JOURNAL OF ECONOMIC THEORY 46, 130–144 (1988)

Repeated Games with Stationary Bounded Recall Strategies

EHUD LEHRER

Graduate School of Management, Northwestern University, Evanston, Illinois 60208

Received December 13, 1986; revised September 29, 1987

We investigate the asymptotic behavior of the set of equilibrium payoffs of repeated games with bounded recall, as the capacity of the memories of both players grow to infinity. *Journal of Economic Literature* Classification Numbers: 021, 026. © 1988 Academic Press, Inc.

1. INTRODUCTION

A great deal of attention has been paid recently to repeated games with bounded complexity. References [3, 5, 6] and others deal with repeated games played by automata. In this case the set of strategies is reduced to the set of those strategies that can be realized by automata.

Here we address ourselves to repeated games played by players with bounded recall who do not know the stage in which they are currently playing. In other words, we restrict ourselves to stationary t-bounded recall (t-SBR) strategies, according to which at each stage when he decides about his action a player may rely only on the t previous signals he has received (see [1] and [2]). Moreover, his decision is not time dependent.

First, we consider zero-sum games which are played sequentially infinitely many times, where player 1 uses t_1-SBR strategies and player 2 uses t_2-SBR strategies. These games are of a finite type and they have a value, denoted by V_{t_1,t_2}. We investigate the asymptotic behavior of V_{t_1,t_2} when t_1 and t_2 grow to infinity.

Assume, for example, that two players play a zero-sum game,where the actions set of player i, denoted by Σ_i, contains only two actions, $i=1, 2$. Assume furthermore that this game is played repeatedly, and that player 2 (the minimizer) is restricted to 1-SBR strategies. Namely, player 2's strategy is a function φ from $\Sigma_1 \times \Sigma_2$ to Σ_2. When player 2 chooses such a function φ, he has to play $\varphi(a_1, a_2)$ if he has observed (a_1, a_2) at the previous stage.

0022-0531/88 $3.00

It is readily shown (see proof of Theorem 2) that if player 1 has 100-bounded recall he can learn about the function φ that has been chosen by player 2 during the game. Consequently, at each stage player 1 can play his best response and thus ensure himself the minmax (with pure stragegies) payoff. Now assume that player 2 increases his ability to remember. The question is: How fast does player 1 have to increase his memory capacity in order to preserve his advantage?

The first result is that if t_1 is a function of t_2, say $t_1 = f(t_2)$, if $t_1 \geqslant t_2$, and if $\log f(t)/t$ tends to zero when t goes to infinity, then V_{t_1, t_2} tends to the value of the one-shot game. This means that the "stronger" player has no advantage in the long run, if his memory capacity grows less than exponentially as a function of his opponent's memory capacity. This result is similar to that of [3], which deals with finite automata. An immediate consequence of this result is that if a non-zero sum game is played by the above players, the set of Nash-equilibrium payoffs tends to the set of all the individually rational and feasible payoffs. This result is described formally in Theorem 4.

The second and the third results are of the following form. If the "stronger" player (here the maximizer) increases his memory capacity exponentially as a function of the other player's memory capacity, then the value of the repeated game is the minmax (with pure strategies) of the one-shot game. This means that in the exponential growth case, the "stronger" player produces the maximum which he can expect from his advantage.

2. Definitions and Notation

A two-player zero-sum game consists of two action sets Σ_1 and Σ_2 and a payoff function $h\colon \Sigma_1 \times \Sigma_2 \to \Re$. Denote this game by $G(\Sigma_1, \Sigma_2, h)$.

Denote $\Sigma = \Sigma_1 \times \Sigma_2$.

Definition. Let t_i be an integer. A *pure stationary* t_i*-bounded recall* (t_i-SBR) strategy of player i is a pair (e, φ), where

(1) e is the initial memory, $e \in \Sigma^{t_i}$, and

(2) φ is a function, $\varphi\colon \Sigma^{t_i} \to \Sigma_i$.

By the following observation, notice that a t_i-SBR strategy is in particular a stragey in the repeated game, which is a sequence $(\varphi_1, \varphi_2, \ldots)$ of functions, where $\varphi_{t+1}\colon \Sigma^t \to \Sigma_i$ for each $t = 0, 1, 2, \ldots$. Fix a t_i-SBR strategy (x, φ), $x = (x_1, \ldots, x_{t_i}) \in \Sigma^{t_i}$. For each $0 \leqslant t \leqslant t_i - 1$ and $(y_1, \ldots, y_t) \in \Sigma^t$, define $\varphi_{t+1}(y_1, \ldots, y_t) = \varphi(x_{t+1}, \ldots, x_{t_i}, y_1, \ldots, y_t)$ and for each $t_i \leqslant t$ and $(y_1, \ldots, y_t) \in \Sigma^t$, define $\varphi_{t+1}(y_1, \ldots, y_t) = \varphi(y_{t-t_i+1}, \ldots, y_t)$.

