Games and Economic Behavior **30,** 228–247 (2000)
doi:10.1006/game.1999.0725, available online at http://www.idealibrary.com on IDEAL®

Repeated Games with Bounded Entropy

Abraham Neyman

Institute of Mathematics, The Hebrew University of Jerusalem, Givat Ram, Jerusalem 91904, Israel and State University of New York at Stony Brook, Stony Brook, New York 11794-4384
E-mail: aneyman@math.huji.ac.il

and

Daijiro Okada

Department of Economics, State University of New York at Stony Brook, Stony Brook, New York 11794-4384
E-mail: dokada@datalab2.sbs.sunysb.edu

Received April 16, 1997

We investigate the asymptotic behavior of the maxmin values of repeated two-person zero-sum games with a bound on the strategic entropy of the maximizer's strategies while the other player is unrestricted. We will show that if the bound $\eta(n)$, a function of the number of repetitions n, satisfies the condition $\eta(n)/n \to \gamma$ $(n \to \infty)$, then the maxmin value $W^n(\eta(n))$ converges to $(\text{cav}\, U)(\gamma)$, the concavification of the maxmin value of the stage game in which the maximizer's actions are restricted to those with entropy at most γ. A similar result is obtained for the infinitely repeated games. *Journal of Economic Literature* Classification Numbers: C73, C72. © 2000 Academic Press

1. INTRODUCTION

We introduced the concepts of strategic entropy for repeated games in Neyman and Okada (1999) for the purpose of studying some strategic complexity issues involving finite automata and bounded recall. Specifically, we considered the repeated game $G^n(c(n))$, the n-fold repetition of a two-person zero-sum stage game G, where player 1, the maximizer, is restricted to strategies of complexity at most $c(n)$, a function of n, while player 2 is unrestricted. The complexity of a (pure) strategy is either the minimum number of states of an automaton or the minimum length of recall required to implement the strategy. Our interest was in how the

0899-8256/00 $35.00

value of such a game depends on the complexity bound, and, in particular, the relationship between the number of repetitions and the complexity bound so that the unrestricted player can take advantage of the other player's strategic limitation. Mathematical formulation of the problem is the asymptotics of the value of $G^n(c(n))$, denoted by $V^n(c(n))$, under some condition on the order of magnitude of $c(n)$. The work thus follows the line of research dealing with models of repeated games with exogenously given strategic complexity bound, e.g., Ben-Porath (1993), Neyman (1985, 1997, 1999), Papadimitriou and Yannakakis (1994, 1998) (finite automata), and Lehrer (1988, 1994) (bounded recall).

We answered in the affirmative the conjecture in Neyman (1997) for the case of automata

$$\lim_{n\to\infty} \frac{c(n)\log c(n)}{n} = 0 \quad \Rightarrow \quad \lim_{n\to\infty} V^n(c(n)) = U_*(G), \tag{1}$$

where $U_*(G)$ is the maximin value in pure actions of the stage game G. The key observation leading to the proof of (1) is that the number of pure strategies, up to the equivalence, implementable by automata of size (number of states) m is of the order $m^{O(m)}$. Thus any probability distribution, i.e., mixed strategy, over such pure strategies has entropy at most $Cm\log m$ for some constant C. This led us to examining repeated games in which there is a direct restriction on player 1's mixed strategies in terms of strategic entropy (again, there is no restriction on player 2's strategies). Strategic entropy of a mixed strategy is the maximum entropy of the play generated by that strategy where the maximum is taken over the other player's pure strategies. We showed that if the strategic entropy bound $\eta(n)$ satisfies the condition $\lim_{n\to\infty}(\eta(n)/n) = 0$, then the maxmin[1] value of the repeated game $G^n(\eta(n))$ converges to $U_*(G)$. We also showed that strategic entropy of a mixed strategy never exceeds its entropy. Hence the condition on the left-hand side of (1) implies that the strategic entropy (in the automata game $G^n(c(n))$ satisfies $\lim_{n\to\infty}(\eta(n)/n) = 0$, which, by what we have proved, implies that the maxmin value of $G^n(c(n))$ converges to $U_*(G)$. But the minimax theorem does apply to the game $G^n(c(n))$, hence the right-hand side of (1).

Above we used strategic entropy as a technical tool. However, there is a sense in which this concept is of an independent interest. Suppose that one tries to encode the possible realizations of a mixed strategy (or any random variable) into sequences of 0s and 1s. Entropy of the mixed strategy is then, roughly, the minimum number of bits required for this task. If one

[1] The set of mixed strategies with bounded strategic entropy is not convex and thus the minimax theorem does not apply.

associates a unit cost to each bit used, the entropy can be considered the (expected) cost of randomization. The strategic entropy in this context is a measure of the cost of randomization in the play generated by the mixed strategy.

In this paper we study the asymptotics of the minimax value, $W^n(\eta(n))$, of $G^n(\eta(n))$ under the condition that the per stage strategic entropy bound $\eta(n)/n$ converges to an arbitrary fixed nonnegative number γ. We will give a characterization of $\lim_{n\to\infty} W^n(\eta(n))$ as a continuous function of γ. This function is derived from the underlying stage game. To be specific, we will look at the stage game in which player 1 is allowed to use mixed actions with entropy at most γ. Then one computes the maxmin value of such a game for each $\gamma \geq 0$, and thus obtains a function $U(\gamma)$. Our main theorem states that $W^n(\eta(n))$ converges to $(\operatorname{cav} U)(\gamma)$ whenever $\eta(n)/n \to \gamma$, where $\operatorname{cav} U$ is the concavification of the function U, i.e., the smallest concave function at least as large as U pointwise. We will give a similar characterization of the maxmin value of the infinitely repeated games in which there is a restriction on player 1's strategies in terms of strategic entropy rate, a concept introduced in Neyman and Okada (1998).

In the next section, the basic notations and terminologies for the stage game and the repeated games are introduced. We review information theoretic concepts in Section 3 and the concepts of strategic entropy in Section 4. Section 5 contains the main results.

2. THE STAGE GAME AND REPEATED GAMES

Let $G = (A, B, r)$ be a two-person zero-sum game in strategic form, where A and B are finite sets of *pure actions* for players 1 and 2, respectively, and r: $A \times B \to \mathbb{R}$ is the payoff matrix of player 1, the maximizer. A *mixed action* is a probability distribution on the set of pure actions. Let $\Delta(A)$ and $\Delta(B)$ be the sets of mixed actions of the two players. We denote the *maxmin value in pure actions and the value* of G by $U_*(G)$ and $\operatorname{Val}(G)$, respectively. That is, $U_*(G) = \max_{a\in A}\min_{b\in B} r(a,b)$ and $\operatorname{Val}(G) = \min_{\beta\in\Delta(B)}\max_{a\in A} E_\beta[r(a,b)]$. By the minimax theorem we can also write $\operatorname{Val}(G) = \max_{\alpha\in\Delta(A)}\min_{b\in B} E_\alpha[r(a,b)]$.

Given a game $G = (A, B, r)$, we next describe a new game in which G is played repeatedly (with complete information and standard signaling).

