TEL-AVIV 7-91

The Sackler Institute of Economic Studies

מכון סאקלר לכלכלה
אוניברסיטת תל-אביב

REPEATED GAMES WITH FINITE AUTOMATA*

by

Elchanan Ben–Porath**

Working Paper No.7–91

August, 1991

* A first version of this paper appeared in 1986 under the name "Repeated Games with Bounded Complexity." The paper is based on my M.S. thesis. My advisor, Professor Abraham Neyman, introduced me to the topic and guided me. I owe him many thanks.

** Department of Economics, Tel–Aviv University and Department of Managerial Economics and Decision Sciences, J.L.Kellogg Graduate School of Management, Northwestern University.

THE SACKLER INSTITUTE OF ECONOMIC STUDIES
Faculty of Social Sciences
Tel–Aviv University, Ramat Aviv, Israel.

ABSTRACT

The paper examines the asymptotic behavior of the set of equilibrium payoffs in a repeated game when there are bounds on the complexity of the strategies players may select. The complexity of a strategy is measured by the size of the minimal automaton that can implement it.

The main result is that in a zero–sum game, when the size of the automata of both players go together to infinity, the sequence of values converges to the value of the one–shot game. This is true even if the size of the automata of one player is a polynomial of the size of the automata of the other player. The result for the zero–sum games gives an estimation for the general case.

1. INTRODUCTION

In this paper I examine the asymptotic behavior of the set of equilibrium payoffs in a repeated game when there are bounds on the complexity of the strategies that players may select. The main part of the paper is devoted to the analysis of zero–sum games. (The extension to general games is then relatively simple.) In particular we are interested in the following problem: Let G be a zero–sum game. If player I is restricted to strategies below complexity n, what is the complexity that is required from player II in order to gain something from player I's limitation?

The interest in putting bounds on the complexity of the strategies stems from the limited computational ability of humans and devices used by humans (see Simon [8], [9]). For example, most of the strategies in a repeated game cannot be implemented by any computer.

It is important to distinguish between the complexity of a strategy and the complexity of the process of selecting a strategy. We will not deal with the selection process directly. We will assume that the limitation of a player is such that he can consider all the strategies below a certain level of complexity. A possible interpretation of this set–up is that the players' abilities are unbounded but they use bounded devices to implement their strategies (for example – computers.)

We use the notion of a finite automaton to define a complexity measure on the strategies. A finite automaton is a machine which has a finite number of states. One of these states is the initial state. The machine has an action function and a transition function. The action function determines the one–shot game action that is played at each state. The transition function specifies the next state as a function of the current state and the current actions of the other players. An automaton induces a strategy as follows: The state at the first stage is the initial state. The state and the actions of the other players at stage t determine the state at stage t+1 (by the transition function). The state determines the one–shot game strategy (by the action function). The size of an automaton is the number of states it has. The complexity of a strategy is defined as the size of the minimal automaton that can implement it.

Let G be a zero–sum game and let V(G) denote its value. Let $V(G_{n,m})$ denote the value of the infinitely repeated game where players I and II are restricted to strategies that can be implemented by automata of size n and size m respectively. The main result (Theorem 2) is that if $Q: N \to N$ is a function such that $Q(n) \geq n$ and $\lim_{n\to\infty} \frac{\ell n[Q(n)]}{n} = 0$ (where $\ell n[Q(n)]$ is the natural log of Q(n)) then $\lim_{n\to\infty} V(G_{n,Q(n)}) = V(G)$. In particular, if Q(n) is a polynomial, then $\lim_{n\to\infty} V(G_{n,Q(n)}) = V(G)$. We also show (Theorem 1) that if Q(n) is big enough $(Q(n) \geq \exp[c\ell n(n)]$, where exp means the exponent and c is a constant) then $V(G_{n,Q(n)})$ is maxmin(G) in pure strategies (which is smaller and typically strictly smaller than the value.) This means that a player can gain from using strategies which are more complicated than his opponent's strategies only if his strategies are much more complicated. The results for zero–sum games provide an estimation for the asymptotic behavior of the set of Nash equilibria payoffs in a general repeated game.

The idea of using a finite automaton in order to distinguish between simple and complicated strategies was proposed by Auman [1]. The first two studies of the model were done (independently of each other) by Neyman [6] and Rubinstein [7]. The two works differ in their interpretations and goals. Neyman shows that cooperation can be achieved in the finitely repeated prisoners' dilemma, if there are bounds (even very large bounds) on the complexity of the strategies that players may use. In Rubinstein's paper, the complexity of the strategies is determined endogenously. Players seek, on the one hand, to maximize their payoff and on the other to minimize the complexity of their strategies. It is shown that in the infinitely repeated prisoners' dilemma this behavior considerably restricts the set of equilibrium strategies.

Since the appearance of the first version of the current paper (Ben–Porath [2]), there have been many studies of complexity in games. We will mention here only two related works, and refer the reader to a survey by Kalai [3]. Lehrer [5] addresses a problem that is similar to the one we study here, but he uses a different notion of complexity. The complexity of a strategy is measured by the length of the recall that the strategy requires. A t–bounded recall strategy (t–BRS) is a strategy where the action of the player at each stage depends only on the play in the preceding t stages. Although the complexity measure in Leher's model is different, his results are similar to the ones obtained here.[1] The main difference between restricting the player to bounded recall strategies and to finite automata is that, while in both cases there is a bound on the amount of information that the player can use at each stage, an automaton gives the player a certain flexibility in deciding what information will be retained. (When a player uses a bounded recall strategy he takes into account only recent information.) In many cases, while there are serious bounds on the amount of information that the decision maker can process, it is not difficult to focus on specific data even if it is not recent. In particular there are strategies in repeated games that are intuitively simple yet are infinitely complex according to the recall measure. (An example is given in the Appendix).

