1, there is $1 \le k \le n$ and a sequence of actions $b = b_1, \ldots, b_n, b_{n-k+1}, \ldots$ with $b_i = b_j$ whenever $i > j > n-k$ and $i - j = 0(mod\ k)$, such that the strategy τ^b of player 2 which plays the sequence b results in payoff $\le \max_{a\in A}\min_{b\in B} h(a,b)$ in seach stage $t > m-k$. Such a strategy is implemented by an automaton of size n, and the number of such strategies σ^b is bounden by $n|B|^n$. The strategy of player 2 immitates choosing at random a pair k, σ^b, and if the resulting payoffs are not sufficiently small, it attempts another randomly chosen pair k, σ^b. The sufficiently large number of states of the automaton of player 2, gaurantees that with high probability the induced play will eventually enter a cycle with payoffs $\le \max_{a\in A}\min_{b\in B} h(a,b)$ in each stage .

Formally, let $b : A \to B$ be a selection from the best reply correspondence of player 2. Construct the following mixed strategy of player 2, τ^*, which is implemented by an automaton with state space

$$M^2 = \{1, \ldots, n\} \times \{1, \ldots, \ell\}.$$

where $\ell = Kn|A|^n$. The initial state of the automaton of player 2 is $(1,1)$. Let $a : M^2 \to A$ be a random function, each such function equally likely, i.e., for every $1 \le i \le n$, and every $1 \le j \le \ell$, $a(i,j)$ is a random element of A each one equally likely, and the various random elements $a(i,j)$ are independent. We define now the random action function of the automaton.

$$f^2(i,j) = b(a(i,j)).$$

The transition function of the automaton depends on a random sequence $k = k_1, \ldots, k_\ell$, $1 \le k_j \le n$, each such sequence equally likely and the sequence k is independent of the function a. We are ready now to define the transition function which depends on the functions b and a and on the random sequence

k.

$$g^2((i,j),c) = \begin{cases} (i+1,j) & \text{if } \; i < n \text{ and } c = a(i,j) \\ (k_j,j) & \text{if } \; i = n \text{ and } c = a(i,j) \\ (1,j+1) & \text{if } \; j < \ell \text{ and } c \neq a(i,j) \\ (1,1) & \quad\text{otherwise.} \end{cases}$$

Let σ be a pure strategy of player 1 that is implemented by an automaton of size n. Let $x_1, x_2, \ldots$ where $x_t = (a_t, b_t)$ be the random play induced by the strategy pair σ and τ^*, and let $q_1^i, q_2^i, \ldots$ be the random sequence of states of the automaton of player 2. Fix $1 \leq j \leq \ell$ and let $t = t_j$ be the random time of the first stage t with $q_t^2 = (1,j)$. Note that

$$Prob(a_{t+s} = a(s+1,j) \;\; \forall\, 0 \leq s < n) = \frac{1}{|A|^n}.$$

and if $a_{t+s} = a(s+1,j) \;\; \forall\, 0 \leq s < n$ then there exists $0 \leq s < n$ such that the state of the automaton of player 1 at stage $t+n$, q_{t+n}^1 coincides with its state at stage $t+s$. Therefore if $k_j = s+1$, the play will enter a cycle in which the payoff to player 1 is at most $\max_{a \in A} \min_{b \in B} h(a,b)$. Therefore the conditional probability, given the history of play up to stage t_j, that the payoff to player 1 in any future stage is at most $\max_{a \in A} \min_{b \in B} h(a,b)$, and that $t_{j+1} = \infty$ is at least $1/(|A|^n n)$. Otherwise, if $t_{j+1} < \infty$, $t_{j+1} \leq t_j + n^2$. Therefore, either $t_\ell = \infty$ and then for every stage $t > \ell n^2$, the payoff to player 1 is at most $\max_{a \in A} \min_{b \in B} h(a,b)$, or $t_\ell < \infty$. However, the previous inequalities imply that,

$$Prob(t_\ell = \infty) \geq 1 - (1 - 1/(n|A|^n))^{\ell - 1} \to 1 \text{ as } K \to \infty,$$

which completes the proof of the lemma. ∎

It is of interest to bridge between the results of this section concerning the infinitely repeated 2-person 0-sum games, by providing asymptotic results of the value $\mathrm{Val}(H(m_1, m_2))$ when $m_1 \to \infty$ and m_2 is approximately a fixed exponential function of m_1. Given a 2-person 0-sum game $H = (A, B, h)$, it will be interesting to find the largest (smallest) monotonic nondecreasing functions $\bar{v} : (0, \infty) \to I\!R$ ($v : (0, \infty) \to I\!R$) such that if $\frac{\ln m_2(m)}{m} \to \alpha > 0$ as $m \to \infty$ then

$$v(\alpha) \leq \liminf_{m \to \infty} \mathrm{Val}(H(m, m_2)) \leq \limsup_{m \to \infty} \mathrm{Val}(H(m, m_2)) \leq \bar{v}(\alpha).$$

Theorems 1 asserts that $\lim_{\alpha\to 0}\bar{v}(\alpha) = \lim_{\alpha\to 0} v(\alpha) = \mathrm{Val}(H)$, and Theorem 3 asserts that for $\alpha > \ln|A|$, $\bar{v}(\alpha) = \max_{a\in A}\min_{b\in B} h(a,b)$. We conjecture that the two functions $\bar{v}$ and v are continuous with $\bar{v} = v$ for all values of $\alpha > 0$ with the possible exception of one critical value.

The next two conjectures address the number of repetitions needed for a unrestricted player to use his advantage over bounded automata. The positive resolutions of each of the conjectures have implications on the equilibrium payoffs of finitely repeated games with automata. A positive resolution of the next conjecture, will provide a positive answer to conjecture 3.