Denote the set of all t_i-SBR strategies of player i by $S_i^{t_i}$.

x is a *word* if $x \in \Sigma^k$ for some integer k. Let x and y be words, and xy will denote concatenation of x and y into one word. If $x = (x_1, ..., x_k)$ is a word and $j \leqslant k - n + 1$, then $x(j, j+n-1)$ will denote the subword $(x_j, x_{j+1}, ..., x_{j+n-1})$.

Let $\tau = (e, \varphi)$ and $\sigma = (e', \psi)$ be two pure strategies in $S_1^{t_1}$ and in $S_2^{t_2}$, respectively. τ and σ determine a string of payoffs $(a_i(\tau, \sigma))_{i=1}^{\infty}$, in the following way:

Denote by z_1 the signal $(\varphi(e), \psi(e'))$:

$$a_1(\tau, \sigma) = h(z_1).$$

Denote by z_2 the signal $(\varphi(e(2, t_1), z_1)\, \psi(e(2, t_2), z_1))$:

$$a_2(\tau, \sigma) = h(z_2)$$

and so on.

Define now the payoff function $H\colon S_1^{t_1} \times S_2^{t_2} \to \Re$, as

$$H(\tau, \sigma) = \lim_T \frac{1}{T} \sum_{i=1}^{T} a_i(\tau, \sigma).$$

H is well defined because the sequence $z_1, z_2, ...$ is periodic. Namely, there are two integers k and d such that, for any integer j, $z(k, k+d-1) = z(k+jd, k+(j+1)d-1)$.

Denote the zero-sum game $G(S_1^{t_1}, S_2^{t_2}, H)$ by G_{t_1,t_2} and its value by V_{t_1,t_2}.

3. Description of the Game in Words

A zero-sum $G(\Sigma_1, \Sigma_2, h)$ is played repeatedly many times. After every stage each player is informed about the actions of his opponent at that stage. In case a player has t-finite recall he remembers this last signal and he forgets the signals he got $(t+1)$ stages before (first in first out). A player can rely on his memory when he chooses his action at the next stage. A strategy is said to be stationary if its choice is independent of the stage of the game. At the first stage of the game a player adopts an initial memory, that is, $e \in (\Sigma_1 \times \Sigma_2)^t$, on which he will rely during the first t stages.

Assume now that player i had t_i-bounded recall, $i = 1, 2$. Let $\tau = (e, \varphi) \in S_1^{t_1}$ and $\sigma = (e', \psi) \in S_2^{t_2}$. At the first stage player 1 will play $\varphi(e)$ and player 2 will play $\psi(e')$. At the first stage player 2 (the minimizer) has to pay $a_1(\tau, \sigma) = h(\varphi(e), \psi(e'))$ to player 1. After stage 1 the players remember the signal $z_1 = (\varphi(e), \psi(e'))$ and forget the first item of their memory, i.e., player 1 forgets e_1 and player 2 forgets e'_1. Now, the memories of player 1 and player 2 are $(e(2, t_1), z_1)$ and $(e'(2, t_2), z_1)$, respectively. The moves that will be made at the second stage are $\varphi(e(2, t_1), z_1)$ for player 1 and $\psi(e'(2, t_2), z_1)$ for player 2. The signal is now $z_2 = (\varphi(e(2, t_1), z_1), \psi(e'(2, t_2) z_1))$ and the payoff is $a_2(\tau, \sigma) = h(z_2)$, and so on. The payoff of the repeated game is the Cesaro limit of $a_1, a_2, \ldots$.

A mixed strategy is defined, as usual, as a probability distribution over the set of pure strategies. If p_i is a probability distribution over $S_i^{t_i}$, $i = 1, 2$, then we define $H(p_1, p_2) = E_{(p_1, p_2)}(H(\tau, \sigma))$, which is the expectation, with respect to the distribution induced by (p_1, p_2), of H.

The game G_{t_1, t_2} is of a finite type (i.e., $S_1^{t_1}$ and $S_2^{t_2}$ are finite). Thus the existence of its value is ensured by the Minmax Theorem.

Putting it precisely, there exists a number V_{t_1, t_2} such that

$$V_{t_1, t_2} = \operatorname*{Min}_{p_2} \operatorname*{Max}_{p_1} H(p_1, p_2) = \operatorname*{Max}_{p_1} \operatorname*{Min}_{p_2} H(p_1, p_2),$$

where the maximum is taken over all the probability distributions p_1 over $S_1^{t_1}$, and the minimum is taken over all the probability distributions p_2 over $S_2^{t_2}$.

4. The Theorems

4.1. Let $G = G(\Sigma_1, \Sigma_2, h)$ be a zero-sum game. Denote by $\Delta(\Sigma_i)$ the set of all the mixed strategies of player i, $i = 1, 2$.

$$V = \text{the value of } G = \operatorname*{Max}_{a \in \Delta(\Sigma_1)} \operatorname*{Min}_{b \in \Delta(\Sigma_2)} h(a, b).$$

$$W = \operatorname*{Min}_{b \in \Sigma_2} \operatorname*{Max}_{a \in \Sigma_1} h(a, b).$$

$$M = \operatorname*{Max}_{a, b} h(a, b).$$

THEOREM 1. *Let f be a function, $f\colon N \to N$, such that* $\lim_n(\log f(n)/n) = \lim_n(\log n/f(n)) = 0$, *then* $\lim_n V_{f(n), n} = V$.