---

[1]Specifically, let $V(\overline{G_{n,Q(n)}})$ denote the value of the repeated game where players I and II are restricted to n–BRS and Q(h)–BRS respectively. The main result is that if $Q(n) \geq n$ and $\lim_{n\to\infty} \frac{\ell n[Q(n)]}{n} = 0$, then $\lim_{n\to\infty} V(\overline{G_{n,Q(n)}}) = V(G)$.

Kalai and Stanford [4] study strategic complexity in repeated games with discounting. They show that any subgame perfect equilibrium in the supergame can be approximated by an ϵ–subgame perfect equilibrium with strategies of finite complexity. They also show that in "robust" equilibria of generic 2–person games, players use strategies of the same complexity. This result appears to contrast with theorem 1 in this paper because it implies that in a zero–sum game a player does not gain from being able to choose complicated strategies. The reason for this difference is that in our model players can choose automata randomly while in KS they cannot. In particular, in zero–sum games there is typically no equilibrium if players cannot randomize. Therefore the results of KS do not apply for such games. The main effort in my paper is in determining the level of complexity that is required of player II in order to beat player I, when player I is restricted to relatively simple strategies and tries to "defend" himself by randomizing.

The paper is organized as follows: The model is described in section 2. Section 3 considers zero–sum games and contains the main theorem. Section 4 extends the results to N–person games and section 5 concludes with a simple extension to a class of complexity measures.

2. THE MODEL

Let G be a zero–sum game, $G = (S^1, S^2, r)$, where S^i is a finite set of actions for player i $(i = 1,2)$ and $r: S^1 \times S^2 \to R$ is the payoff function of player I. Let $V(G)$ denote the value of the game and let maxmin(G) and minmax(G) denote $\max_{s^1 \in S^1} \min_{s^2 \in S^2} r(s^1,s^2)$ and $\min_{s^2 \in S^2} \max_{s^1 \in S^1} r(s^1,s^2)$ respectively.

An automaton A^i for player i is a four–tuple $<M^i, \bar{q}^i, f^i, g^i>$ where M^i is a set, $\bar{q}^i \in M^i$, $f^i: M^i \to S^i$, and $g^i: M^i \times S^j \to M^i (j \neq i)$. These symbols have the following interpretation: M^i is the set of states of the automaton, $\bar{q}^i$ is the initial state, $f^i(q^i)$ is the action the player chooses when the automaton is at state q^i, and $g^i(q^i, s^j)$ is the state of the automaton at the next stage, if at the current stage the automaton is at state q^i and the other player plays the action s^j. An automaton is finite if the set of states is finite. We will consider only finite automata. The size of an automaton is the number of the states it has. An automaton of player i induces a (pure) strategy in the repeated game as follows: The action at stage t is $f^i(q^i_t)$ where q^i_t is the state of the automaton at stage t. The sequence of states is determined inductively by $q^i_1 = \bar{q}^i$, $q^i_{t+1} = g^i(q^i_t, s^j_t)$ where s^j_t is the action of the other player at stage t. For example, consider the game described in figure 1.

The strategy of player I which begins with T, continues with it as long as player II chooses L and plays B forever if player II plays R, is induced by the automaton $A = \langle M,q,f,g \rangle$; where

$$M = \{1,2\};\ q = 1;\ f(1) = T\ \ f(2) = B$$

$$g(1,L) = 1\ \ g(1,R) = 2\ \ g(2,L) = g(2,R) = 2.$$

Given that the automata of the players are A^1 and A^2, the corresponding strategies in the repeated game determine a sequence of actions and payoffs. Denote by $R_t(A^1, A^2)$ the payoff at stage t. The payoff when player I chooses A^1 and player II chooses A^2 is defined to be the limit of the means:[2]

---

[2]Since the set of the states of each automaton is finite the automata enter a cycle, i.e., there exists numbers $c, k \leq |M^1||M^2|$ such that for every $t \geq c$ $(q^1_t, q^2_t) = (q^1_{t+k}, q^2_{t+k})$ and so the limit (1) exists.

$$(1) \qquad R(A^1,A^2) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} R_t(A^1,A^2).$$

We are interested in the value of the game where each player is restricted to strategies that can be implemented by an automaton of a given size.

Formally define:

$$\mathcal{A}_n^1 = \{A^1 \mid A^1 \text{ is an automaton of size n for player I}\}$$
$$\mathcal{A}_m^2 = \{A^2 \mid A^2 \text{ is an automaton of size m for player II}\}$$
$$G_{n,m} = (\mathcal{A}_n^1, \mathcal{A}_m^2, R).$$

Thus $G_{n,m}$ is the game induced by restricting player I and player II to strategies that can be implemented by automata of size n and m respectively. Note that $G_{n,m}$ is also a zero–sum game.

I assume without loss of generality that the set of states of an automaton of size n is $\{1,\dots,n\}$. With this identification $\mathcal{A}_n^1$ and $\mathcal{A}_m^2$ are finite. Let $k = |S^1|$ and $h = |S^2|$. Then $|\mathcal{A}_n^1| = n \cdot k^n \cdot n^{nxh}$ and $|\mathcal{A}_m^2| = m \cdot h^m \cdot m^{mxk}$.

3. THE ASYMPTOTIC BEHAVIOR OF $V(G_{n,m})$

The main result in this section is that if $p(n)$ is a polynomial, then $\lim_{n\to\infty} V(G_{n,p(n)}) = V(G)$.