A *play* of a repeated game is an infinite sequence $\omega = (\omega_k)_{k=1}^\infty$, where $\omega_k = (a_k, b_k) \in A \times B$. We denote the set of all plays by Ω_∞, i.e., $\Omega_\infty = (A \times B)^\infty$. This will be our basic space throughout the paper.

Two plays $\omega = (\omega_k)_{k=1}^\infty$ and $\omega' = (\omega'_k)_{k=1}^\infty$ are said to be n-equivalent if $\omega_k = \omega'_k$ for $k = 1, \ldots, n$. The n-equivalence is clearly an equivalence relation on Ω_∞. Denote by $\mathscr{H}_n$ the finite partition of Ω_∞ into the n-equivalence classes. Each n-equivalence class of plays is called an *n-history* and it represents an information available to the players at the end of stage n. We sometimes represent an n-history by exhibiting the first n coordinates, e.g., $(\omega_1, \ldots, \omega_n)$. We denote by $\mathscr{A}_n$ the algebra on Ω_∞ induced by $\mathscr{H}_n$. Set $\mathscr{A}_0 = \{\phi, \Omega_\infty\}$. Clearly, $\mathscr{A}_{n-1} \subset \mathscr{A}_n$ for every $n = 1, 2, \ldots$. The σ-algebra generated by $\bigcup_{n \geq 0} \mathscr{A}_n$ is denoted by $\mathscr{A}_\infty$.

Denote by S_n and T_n the sets of measurable mappings from $(\Omega_\infty, \mathscr{A}_{n-1})$ to A and B, respectively. Each element of S_n and T_n represents a "strategy at stage n." For each $s_n \in S_n$, since $s_n(\omega)$ depends only on the first $n-1$ coordinates of $\omega = (\omega_k)_{k=1}^\infty$, we sometimes write $s_n(\omega_1, \ldots, \omega_{n-1})$. Similarly for $t_n \in T_n$. A *pure strategy* of player 1 (resp. 2) is a sequence $s = (s_n)_{n=1}^\infty$ with $s_n \in S_n$ (resp. $t = (t_n)_{n=1}^\infty$ with $t_n \in T_n$). Thus the sets of pure strategies of the two players are $S = \times_{n \geq 1} S_n$ and $T = \times_{n \geq 1} T_n$. We consider S and T to be endowed with the product topologies with the discrete topology on each factor. Denote by $\mathscr{S}$ and $\mathscr{T}$ the Borel σ-algebra of S and T, respectively. A *mixed strategy* of player 1 (resp. 2) is then a probability on $(S, \mathscr{S})$ (resp. $(T, \mathscr{T})$).

A *behavioral strategy* of player 1 (resp. 2) is a sequence $\sigma = (\sigma_n)_{n=1}^\infty$ (resp. $\tau = (\tau_n)_{n=1}^\infty$), where σ_n (resp. τ_n) is a measurable mapping from $(\Omega_\infty, \mathscr{A}_{n-1})$ to $\Delta(A)$ (resp. $\Delta(B)$). By a slight abuse of notation we denote by $\Delta(S)$ and $\Delta(T)$ the sets of all mixed and behavioral strategies of players 1 and 2, respectively. A pure strategy is considered to be a special (degenerate) case of mixed and behavioral strategies.

Every pair of pure strategies (s, t) induces a play $\omega = (\omega_k)_{k=1}^\infty \in \Omega_\infty$, where ω_k is defined inductively as

$$\omega_k = (a_k, b_k) = \begin{cases} (s_1, t_1) & \text{for } k = 1 \\ (s_k(\omega_1, \ldots, \omega_{k-1}), t_k(\omega_1, \ldots, \omega_{k-1})) & \text{for } k > 1. \end{cases}$$

Note that s_1 and t_1, being $\mathscr{A}_0$-measurable, are constant everywhere. Accordingly, every pair $(\sigma, \tau) \in \Delta(S) \times \Delta(T)$ induces a probability $P_{\sigma,\tau}$ on $(\Omega_\infty, \mathscr{A}_\infty)$. Equivalently, (σ, τ) induces a sequence of random actions, or a *random play*, $(X_k)_{k=1}^\infty$, where $X_k = (a_n, b_n)$ is a $(A \times B)$-valued random variable. The expectation operator with respect to $P_{\sigma,\tau}$ is denoted by $E_{\sigma,\tau}$.

For each positive integer n, we define the n-average payoff function r_n: $S \times T \to \mathbb{R}$ by $r_n(s,t) = (1/n)\sum_{k=1}^n r(a_k, b_k)$. Its bilinear extension to $\Delta(S) \times \Delta(T)$ is also denoted by r_n, i.e., $r_n(\sigma, \tau) = E_{\sigma,\tau}[(1/n) \sum_{k=1}^n r(a_k, b_k)]$.

In this paper we study two classes of repeated games:

Finitely repeated game G^n with n-average payoff r_n.

Undiscounted Game G^∞, where the payoff to player 1 from (s,t) is evaluated by the Cesàro limit, if it exists, of the induced sequence of stage payoffs, i.e., $\lim_{n \to \infty} r_n(s,t)$.

For the undiscounted game G^{∞}, the above Cesàro limit does not necessarily exist. We will supplement this loose end by being explicit in the description of the solution concepts in Section 5.2.

For each positive integer n, two pure strategies of a player are said to be n-equivalent if, against any pure strategy of the other player, they induce n-equivalent plays. If two pure strategies induce the same play against any strategy of the other player, they are said to be *equivalent*. Extending this notion to mixed and behavioral strategies, we say that two strategies of a player are n-equivalent if, against any strategy of the other player, they induce the same probability on $(\Omega_{\infty}, \mathscr{A}_n)$. The equivalence of two strategies is similarly defined by replacing $\mathscr{A}_n$ by $\mathscr{A}_{\infty}$ above. Perfect recall implies that every mixed strategy has an equivalent behavioral strategy and vice versa (Kuhn's theorem).

3. INFORMATION THEORETIC CONCEPTS

Let X be a random variable which takes values in a finite set Θ and whose distribution is $p \in \Delta(\Theta)$, i.e., $p(\theta) = \text{Prob}(X = \theta)$ for each $\theta \in \Theta$.

DEFINITION 3.1. The entropy $H(X)$ of X is defined by

$$H(X) = -\sum_{\theta \in \Theta} p(\theta) \log p(\theta) = -E_X[\log p(X)],$$

where $0 \log 0$ is defined to be 0.

To fix the unit, we take the logarithm to the base 2, and we say that entropy is measured in *bits*. The entropy of a random variable depends only on its distribution and not on the particular values it takes. Thus we also write $H(p)$ for the quantity in the above definition and regard H as a function on $\Delta(\Theta)$.

Entropy possesses a number of desirable properties as a measure of uncertainty of a random variable or a probability distribution.

EXAMPLE 3.1. Let $\Theta = \{0, 1\}$. Then for each $p = (p_0, 1 - p_0)$ in $\Delta(\Theta)$ (or any random variable X with this distribution),

$$H(p)(= H(X)) = -p_0 \log p_0 - (1 - p_0)\log(1 - p_0).$$

Figure 1 shows the graph of $H(p)$ which is parameterized by p_0.