Conjecture 1 *For every $\varepsilon > 0$, if $m : I\!N \to I\!N$ satisfies $m(T) \geq \varepsilon T$, then*

$$\lim_{T\to\infty} \mathrm{Val}\left(H^T(m(T),\infty)\right) = \mathrm{Val}\left(H\right).$$

The truth of the above conjecture implies that there is a function $m : I\!N \to I\!N$ with $\lim_{T\to\infty} m(T)/T = 0$, and such that

$$\lim_{T\to\infty} \mathrm{Val}(H^T(m(T),\infty) = \mathrm{Val}(H).$$

An interesting open problem is to find the "smallest" such function. The next conjecture specifies a domain for such a function.

Conjecture 2 *If $m : I\!N \to I\!N$ obeys $\lim_{T\to\infty}(T/\log T)/m(T) = 0$, then*

$$\lim_{T\to\infty} \mathrm{Val}\left(H^T(m(T),\infty)\right) = \mathrm{Val}\left(H\right).$$

If $m : I\!N \to I\!N$ obeys $\lim_{T\to\infty} m(T)/(T/\log T) = 0$, then

$$\lim_{T\to\infty} \mathrm{Val}\left(H^T(m(T),\infty)\right) = \max_{a^1\in A_1}\min_{a^2\in A_2} h^1(a^1,a^2).$$

4 Equilibrium payoffs of the supergame G^*_∞

We state here a result, due to Ben-Porath, which is a straightforward corollary of his result in the 2-person 0-sum case. All convergence of sets is with respect to the Hausdorff topology. Recall that for a sequence of subsets, E_n, of a Euclidean space $I\!R^k$,

$$\liminf_{n\to\infty} E_n = \{x \in I\!R^k \mid \forall\varepsilon > 0, \exists N \text{ s.t. } \forall n \geq N,\ d(x,E_n) < \varepsilon\}$$

where $d(x, E)$ denotes the distance of the point x from the set E,

$$\limsup_{n\to\infty} E_n = \{x \in I\!R^k \mid \forall \varepsilon > 0,\ \forall N,\ \exists n \geq N \text{ with } d(x, E_n) < \varepsilon\},$$

and $\lim_{n\infty} E_n = E$ if $E = \liminf E_n = \limsup E_n$.

Theorem 4 *(Ben-Porath 1986, 1993). Let $G = (N; (A_i)_{i\in N}; (r^i)_{i\in N})$ be a strategic game, and $m_i(k)$, $i \in N$, sequences with $\lim_{k\to\infty} m_i(k) = \infty$ and*

$$\lim_{k\to\infty} \frac{\log(\max_{i\in N} m_i(k))}{\min_{i\in N} m_i(k)} = 0.$$

Then,

$$\{x \in F \mid x^i \geq v^i(G)\} \subseteq \liminf_{k\to\infty} E(G(m_1(k), \ldots, m_n(k))),$$

and

$$\limsup_{k\to\infty} E(G(m_1(k), \ldots, m_n(k))) \subseteq \{x \in F \mid x^i \geq w^i(G)\}.$$

Note that in two-person games $v^i(G) = w^i(G)$ and therefore the above theorem provides exact asymptotics for two-person games. An interesting open problem is to find the asymptotic behavior of $E(G(m_1(k), \ldots, m_n(k)))$ as $k \to \infty$ and $\lim_{k\to\infty}\{\log(\max_{i\in N} m_i(k))/\min_{i\in N} m_i(k)\} = 0$. Such questions lead to the study of the asymptotics of

$$v^i(G(m_1(k), \ldots, m_n(k))) = \min_{\tau^{-i}} \max_{\sigma^i} r^i_\infty(\sigma^i, \tau^{-i}),$$

where the min ranges over all $\tau^{-i} \in \times_{j\neq i}\Delta(\Sigma^j(m_j(k)))$ and the max is over $\sigma^i \in \Sigma^i(m_i(k))$ and where $m_i(k)$, $i \in N$, is a sequence with $\lim_{k\to\infty} m_i(k) = \infty$ and $\lim_{k\to\infty}(\log(\max_{i\in N} m_i(k)))/(\min_{i\in N} m_i(k)) = 0$. W.l.o.g. assume that $m_1(k) \leq m_2(k) \leq \ldots \leq m_n(k)$ and $\lim_{n\to\infty} \log m_n(k)/m_1(k) = 0$, and let $i < n$. Set $v^i(k) = v^i(G(m_1(k), \ldots, m_n(k)))$. We denote by $Q(i)$, or Q for short, the set of all probability measures on A_{-i} whose marginal distribution on $\times_{j<i}A_j$ is a product measure. The following is a partial answer to the study of the asymptotics of $v^i(k)$.

Proposition 3 *(a) If* $\lim_{k\to\infty} \frac{m_1(k)\log m_1(k)}{m_2(k)} = 0$, *then*

$$\limsup_{k\to\infty} v^1(k) \le \min_{q\in Q} \max_{a^1\in A_1} \sum_{a^{-1}\in A_{-1}} q(a^{-1}) r^1(a^1, a^{-1}).$$

(b) If for a fixed player $1 < i < n$, $\lim_{k\to\infty} \frac{\log m_{i+1}(k)}{\log m_i(k)} = \infty$, *then*

$$\limsup_{k\to\infty} v^i(k) \le \min_{q\in Q} \max_{a^i\in A_i} \sum_{a^{-i}\in A_{-i}} q(a^{-i}) r^i(a^i, a^{-i}).$$