First note that player I can get at least the maxmin (G) by playing constantly the maxmin action. An automaton of size one can implement this strategy and, of course, any larger automaton can do it as well. Similarly player II can get minmax (G). Thus the following inequalities hold

$$\text{maxmin}\,(G) \le V(G_{n,m}) \le \text{minmax}\,(G).$$

The first result states that for any number n there exists a larger number m such that $V(G_{n,m}) = \text{maxmin } G$.

First we need a definition and a lemma.

DEFINITION:

(a) A^i is a partial automaton for player i $(i = 1,2)$ if the transition function, g^i, is defined on a subset of $M^i \times S^j$.

(b) For two partial automata A and A', A' is an extension of A if its transition function is an extension of the transition function of A. If A' is an automaton it will be called a completion of A.

LEMMA (1.1): For every $A^1 \in \mathcal{A}^1_n$ there exists a partial automaton, $\overline{A^1}$, of size n for player II such that for any extension, A^2, of $\overline{A^1}$ $R(A^1,A^2)$ is well defined (i.e. when A^1 and A^2 play the transition of A^2 is defined at every stage) and:

(2) $$R(A^1, A^2) \le \text{maxmin } G$$

PROOF: Define

$$h\colon\ S^1 \to S^2$$
$$h(s^1) = \text{argmin}_{s^2 \in S^2}\, r(s^1,s^2).$$

$h(s^1)$ is the best reaction for player II to the action s^1 of player I. Denote the different states of A^1 by $\{1,\ldots,n\}$ and assume that the initial state is 1. Define the partial automaton $\overline{A^1} = \langle M^2, \bar{q}^2, f^2, g^2 \rangle$ as follows:

$$M^2 = \{1,\ldots,n\} \quad \bar{q}^2 = 1 \quad f^2(i) = h(f^1(i)) \quad 1 \le i \le n.$$

g^2 is a partial function satisfying:

$$g^2(i, f^1(i)) = g^1(i, f^2(i)) \quad 1 \le i \le n.$$

It is easy to see that g^2 can be defined in such a way and that every extension A^2 of $\overline{A^1}$ satisfies (2).

<u>Theorem 1</u>: If $L(n) \ge n|\mathcal{A}^1_n| = n^2 \cdot k^n \cdot n^{nxh}$ there exists an automaton $A^2 \in \mathcal{A}^2_{L(n)}$ such that for every $A^1 \in \mathcal{A}^1_n$ $R(A^1, A^2) \le \text{maxmin}(G)$.

We will construct A^2. Roughly, A^2 operates as follows: it first identifies the automaton A^1 and then uses a subautomaton from the type described in Lemma 1.1.

<u>Proof</u>: Let A^1_ℓ, $1 \le \ell \le |\mathcal{A}^1_n|$, be an ordering of the automata of player I. With every automaton A^1_ℓ associate n similar partial automata $A^2_{\ell,1},\ldots,A^2_{\ell,n}$ from the type defined in lemma 1.1. (Therefore each partial automaton $A^2_{\ell,r}$ has n states and $R(A^1_\ell, A^2_{\ell,r}) \le$ maxmin(G)$\cdot$).

The set of states of any two different partial automata, $A^2_{i,r}$ and $A^2_{j,z}$ ($i \neq j$ or $r \neq z$), are disjoint. Together they form a partial automaton with $n^2 |\mathcal{A}^1_n|$ states. Define the initial state to be the initial state of $A^2_{1,1}$. Denote this partial automaton by A^2_1. We will extend the automaton inductively so that the partial automaton at stage p, A^2_p will satisfy

(3) $$R(A^1_s, A^2_p) \leq \text{maxmin}(G)$$
for $1 \leq s \leq p$.

The extension is complete when we define $A^2_{|\mathcal{A}^1_n|}$. We then show that only $n|\mathcal{A}^1_n|$ states are actually used and we can discard the rest.

So assume that A^2_p has been defined and define A^2_{p+1}. Play A^1_{p+1} against A^2_p. If there is some $1 \leq s \leq p$ such that the sequence of actions of A^1_{p+1} equals to the sequence of actions when A^1_s plays against A^2_p then $R(A^1_{p+1}, A^2_p) = R(A^1_s, A^2_p) \leq$ maxmin(G) and we define $A^2_{p+1} = A^2_p$. Otherwise, A^1_{p+1} "reveals its identity" at some time, i.e. there exists a time t such that the sequence of actions played by A^1_{p+1} – $(s^1_1,\ldots,s^1_t)$ – is different from the action sequences of the automata $A^1_1,\ldots,A^1_p$. Assume that t is the first time when this happens. Let q^1_{t+1} denote the state of A^1_{p+1} at time t+1 (this state is determined at time t) and let $A^1_{p+1}(q^1_{t+1})$ denote the automaton which is identical to A^1_{p+1} but has q^1_{t+1} as an initial state. We extend A^2_p to A^2_{p+1} so that starting from stage t+1, A^2_{p+1} will play against A^1_{p+1} like $\overline{A^1_{p+1}(q^1_{t+1})}$ (the partial automaton defined in Lemma 1.1.) Let i(p+1) denote the index of $A^1_{p+1}(q^1_{t+1})$ in the order. Let c(p+1) denote the index of the first copy among $A^2_{i(p+1),1},\ldots,A^2_{i(p+1),n}$

that hasn't been used yet (in the process of the extension.) Finally let q_t^2 denote the state of A_p^2 at time t. The inductive step is to set $g^2(q_t^2, s_t^1)$ to be equal to the initial state of $A^2_{i(p+1),c(p+1)}$. A simple induction shows that for every p, $1 \leq p \leq |\mathscr{A}_n^1|$, A_{p+1}^2 is an extension of A_p^2 and (3) is satisfied. So for every $1 \leq \ell \leq |\mathscr{A}_n^1|$, $R(A_\ell^1, A^2_{|\mathscr{A}_n^1|}) \leq$ maxmin(G).