Note that if $p_0 = 0$ or 1, then there is no uncertainty involved, and the entropy is indeed 0. On the other hand, the highest value of the entropy, $\log 2 = 1$, is achieved at $p_0 = \frac{1}{2}$. Also observe that $H(p)$ is strictly concave and continuous.

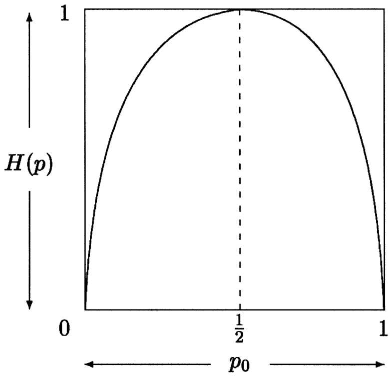

FIG. 1. The graph of an entropy function.

The next proposition summarizes the properties of the entropy. They are easily verified. See, e.g., Cover and Thomas (1991, Chap. 2).

PROPOSITION 3.1. *Let X be a random variable with a finite range Θ. Then*

(1) $0 \leq H(X) \leq \log|\Theta|$,

(2) $H(X) = 0$ *if and only if* $\text{Prob}(X = \theta) = 1$ *for some* $\theta \in \Theta$; $H(X) = \log|\Theta|$ *if and only if* $\text{Prob}(X = \theta) = 1/|\Theta|$ *for every* $\theta \in \Theta$,

(3) *H is continuous and strictly concave as a function on* $\Delta(\Theta)$.

Let $X = (X_1, \dots, X_n)$ be a vector of finite random variables with the range $\times_{k=1}^{n} \Theta_k$. Then by Definition 3.1,

$$H(X_1, \dots, X_n) = -\sum_{\theta_1 \in \Theta_1} \cdots \sum_{\theta_n \in \Theta_n} p(\theta_1, \dots, \theta_n) \log p(\theta_1, \dots, \theta_n),$$

where $p(\theta_1, \dots, \theta_n) = \text{Prob}(X_1 = \theta_1, \dots, X_n = \theta_n)$.

Given a pair of random variables (X_1, X_2) taking values in $\Theta_1 \times \Theta_2$ with the joint distribution $p(\theta_1, \theta_2)$, denote by $p(\theta_2|\theta_1)$ the conditional probability that $X_2 = \theta_2$ given $X_1 = \theta_1$. Define

$$h(X_2|\theta_1) = -\sum_{\theta_2 \in \Theta_2} p(\theta_2|\theta_1) \log p(\theta_2|\theta_1).$$

Thus $h(X_2|\theta_1)$ is the entropy of X_2 when the realization $X_1 = \theta_1$ is known. Consider $h(X_2|\cdot)$ as a random variable on Θ_1 equipped with the marginal distribution of X_1, $p(\theta_1) = \sum_{\theta_2 \in \Theta_2} p(\theta_1, \theta_2)$.

DEFINITION 3.2. The conditional entropy $H(X_2 \mid X_1)$ of X_2 given X_1 is defined by

$$H(X_2|X_1) = E_{X_1}\left[h(X_2|X_1)\right] = \sum_{\theta_1 \in \Theta_1} p(\theta_1) h(X_2|\theta_1).$$

The conditional entropy $H(X_2|X_1)$ measures the (ex ante) average uncertainty of X_2 when the realization of X_1 is observable.

Since $\log p(\theta_1, \theta_2) = \log p(\theta_1) + \log p(\theta_2|\theta_1)$, we have

$$H(X_1, X_2) = H(X_1) + H(X_2|X_1).$$

An application of this equality to $(X_1, \ldots, X_{n-1})$ and X_n yields

$$H(X_1, \ldots, X_n) = H(X_1, \ldots, X_{n-1}) + H(X_n|X_1, \ldots, X_{n-1}).$$

Hence by induction we have the following equality.

PROPOSITION 3.2. $H(X_1, \ldots, X_n) = H(X_1) + \sum_{k=2}^{n} H(X_k| X_1, \ldots, X_{k-1})$.

An extension of the entropy measure to stochastic processes is called *entropy rate* and is defined as follows.

DEFINITION 3.3. Let $(X_n)_{n=1}^{\infty}$ be a stochastic process where each X_n takes values in a finite set Θ. The entropy rate $\overline{H}((X_n))$ of $(X_n)_{n=1}^{\infty}$ is defined by

$$\overline{H}((X_n)) = \limsup_{k \to \infty} \frac{1}{k} H(X_1, \ldots, X_k).$$

Thus the entropy rate is the upper limit of the average uncertainty per bit of the process. For example, for an i.i.d. process the entropy rate coincides with the entropy of the common distribution. This follows immediately from Proposition 3.2 with the i.i.d. assumption. Since $H(X_1, \ldots, X_k) \leq \log|\Theta|^k$ by Proposition 3.1(1), the entropy rate $\overline{H}((X_n))$ is bounded by $\log|\Theta|$.

4. STRATEGIC ENTROPY

We are now ready to define the concepts of strategic entropy and strategic entropy rate. The following discussion focuses on player 1's strategy.

For each positive integer n, define a function $H^n(\cdot : \cdot): \Delta(S) \times T \to \mathbb{R}$ as follows. Given $(\sigma, t) \in \Delta(S) \times T$, let $(X_k)_{k=1}^{\infty}$ be the random play induced by (σ, t). Then $H^n(\sigma : t)$ is defined to be the entropy of this random play *up to stage n*, that is,

$$H^n(\sigma : t) = H(X_1, \dots, X_n) = -\sum_{C \in \mathscr{H}_n} P_{\sigma, t}(C) \log P_{\sigma, t}(C).$$

Recall that $\mathscr{H}_n$ is the partition of Ω_∞ with respect to actions in the first n stages. Thus $H^n(\sigma : t)$ is the uncertainty about the play up to stage n that player 2 faces when player 1 uses σ and player 2 uses t. We summarize the properties of $H^n(\cdot : \cdot)$ as a proposition. See Neyman and Okada (1999) for the proof.

PROPOSITION 4.1. (1) *$H^n(\cdot : \cdot)$ is continuous on $\Delta(S) \times T$.*

(2) *For each $t \in T$, $H^n(\cdot : t)$ is concave on $\Delta(S)$.*

(3) *For each $t \in T$, $H^n(\cdot : t)$ is constant on each n-equivalence class of $\Delta(S)$.*

(4) *For each $\sigma \in \Delta(S)$, $H^n(\sigma : \cdot)$ is constant on each n-equivalence class of T.*

(5) $0 \le H^n(\sigma : t) \le n \log|A|$ *for every $(\sigma, t) \in \Delta(S) \times T$.*

The next lemma follows directly from Proposition 3.2.