Proof. Part (a) follows from Proposition 2. We turn to the proof of part (b). Let $(N(k))_{k=1}^{\infty}$ be a sequence of positive integers with $\lim_{k\to\infty} N(k)/\log m_i(k) = \infty$ and $\lim_{k\to\infty} N(k)/\log m_{i+1}(k) = 0$. The constructed $N \setminus \{i\}$ tuple of minimax strategies, $(\sigma^j)_{j\neq i}$, will enter a cycle of length $N(k)$, following the first $N(k)(n-1)$ stages. The cycle play, $X_1, \ldots, X_{N(k)}$, is a sequence of i.i.d. actions in A_{-i} with each X_t distributed according to a minimizing probability $q \in Q$. For every k let $q^*(k) \in Q$, or q for short, attain the minimum and let q_j be the marginal distribution of q on A_j, $j < i$. Let σ^j, $j < i$, be the strategy $\sigma^{j,X}$ which plays a random $N(k)$-periodic sequence of actions $X_1^j, X_2^j, \ldots$ where $X_1^j, \ldots, X_{N(k)}^j$ are i.i.d. and the distribution of each X_t^j is q_j. We define next the strategy σ^j for $j > i$, which is a mixture of pure strategies, each implemented by an automaton of size $iN(k)|A^{N(k)}|$ which for sufficiently large k is $\le m_{i+1}(k)$. For every $b = (b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_n) \in \times_{j<i} A_j$, denote by b^j the projection of b on $\times_{i\neq j'<j} A_{j'}$, and let $q^{i+1}(b)$ denote the marginal of the conditional probability $(q|b^j)$ on A_j. The automaton includes $|A|^{N(k)} + (j-i-1)N(k)$ states which are used to record the realization of the choices of all players in stages $N(k)(j-i-1)+1, \ldots, N(k)(j-i)$ stages. Thereafter, player j plays an $N(k)$-periodic sequence $a_1^j, \ldots, a_{N(k)}^j$ which is a realization of a sequence of independent actions $X_1^j, \ldots, X_{N(k)}^j$ with the distribution of X_t^j being $(q^{i+1}|b^j_{t+(j-i-1)N(k)})$. One verifies that following the first $(n-1)N(k)$ stages the play of players $j \neq i$ enters an $N(k)$ cycle of i.i.d actions each distributed according to q. ∎

In the above proof we can also construct the minimaxing strategies of players $j > i$ to be pure strategies, as in Proposition 2. Consider the following 3 player game G.

0, 0, 0	8, 0, 4
0, 0, 8	0, 0, 8

0, 0, 8	0, 0, 8
0, 8, 4	0, 0, 0

Player 1 chooses the row, player 2 the column, and player 3 chooses the matrix. Note that $v^1(G) = 0 = v^2(G)$ and $v^3(G) = 5$. However, $w^3(G) = 4$ (and $w^1(G) = w^2(G) = 0$). Therefore we can not deduce from Theorem 4 whether or not the vector payoff $(4, 4, 4)$ is approximated by equilibrium payoffs of the restricted games $G(m_1(k), m_2(k), m_3(k))$ for sufficiently large k and where $k < m_i(k)$ are sequences with $\log \max m_i(k)/ \min m_i(k) \to 0$ as $k \to \infty$. However, Proposition 3 characterizes for this game the limit of the equilibrium payoffs provided that we assume in addition that $\lim_{k\to\infty} \log \max(m_1(k), m_2(k))/ \log m_3(k) = \infty$. In particular, it follows in this case that $(4, 4, 4)$ is in the limit of the equilibrium payoffs.

We state now a result which provides a partial answer to the asymptotic behavior of the set of equilibrium payoffs of repeated games with bounded automata. Denote by

$$\mathbf{d}^i = \min_{q\in Q(i)} \max_{a^i\in A_i} \sum_{a^{-i}\in A_{-i}} q(a^{-i})r^i(a^i, a^{-i})$$

and

$$\mathbf{F} = \{x \in F(G) | x^i > \mathbf{d}^i\}$$

Proposition 4 *Assume that* $m_1(k) \leq \ldots \leq m_n(k)$, $\lim_{k\to\infty} \frac{\log m_n(k)}{m_1(k)} = 0$, $\lim_{k\to\infty} \frac{m_1(k)\log m_1(k)}{m_2(k)} = 0$ *and that for* $i > 1$ $\lim_{k\to\infty} \frac{\log m_i(k)}{\log m_{i+1}(k)} = 0$*. Then,*

$$\liminf_{k\to\infty} E(G(m_1(k), \ldots, m_n(k))) \supset \mathbf{F}$$

5 Cooperation in finitely repeated games

The results in this section address the asymptotic behavior of the sets of equilibrium payoffs, $E(G^T(m_1, m_2))$, of the games $G^T(m_1, m_2)$, as T, m_1 and m_2 go to ∞. All convergence of sets is with respect to the Hausdorff topology. In each one of the theorems in the present section we assume that $G = (\{1, 2\}, A, r)$ is a fixed 2-person strategic game, $F = F(G)$ stands for the feasible payoffs in the infinitely repeated game, i.e., $F = co(r(A))$ and that $(T(n), m_1(n), m_2(n))_{n=1}^{\infty}$ is a sequence of triples. For simplicity, the statements of the theorems are nonsymmetric with respect to the two players, and therefore we assume in addition that $m_2(n) \geq m_1(n)$. We also suppress often the dependence on n; no confusion should result.

Theorem 5 *Let $G = (\{1,2\}, A, r)$ be a two person game in strategic form, and assume that there is $x \in F(G))$ with $x^1 > v^1(G)$, and $x^2 > u^2(G)$. Then, if $m_1(n) \to \infty$ and $\frac{\log m_1(n)}{T(n)} \to 0$ as $n \to \infty$,*

$$\liminf_{n\to\infty} E(G^T(m_1, m_2)) \supseteq \{x \in F \mid x^1 \geq v^1(G) \text{ and } x^2 \geq u^2(G)\}.$$

Special cases of the above theorem have been stated in previous publications. Neyman (1985) states that in the case of the finitely repeated Prisoner's dilemma G, for any positive integer k, there is T_0 such that if $T \geq T_0$ and $T^{1/k} \leq \min(m_1, m_2) \leq \max(m_1, m_2) \leq T^k$, then there is a mixed strategy equilibrium of $G^T(m_1, m_2)$ in which the payoff is $1/k$-close to the "cooperative" payoff of G. Papadimitriou and Yannakakis (1994) state the special case of Theorem 5 obtained by assuming that the payoffs of the underlying game are rational numbers and replacing $F(G)$ in the statement of the theorem with $\{x \in r(A) \text{ with } x^i > v^i(G)\}$. They also state a result for a subset of F with the additional assumption that the bounds on both automata are subexponential in the number of repetitions.