Now since there are only $|\mathscr{A}_n^1|$ automata for player I at most $|\mathscr{A}_n^1|$ partial automata from the set $\{A^2_{\ell,r} | 1 \leq \ell \leq |\mathscr{A}_n^1|, 1 \leq r \leq n\}$ are actually used in the extension that we described. Let B denote this set of partial automata and let $\hat{A}^2$ denote the partial automaton that is derived from $A^2_{|\mathscr{A}_n^1|}$ by discarding all the states that do not belong to a member in B. (The action and transition functions of $\hat{A}^2$ are restrictions of the respective functions of $A^2_{|\mathscr{A}_n^1|}$.) $\hat{A}^2$ has at most $n|\mathscr{A}_n^1|$ states and it is straightforward to check that if A^2 is an extension of $\hat{A}^2$ then $R(A_p^1, A^2) = R(A_p^1, A^2_{|\mathscr{A}_n^1|}) \leq$ maxmin(G) for every $1 \leq p \leq |\mathscr{A}_n^1|$. This implies the result.

<u>THEOREM 2</u>:

Let $Q(n)$ be a function which satisfies $Q(n) \geq n$ and $\lim_{n\to\infty} \frac{\ln[Q(n)]}{n} = 0$. Then $\lim_{n\to\infty} V(G_{n,Q(n)}) = V(G)$.

<u>PROOF</u>: We will show that $\underline{\lim}_{n\to\infty} V(G_{n,Q(n)}) \geq V(G)$. A similar argument proves that $\overline{\lim}_{n\to\infty} V(G_{n,Q(n)}) \leq V(G)$. The two inequalities imply the result.

For each n we will define a mixed strategy P^n for player I (i.e., a probability distribution on $\mathscr{A}_n^1$), such that for every $A^2 \in \mathscr{A}^2_{Q(n)}$ $\sum_{A^1 \in \mathscr{A}_n^1} P^n(A^1) \cdot R(A^1, A^2) \geq$

$V(G) - \epsilon_n$ where $\lim_{n\to\infty} \epsilon_n = 0$.

Assume $S^1 = \{1,\dots,k\}$ and $S^2 = \{1,\dots,h\}$. Define $\Omega_n = \{S^1\}^n$. Let $\omega \in \Omega_n$, $\omega = (\omega_1,\dots,\omega_n)$. Denote by A^1_ω ($A^1_\omega \in \mathcal{A}^1_n$) the following automaton:

$$M^1 = \{1,\dots,n\},\ \bar{q}^1 = 1,\ f^1(i) = \omega_i,\ g(i,s^2_j) = \begin{cases} i+1 & i < n \\ 1 & i = n \end{cases}.$$

Let $(P_1,\dots,P_k)$ be a mixed strategy for player I in the game G, such that $\sum_{i=1}^{k} P_i \cdot r(i,j) \geq V(G)$ for every $j \in S^2$. Define a probability measure on $(\Omega_n, 2^{\Omega_n})$ by

$$\mu_n(\omega) = \Pi_{i=1}^{n} P_{\omega_i}.$$

The strategy P^n of player I in the game $G_{n,Q(n)}$ is defined by

$$P^n(A^1_\omega) = \mu_n(\omega).$$

We have to show that,

(1) For every $\epsilon > 0$ there exists $N(\epsilon)$ such that for every $n > N(\epsilon)$ and for every $A^2 \in \mathcal{A}^2_{Q(n)}$ the following is satisfied:

(1′) $$\sum_{\omega\in\Omega_n} \mu_n(\omega) R(A^1_\omega, A^2) \geq V(G) - \epsilon.$$

It suffices to show that:

(2) For every $\epsilon > 0$ there exists $N(\epsilon)$ such that for every $n > N(\epsilon)$, $A^2 \in \mathscr{A}^2_{Q(n)}$ and $c \in N$ (2′) is satisfied.

$$(2') \qquad \sum_{\omega\in\Omega_n} \mu_n(\omega) \cdot \frac{1}{n} \sum_{t=c\cdot n+1}^{(c+1)n} R_t(A^1_\omega, A^2) \geq V(G) - \epsilon.$$

Let $A^2 \in \mathscr{A}^2_{Q(n)}$ and assume $M^2 = \{1,\ldots,Q(n)\}$. Let $A^2(p)$, $p = 1,\ldots,Q(n)$, denote the automaton which is identical to A^2 but with initial state p. Let $q(A^2,\omega,c)$ denote the state of the automaton A^2 in the stage $c \cdot n+1$, when player I chooses A_ω. Finally define:

$$(3) \qquad X_{A^2,c}(\omega) \overset{def}{=} \frac{1}{n} \sum_{t=c\cdot n+1}^{(c+1)\cdot n} R_t(A^1_\omega, A^2) = \frac{1}{n} \sum_{t=1}^{n} R_t(A^1_\omega, A^2(q(A^2,\omega,c))).$$

Note that the left expression in (2′) is the expectation of $X_{A^2,c}$. To estimate the expectation and compare it to $V(G)$, we will estimate:

$$(4) \qquad \mu_n\{\omega : X_{A^2,c}(\omega) < V(G) - \epsilon\}.$$

The important point in the proof is that the state of an automaton $\hat{A}^2$ for player II in stage ℓ, $1 \leq \ell \leq n$, is determined by the first $\ell - 1$ actions of player I. Thus, given an automaton $\hat{A}^2$ we can associate with every action $j = 1,\ldots,h$ of player II a sequence of random variables $(f_{j,\ell})_{\ell=1,\dots,n}$

$$f_{j,\ell}(\omega) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{the action of } \hat{A}^2 \text{ at stage } \ell \text{ when it plays against } A_\omega \text{ is } j \\ 0 & \text{otherwise} \end{cases}$$

and $f_{j,\ell}$ are measurable w.r.t. the algebra which is generated by the first $\ell-1$ elements. We need two lemmas. Call a sequence $\omega \in \Omega_n$ "nice", if the number of times the actions (i,j) are played divided by the number of times j is played, is 'close' to P_i (the probability of the action i in a mixed strategy that ensures V(G) for player I.) Lemma 2.1 (the main lemma) says that almost all the sequences are "nice". Lemma 2.2 states that in a "nice" sequence the average payoff is 'close' to V(G).