LEMMA 4.1. *If $(X_k)_{k=1}^{\infty}$ is the random play induced by (σ, t), then*

$$H^n(\sigma : t) = H(X_1) + \sum_{k=2}^{n} H(X_k | X_1, \dots, X_{k-1}).$$

Suppose that $\sigma = (\sigma_k)_{k=1}^{\infty}$ is a sequence of independent mixed actions, i.e., for each k, there is $\alpha_k \in \Delta(A)$ such that $\sigma_k(\omega) = \alpha_k$ for every $\omega \in \Omega_\infty$. Let $t \in T$ and $(X_k)_{k=1}^{\infty}$ be the random play induced by (σ, t). Then $h(X_k | \omega_1, \dots, \omega_{k-1}) = H(\alpha_k)$. Hence by the definition of the conditional entropy, $H(X_k | X_1, \dots, X_{k-1}) = H(\alpha_k)$ for every k. Therefore, by Lemma 4.1, we have the following lemma.

LEMMA 4.2. *If $\sigma = (\sigma_k)_{k=1}^{\infty}$ is a sequence of independent actions where $\sigma_k \in \Delta(A)$, then*

$$H^n(\sigma : t) = \sum_{k=1}^{n} H(\alpha_k)$$

for every $t \in T$.

For each $\sigma \in \Delta(S)$, we now define the *n-strategic entropy* of σ to be the maximum of $H^n(\sigma : t)$, where the maximum is taken over all pure strategies $t \in T$ of player 2.

DEFINITION 4.1. The n-strategic entropy $H^n(\sigma)$ of $\sigma \in \Delta(S)$ is defined by

$$H^n(\sigma) = \max_{t \in T} H^n(\sigma : t).$$

By Proposition 4.1(1), $H^n(\sigma : \cdot)$ is continuous on the compact set T, and so the above maximum does exist. Alternatively, by Proposition 4.1(4), $H^n(\sigma : \cdot)$ is constant on each n-equivalence classes of T, and since there are only a finite number of the n-equivalence classes, the maximum exists. The next proposition summarizes the properties of the n-strategic entropy.

PROPOSITION 4.2. (1) *$H^n(\sigma)$ is continuous as a function on $\Delta(S)$.*
(2) *If σ and σ' are n-equivalent, then $H^n(\sigma) = H^n(\sigma')$.*
(3) *$0 \le H^n(\sigma) \le n \log|A|$ for every $\sigma \in \Delta(S)$.*
(4) *$H^n(\sigma) = 0$ if and only if σ is n-equivalent to a pure strategy.*

If $(X_n)_{n=1}^{\infty}$ is the random play induced by (σ, t), we denote its entropy rate by $H^{\infty}(\sigma : t)$. By the definition of entropy rate and that of $H^n(\sigma : t)$ we have

$$H^{\infty}(\sigma : t) = \limsup_{n \to \infty} \frac{1}{n} H(X_1, \dots, X_n) = \limsup_{n \to \infty} \frac{1}{n} H^n(\sigma : t).$$

The *strategic entropy rate* of σ is defined to be the supremum of $H^{\infty}(\sigma : t)$ over $t \in T$.

DEFINITION 4.2. The strategic entropy rate $H^{\infty}(\sigma)$ of $\sigma \in \Delta(S)$ is defined by

$$H^{\infty}(\sigma) = \sup_{t \in T} H^{\infty}(\sigma : t).$$

For example, every behavioral strategy that plays pure actions after a finite number of stages has zero strategic entropy rate. An alternative definition of strategic entropy rate is $\hat{H}^{\infty}(\sigma) = \limsup_{n \to \infty}(1/n)H^n(\sigma)$. Clearly $H^{\infty}(\sigma) \le \hat{H}^{\infty}(\sigma)$. This inequality can be strict. See Neyman and Okada (1998, Example 4.4).

To conclude this section we mention an additional concept of strategic entropy. As we remarked in the Introduction, one can view the strategic entropy as a cost of randomization in repeated play. As such, the definition of the strategic entropy in Definition 4.1 has a disadvantage. Recall that $H^n(\sigma : t)$ is the entropy of the random play averaged over all histories. Thus it is possible that after some history with positive, but small, probability, the conditional entropy of the play for the rest of the game is quite large. The definition below would be more adequate if one wishes to bound the entropy of the play conditional on any history with a positive probability.

Given $(\sigma, t) \in \Delta(S) \times T$, let $(X_n)_{n=1}^{\infty}$ be the play generated by (σ, t). Define $\bar{h}^n(\sigma : t)$ by

$$\bar{h}^n(\sigma : t) = \max_{0 \le k \le n} \max_{\substack{\omega_1, \dots, \omega_k \\ P_{\sigma,\tau}(\omega_1, \dots, \omega_k) > 0}} h\big(X_{k+1}, \dots, X_n | \omega_1, \dots, \omega_k\big).$$

Recall that, for random variables X and Y taking values in sets Θ_1 and Θ_2, respectively, $h(Y|\theta_1)$ denotes the entropy of Y given $X = \theta_1 \in \Theta_1$, i.e., $h(Y|\theta_1) = -\sum_{\theta_2 \in \Theta_2} p(\theta_2|\theta_1)\log p(\theta_2|\theta_1)$.

For $k = 0$,

$$\begin{aligned} h\big(X_{k+1}, \dots, X_n | \omega_1, \dots, \omega_k\big) &= h(X_1, \dots, X_n \mid \varnothing) \\ &= H(X_1, \dots, X_n) \\ &= H^n(\sigma : t). \end{aligned}$$

Hence $\bar{h}^n(\sigma : t) \ge H^n(\sigma : t)$ for all n. This inequality can be strict.

EXAMPLE 4.1. Let $A = \{T, B\}$. Fix an $\varepsilon \in (0,1)$ and define a behavioral strategy $\sigma = (\sigma)_{n=1}^{\infty}$ as

$$\sigma_1(\varnothing) = \begin{cases} T & \text{with probability } \varepsilon \\ B & \text{with probability } 1 - \varepsilon, \end{cases}$$

and, for $n > 1$ and $\omega_1, \dots, \omega_{n-1} \in (A \times B)^{n-1}$, if $\omega_1 = (T, \cdot)$, then

$$\sigma_n(\omega_1, \dots, \omega_{n-1}) = \begin{cases} T & \text{with probability } \frac{1}{2} \\ B & \text{with probability } \frac{1}{2}; \end{cases}$$

otherwise, $\sigma_n(\omega_1, \dots, \omega_{n-1}) = T$ with probability 1.

It is easy to see that $H^n(\sigma : t) = H(\varepsilon, 1 - \varepsilon) + \varepsilon(n-1)$ while $\bar{h}^n(\sigma : t) = h(X_2, \dots, X_n | (T, \cdot)) = n - 1$.