The conclusion of the theorem fails if we replace in the assumptions of the theorem the strict inequality $x^1 > v^1(G)$ by the weak inequality $x^1 \geq v^1(G)$. For example in the game

$0,4$	$1,3$
$1,1$	$1,0$

the only equilibrium payoff in $G^T(m_1, m_2)$ with $m_2 \geq 2^T$ is (1,1).

The next theorem relates the equilibrium payoffs of $G^T(m_1, m_2)$ to the equilibrium payoffs of the undiscounted infinitely repeated game G^*_∞. Recall that the Folk Theorem asserts that

$$E(G^*_\infty) = \{x \in F \mid x^1 \geq v^1(G) \text{ and } x^2 \geq v^2(G)\}.$$

Theorem 6 *Let $G = (\{1,2\}, A, r)$ be a two person game in strategic form , and let $(T, m_1(T), m_2(T))_{T=1}^\infty$ be a sequence of triples of positive integers with $m_1(T) \leq m_2(T)$ and $m_1(T) \to \infty$ as $T \to \infty$, and*

$$\lim_{T\to\infty} \frac{\log m_2(T)}{\min(m_1(T), T)} = 0.$$

Then,
$$\lim_{T\to\infty} E(G^T(m_1(T), m_2(T))) = E(G^*_\infty).$$

The limiting assumption $\lim_{T\to\infty} \frac{\log m_2(T)}{m_1(T)} = 0$ in Theorem 6, could probably be replaced by an alternative lower bound, as a function of T, on $m_1(T)$, provided that we also assume that there is $x \in F$ with $x^1 > v^1(G)$. One example of such a result is presented in the following conjecture.

Conjecture 3 *Let $G = (\{1,2\}, A, r)$ be a two person game in strategic form, and assume that there is $x \in F$ with $x^1 > v^1(G)$. Then, if*

$$\liminf_{T\to\infty} m_1(T)/T > 0,$$

and
$$\lim_{T\to\infty} \frac{\log m_1(T)}{T} = 0,$$

Then
$$\lim_{T\to\infty} E(G^T(m_1, m_2)) = \{x \in F \mid x^i \geq v^i(G)\}.$$

The next theorem is straightforward and very easy. We state it as a contrast to the previous results. It shows that the subexponential bounds on the sizes of the automata as a function of the number of repetitions is essential to obtain equilibrium payoffs that differ from those of the finitely repeated game G^T.

Theorem 7 *For every game G in strategic form there exists a constant c such that if $m_i \geq \exp(cT)$ then*

$$E(G^T(m_1, \ldots, m_n)) = E(G^T).$$

6 Repeated games with bounded recall

Aumann (1981) mentioned two ways of modeling a player with bounded rationality: with finite automata and with bounded recall strategies. There are two alternatives to define strategies with bounded recall. The first one (see e.g. Aumann and Sorin, 1989) considers strategies with bounded recall which choose an action as a function of the recalled opponents' actions, and the second alternative (see e.g. Kalai and Stanford, 1988, or Lehrer 1988) allows a player to rely on his opponents' actions as well as on his own. The following are results on repeated games with bounded recall of the second type which are closely related to those presented for finite automata. Let $BR^i(m)$ denote all strategies of player i in a repeated game that choose an action as a function of all players action in the last m stages. Each pure strategy $\sigma^i \in BR^i(m)$ is thus represented by a function $f^i : A^m \to A_i$ and a fixed element, initial memory, $e = (e_1, \ldots, e_m) \in A^m$; for $t > m$, $\sigma^i(a_1, \ldots, a_{t-1}) = f^i(a_{t-m}, \ldots, a_{t-1})$ and for $t \leq m$, $\sigma^i(a_1, \ldots, a_{t-1}) = f^i(e_t, \ldots, e_m, a_1, \ldots, a_{t-1})$. Given a strategy $\sigma^i = (e, f^i) \in BR^i(m)$, the automaton $\langle A^m, e, f^i, g^i \rangle$ where $g^i(x, y)$, $x = (x_1, \ldots, x_m) \in A^m$ and $y \in A_{-i}$, equals $(x_2, \ldots, x_m, (f^i(x), y))$, implements the strategy σ^i. Thus each strategy in $BR^i(m)$ is implemented by an automaton of size $|A|^m$, or in symbols and identifying a strategy with its equivalence class, $BR^i(m) \subset \Sigma^i(|A|^m)$. Given a fixed two-person zero-sum game $G = (A, B, h)$, we denote by V_{m_1,m_2} the value of the undiscounted infinitely repeated game G where player i is restricted to mixed strategies with support in $BR^i(m_i)$. Lehrer (1988) proves the following result which is related and has a spirit similar to the result of Ben-Porath (1986,1993).

Theorem 8 *(Lehrer, 1988). For every function* $m : \mathbb{N} \to \mathbb{N}$ *with* $\log m(n)/n \to 0$ *as* $n \to \infty$*,*

$$\liminf V_{n,m(n)} \geq \ \mathit{Val}\ (G).$$

Proof. Note that, by identifying a strategy with its equivalence class, $BR^2(m(n)) \subset \Sigma^2(|A \times B|^{m(n)})$. Let $k : \mathbb{N} \to \mathbb{N}$ with $\lim_{n\to\infty} m(n)/k(n) = 0$ and $\lim_{n\to\infty} \log k(n)/n = 0$. E.g., $k = m^2$. Let $X = X_1, X_2, \ldots, X_{k(n)}, \ldots$ be a random $k(n)$-periodic sequence of actions, with $X_1, \ldots, X_{k(n)}$ i.i.d and X_t an optimal strategy of player 1 in the one shot game. W.l.o.g. we assume that the support of X_t has at least two elements. Consider the mixed strategy $\sigma^1(k(n))$ of player 1 which plays the realization of X. The proof of Theorem 1 shows that for any sequence of strategies $\tau(n) \in \Sigma^2(|A \times B|^{m(n)})$, $\liminf_{n\to\infty} r_\infty(\sigma^1(k(n)), \tau(n)) \geq \mathrm{Val}(G)$ as $n \to \infty$. (Note that $\sigma^1(k(n)) \notin \Delta(BR^1(n))$). It is thus sufficient to prove that the norm distance of $\sigma^1(k(n))$ from $\Delta(BR^1(n))$ tends to zero as $n \to \infty$, i.e., that for most realizations of X, the implied pure strategy is in $BR^1(n)$. Note that for any $0 \leq s < t < k(n)$ there are positive integers s' and t' with $t \leq t' \leq t + n - [n/3]$, and $s \leq s' \leq$