<u>LEMMA 2.1</u>. Let $A = \{a_1,\dots,a_k\}$ be a finite set. Let $(A, 2^A, P)$ be a probability space and let $(\Omega, 2^\Omega, \mu)$ be the product space $(A, 2^A, P)^n$ i.e., $\Omega = A^n$ and for every $\omega = (\omega_1,\dots,\omega_n) \in A^n$ $\mu(\omega) = \Pi_{i=1}^n P(\omega_i)$. Let H_ℓ be the partial algebra of 2^Ω which is generated by $\omega_1,\dots,\omega_{\ell-1}$ $H_1 = \{\phi,\Omega\}$. Let $\{f_{j,\ell}\}_{\ell=1,\dots,n}^{j=1,\dots,h}$ be a set of random variables which are adapted to $(H_\ell)_{\ell=1,\dots,n}$ (i.e., $f_{j,\ell}$ is measurable w.r.t. (Ω, H_ℓ)) and with values in the set $\{0,1\}$. Define

$$S_\epsilon = \left\{ \begin{array}{l} \omega: \left| \frac{1}{n} \sum_{\ell=1}^n I(\omega_\ell = a_i)(\omega) f_{j,\ell}(\omega) - \frac{1}{n} \sum_{\ell=1}^n P(a_i) f_{j,\ell}(\omega) \right| > \epsilon \\ \text{for some} \quad 1 \le i \le k \ 1 \le j \le h. \end{array} \right\}$$

There exists $b > 0$ such that for every $b > \epsilon > 0$ $\mu(S_\epsilon) \leq 2 \cdot k \cdot h \cdot e^{-\frac{\epsilon^2 n}{4}}$.

In Lemma 2.2 Ω_n and $f_{j,\ell}$ refer to $\{S^1\}^n$ and the indicator functions of the actions respectively, as defined before Lemma 2.1. Let $W(G)$ denote $\max_{s^1 \in S^1, s^2 \in S^2} |r(s^1,s^2)|$.

LEMMA 2.2: Define

$$S_{\epsilon,\hat{A}^2} = \left\{ \begin{array}{l} \omega: \left| \frac{1}{n} \sum_{\ell=1}^{n} I(\omega_\ell = i)(\omega) f_{j,\ell}(\omega) - \frac{1}{n} \sum_{\ell=1}^{n} P_i f_{j,\ell}(\omega) \right| > \epsilon \\ \text{for some} \quad 1 \leq j \leq k \; 1 \leq i \leq h. \end{array} \right\}$$

If $\omega \in \Omega_n - S_{\epsilon,\hat{A}^2}$ then $V(G) - \frac{1}{n}\sum_{t=1}^{n} R_t(A^1_\omega, \hat{A}^2) < W(G) \cdot k \cdot h \cdot \epsilon$.
The proofs of the lemmas are in the Appendix.

We can now evaluate (4). It follows from (3) that:

(5) $$\{\omega : X_{A^2,c}(\omega) < V(G) - \epsilon\} \subseteq \cup_{p=1}^{Q(n)} \{\omega : X_{A^2(p),0}(\omega) < V(G) - \epsilon\}.$$

Let $\eta = \frac{\epsilon}{W(G) \cdot k \cdot h}$. Note that by lemma 2.2 if $\omega \in \Omega_n - S_{\eta,A^2(p)}$ then $X_{A^2(p),0}(\omega) > V(G) - \epsilon$. This and lemma 2.1. imply that for every $p = 1,\ldots,Q(n)$

(6) $$\mu_n\{\omega | X_{A^2(p),0}(\omega) < V(G) - \epsilon\} \leq \mu_n(S_{\eta,A^2(p)}) \leq 2 \cdot k \cdot h \cdot e^{\frac{-\eta^2 n}{4}}.$$

From (5) and (6) we have

$$\mu_n\{\omega : X_{A^2,c}(\omega) < V(G) - \epsilon\} \le \sum_{p=1}^{Q(n)} \mu_n\{\omega : X_{A^2(p),0}(\omega) < V(G) - \epsilon\}$$

$$\le Q(n) \cdot 2 \cdot k \cdot h \cdot e^{\frac{-\eta^2 \cdot n}{4}} .$$

Since $\lim_{n\to\infty} \frac{1}{n} \ell n\, Q(n) = 0$, for every $\eta > 0$ $\lim_{n\to\infty} Q(n) \cdot 2 \cdot k \cdot h \cdot e^{\frac{-\eta^2 n}{4}} = 0$. Hence, for every $\epsilon > 0$ there exists $N(\epsilon)$ such that for any $n \ge N(\epsilon)$, $A_2 \in \mathscr{A}^2_{A(n)}$ and $c \in N$, (7) is satisfied.