In this example $\bar{h}^n(\sigma : t) = O(H^n(\sigma : t))$. It is also possible that $H^n(\sigma : t)$ is bounded uniformly in n while $\bar{h}^n(\sigma : t) \to \infty$ as $n \to \infty$

EXAMPLE 4.2. Let A as in the previous example. Fix a sequence $(\varepsilon_n)_{n=1}^{\infty}$ with $\varepsilon_n \in (0,1)$ for all n. We divide the stages into two classes, the experimental stages and the rest. An experimental stage is a stage of the form $e(l) = \sum_{j=1}^{l} j$ for some l, i.e., $1, 3, 6, 10, 15, \dots$. Note that between the lth and the $(l+1)$th experimental stages, there are l nonexperimental

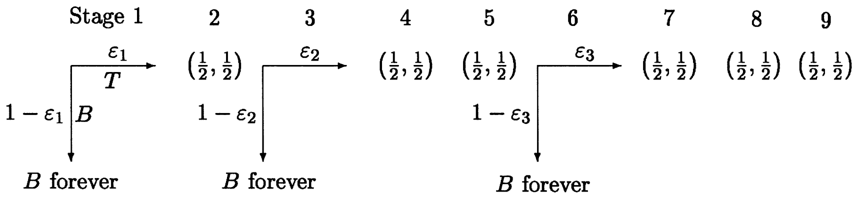

FIG. 2. The strategy of Example 4.2.

stages. Now we define $\sigma = (\sigma_n)_{n=1}^{\infty}$ as follows: if $n = e(l)$ for some l, then

$$\sigma_n(\omega_1, \dots, \omega_{n-1}) = \begin{cases} (\varepsilon_l, 1 - \varepsilon_l) & \text{if } \omega^1_{e(j)} = T \text{ for all } j < l \\ B & \text{otherwise,} \end{cases}$$

and, for n with $e(l) < n < e(l+1)$,

$$\sigma_n(\omega_1, \dots, \omega_{n-1}) = \begin{cases} \left(\frac{1}{2}, \frac{1}{2}\right) & \text{if } \omega^1_{e(j)} = T \text{ for all } j \le l \\ B & \text{otherwise.} \end{cases}$$

See Fig. 2. Let $\omega_1, \dots, \omega_k$ be such that $\omega^1_j = T$ for all $j = 1, \dots, k$. Then, for every positive integer m, $h(X_{k+1}, \dots, X_{k+m} | \omega_1, \dots, \omega_k) = m$, and thus $\bar{h}^n(\sigma : t) \to \infty (n \to \infty)$ for all $t \in T$.

To compute $H^n(\sigma : t)$, note that $H(X_1) = H(\varepsilon_1, 1 - \varepsilon_1)$, $H(X_2 \mid X_1) = \varepsilon_1$, $H(X_3 \mid X_1, X_2) = \varepsilon_1 H(\varepsilon_2, 1 - \varepsilon_2)$, $H(X_4 \mid X_1, X_2, X_3) = H(X_5 \mid X_1, \dots, X_4) = \varepsilon_1 \varepsilon_2$, $H(X_6 \mid X_1, \dots, X_5) = \varepsilon_1 \varepsilon_2 H(\varepsilon_3, 1 - \varepsilon_3)$, and so on. In general, we have

$$H(X_k | X_1, \dots, X_{k-1}) = \begin{cases} \left(\prod_{l=1}^{l-1} \varepsilon_i\right) H(\varepsilon_l) & \text{if } k = e(l) \text{ for some } l \\ \prod_{l=1}^{l-1} \varepsilon_i & \text{if } e(l) < k < e(l+1) \\ & \quad \text{for some } l. \end{cases}$$

Since $H(\varepsilon_l) \le 1$, $H^n(\sigma : t)$ is bounded by $\sum_{l=1}^{\infty} (\prod_{i=1}^{l-1} \varepsilon_i) l$. One can choose ε_l inductively so that $(\prod_{i=1}^{l-1} \varepsilon_i) l = 2^{-l}$, i.e., $\varepsilon_1 = \frac{1}{2}$, $\varepsilon_2 = \frac{1}{4}$, $\varepsilon_3 = \frac{1}{3}$, $\varepsilon_4 = \frac{3}{8}$, etc. In this case $H^n(\sigma : t) \le 1$.

It can be shown that if a mixed strategy σ has a finite support, then $\bar{h}^n(\sigma : t) \le \log |\mathrm{Supp}(\sigma)|$. In particular, if σ is a mixture over finite automata with m states, then for every n and $t \in T$, $\bar{h}^n(\sigma : t) \le Km \log m$, where K is a constant. If $\sigma = (\sigma_n)_{n=1}^{\infty}$ is a sequence of independent actions, i.e., for each n there is a mixed action $\alpha_n \in \Delta(A)$ such that

$\sigma_n(\cdot) = \alpha_n$, then $H^n(\sigma : t) = \sum_{k=1}^n H(\alpha_k)$. In this case, for any two histories $\omega_1, \ldots, \omega_k$ and $\omega'_1, \ldots, \omega'_k$, we have $h(X_{k+1}, \ldots, X_n | \omega_1, \ldots, \omega_k) = h(X_{k+1}, \ldots, X_n | \omega'_1, \ldots, \omega'_k) = H(\alpha_{k+1}) + \cdots + H(\alpha_n)$. So $\underline{h}^n(\sigma : t) = \sum_{k=1}^n H(\alpha_k) = H^n(\sigma : t)$. The strategies constructed in the proofs of Theorems 5.1 and 5.2 are of this type.

5. REPEATED GAMES WITH STRATEGIC ENTROPY BOUND

5.1. *The Finitely Repeated Game $G^n(\eta)$*

Given a finitely repeated game $G^n = (S, T, r_n)$, we consider a modified version in which player 1's strategy is restricted to strategies whose n-strategic entropy does not exceed a prespecified bound. Player 2's set of strategies will be left intact. For $\eta \geq 0$, define $\Sigma^n(\eta)$ to be the subset of $\Delta(S)$ consisting of all strategies whose n-strategic entropy is at most η, i.e., $\Sigma^n(\eta) = \{\sigma \in \Delta(S) | H^n(\sigma) \leq \eta\}$. Let $G^n(\eta)$ be the n-fold repetition of G in which player 1 is restricted to strategies in $\Sigma^n(\eta)$ while player 2's strategy set remains $\Delta(T)$.

The set $\Sigma^n(\eta)$ is in general not convex and therefore $G^n(\eta)$ may not have the value. Consider, for example, the one-shot game of Example 5.1 below. The subsequent discussion will focus on the maximin value of $G^n(\eta)$, $W^n(\eta) = \max_{\sigma \in \Sigma^n(\eta)} \min_{t \in T} r_n(\sigma, t)$. Recall that r_n is continuous on $\Delta(S) \times T$. Thus for each $\sigma \in \Delta(S)$, $r_n(\sigma, \cdot)$ is continuous on the compact set T and hence $\min_{t \in T} r_n(\sigma, t)$ is well defined and continuous in σ. Proposition 4.2 (1) implies that $\Sigma^n(\eta)$ is compact. Therefore, the above expression of the maxmin value of $G^n(\eta)$ is well defined.

We are interested in the asymptotics of this maxmin value when we allow the entropy bound η to be the function of n. Let us also write $\overline{\eta}(n)$ for $(\eta(n))/n$. In Neyman and Okada (1999) it has been shown that if $\overline{\eta}(n) \to 0$ as $n \to \infty$, then $W^n(\eta(n)) \to U_*(G)$ as $n \to \infty$. In other words, if the entropy bound grows more slowly than the number of repetitions, then the most player 1 can guarantee in the long run is the one-shot game maxmin payoff in pure actions. In what follows we will provide a full characterization of the asymptotics of $W^n(\eta(n))$ as $\overline{\eta}(n)$ tends to an arbitrary nonnegative number.