THEOREM 2. *There is a constant c, depending on G, such that $V_{c^n, n} = W$, and player 1 can achieve it by adopting a pure strategy.*

THEOREM 3. *Let $d = |\Sigma_1 \times \Sigma_2|$ and $g\colon N \to N$, such that* $\lim_n g(n) = \infty$, *then*

$$\lim_n V_{g(n) d^n, n} = W.$$

4.2. *Sketch of the Proof of Theorem* 1.

DEFINITION. Let $x = (x_1, ..., x_k)$ be a word. x has an n-cycle if there are two integers $t < t' \leqslant k$ such that $x(t, t+n-1) = x(t', t'+n-1)$, where $j > k$ is identified with $j \pmod k$.

PROPOSITION 1. *Let $X_1, X_2, ...$ be a sequence of identically distributed, mutually independent non-trivial Bernoulli random variables, and let f be as in Theorem* 1. *Define*

$$c_n(f) = \text{prob}\{\textit{There is an n-cycle in } X_1, ..., X_{f(n)}\}.$$

Then, $\lim_n f(n)\, c_n(f) = 0$.

Proof. See the Appendix. ∎

For each n player 2 constructs an n-SBR strategy by the following way. He generates a finite sequence of i.i.d random variables (the length of which is a function of the recalls of both players) using his optimal strategy in the one-shot game. To each sequence with no cycle is associated a pure SBR strategy. Proposition 1 shows that this can be done with high probability. Proposition 3 is then used to prove that the recall of player 1 is too small to allow correlation with player 2's moves, hence the result.

In [3] player 2 generates n i.i.d. random variables with the same distribution as above, one for each state. To *each* sequence is associated an automaton with n states.

5. PROOF OF THEOREM 1

5.1. Some notations from Information Theory must be introduced before proceeding to the proof. Let X be a finite random variable (f.r.v) that ranged to $\{x_1, ..., x_n\}$. Define the *entropy* of X to be

$$H(X) = -\sum_i p(x_i) \log p(x_i), \qquad \text{where} \quad p(x_i) = \text{prob}(X = x_i).$$

If Y is another f.r.v. that ranged to $\{y_1, ..., y_m\}$, then $H(X, Y) = -\sum_{ij} p(x_i, y_j) \log p(x_i, y_j)$, where $p(x_i, y_j) = \text{prob}(X = x_i \text{ and } Y = y_j)$.

Denote by $|X|$ the number of atoms of X. $H(X)$ is always less or equal to $\log |X|$.

DEFINITION. A f.r.v. Y is said to be ε-independent of a f.r.v. X (write $Y \perp^{\varepsilon} X$) if

$$\sum_j |\text{prob}(Y = y_j \mid X = x_i) - \text{prob}(Y = y_j)| < \varepsilon$$

for all i except a set of atoms which union has a measure less than ε.

Define

$$I(Y \mid X) = H(X) + H(Y) - H(X, Y).$$

Theorem 4.6 of [4] states that $I(Y \mid X) \geqslant 0$ with equality if and only if Y and X are independent.

PROPOSITION 2 [7, p. 22]. *Given $\varepsilon > 0$ there exists $\delta(\varepsilon) > 0$ s.t. for any f.r.v. X and Y, $I(Y \mid X) < \delta$ implies $Y \perp^{\varepsilon} X$.*

The final statement we need is that if Z is Y-measurable, then $I(Y \mid X) \geqslant I(Z \mid X)$ [4, p. 26].

5.2. The strategy defined below is defined as if the information of player 2 is trivial. Namely, that after each stage player 2 gets information only about his action and not about his opponent's action. Such a strategy can be easily modified in order to get a strategy of standard information, without losing its "good" properties.

In the case where player 2 has a pure optimal strategy in the one-shot game, then the result is obvious.

Fix and n and let $k = n^3 f(n)$. Now, let $X_1, X_2, ..., X_k$ be a sequence of mutually independent identically distributed random variables, which are defined on the sample space Ω and range to Σ_2. Assume that $\text{prob}(X_1 = s) = p_s$ for all $s \in \Sigma_2$, where $p = (p_1, ..., p_u) \in \Delta(\Sigma_2)$ satisfies $V = \text{Max}_{q \in \Delta(\Sigma_1)} h(q, p)$. We can assume that the optimal strategy is not pure so $k \cdot c_n(k) \to 0$.

Define the event

$$A = \{\omega \in \Omega \mid (X_1(\omega), ..., X_k(\omega)) \text{ has no } n\text{-cycles}\}.$$

For every $\omega \in A$ define $e(\omega) = (X_{k-n+1}(\omega), ..., X_k(\omega))$. $\varphi(\omega)(X_i(\omega)), ..., X_{i+n-1}(\omega)) = (X_{i+1}(\omega), ..., X_{i+n}(\omega))$, $i = 1, ..., k$, where $t > k$ is identified with $t \pmod k$. On all the remaining tuples of Σ_2^n, $\varphi(\omega)$ is defined arbitrarily. Denote this mixed strategy by σ.