$s+n-[n/3]$ such that $X_{s'+1},\dots,X_{s'+[n/3]},X_{t'+1},\dots,X_{t'+[n/3]}$ are independent and $(X_{s+1},\dots,X_{s+n})=(X_{t+1},\dots,X_{t+n})$ only if $(X_{s'+1},\dots,X_{s'+[n/3]})=(X_{t'+1},\dots,X_{t'+[n/3]})$. Indeed, if $\min\{t-s,s+k(n)-t\}\geq[n/3]$ set $s'=s$ and $t'=t$; if $t<s+[n/3]$ set $s=s'$ and $t'-s$ is the smallest multiple of $t-s$ which is $\geq[n/3]$; and if $s+k(n)<t+[n/3]$ set $t=t'$ and $s'-t$ is the smallest multiple of $s+k(n)-t$ which is $\geq[n/3]$. There is a constant $0<\alpha<1$ that depends on the optimal strategy of player 1 in the one shot game, such that $\Pr(X_{s'+i}=X_{t'+i})\leq\alpha$. (e.g., if p is the probability vector associated with the optimal strategy in the one shot game $\alpha=\sum p_i^2$).Therefore, $\Pr(\forall i,\ 1\leq i\leq[n/3],\ \ X_{s'+i}=X_{t'+i})\leq\alpha^{[n/3]}$. Therefore

$$\Pr(\exists\ s,t,\ 0\leq s<t<k(n)\ \text{s.t}\ \forall\ i,0<i\leq n,\ X_{s+i}=X_{t+i}))$$
$$<k^2(n)\alpha^{[n/3]}\to_{n\to\infty}0.$$

Note that the strategy $\sigma^*(n)$ which is defined as the strategy $\sigma^1(k(n))$ conditional on $\{\forall\ s,t,\ 0\leq s<t<k(n),\exists 1\leq i\leq n\ \text{s.t.}X_{s+i}\neq X_{t+i}\}$ is in $\Delta(BR^1(n))$ with $d(\sigma^*(n),\sigma(k(n)))\to 0$ as $n\to\infty$, where $d(\sigma^*(n),\sigma(k(n)))$ denotes the norm distance between the mixed strategies (viewed as distributions) $\sigma^*(n)$ and $\sigma(k(n))$. Therefore

$$\liminf_{n\to\infty}\ \text{Val}(BR^1(n),\Sigma^2(|A\times B|^{m(n)}),h_\infty)\geq\ \text{Val}\ (G).$$

which completes the proof. ∎

The next result is an analog of Proposition 2.

Proposition 5 *For every 2-player 0-sum game $H=(A,B,h)$ there is a positive constant K such that if $m:I\!N\to I\!N$ with $n>Km(n)$, then for every n there exists a strategy $\sigma(n)\in BR^1(n)$ such that*

$$\lim_{n\to\infty}(\inf\{h_t(\sigma(n),\tau)\mid\tau\in BR^i(m(n)),\ t\geq\exp(n)\})=\text{Val}(H).$$

An interesting straightforward corollary of this proposition is the following. Let $G=(A,r)$ be an n-person game, $A=\times_{1\leq i\leq n}A_i$ and $r=(r^i)_{1\leq i\leq n}$, and assume that $m_1(k)\leq\ldots\leq m_n(k)$.

Corollary 2 *There is a constant K such that if $Km_1(k)\leq m_2(k)$, then for every k there exist an $(n-1)$-tuple of strategies $\tau(k)=(\tau_2,\dots,\tau_n)\in\times_{1<i\leq n}BR^i(m_i(k))$ such that for any strategy $\sigma(k)\in BR^1(m_1(k))$,*

$$\limsup_{k\to\infty}r^1_\infty(\sigma(k),\tau(k))\leq\max_{q\in\Delta(A^1)}\ \min_{a^{-1}\in A_{-1}}\sum_{a^i\in A^i}q(a^i)r^1(a^1,a^{-1}).$$

The next result is an analog of Proposition 3. Assume that $m_1(k) \leq \ldots \leq m_n(k)$ with $\lim_{k\to\infty} \log m_n(k)/m_1(k) = 0$. For every $1 \leq i \leq n$ denote by $v^i(m_1(k), \ldots, m_n(k))$, or $v^i(k)$ for short, the minimax payoff to player i, i.e., $\min_{\tau^{-i}} \max_{\sigma^i} r^i_\infty(\sigma^i, \tau^{-i})$, where the min ranges over all $\tau^{-i} \in \times_{j\neq i}\Delta(BR^j(m_j(k)))$ and the max ranges over all $\sigma^i \in BR^i(m_i(k))$. As in section 4, we denote by $Q(i)$, or Q for short, the set of all probability measures on A_{-i} whose marginal distribution on $\times_{j<i}A_j$ is a product measure.

Proposition 6 *If for a fixed player* $1 \leq i < n$, $\lim_{k\to\infty} m_{i+1}/m_i(k) = \infty$, *then*

$$\limsup_{k\to\infty} v^i(k) \leq \min_{q\in Q} \max_{a^i\in A_i} \sum_{a^{-i}\in A_{-i}} q(a^{-i}) r^i(a^i, a^{-i}).$$

We state now a result which provides a partial answer to the asymptotic behavior of the set of equilibrium payoffs of repeated games with bounded recall. It is an analog of Proposition 4.