In the stage game $G = (A, B, r)$, and for $\gamma \geq 0$, define

$$U(\gamma) = \max_{\substack{\alpha \in \Delta(A) \\ H(\alpha) \leq \gamma}} \min_{b \in B} r(\alpha, b).$$

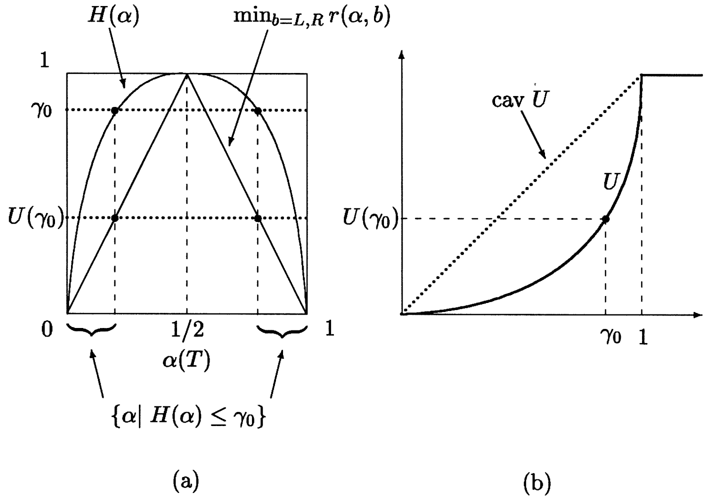

FIG. 3. cav U for Example 5.1.

Thus $U(\gamma)$ is what player 1 can secure in G using a mixed action of entropy at most γ. Note that $H(\alpha) \leq 0$ if and only if α is a pure action. So $U(0) = U_*(G)$. On the other hand, if we allow γ to be at least as large as the entropy of some optimal action of player 1 in G, then he can use that optimal action and achieve $\mathrm{Val}(G)$. Let us define $\bar{\gamma} = \min H(\alpha^*)$, where the minimum is taken over all optimal actions α^* of player 1 in G.[2] Then $U(\gamma) = \mathrm{Val}(G)$ for every $\gamma \geq \bar{\gamma}$. We regard U as a function of nonnegative reals and define cav U to be its concavification, i.e., the smallest concave function which is at least as large as U at every point of its domain.

EXAMPLE 5.1. Consider the 2×2 game

2	0
0	2

Figure 3(a) shows the security level $\min_{b \in B} r(\alpha, b)$ for player 1 superimposed with the entropy of his mixed actions. The horizontal axis measures the probability that player 1 chooses the top row, $\alpha(T)$. In this game $U_*(G) = 0$, $\mathrm{Val}(G) = 1$, and the unique optimal strategy of player 1, $(\frac{1}{2}, \frac{1}{2})$, has entropy $\log 2 = 1$.

[2] It is well defined because the entropy H is a continuous function on $\Delta(A)$ and the optimal actions in G form a compact convex subset of $\Delta(A)$.

To illustrate how one can obtain U, take an antropy level γ_0 in Fig. 3(a). Drop two vertical lines from the two intersections of the graph of the entropy function and the horizontal line at the level γ_0. Each of the two vertical lines meets the graph of $\min_{b \in B} r(\alpha, b)$ at some level of payoff, and the maximum of these two levels is $U(\gamma_0)$. In this example the two levels are the same because $\min_{b \in B} r(\alpha, b)$ is symmetric with respect to $\alpha(T) = \frac{1}{2}$.

From this picture one can infer the shapes of U and cav U as depicted in Fig. 3(b): up to $\bar{\gamma} = 1$, $U(\gamma)$ is the inverse image of the concave function H along the (piecewise) linear function, hence it is a convex function for $\gamma \leq \bar{\gamma}$.

In general, $\min_{b \in B} r(\alpha, b)$ is, as a function of α, piecewise linear and concave. Figure 4(a) depicts a typical case for a $2 \times k$ game along with the entropy of mixed actions. (Player 1 has pure actions, say, T and B, and the probability of choosing T is $\alpha(T)$.)

To obtain the picture of U, note that we can write $U(\gamma) = \max_{i=1,2} U_i(\gamma)$, where

$$U_1(\gamma) = \max_{\substack{\alpha(T) \leq 1/2 \\ H(\alpha) \leq \gamma}} \min_{b \in B} (\alpha, b)$$

and

$$U_2(\gamma) = \max_{\substack{\alpha(T) \geq 1/2 \\ H(\alpha) \leq \gamma}} \min_{b \in B} (\alpha, b).$$

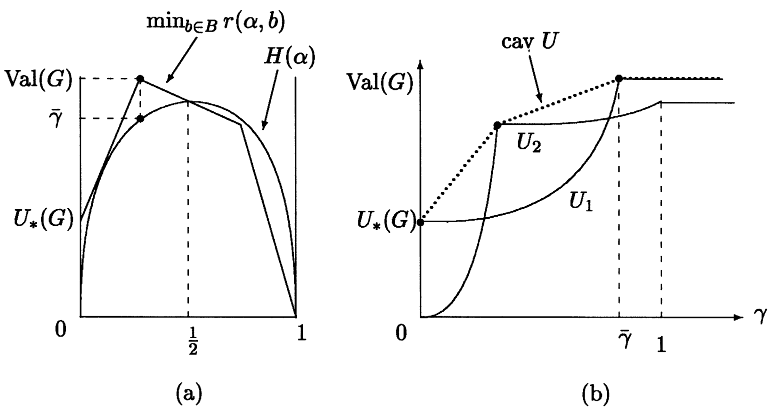

FIG. 4. cav U for a $2 \times k$ game.

In Fig. 4(b), the upper envelope of U_1 and U_2 is U. Thus the graph of U consists of four pieces of strictly increasing convex parts and a horizontal part. In general, U is strictly increasing and piecewise convex up to $\bar{\gamma}$ and then becomes a constant, $\mathrm{Val}(G)$, thereafter. Hence $\mathrm{cav}\, U$ is nondecreasing and piecewise linear.

Note that if $(X_k)_{k=1}^{\infty}$ is a random play induced by a pair of strategies, then $h(X_k|X_1,\ldots,X_{k-1})$ is $\mathscr{A}_{k-1}$-measurable. In the proof of the next theorem we denote $h(X_k|X_1,\ldots,X_{k-1})$ (resp. $H(X_k|X_1,\ldots,X_{k-1})$) by $h(X_k|\mathscr{A}_{k-1})$ (resp. $H(X_k|\mathscr{A}_{k-1})$).