(7) $$\mu_n\{\omega \mid X_{A^2,c}(\omega) < V(G) - \epsilon\} < \epsilon.$$

Since $X_{A^2,c}$ are uniformly bounded (i.e. $\forall c, \forall n, \forall A^2 \in \mathscr{A}^2_{Q(n)} \mid X_{A^2,c}(\omega) \mid \le W(G)$) we obtain (2).

We have proved $\underline{\lim}_{n\to\infty} V(G_{n,Q(n)}) \ge V(G)$. Since $Q(n) \ge n$ a similar argument shows that $\lim_{n\to\infty} V(G_{n,Q(n)}) \le V(G)$. These inequalities imply the result.

4. N–PERSON GAMES

In this section we study N–person games. Given a one–shot game G, we can associate with each vector of N natural numbers, $(\alpha_1,...,\alpha_N)$, the set of Nash equilibrium payoffs in the repeated game where player i is restricted to strategies that can be implemented by automata of size α_i.[3] Let $S(\alpha_1,\ldots,\alpha_N)$ denote this set. Let S_m be a sequence of such sets where the sizes of the automata of all the players tend together to infinity. We will use theorem 2 to derive an estimate for the liminf and limsup of the sequence S_m. Let co G denote the convex hull of the vectors in the payoff matrix of G. Let $\underline{S}$ denote the set of individually rational and feasible payoffs in mixed strategies. Formally $\underline{S}$ is defined as follows: Let $\Delta(B)$ denote the set of probability distributions on a set B. Let $z_i = \min_{\tau^{-i}} \max_{\tau^i} r^i(\tau^i, \tau^{-1})$ where $\tau^i \in \Delta(S^i)$ and $\tau^{-i} \in \Pi_{j\neq i} \Delta(S^j)$. Define $\underline{S} = \{x : x \in \text{co } G\ x \geq z\}$. Let $y = (y_1,\ldots,y_N)$ denote the vector of maximal payoffs that each player can get regardless of what the other players do (in mixed strategies), i.e., $y_i = \max_{\tau^i} \min_{\tau^{-i}} \tau^i(\tau^i, \tau^{-i})$ where $\tau^i \in \Delta(S^i)$ and $\tau^{-i} \in \Pi_{j\neq i} \Delta(S^j)$. Define $\overline{S} = \{x : x \in \text{co } G\ x \geq y\}$.

By the Folk Theorem, $\underline{S}$ is the set of Nash equilibrium payoffs in the infinitely repeated game (with mixed strategies). Note that $\underline{S} \subseteq \overline{S}$ and in two–person games $\underline{S} = \overline{S}$.

THEOREM 3:

Let $Q_2(n),\ldots,Q_N(n)$ be functions such that $Q_i(n) \geq n$, $i = 2,\ldots,N$ and $\lim_{n\to\infty} \frac{\ln[Q_i(n)]}{n} = 0$. Then

$$\underline{S} \subseteq \underline{\lim}_{n\to\infty} S(n,Q_2(n),\ldots,Q_N(n)) \subseteq \overline{\lim}_{n\to\infty} S(n,Q_2(n),\ldots,Q_N(n)) \subseteq \overline{S}.$$

It is easy to see that every rational convex combination of the payoff vectors in G can be implemented by automata which are large enough and that conversely, each payoff

---

[3] An automaton for player i is the same as was previously defined, except that the transition function is defined on $M^i \times S^{-i}$ where S^{-i} is the set of action tuples of the other players.

vector in the repeated game with automata is a rational convex combination of the payoff vectors in G. Player i can get y_i by using the strategy that was defined in Theorem 2. Therefore $\overline{\lim}_{n\to\infty} S(n,Q_2(n),...,Q_N(n)) \subseteq \overline{S}$. The other players can bring player i down to his individual rational payoff by using the same type of strategies. More specifically, each pure strategy (i.e. automaton) of player i is composed of an equilibrium phase and punishment phases for each player j, $j \neq i$ which will be implemented if j deviates. It is easy to see that player i can randomize his automata so that if j deviates i will play $\tau_i^{-j}(z_j)$, where $\tau_i^{-j}(z_j)$ is i's part in the combination of mixed strategies that bring j's payoff down to z_j. This implies that $\underline{S} \subseteq \underline{\lim}_{n\to\infty} S(n,Q_2(n),...,Q_N(n))$. The formal proof goes along the lines of the proof of Theorem 2. I omit the details.

<u>COROLLARY</u>: Let G be a 2–person game and let Q(n) be a function that satisfies $Q(n) \geq n$ and $\lim_{n\to\infty} \frac{\ln[Q(n)]}{n} = 0$. Then

$$\lim_{n\to\infty} S(n,Q(n)) = \underline{S} = \overline{S}.$$

So in two–person games there exists a limit set and it is equal to the set of equilibrium payoffs in the super–game.

5. <u>CONCLUSION</u>

So far we have considered a specific measure of complexity. However, a version of Theorem 2 is true for a large class of measures.

DEFINITION: A function $g: N \to N$ is log–polynomial bounded (henceforth l.p.b.) if there exists a polynomial p such that $\ln(g(n)) \leq p(\ln(n))$.

Note that every polynomial is an l.p.b.

Consider a complexity measure as a function from the set of strategies to the natural numbers.

DEFINITION: Two measures of complexity C_1, C_2 are log–polynomial equivalent if there exists a pair of functions $g_1(n), g_2(n) \geq n$ which are increasing and l.p.b. such that

$$C_1^{-1}\{x : x \leq n\} \subseteq C_2^{-1}\{x : x \leq g_1(n)\}$$

$$C_2^{-1}\{x : x \leq n\} \subseteq C_1^{-1}\{x : x \leq g_2(n)\}.$$

Let C be a complexity measure. $G^c_{n,m}$ denotes a game that is similar to $G_{n,m}$ except that the complexity of the strategies is measured by C. (So player I, for example, is restricted to strategies that according to C have a complexity that is less than n.)