5.3. Let Ω' be the sample space A, with the probability $\text{prob}'(\cdot)$ defined by

$$\text{prob}'(D) = \text{prob}(D \mid A) \qquad \text{for every } D \subseteq A.$$

Define $\bar{X}_i = X_i$ on A.

PROPOSITION 3. *Let $\bar{X}_1, ..., \bar{X}_k$ and $f(\cdot)$ be as previously defined, and let $d \in N$. Then, for every f.r.v. Y, defined on Ω', with $|Y| \leqslant d^{f(n)}$ we have*

$$I(Y, \bar{X}_1, ..., \bar{X}_{i-1} \mid \bar{X}_i) \leqslant 1/n$$

for at least $(1-1/n)k$ i's, providing that n is big enough.

Proof. See the Appendix. ▮

5.4. Assume[1] that the payoff function h is non-negative, and let $\tau = (e', \psi)$ be a pure strategy of player 1.

$$\begin{aligned} H(\tau, \sigma) &= E\left(\lim_T \frac{1}{T} \sum_{i=1}^{T} a_i\right) \\ &= \lim_T E\left(\frac{1}{T} \sum_{i=1}^{T} a_i\right) = \lim_T \frac{1}{T} \sum_{i=1}^{T} E(a_i) \\ &= \lim \frac{1}{T} \sum_{i=1}^{T} E(a_i \mid A) \operatorname{prob}(A) + E(a_i \mid A^c) \operatorname{prob}(A^c)] \\ &\leqslant \lim \frac{1}{T} \sum_{i=1}^{T} E(a_i \mid A) + c_n \cdot M. \end{aligned} \tag{6}$$

To evaluate the first term let us consider $\sum_{i=tk+1}^{(t+1)k} E(a_i \mid A)$ for some t. Assume from now on that A is given. Denote by Y_i the f.r.v. which is defined to be the action of player 1 at stage $tk+i$, and denote by Y the signal of player 1 after stage tk. Recall that the action of player 2 at stage $tk+i$ is $\bar{X}_i$. Always $|Y| \leqslant d^{f(n)}$, where $d = |\Sigma_1 \times \Sigma_2|$.

By Proposition 3, $I(Y, \bar{X}_1, ..., \bar{X}_{i-1} \mid \bar{X}_i) < 1/n$ for at least $(1-1/n)k$ i's in the block $(tk+1, ..., (t+1)k)$. Y_i is $(Y, \bar{X}_1, ..., \bar{X}_{i-1})$-measurable, thus $I(Y_i \mid \bar{X}_i) \leqslant I(Y, \bar{X}_1, ..., \bar{X}_{i-1} \mid \bar{X}_i)$. By Proposition 2 we get $Y_i \perp^{\xi(1/n)} \bar{X}_i$, where $\xi(1/n) \to 0$ whenever $n \to \infty$.

Hence,

$$\frac{1}{k} E\left(\sum_{i=tk+1}^{(t+1)k} a_i \mid A\right) \leqslant \frac{1}{k}\left(k \cdot \frac{1}{n}\right) M + 2\xi\left(\frac{1}{n}\right) M + V. \tag{7}$$

Inequalities (6) and (7) lead to the conclusion that

$$H(\tau, \sigma) \leqslant M(1/n + 2\xi(1/n) + c_n) + V.$$

The first term tends to zero as $n \to \infty$. In order to conclude the proof of Theorem 1 we can use the previous procedure, exchanging the roles of the players.

[1] Obviously there is no loss of generality in this assumption, and it is taken for the sake of simplicity.

5.5 *Remark.* Actually, the proof proves more than desired. It is not assumed that the strategy of player 1 is stationary (Y_i is defined to be the action of player 1 at stage $tk+i$). Thus, by using τ, player 2 can ensure V (at the limit), even that his opponent does not use SBR strategies, but rather BR strategies, which can be time dependent.

In the non-zero-sum case, the two players can agree to act on any cooperative action. Since $V_{f(n),n} \to V$, each player has a threat to his opponent, by which he forces the opponent not to deviate from the cooperative action. Thus, if $(\Sigma_1, \Sigma_2, h_1, h_2)$ is a non-zero-sum game and $N_{f(n),n}$ is the set of Nash-equilibrium payoffs when players 1 and 2 use n-SBR and $f(n)$-SBR strategies, respectively, then $N_{f(n),n}$ tends to the set of all the individually rational and feasible payoffs. In order to demonstrate it precisely we need the following notations.

Notations. (i) Let $i \in \{1, 2\}$.

$$r_i = \operatorname*{Min}_{p \in \Delta(\Sigma_{3-i})} \operatorname*{Max}_{q \in \Delta(\Sigma_i)} h_i(p, q).$$

(ii) $(x_1, x_2) \in \mathfrak{R}^2$ is individually rational if $x_1 \geqslant d_1$ and $x_2 \geqslant d_2$.

(iii) $h = (h_1, h_2)$.

(iv) $x \in \mathfrak{R}^2$ is feasible if $x \in \operatorname{conv} h(\Sigma)$, i.e., x is contained in the convex hull of all the payoffs.