THEOREM 5.1. *For every n,*

$$(\mathrm{cav}\, U)(\bar{\eta}(n)) - \frac{\mathrm{Val}(G) - U_*(G)}{n} \le W^n(\eta(n)) \le (\mathrm{cav}\, U)(\bar{\eta}(n)).$$

And therefore, for every $\gamma \ge 0$, if $\lim_{n\to\infty} \bar{\eta}(n) = \gamma$, then

$$\lim_{n\to\infty} W^n(\eta(n)) = (\mathrm{cav}\, U)(\gamma).$$

Proof. First we describe player 1's strategy that guarantees him a payoff at least as large as the left-hand side of the above inequality. Since U is a continuous function defined on nonnegative reals, there exist $\gamma_0 \ge 0$, $\gamma_1 \ge 0$, and $p \in [0,1]$ such that

(1) $\gamma_0 \le \bar{\eta}(n) \le \gamma_1$,

(2) $\bar{\eta}(n) = (1-p)\gamma_0 + p\gamma_1$,

(3) $(\mathrm{cav}\, U)(\gamma_i) = U(\gamma_i)$, $i = 0, 1$, and

(4) $(\mathrm{cav}\, U)(\bar{\eta}(n)) = (1-p)U(\gamma_0) + pU(\gamma_1)$.

See Fig. 5.

Let α_i, $i = 0,1$, be player 1's actions with $H(\alpha_i) \le \gamma_i$ that guarantees him $U(\gamma_i)$, i.e., $\min_{b\in B} E_{\alpha_i}[r(a,b)] = U(\gamma_i)$.

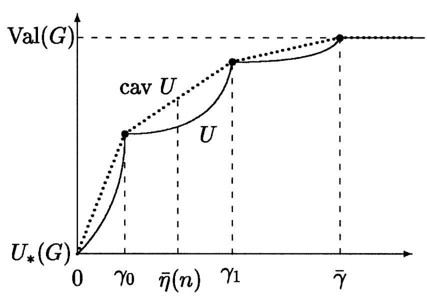

FIG. 5. Splitting $\bar{\eta}(n)$ into γ_0 and γ_1.

Now let $\sigma = (\sigma_k)_{k=1}^{\infty}$ be a behavioral strategy of player 1 such that for every k and every $\omega \in \Omega_\infty$;

$$\sigma_k(\omega) = \begin{cases} \alpha_1 & \text{for } k = 1, \dots, [pn] \\ \alpha_0 & \text{for } k = [pn] + 1, \dots, n, \end{cases}$$

where $[pn]$ is the integer part of pn. Then, the strategic entropy of σ is simply the entropy of the sequence of independent random actions in which the first $[pn]$ terms are α_1's followed by $n - [pn]$ terms of α_0's. Therefore, σ indeed has the strategic entropy within the specified bound $\eta(n)$,

$$\begin{aligned} H^n(\sigma) &= [pn]H(\alpha_1) + (n - [pn])H(\alpha_0) \\ &\le pn\gamma_1 + (1-p)n\gamma_0 \\ &= n\bar{\eta}(n) \\ &= \eta(n), \end{aligned}$$

where the first equality follows from the definition of $H^n(\sigma)$ and Lemma 4.2. By playing σ, player 1 guarantees himself $U(\gamma_1)$ at every stage in the first $[pn]$ rounds and then $U(\gamma_0)$ at every stage for the rest of the game, and hence for any pure strategy t of player 2,

$$\begin{aligned} r_n(\sigma, t) &\ge \frac{[pn]U(\gamma_1) + (n - [pn])U(\gamma_0)}{n} \\ &\ge pU(\gamma_1) + (1-p)U(\gamma_0) - \frac{\mathrm{Val}(G) - U_*(G)}{n} \\ &= (\mathrm{cav}\, U)(\bar{\eta}(n)) - \frac{\mathrm{Val}(G) - U_*(G)}{n}. \end{aligned}$$

Next we show that player 2 can prevent player 1 from achieving a payoff greater than $(\mathrm{cav}\, U)(\bar{\eta}(n))$. It suffices to consider behavioral strategies of player 1. Take $\sigma = (\sigma_k)_{k=1}^{\infty} \in \Sigma^n(\eta(n))$. Define a strategy $t = (t_k)_{k=1}^{\infty} \in T$ of player 2 in such a way that, at every stage k and for every history $\omega \in \Omega_\infty$, $t_k(\omega)$ minimizes player 1's stage payoff $E_{\sigma_k(\omega)}[r(a,b)]$. Denote the random play induced by (σ, t) by $(X_k)_{k=1}^{\infty}$. By the definition of U and t, and since $(\mathrm{cav}\, U)(\gamma) \ge U(\gamma)$ for every $\gamma \ge 0$, we have

$$E_{\sigma,t}[r(X_k)|\mathscr{A}_{k-1}] \le U(h(X_k|\mathscr{A}_{k-1})) \le (\mathrm{cav}\, U)(h(X_k|\mathscr{A}_{k-1})).$$

Thus by the concavity of cav U and Jensen's inequality,

$$E_{\sigma,t}[r(X_k)] \le E_{\sigma,t}[(\text{cav}\, U)(h(X_k|\mathscr{A}_{k-1}))]$$
$$\le (\text{cav}\, U)(E_{\sigma,t}[h(X_k|\mathscr{A}_{k-1})]). \quad (2)$$

Finally, we obtain

$$\begin{aligned} r_n(\sigma,t) &= \frac{1}{n}\sum_{k=1}^{n} E_{\sigma,t}[r(X_k)] \\ &\le^{(a)} \frac{1}{n}\sum_{k=1}^{n} (\text{cav}\, U)(E_{\sigma,t}[h(X_k|\mathscr{A}_{k-1})]) \\ &\le^{(b)} (\text{cav}\, U)\left(\frac{1}{n}\sum_{k=1}^{n} E_{\sigma,t}[h(X_k|\mathscr{A}_{k-1})]\right) \\ &=^{(c)} (\text{cav}\, U)\left(\frac{1}{n}\sum_{k=1}^{n} H(X_k|\mathscr{A}_{k-1})\right) \\ &=^{(d)} (\text{cav}\, U)\left(\frac{H^n(\sigma:t)}{n}\right) \\ &\le^{(e)} (\text{cav}\, U)\left(\frac{H^n(\sigma)}{n}\right) \\ &\le^{(f)} (\text{cav}\, U)\left(\frac{\eta(n)}{n}\right), \end{aligned}$$

where (a) is from (2), (b) is by the concavity of cav U and Jensen's inequality, (c) is from the definition of conditional entropy, (d) is from the definition of $H^n(\sigma:t)$ and Lemma 4.1, (e) is from the definition of the strategic entropy, and (f) is by the assumed bound on the n-entropy. This completes the proof of the theorem. Q.E.D.

5.2. *The Infinitely Repeated Game $G^\infty(\gamma)$*

The game $G^\infty(\gamma)$, $\gamma \ge 0$, is obtained from G^∞ by restricting player 1's set of strategies to $\Sigma^\infty(\gamma) = \{\sigma \in \Delta(S) \mid H^\infty(\sigma) \le \gamma\}$. Again, player 2's strategy set remains intact.

The next theorem asserts two things which are considered to be natural requirements for the payoff, $(\text{cav}\, U)(\gamma)$, to be the maxmin value of $G^\infty(\gamma)$. First, player 1 can "guarantee" $(\text{cav}\, U)(\gamma)$ in a rather strong sense that he has such a strategy that, regardless of player 2's strategy, the expected average payoff in the first n stages is at least $(\text{cav}\, U)(\gamma)$ for every n.