Proposition 7 *There is a constant* $K > 0$ *such that if* $m_1(k) \leq \ldots \leq m_n(k)$, $\lim_{k\to\infty} \frac{\log m_n(k)}{m_1(k)} = 0$, $m_2(k) > Km_1(k)$, *and* $\lim_{k\to\infty} \frac{m_i(k)}{m_{i+1}(k)} = 0$ *for* $i > 1$, *then*

$$\liminf_{k\to\infty} E(G(BR^1(m_1(k)), \ldots, BR^n(m_n(k)))) \supset \mathbf{F}$$

The next proposition and conjecture address the advantage of an unrestricted player over a player restricted to bounded recall strategies in finitely repeated 2-player 0-sum games. For a fixed two-person zero-sum game $H = (A, B, h)$, we denote by $V^T_{n,\infty}$, or V^T_n for short, the value of the finitely repeated game $H^T(BR^1(n), \Sigma^2)$, i.e. the value of the T-repeated game in which player 1 is restricted to use strategies in $BR^1(n)$ while player 2 can use any strategy in Σ^2. The following proposition asserts that if the duration T is shorter then some exponential function of n then the unrestricted player has no advantage.

Proposition 8 *There exists a constant* $K > 0$ *such that if* $T : \mathbb{N} \to \mathbb{N}$ *satisfies* $T(n) \leq \exp(Kn)$, *then*

$$\lim_{n\to\infty} V^{T(n)}_n = \mathit{Val}\ H.$$

Proof. It is sufficient to prove the result in the case that any optimal strategy of player 1 in the one shot game is not pure. Let $X_1, X_2, \ldots$ be a sequence of i.i.d optimal strategies in the one shot game. The stochastic process $X_1, \ldots$ induces a strategy $\sigma \in \Delta(BR^1(n))$ as follows. For each realization of the random sequence define the initial memory $e = (X_1, \ldots, X_n)$ and the action function $f : (A \times B)^n \to A$ is defined as follows: for every $(a_1, b_1), \ldots, (a_n, b_n)$ define the stopping time S as the smallest value of t such that $(a_1, \ldots, a_n) = (X_{t-n}, \ldots, X_{t-1})$ and define

$$f((a_1, b_1), \ldots, (a_n, b_n)) = X_S.$$

Note that the strategy induced by each realization of the random sequence $X_1, \ldots$ consists of a deterministic sequence (which enters eventually a cycle) of actions of player 1 and that the sequence is independent of the strategy of player 2. It is easy to verify that there is a positive constant $K > 0$ such that

$$\lim_{n\to\infty} \text{Prob}_\sigma(a_t(\sigma) = X_{t+n} \ \forall t \le \exp Kn) = 0$$

and therefore the norm distance between the strategies σ and σ^X goes to zero as n goes the infinity. As σ^X is an optimal strategy in H^T the result follows. ∎

The conjecture below claims that there is an exponential function of n, $\exp(Kn)$ such that if the number of repeatition T is larger then $\exp(Kn)$, the values V_n^T converge to the maxmin in pure strategies.

Conjecture 4 *There is a constant K such that if $T : I\!N \to I\!N$ satisfies $T(n) \ge \exp(Kn)$, then*

$$\lim_{n\to\infty} V_n^{T(n)} = \max_{a\in A} \min_{b\in B} h(a, b).$$

As $BR^1(n) \subset \Sigma^1(|A \times B|^n)$ a positive answer to the second part of Conjecture 2 provides also a positive answer to Conjecture 4. It is of interest to study the asymptotics of $V_n^{T(n)}$ where $T(n)$ is approximately a fixed exponential function of n. This would close the gap between Proposition 8 and Conjecture 4. Given a 2-person 0-sum game $H = (A, B, h)$, it will be interesting to find the largest (smallest) function $\bar{u} : (0, \infty) \to I\!R$ ($u : (0, \infty) \to I\!R$) such that if $\frac{\ln T(n)}{n} \to \alpha$ as $n \to \infty$ then

$$u(\alpha) \le \liminf_{n\to\infty} V_n^{T(n)} \le \limsup_{n\to\infty} V_n^{T(n)} \le \bar{u}.$$

Proposition 8 asserts that there is a constant $K_1 > 0$ such that $u(K_1) = \text{Val } H$ and the conjecture claims that there is a constant K_2 with $\bar{u}(K_2) = \max_{a\in A}\min_{b\in B} h(a,b)$. It is interesting to find the sup and inf of K_1 and K_2 respectively. We conjecture that the two functions $\bar{u}$ and u are continuous with $\bar{u} = u$ for all values of α with the possible exception of one critical value. We do not exclude the possibility of the existence of a positive constant K such that $u(K) = \text{Val } (H)$ and $\bar{u}(K) = \max_{a\in A}\min_{b\in B} h(a,b)$.

7 Variations and extensions

We have studied here some topics in the theory of repeated games with deterministic automata. There are several variants of the concept of automata which merits study in the context of repeated games broadly conceived, i.e., including repeated games with incomplete information and stochastic games. The variations of the concept of an automaton are in several independent dimensions. E.g., we can allow transitions that depend on the actions of all players and also allow for probabilistic actions and/or transitions, and moreover we can consider transition and/or action functions which are time dependent. A *full automaton* for player i is a 4-tuple $\langle M, q_1, f, g\rangle$ where the set of states M , the initial state $q_1 \in M$ and the action function $f: M \to A_i$ are as in a (standard) automaton , and the transition function $g: M \times A \to M$ specifies the next state as a function of the current state and the n-tuple of actions of all players. The strategy σ^i_α induced by a full automaton α= $\langle M, q_1, f, g\rangle$ is defined naturally by $\sigma^i_1 = f(q_1)$ and for every $a_1, \ldots, a_{t-1}$ in A, $\sigma^i(a_1, \ldots, a_{t-1}) = f(q_t)$ where q_t is defined inductively for $t > 1$ by $q_t(a_1, \ldots, a_{t-1}) = g(q_{t-1}, a_{t-1})$. Obviously, every strategy which is induced by a full automaton of size m is equivalent to a strategy induced by an automaton of size m. Therefore when the actions and transitions are deterministic, allowing transitions to depend on your own action is not affecting the equilibrium theory. However it does have implications in the study of subgame perfect equilibrium of repeated games (see e.g. Kalai and Stanford (1988) and Ben-Porath and Peleg (1987)). Kalai and Stanford (1988) show that given a pure strategy σ^i of a player in the repeated game, the number of different strategies induced by it and any finite history $(a_1, \ldots, a_t)$ equals the size of the full automaton that induces σ^i. A *mixed action automaton* for player i is a 4-tuple $\langle M, q_1, f, g\rangle$ where M is a finite set, $q_1 \in M$ is the initial state, $f : M \to \Delta(A_i)$ is a function specifying a mixed action as a function of the state, and $g : M \times A \to M$ is the transition function. Each mixed action automaton induces a behavioral strategy σ^i as follows. $\sigma_1^i = f(q_1)$. Define inductively $q_t(a_1, \ldots, a_{t-1}) = g(q_{t-1}, a_{t-1})$, and