(v) $FIR = \{x \in \mathfrak{R}^2 \mid x \text{ is feasible and individually rational}\}$.

Let $\{A_n\}$ be a sequence of subsets of $\mathfrak{R}^2$. We denote by $\lim A_n$ the set $\{x \mid \text{ there is a sequence } x_n \in A_n \text{ such that } \lim x_n = x\}$.

THEOREM 4. *Let f be a function as in Theorem* 1. *Then* $\lim_n N_{f(n),n} = FIR$.

Proof. It is clear that $x \in \lim N_{f(n),n}$ is feasible.

By Theorem 1 it is ensured that any $x \in \lim N_{f(n),n}$ is individually rational. Thus $\lim N_{f(n),n} \subseteq FIR$.

In order to prove the inverse inclusion we will describe two strategies, one for each player, consisting of two plans: the master plan and the punishment plan. The punishment plan will be executed in a case where a deviation had occurred. However, we have to take care that the histories used in the master plan (in order to achieve a cooperative payoff) do not interfere with the one that could appear in the punishment plan.

Assume first that $\operatorname{int} FIR \neq \varnothing$. We will use as punishing strategies the same kind of strategies as described in the previous proof and assume that they have $n/2$ or $f(n)/2$ memory and induce histories with no $\frac{1}{6}\operatorname{Min}\{n, f(n)\}$ cycles. Denote $u(n) = [\operatorname{Min}\{n, f(n)\}/6]$. By using histories with $u(n)$-cycles we will achieve (at the limit) any cooperative payoff in int FIR. Let $x \in \text{int } FIR$. $x = \sum_{a \in \Sigma} \alpha_a h(a)$, where $\alpha_a \geqslant 0$ and $\Sigma\, \alpha_a = 1$. Since $u(n) \to_{n \to \infty} \infty$, for each $\varepsilon > 0$ there are integers n and j_a, $a \in \Sigma$, so that $\sum j_a = u(n)$ and $\|x - \sum_{a \in \Sigma} (j_a/u(n))\, h(a)\|_\infty < \varepsilon$. It means that we can get x as a limit of rational combinations of the payoffs.

Denote $\Sigma = \{a^1, ..., a^{|\Sigma|}\}$ and for each $1 \leqslant k \leqslant |\Sigma|$, $a^k = (a_1^k, a_2^k)$. The master plan for player i is defined as follows. Play a_i^1 for j_{a^1} times, then play a_i^2 for j_{a^2} times and so on till $a_i^{|\Sigma|}$. Then again play a_i^1 for j_{a^1} times and so on. It is obvious that this strategy can be implemented by $6u(n)$-SBR strategy.

We will describe the punishment plan of player 2. A similar description holds for player 1. The punishment plan uses $n/2$-SBR. Denote by $\sigma = (x, \varphi)$ the strategy of player 2 that was described in the previous proof, for $n/2$-bounded recall (not n-bounded recall). Immediately after observing a deviation, player 2 plays according to the initial memory of σ. That is, he plays x_1 then x_2 and so on for $n/2$ stages ($x = (x_1, ..., x_{n/2})$). Notice that after these $n/2$ stages the memory of player 2 still contains a $u(n)$-cycle and also the deviation. From the moment in which the terminal $n/2$-word that appears in player 2's memory coincides with a possible $n/2$-history of the punishment strategy σ, player 2 continues playing according to σ. By this description it is ensured that from the $n/2+1$st stage after the deviation on, player 2 plays according to σ.

Since $x \in \text{int } FIR$, and by Theorem 1, the pair of strategies described above is a Nash equilibrium. We have got that $\text{int } FIR \subseteq \lim N_{f(n), n}$ and since $\lim N_{f(n), n}$ is a closed set we have $FIR \subseteq \lim N_{f(n), n}$.

The case where $\text{int } FIR = \varnothing$ is much more simpler, and we leave it to the reader. ▌

6. The Proof of Theorem 2

The main idea of the proof is the following. After each n-word that already appears in his memory player 1 knows the next move of player 2, and thus he can play a one-shot game best reply. Since player 1's recall is much bigger than that of his opponent's he can learn more and more about player 2's moves, and thus ensure himself at least W for longer and longer sequences of stages. It is proved that from a certain stage there is an n-cycle in which player 1 knows all player 2's moves.

Let $c = |\Sigma|^2 + 1$, and let $b = c^n$. The pure strategy $\tau = (e, \varphi)$ will be described as follows. e is an arbitrary tuple of Σ^b. For any $z \in \Sigma^b$ denote $t(z) = z(b-n+1, b)$, i.e., $t(z)$ is the last n-subword of z. Now define $\varphi(z)$ arbitrarily if $t(z)$ appears only once in z. And define $\varphi(z)$ to be the best response of player 1 against z_k^2, where $z(k-n, k-1)$ is the antecedent appearance of $t(z)$ to $z(b-n+1, b)$, and $z_k = (z_k^1, z_k^2)$.