THEOREM 4: Let C be a complexity measure that is log–polynomial equivalent to the automata measure and let Q(n) be a l.p.b. function that satisfies $Q(n) \geq n$. Then

$$\lim_{n\to\infty} V(G^c_{n,Q(n)}) = V(G).$$

PROOF: There exists an l.p.b. function g_1 such that player II is restricted to strategies that can be implemented by an automaton of size $g_1(Q(n))$. Since g_1 and Q are l.p.b. there exists polynomials p_1 and p_2 such that $\ln[g_1(Q(n))] \leq p_1[\ln(Q(n))]$ and $\ln[Q(n)]$

$\leq p_2[\ell n(n)]$. Together these inequalities yield:

(8) $$\ell n[g_1(Q(n))] \leq p_1[p_2(\ell n(n))].$$

There exists an l.p.b. function g_2 such that for every $x \in N$ that satisfies $g_2(x) \leq n$ player I can use any strategy that can be implemented by an automaton of size x. Let m be the largest number such that $g_2(m) \leq n$. We have $g_2(m + 1) \geq n$. Since g_2 is l.p.b. there exists a polynomial p_3 such that $\ell n[g_2(m+1)] \leq p_3[\ell n(m)]$. Putting together with (1) we get:

$$\ell n[g_1(Q(n))] \leq p_1(p_2(p_3(\ell n(m)))).$$

A composition of polynomials is a polynomial. Hence $\lim_{m \to \infty} \frac{p_1[p_2[p_3[\ell n(m]]]}{m} = 0$. Since player I can use any automaton of size m while player II is restricted to a subset of the automata of size $g_1(Q(n))$, and because $m \to \infty$ when $n \to \infty$ Theorem 2 implies that $\underline{\lim}_{n \to \infty} V(G^c_{n,Q(n)}) \geq V(G)$. Since $Q(n) \geq n$ a similar calculation gives $\overline{\lim}_{n \to \infty} V(G^c_{n,Q(n)}) \leq V(G)$. The last two inequalities imply the result.

APPENDIX

An example of an unbounded recall strategy that is intuitively simple.

The game that is described in Figure 2 is the prisoners' dilemma. Consider the following strategy for the row player. Start by cooperating (playing C) and continue to do so as long as the opponent cooperates. If the opponent defects in the first stage (by playing D) punish him (play D) forever (regardless of what he does.) If the opponent defects at a stage different from the first, punish him as long as he defects, but if and when he plays cooperatively "forgive" him and continue to play as though the game has just started. In order to play this strategy the player does not need to remember much (in particular this strategy can be implemented by an automaton of size 3) but he has to remember what happened in the first stage. Therefore this is not a finite recall strategy.

PROOF of LEMMA 2.1: We will show that if f_ℓ is measurable w.r.t. H_ℓ then:

$$\forall a \in A \quad \mu\{\omega : \left| \frac{1}{n} \sum_{\ell=1}^{n} I(\omega_\ell = a)(\omega) f_\ell(\omega) - \frac{1}{n} \sum_{\ell=1}^{n} P(a) g_\ell(\omega) \right| > \epsilon\} \leq 2 \cdot e^{\frac{-\epsilon^2 n}{4}} .$$

It is easy to see that this implies the lemma.

PROOF: Define

$$Z_\ell = I(\omega_\ell = a) - P(a)$$

$$Y_\ell = Z_\ell \cdot f_\ell.$$

It suffices to show that there exists $\lambda > 0$ such that

(1) $$\mu\{\omega: |\lambda \sum_{\ell=1}^{n} Y_\ell| > \lambda\epsilon n\} \le 2 \cdot e^{\frac{-\epsilon^2 n}{4}}.$$

We will show that there exists $\lambda > 0$ which satisfies:

(2) $$\mu\{\omega: \lambda \sum_{\ell=1}^{n} Y_\ell > \lambda\epsilon n\} \le e^{\frac{-\epsilon^2 n}{4}}.$$

(3) $$\mu\{\omega: \lambda \sum_{\ell=1}^{n} Y_\ell < -\lambda\epsilon n\} \le e^{\frac{-\epsilon^2 n}{4}}.$$

This implies (1). We will show that (2) and (3) can be proved in a similar way.

Z_ℓ are independent w.r.t. H_ℓ, hence $E(\exp(\lambda Z_\ell) \mid H_\ell) = E(\exp(\lambda Z_\ell))$. Exp is a convex function and therefore:

(4) $$1 = \exp(E(\lambda Z_\ell)) \le E(\exp(\lambda Z_\ell))).$$

By the Taylor expansion

$$\exp(\lambda Z_\ell) = 1 + \lambda Z_\ell + \frac{\lambda^2 Z_\ell^2}{2} + R(\lambda Z_\ell)$$

where

$$\lim_{\lambda \to 0} \frac{R(\lambda Z_\ell)}{\lambda^2 Z_\ell^2} = 0.$$

Hence there exists $d > 0$ such that if $d > \lambda \geq 0$

$$\exp(\lambda Z_\ell) \leq 1 + \lambda Z_\ell + \lambda^2 Z_\ell^2$$

which implies

$$(5) \qquad E(\exp(\lambda \cdot Z_\ell)) \leq 1 + \lambda E(Z_\ell) + \lambda^2 E(Z_\ell^2) \leq 1 + \lambda^2.$$

$Z_1,\ldots,Z_n$ are i.i.d. from (4) and (5) we have

$$1 \leq E(\exp(\lambda \sum_{\ell=1}^{n} Z_\ell)) \leq (1 + \lambda^2)^n.$$

CLAIM: $\forall \ell \;\; 1 \leq \ell \leq N$.

$$(6) \qquad E(\exp(\lambda Y_\ell) \mid H_\ell) \leq E(\exp(\lambda Z_\ell) \mid H_\ell) = E(\exp(\lambda Z_\ell)).$$

PROOF: When $f_\ell = 1$ $Y_\ell = Z_\ell$. When $f_\ell = 0$ $Y_\ell = 0$ and thus the left expression equals one while the right expression is greater or equal to one.