$$\sigma_t^i(a_1, \ldots, a_{t-1}) = f(q_t(a_1, \ldots, a_{t-1})).$$

Denote by $\Sigma_p^i(m_i)$ all equivalence classes of behavioral strategies which are induced by a mixed action automata of size m_i. Two mixed (or behavioral) strategies, σ^i and τ^i, of player i are *equivalent* if for any $N \setminus \{i\}$-tuple of pure strategies σ^{-i}, (σ^i, σ^{-i}) and (τ^i, σ^{-i}) induce the same distribution on the play of the repeated game. Note that $\Sigma^i(m_i) \subset \Sigma_p^i(m_i)$ and that $\Sigma_p^i(1) \setminus \Delta(\cup_{m=1}^{\infty} \Sigma^i(m)) \neq \emptyset$. A stationary behavioral strategy in a repeated game with complete information is induced by a mixed action automaton with one state. Given a behavioral strategy $\sigma^i = (\sigma_t^i)_{t=1}^{\infty}$, the number of equivalence classes of the behavioral strategies of the form $(\sigma^i \mid a_1, \ldots, a_t)$ where $a_1, \ldots, a_t$ ranges over all histories which are consistent with σ^i, (i.e., $(\sigma^i \mid a_1, \ldots, a_s)(a_{s+1}^i) > 0$ for every $s < t$), equals the size of the smallest mixed action automaton that implements σ^i. A *probabilistic transition automaton* is a 4-tuple $\langle M, q_1, f, g \rangle$ where M is a finite set, $q_{\in} M$ is the initial state, $f : M \to A_i$ is the action, and $g : M \times A_{-i} \to \Delta(M)$. Each probabilistic transition automaton induces a mixed strategy σ^i as follows. $\sigma_1^i = f(q)$. Then the automaton changes its states stochastically in the course of playing the repeated game. If its state in stage t is q_t and the other players action in stage t is a_t^{-i} the conditional probability of q_{t+1}, given the past is $g(q_t, a_t^{-i})$, and its action in stage t is $f(q_t)$. Denote by $\Sigma_t^i(m_i)$ all equivalence classes of strategies which are induced by probabilistic transition automata of size m_i. Note that $\Sigma_p^i(m) \subset \Sigma_t^i(m|A_i|)$.

Repeated games with complete information. The theory of finitely or infinitely repeated 2-person 0-sum games with complete information and either mixed action or probabilistic transition automata is trivial and not of much interest. However, the asymptotic behavior of the set of equilibrium payoffs of n-person ($n \geq 3$) infinitely repeated or 2-person finitely repeated games with either mixed action or probabilistic transition automata is unknown and of interest. The difficulties in the study of equilibrium payoffs of n-person infinitely repeated games is the asymptotics of the minmax payoffs which is unknown. As for 2-player finitely repeated games with either mixed action or probabilistic transition automata, it seems that our constructed equilibrium (Neyman 1995) in the finitely repeated games, remains an equilibrium in the game in which players are restricted to play with either mixed action or probabilistic automata with the same bounds. Note that as $\Sigma^i(m)$ is a proper subset of $\Sigma^i_p(m)$ (and of $\Sigma^i_t(m)$), the assertion that $\sigma^* \in \Delta(\Sigma^1(m_1)) \times \Delta(\Sigma^2(m_2))$ is an equilibrium in $(\{1,2\}; \Sigma^1_p(m_1), \Sigma^2_p(m_2); r_T)$ (in $(\{1,2\}; \Sigma^1_t(m_1), \Sigma^2_t(m_2); r_T)$) is stronger than the assertion that it is an equilibrium in $G^T(m_1, m_2)$. Moreover, in this setup, holding a player down to his individual rational payoff requires just one state (a fixed finite number of states) and therefore in the theorems there is only a need to bound the size of one of the automata.

Repeated games with incomplete information. The theory of repeated games with incomplete information and either probabilistic or deterministic action function is of interest. Here the initial state is allowed to be a function of the initial information, or equivalently, the initial move of nature is part of the input at stage 0. Alternatively, one allows the action function to depend on the state of the machine and the information about the state of nature. It is relatively easy to verify that in the case of 2-person 0-sum repeated games with incomplete information on one side the value of the "restricted game" $?(p, m_1, m_2)$ converges to $\lim v_n(p)$ as $m_i \to \infty$ and $(\log \max\{m_1, m_2\})/\min\{m_1, m_2\} \to 0$. It is of interest to find whether in repeated games with incomplete information on both sides and under the above asymptotic conditions on m_1 and m_2 the values of $?(p, m_1, m_2)$ converge to a limit and whether this limit equals $\lim v_n(p)$.

Stochastic Games. My initial interest in the theory of repeated games with finite automata stemmed from my work with J.-F. Mertens on the existence of a value in stochastic games. The ε-optimal strategies exhibited there are behavioral strategies which are not implemented by any finite state mixed action automaton. Blackwell and Ferguson (1968) show that in the "Big Match" there are no stationary (i.e. implemented by a mixed action automaton of size 1) ε-optimal strategies, and it can be shown further that there are no ε-optimal strategies which are implemented by a mixture of mixed action automata of finite size. However, when both players are restricted to strategies that are implemented by either deterministic or mixed action automata of sizes m_1 and m_2 we are faced with a 2-person 0-sum game $G(m_1, m_2)$ in normal form which has a value $V(m_1, m_2)$. It is of interest to study the asymptotic behavior of $V(m_1, m_2)$ as $m_i \to \infty$. Consider the "Big

Match" (Blackwell and Ferguson 1968) which is an example of a 2-person 0-sum stochasic game.