In any string z of length b there is at least one n-subword, which appears

$t = |\Sigma|^n$ times. Denote one of these n-subwords by $v(z)$. Let $\sigma = (e', \psi)$ be a pure strategy of player 2. Denote by z the memory of player 1 at stage b, when player 1 plays τ and player 2 plays σ, and denote $(v_1, ..., v_n) = v = v(z)$. After the first appearance of v player 1 knows $\psi(v)$. Denote by $B(\psi(v))$ the best response of player 1 agains $\psi(v)$, and by ω_1 the pair $(B(\psi(v)), \psi(v))$. After the second appearance of v player 2 knows $\psi(v_2, ..., v_n, \omega_1)$. Let $B(\psi(v_2, ..., v_n, \omega_1))$ be the best response of player 1 against it, and so on. After the $|\Sigma|^n$-th appearance of v the memory of player 1 will contain the word $z' = (v_1, ..., v_n, \omega_1, ..., \omega_t)$, which satisfies the following. For every $(n+1)$-subword of z', $y = (y_1, ..., y_n, y_{n+1})$, we have $y_{n+1} = (B(\psi(y_1, ..., y_n)), \psi(y_1, ..., y_n))$. z' is of length $t+n$, so it has an n-cycle.

We have proved that there is an n-cycle in the memory of player 1 at stage $b+t$. At each stage of this cycle player 1 ensures himself at least W, thus he can ensure W as a payoff of the repeated game.

Player 2 can easily ensure himself W, so the proof is complete.

7. Proof of Theorem 3

Let τ be the strategy of player 1 described in the previous section, and let σ be any pure strategy of player 2. We shall show that the number of n-words that do not appear at least two times in a long sequence is small, and hence their relative weight for the payoff is negligible when n is sufficiently large.

Divide the set of stages into blocks $B_0 = \{1, ..., g(n)\, d^n\}$, $B_1 = \{g(n)\, d^n + 1, ..., 2g(n)\, d^n\}$, and so on. Fix an integer t. We shall evaluate $\Sigma_{i \in B_t} a_i(\tau, \sigma)$. Denote by z_i the pair of actions played at stage i, $i \in B_t$, and by z the $g(n)\, d^n$-word $(z_{tg(n)d^n}, z_{tg(n)d^n+1}, ..., z_{(t+1)g(n)d^n-1})$. Let, for each integer j, k_j be the number of n-words appearing j times in the word z. We have

$$\sum_{j=1}^{\infty} jk_j = g(n)\, d^n - n + 1 \qquad \text{and} \qquad \sum k_j \leqslant d^n.$$

Now, if for $i \in B_t$ the n-subword of z, which terminates at z_{i-1}, has already appeared in z, then player 1 gets at least W at the stage i. There are $\sum (j-1)\, k_j$ such stages. So,

$$\begin{aligned}\frac{1}{|B_t|} \sum_{i \in B_t} a_i(\tau, \sigma) &\geqslant \frac{W}{|B_t|} \sum (j-1)\, k_j - M \sum k_j \\ &\geqslant \frac{W}{g(n) d^n} (g(n)\, d^n - n) - 2Md^n \\ &= W - M\left(\frac{n}{g(n) d^n} - \frac{2}{g(n)}\right) \to W,\end{aligned}$$

where M is the largest absolute value of the payoffs in the game. Hence,

$$\liminf V_{g(n)d^n,n} \geqslant W.$$

Since player 2 can always ensure himself W, the theorem follows.

Appendix

Proof of Proposition 1. We can assume that X_i takes only two values with positive probability, 0 and 1, where $\text{prob}(X_1=0)=p$ and $\text{prob}(X_1=1)=q$ $(p+q=1)$. Let $m>n$, then

$$\begin{aligned}\text{prob}\{(X_1, ..., X_n) = (X_m, ..., X_{m+n-1})\}\\ = \sum_{k=0}^{n} \binom{n}{k} (p^k q^{n-k})^2 = (p^2+q^2)^n.\end{aligned} \tag{1}$$

The event $C = \{\text{There is an } n\text{-cycle in } X_1, ..., X_{f(n)}\}$ is included in the union of A_i, where

$$A_i = \left\{\begin{matrix}\text{There is } j \neq i \text{ and } j+n-1 \notin \{i, ..., i+n-1\} \text{ such that}\\ (X_i, X_{i+1}, ..., X_{i+n-1}) = (X_j, X_{j+1}, ..., X_{j+n-1})\end{matrix}\right\}.$$

Define

$$A_i' = \left\{\begin{matrix}\text{There is } j \text{ s.t. } j, j+n-1 \notin \{i, ..., i+n-1\} \text{ such that}\\ (X_i, X_{i+1}, ..., X_{i+n-1}) = (X_j, X_{j+1}, ..., X_{j+n-1})\end{matrix}\right\},$$

and

$$A_i'' = \left\{\begin{matrix}\text{There is } j \in \{i+1, ..., i+n-1\} \text{ such that}\\ (X_i, ..., X_{i+n-1}) = (X_j, ..., X_{j+n-1})\end{matrix}\right\}.$$

Thus, $A_i = A_i' \cup A_i''$. By (1),

$$\text{prob}(A_i') \leqslant f(n) \cdot (p^2+q^2)^n. \tag{2}$$

Note that $\text{prob}(A_1'') = \text{prob}(A_i'')$. Denote by B_t the event

$$\{(X_1, ..., X_n) = (X_t, ..., X_{t+n-1})\}.$$

$A_1'' \subseteq \bigcup_{t=1}^{n-1} B_t$, and thus $\text{prob}(A_1'') \leqslant \sum_{t=1}^{n-1} \text{prob}(B_t)$.