<u>CLAIM</u>: $\forall_j \; 1 \le j \le n.$

(7) $$E(\exp(\lambda \sum_{\ell=1}^{j} Y_\ell)) \le E(\exp(\lambda \sum_{\ell=1}^{j} Z_\ell)).$$

<u>PROOF</u>: By induction. For $j = 1$ the claim follows from (6). Let $j > 1$:

$$E(\exp(\lambda \sum_{\ell=1}^{j} Y_\ell)) = E(E(\exp(\lambda \sum_{\ell=1}^{j=1} Y_\ell) \cdot \exp(\lambda Y_j) | H_j)).$$

Since $\sum_{\ell=1}^{j-1} Y_\ell$ is measurable w.r.t. H_j

$$= E(\exp(\lambda \sum_{\ell=1}^{j-1} Y_\ell) \cdot E(\exp(\lambda Y_j) \mid H_j))$$

$$\le E(\exp(\lambda Z_j)) \cdot E(\exp(\lambda \sum_{\ell=1}^{j-1} Y_\ell)) \text{ by (6)}$$

$$\le E(\exp(\lambda Z_j)) \cdot E(\exp(\lambda \sum_{\ell=1}^{j-1} Z_\ell)) \text{ by the induction hypothesis}$$

$$= E(\exp(\lambda \sum_{\ell=1}^{j} Z_\ell)) \text{ since } Z_1,..,Z_j \text{ are i.i.d.}$$

From (7) and (5) we get

$$E(\exp(\lambda \sum_{\ell=1}^{n} Y_\ell)) \le (1+\lambda^2)^n.$$

By Chebyshev inequality

$$\mu\{\omega : \lambda \sum_{\ell=1}^{n} Y_\ell > \lambda\epsilon n\} \le (1+\lambda^2)^n \cdot \exp(-\lambda\epsilon n).$$

For $\lambda = \frac{\epsilon}{2}$ the right expression is less than $e^{\frac{-\epsilon^2 n}{4}}$ (since $(1+\lambda^2)^n \le e^{\frac{\epsilon^2 n}{4}}$), hence for $0 < \epsilon \le 2d$ (d is the constant derived from the Taylor expansion of Exp) (2) is satisfied.

<u>Proof of Lemma 2.2</u>: Denote:

$\ell_{ij}(\omega)$ – the number of times (i,j) was played

$x_j(\omega)$ – the number of times A^2 played j. $x_j = \sum_{i=1}^{k} \ell_{ij}(\omega)$.

From the definitions

$$\ell_{ij}(\omega) = \sum_{\ell=1}^{n} I(\omega_\ell = i)(\omega) f_{j\ell}(\omega)$$

$$x_j(\omega) \cdot P_i = \sum_{\ell=1}^{n} P_i f_{j\ell}(\omega).$$

Hence if $\omega \in \Omega_n - S_{\epsilon,A^2}$ then for every $1 \le i \le k$ $1 \le j \le h$

$$\left| \frac{\ell_{ij}(\omega)}{x_j(\omega)} - P_i \right| \leq \frac{n}{x_j(\omega)} \cdot \epsilon .$$

From this:

$$\left| \frac{x_j(\omega)}{n} \left(\sum_{i=1}^{k} \frac{\ell_{ij}(\omega)}{x_j(\omega)} r(i,j) - \sum_{i=1}^{k} P_i r(i,j) \right) \right|$$

$$\leq \frac{x_j(\omega)}{n} \sum_{l=1}^{k} \left| \frac{\ell_{ij}(\omega)}{x_j(\omega)} - P_i \right| \cdot \left| r(i,j) \right| \leq k \cdot W(G) \cdot \epsilon .$$

For every $1 \leq j \leq h$ $\sum_{i=1}^{k} P_i \, r(i,j) \geq V(G)$.

Hence, summing over j gives the result.

REFERENCES

1. R.Aumann, Survey of Repeated Games, in "Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern", Bibliographisches Institut, Zurich, 1981, 11–42.
2. E.Ben–Porath, Repeated Games with Bounded Complexity, Mimeo, Stanford University, 1986.
3. E.Kalai, Bounded Rationality and Strategic Complexity in Repeated Games, Discussion Paper No.783, MEDS, Northwestern University, 1988.
4. E.Kalai and W.Stanford, Finite Rationality and Interpersonal Complexity in Repeated Games, Econometrica, 56 (1988), 397–410.
5. E.Lehrer, Repeated Games with Stationary Bounded Recall Strategies, J.E.T. 46 (1988), 130–144.
6. A.Neyman, Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoners' Dilemma, Economic Letters 19 (1985), 227–229.
7. A.Rubinstein, Finite Automata Play – The Repeated Prisoners' Dilemma, J.E.T. 39 (1986), 83–96.
8. H.Simon, "Theories of Bounded Rationality" in Decision and Organization ed. by C.McGuire and R.Radner, North–Holland, Amsterdam, 1972.
9. H.Simon, "Models of Man". John Wiley & Sons, 1957.

	L	R
T	1	–1
B	–1	1

Figure 1

	C	D
C	3,3	0,4
D	4,0	1,1

Figure 2