1	0
0^*	1^*

The value of the (unrestricted) undiscounted game is 1/2. Mor Amitai (1989) showed that for this game there is a polynomial function $m : I\!N \to I\!N$ such that the value of the restricted "Big Match" where player 1 and player 2 are resricted to strategies which are implemented by automata of sizes $m(n)$ and n respectively equals 1.

Another generalization of automata suggested by the theory of stochastic games is a time dependent probabilistic action automaton. In a time dependent automaton the action of the automaton depends on its internal state and the stage. This generalizes also the concept of a Markov strategy. Blackwell and Ferguson (1968) have shown that in the "Big Match" there are no ε-optimal Markov strategies. This leads to the natural question as to whether or not there are ε-optimal strategies which are implemented by a finite state time dependent probabilistic automaton. When raising this question in a seminar in Stanford University, Jerry Green has pointed out to me the work of T. Cover (1969) which illustrates a statistical decision problem in which a difference between the stationary finite automata and the time dependent automata emerges. My attention to the topic of repeated games with bounded automata was recalled in discussions I had with Alan Hoffman during his visit to Jerusalem in the spring of 1983. Hoffman informed me that in the early fifties, when engineers at SEAC were actually playing tick-tack-toe on the SEAC, he was concerned on how game theorists will view/study the fact that a 2-person 0-sum fair (value 0) game, becomes an unfair game when players are restricted by their "programs". This triggered my attention to pose the problem settled by Ben-Porath and later to the study of the possible cooperation in finitely repeated games with bounded automata. I am indepted to each one of the above mentioned individuals for their influence, either directly or indirectly, on my working on repeated games with finite automata.

Mor Amitai proved the following interesting result concerning the maxmin of repeated games with absorbing states and probabilistic transition automata: for any repeated game with absorbing states and $\varepsilon > 0$, there exists a constant m such that for any m_1 and any strategy $\sigma^1 \in \Sigma_t^1(m_1)$ there exists a strategy $\sigma^2 \in \Sigma_t^2(m)$ such that

$$r_\infty^1(\sigma^1, \sigma^2) \leq \sup \inf r_\infty^1(\sigma, \tau) + \varepsilon$$

where the sup ranges over all stationary strategies of player 1 and the inf ranges over all strategies of player 2.

References

Amitai, M. (1989). "Stochastic Games with Automata", (Hebrew) MSc. Thesis, Hebrew University, Jerusalem.

Aumann, R. J. (1981). "Survey of Repeated Games" in *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern*, 11-42. Mannheim/Vienna/Zurich:Bibliographisches Institute.

Aumann, R. J., and S. Sorin (1990). "Cooperation and Bounded Recall," *Games and Economic Behavior* **1** 5-39.

Ben-Porath, E., (1986). "Repeated Games with Bounded Complexity," (Hebrew) MSc. Thesis, Hebrew University, and Stanford University (mimeo).

Ben-Porath, E., (1993). "Repeated games with finite automata," *Journal of Economic Theory* **59** 17-32.

Ben-Porath, E., and B. Peleg (1987). "On the Folk Theorem and Finite Automata." The Hebrew University, Research Memorandum 77, Center for Research in Mathematical Economics and Game Theory.

Benoit, J.-P. and V. Krishna (1985), "Finitely Repeated Games," *Econometrica* 53, 905-922.

Benoit, J.-P. and V. Krishna (1987), "Nash Equilibria of Finitely Repeated Games," *International Journal of Game Theory* 16, 197-204.

Blackwell, D. and T. S. Ferguson (1968). "The Big Match,"*Ann. Math. Stat.*, **39** 159-163.

Cover, T. M. (1969). "Hypothesis Testing with Finite Statistics," *Annals of Mathematical Statistics* **40** 828-835.

Forges, F., Mertens J.-F., and A. Neyman (1986), "A Counter Example to the Folk Theorem with Discounting," *Economics Letters* **20** 7.

Hopcroft, J., and J. Ullman (1979). "Introduction to Automata Theory, Languages, and Computation." Reading, Massachusetts: Addison-Wesley.

Kalai, E. (1990). "Bounded Rationality and Strategic Complexity in Repeated Games," *Game Theory and Applications*, edited by T. Ichiishi, A. Neyman and Y. Tauman, Academic Press, 131-157.

Kalai, E., and W. Stanford (1988). "Finite Rationality and Interpersonal Complexity in Repeated Games," *Econometrica* **56** 397-410.

Lehrer, E. (1988). "Repeated Games with Stationary Bounded Recall Strategies," *Journal of Economic Theory* **46** 130-144.

Lehrer, E. (1994). "Many Players with Bounded Recall in Infinite Repeated Games." *Games and Economic Behavior* **7** 390-405.

Megiddo, N., and A. Wigderson (1986). "On Play by Means of Computing Machines," in *Proc. 1st Cong. on Theoretical Aspects of Reasoning and Knowledge"* (J. Halpern, ed.), 259-274. Los Altos, California, Morgan Kaufman.

Mertens, J. F., and A. Neyman (1981). "Stochastic games," *International Journal of Game Theory.* 53-66.

Neyman, A. (1985). "Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoner's Dilemma," *Economics Letters.* **19**, 227-229.

Neyman, A. (1995). "Finitely Repeated Games with Finite Automata", Discussion Paper no. 69, Center for Rationality and Interactive Decision Theory. Submitted to *Mathematics of Operations Research.*

Papadimitriou, C. H., and M. Yannakakis (1994). "On Complexity as Bounded Rationality" (extended abstract). STOC - 94 , 726-733.

Papadimitriou,C. H., and M. Yannakakis (1995). "On Bounded Rationality and Computational Complexity" Mimeo.

Zemel, E. (1989). "Small Talk and Cooperation: A Note on Bounded Rationality," *Journal of Economic Theory* **49**, 1-9.