Let $t > [n/2]$, and apply (1) to t-cycles in order to get

$$\text{prob}(B_t) \leqslant n(p^2+q^2)^t \leqslant n(p^2+q^2)^{n/2}. \tag{3}$$

Since

$$\operatorname{prob}\{(X_1, ..., X_n) = (X_t, ..., X_{t+n-1})\}$$
$$\sum_{j=0}^{t} \binom{t}{j} (p^j q^{t-j})^{[n/t]+1} \leqslant \sum_{j=0}^{t} \binom{t}{j} (p^j q^{t-j})^{n/t}$$
$$= (p^{n/t} + q^{n/t})^t,$$

we have $\operatorname{prob}(B_t) \leqslant (p^{n/t} + q^{n/t})^t$.

Thus,

$$\sum_{t=1}^{[n/2]} \operatorname{prob}(B_t) \leqslant \sum_{t=1}^{[n/2]} (p^{n/t} + q^{n/t})^t$$
$$= p^n \sum_{t=1}^{[n/2]} \left(1 + \left(\frac{q}{p}\right)^{n/t}\right)^t.$$

Without loss of generality $p \geqslant q$, so $q/p \leqslant 1$. Since the function $(1 + a^{1/x})^x$ is increasing whenever $0 < a \leqslant 1$, we get

$$\sum_{t=1}^{[n/2]} \operatorname{prob}(B_t) \leqslant p^n \frac{n}{2} \left(1 + \left(\frac{q}{p}\right)^{n/(n/2)}\right)^{n/2}$$
$$= \frac{n}{2} (p^2 + q^2)^{n/2}. \tag{4}$$

Combine (3) and (4) in order to get

$$\operatorname{prob}(A_1'') \leqslant \left(\frac{n}{2} + 1\right) \cdot n(p^2 + q^2)^{n/2} + \frac{n}{2} (p^2 + q^2)^{n/2}. \tag{5}$$

Since $(p^2 + q^2) < 1$, by (2) and (5) we get

$$\operatorname{prob}(A_i) \leqslant (p^2 + q^2)^{n/2} \left(f(n) + \left(\frac{n}{2} + 1\right) n + \frac{n}{2}\right).$$

Moreover,

$$c_n(f) = \operatorname{prob}(C) \leqslant f(n) \cdot (p^2 + q^2)^{n/2} \left(f(n) + \left(\frac{n}{2} + 1\right) n + \frac{n}{2}\right)$$
$$\leqslant (p^2 + q^2)^{n/2} f(n)(f(n) + n^3).$$

$\log f(n)/n$ tends to zero when n goes to infinity and thus $c_n(f) \cdot f(n)$ also tends to zero. ▌

Proof of Proposition 3. By direct computation,

$$I(Y \mid \bar{X}_1) + I(Y, \bar{X}_1 \mid \bar{X}_2) + \cdots + I(Y, \bar{X}_1, ..., \bar{X}_{k-1} \mid \bar{X}_k)$$
$$= H(Y) + H(\bar{X}_1) + \cdots + H(\bar{X}_k) - H(Y, \bar{X}_1, ..., \bar{X}_k)$$

(since $(\bar{X}_1, ..., \bar{X}_k)$ is $(Y, \bar{X}_1, ..., \bar{X}_k)$-measurable, and $H(\bar{X}_k) \geqslant 0$)

$$\leqslant H(Y) + H(\bar{X}_1) + \cdots + H(\bar{H}_k) - H(\bar{X}_1, ..., \bar{X}_k).$$

We will show that $\sum_{i=1}^{k} H(\bar{X}_i) - H(\bar{X}_1, ..., \bar{X}_k) \leqslant c$, for some constant c, which is independent of k. Define

$$F = \{(i_1, ..., i_k) \in \Sigma_2^k \mid \text{There is no } n\text{-cycle in } (i_1, ..., i_k)\}.$$

Recall $c_n = c_n(k) = \text{prob}(\Omega \setminus A)$. Denote $\text{prob}(X_1 = i) = p_i$.

$$H(\bar{X}_1, ..., \bar{X}_k) = \sum_{(i_1, ..., i_k) \in F} -\left(\frac{1}{1-c_n} \prod_{j=1}^{k} p_{i_j}\right) \log\left(\frac{1}{1-c_n} \prod_{j=1}^{k} p_{i_j}\right)$$
$$= \frac{-1}{1-c_n} \sum_F \left(\prod p_{i_j}\right) \log\left(\prod p_{i_j}\right) - \frac{1}{1-c_n} \sum_F \left(\prod p_{i_j}\right) \log \frac{1}{1-c_n}$$
$$= \frac{-1}{1-c_n} \sum_F \left(\prod p_{i_j}\right) \log\left(\prod p_{i_j}\right) + \frac{1}{1-c_n} \log(1-c_n) \sum_F \left(\prod p_{i_j}\right),$$

$X_1, ..., X_k$ are independent, so

$$H(X_1 + \cdots + H(X_k) - H(\bar{X}_1, ..., \bar{X}_k)$$
$$= -\sum_{(i_1, ..., i_k) \in \Sigma_2^k} \left(\prod p_{i_j}\right) \log\left(\prod p_{i_j}\right) + \frac{1}{1-c_n} \sum_F \left(\prod p_{i_j}\right) \log\left(\prod p_{i_j}\right)$$
$$- \frac{1}{1-c_n} \log(1-c_n) \sum_F \left(\prod p_{i_j}\right)$$
$$\leqslant \frac{c_n}{1-c_n}\left[-\sum_F \left(\prod p_{i_j}\right) \log\left(\prod p_{i_j}\right)\right] - \sum_{F^c} \left(\prod p_{i_j}\right) \log\left(\prod p_{i_j}\right)$$
$$+ \frac{1}{1-c_n} \log \frac{1}{1-c_n}$$
$$\leqslant \frac{c_n}{1-c_n} k \cdot H(X_1) - c_n \sum_{F^c} \left(\frac{\prod p_{i_j}}{c_n}\right) \log\left(\prod p_{i_j}\right) + \frac{1}{1-c_n} \log \frac{1}{1-c_n}$$

(by Jensen inequality)

$$\leqslant \frac{c_n}{1-c_n} kH(X_1) + c_n \log \sum_{F^c} \frac{1}{c_n} + \frac{1}{1-c_n} \log\left(\frac{1}{1-c_n}\right)$$

$$\leqslant \frac{1}{1-c_n} \left(c_n \cdot kH(X_1) + \log\left(\frac{1}{1-c_n}\right)\right) + c_n \cdot \log \frac{|F^c|}{c_n}.$$

This sum tends to zero because $k \cdot c_n \to_k 0$, $1/(1-c_n)\log(1/(1-c_n)) \to_k 0$, and because

$$c_n \log \frac{|F^c|}{c_n} \leqslant c_n \log |\Sigma_2^k| - c_n \log c_n = c_n \cdot k \log |\Sigma_2| - c_n \log c_n \xrightarrow[k \to \infty]{} 0.$$

Second, denote $\text{prob}'(\bar{X}_1 = i) = \bar{p}_i$.

$$k(H(\bar{X}_1) - H(X_1)) = k\left(-\sum_i \bar{p}_i \log \bar{p}_i + p_i \log p_i\right)$$

$$= k \sum_i \bar{p}_i \left(\log \frac{p_i}{\bar{p}_i}\right) + k \sum_i (p_i - \bar{p}_i) \log p_i$$

$$\leqslant k \sum_i \bar{p}_i \log\left(\frac{\bar{p}_i + c_n}{\bar{p}_i}\right) + k \sum_i c_n \log p_i$$

(again by Jensen equality)

$$\leqslant k \log(1 + |\Sigma_2|\, c_n) + kc_n \sum_i \log p_i \to 0.$$

We have got that $H(\bar{X}_1) + \cdots + H(\bar{X}_k) - H(\bar{X}_1, ..., \bar{X}_k) \to_k 0$. Hence for n big enough we have

$$I(Y \mid \bar{X}_1) + I(Y, \bar{X}_1 \mid \bar{X}_2) + \cdots + I(Y, \bar{X}_1, ..., \bar{X}_{k-1} \mid \bar{X}_k) \leqslant H(Y) + 1$$

$$\leqslant \log d^{f(n)} + 1 = f(n) \log d + 1.$$

If, contrary to the proposition, at least $[k/n]$ i's satisfy $I(Y, \bar{X}_1, ..., \bar{X}_{i-1} \mid \bar{X}_i) > 1/n$, then we would get that $1/n \cdot k/n = n^3 f(n)/n^2 = nf(n)$ is less than $f(n) \log d + 1$ for infinitely many n's. However $nf(n)/(f(n) \log d + 1) \to \infty$, and thus the proof is complete. ∎

Acknowledgments

I am gratefully indebted to Professor A. Neyman for helpful comments, especially for Section 7. I am also indebted to the anonymous referee of the Journal of Economic Theory for his or her valuable suggestions.

References

1. R. Aumann, Survey of repeated games, *in* "Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgensten," pp. 11–42, Bibliographisches Institut, Mannheim/Vienna/Zurich, 1981.
2. R. Aumann and S. Sorin, Cooperation and Bounded Rationality, manuscript, 1986.
3. E. Ben Porath, Repeated games with bounded complexity, manuscript, 1986.
4. R. J. McEliece, "The Theory of Information and Coding," Addison–Wesley, Reading, MA, 1977.
5. A. Neyman, Bounded complexity justifies cooperation in the finitely repeated prisoner's dilemma, *Econ. Lett.* **19** (1985), 227–229.
6. A. Rubinstein, Finite automata play—The repeated prisoners' dilemma, *J. Econ. Theory* **39** (1986), 83–96.
7. M. Smorodinsky, "Ergodic Theory, Entropy," Springer-Verlag, New York/Berlin, 1971.