Second, for every strategy of player 1, player 2 has a counterstrategy that keeps player 1's payoff arbitrarily close to $(\operatorname{cav} U)(\gamma)$ in the long run.

THEOREM 5.2. (1) $\exists \sigma \in \Sigma^{\infty}(\gamma)$ *such that* $\forall t \in T, \forall n$,

$$E_{\sigma,t}\left[\frac{1}{n}\sum_{k=1}^{n} r(a_k, b_k)\right] \geq (\operatorname{cav} U)(\gamma).$$

(2) $\forall \sigma \in \Sigma^{\infty}(\gamma)$, $\exists t \in T$ *such that* $\forall \varepsilon > 0$, $\exists n(\varepsilon)$ *such that* $\forall n > n(\varepsilon)$,

$$E_{\sigma,t}\left[\frac{1}{n}\sum_{n=1}^{n} r(a_k, b_k)\right] \leq (\operatorname{cav} U)(\gamma) + \varepsilon.$$

Proof. Fix $\gamma \geq 0$. As in the proof of Theorem 5.1, let $\gamma_i \geq 0$, $i = 0, 1$, and $0 \leq p \leq 1$ be such that $\gamma_0 \leq \gamma \leq \gamma_1$, $\gamma = (1-p)\gamma_0 + p\gamma_1$, $(\operatorname{cav} U)(\gamma_i) = U(\gamma_i)$, $i = 0, 1$, and $(\operatorname{cav} U)(\gamma) = (1-p)U(\gamma_0) + pU(\gamma_1)$. Let $\alpha_i \in \Delta(A)$, $i = 0, 1$, be such that $H(\alpha_i) \leq \gamma_i$ and $\min_{b \in B} E_{\alpha_i}[r(a, b)] = U(\gamma_i)$.

We now define a behavioral strategy of player 1, $\sigma = (\sigma_k)_{k=1}^{\infty} \in \Sigma^{\infty}(\gamma)$ that satisfies the first statement of the theorem. Set $\sigma_1 = \alpha_1$. For $k > 1$ define by induction

$$\sigma_k = \begin{cases} \alpha_0 & \text{if } \dfrac{1}{k}\displaystyle\sum_{m=1}^{k-1} I(\sigma_m = \alpha_1) \geq p \\ \alpha_1 & \text{otherwise.} \end{cases}$$

That is, at each stage player 1 asks himself, "If I do not play α_1 at this stage, does the average frequency of $\sigma_m = \alpha_1$ still exceed p?" If the answer is yes, then he plays α_0, otherwise he plays α_1. It is easily verified by induction that for every $n = 1, 2, \ldots,$

$$p \leq \frac{1}{n}\sum_{k=1}^{n} I(\sigma_k = \alpha_1) < p + \frac{1}{n}. \tag{3}$$

We verify that σ has strategic entropy rate at most γ. Take $t \in T$ arbitrarily and let $(X_k)_{k=1}^{\infty}$ be the random play induced by (σ, t). Then by the construction of σ and Lemma 4.2, we have, for every n,

$$\begin{aligned}\frac{1}{n}H(X_1,\dots,X_n) &= \frac{1}{n}\sum_{k=1}^{n} H(\sigma_k)\\ &= \frac{1}{n}\sum_{k=1}^{n}\left(H(\alpha_0)I(\sigma_k=\alpha_0)+H(\alpha_1)I(\sigma_k=\alpha_1)\right)\\ &\le \gamma_0\frac{1}{n}\sum_{k=1}^{n} I(\sigma_k=\alpha_0)+\gamma_1\frac{1}{n}\sum_{k=1}^{n} I(\sigma_k=\alpha_1)\\ &\le \left(1-p-\frac{1}{n}\right)\gamma_0+\left(p+\frac{1}{n}\right)\gamma_1\\ &= \gamma+\frac{\gamma_1-\gamma_0}{n}.\end{aligned}$$

The last inequality follows from the second inequality in (3). Hence by the definition of strategic entropy rate, $H^\infty(\sigma)\le\gamma$.

As for the payoff resulting from (σ, t), we have

$$\begin{aligned}&E_{\sigma,t}\left[\frac{1}{n}\sum_{k=1}^{n} r(X_k)\right]\\ &= \frac{1}{n}\sum_{k=1}^{n}\left(E_{\alpha_0}[r(a_k,b_k)]I(\sigma_k=\alpha_0)+E_{\alpha_1}[r(a_k,b_k)]I(\sigma_k=\alpha_1)\right)\\ &\ge U(\gamma_0)\frac{\sum_{k=1}^{n} I(\sigma_k=\alpha_0)}{n}+U(\gamma_1)\frac{\sum_{k=1}^{n} I(\sigma_k=\alpha_1)}{n}\\ &\ge (1-p)U(\gamma_0)+pU(\gamma_1)\\ &= (\operatorname{cav} U)(\gamma).\end{aligned}$$

The last inequality follows from the first inequality in (3). This completes the proof of the first part.

The proof for the second part proceeds in the same way as in the second part of the proof of Theorem 1. We omit the details. Q.E.D.

For the strategy constructed in the above proof the average frequency that player 1 plays α_1 exceeds p at every stage. As a consequence the average frequency that he receives at least $(\operatorname{cav} U)(\gamma_1)$ exceeds p at every stage, and hence his expected average payoff is at least $(\operatorname{cav} U)(\gamma) = (1-p)(\operatorname{cav} U)(\gamma_0)+p(\operatorname{cav} U)(\gamma_1)$ at every stage. Such strategy is permissible because the definition of strategic entropy rate does not place any bound on the entropy of a play in finite stages. In particular, in the above proof, $(1/n)H(X_1, \ldots, X_n)$ may exceed γ for every n. By modifying the sequence $(\sigma_k)_{k=1}^{\infty}$, where $\sigma_k = \alpha_0$ or α_1, one can prove a similar result as the first part of the above theorem but with a bound on the entropy of the play up to each finite stage. More precisely, for every $\eta(\cdot)$ with $\lim_{n\to\infty}(\eta(n)/n) = \gamma \geq 0$, there is $\sigma = (\sigma_n)_{n=1}^{\infty}$ such that (1) for every $t = (t_k)_{k=1}^{\infty} \in T$ and every n, $H(X_1, \ldots, X_n) \leq \eta(n)$, and (2) for every $\varepsilon > 0$ there is $n(\varepsilon)$ such that for every $t \in T$ and every $n > n(\varepsilon)$, $E_{\sigma,t}[(1/n)\sum_{k=1}^{n} r(a_k, b_k)] \geq (\text{cav}\, U)(\gamma) - \varepsilon$.

REFERENCES

Ben-Porath, E. (1993). "Repeated Games with Finite Automata," *J. Econom. Theory* **59**, 17–32.

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

Lehrer, E. (1988). "Repeated Games with Stationary Bounded Recall Strategies," *J. Econom. Theory* **46**, 130–144.

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

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

Neyman, A. (1997). "Cooperation, Repetition, and Automata," in *Cooperation: Game-Theoretic Approaches* (S. Hart and A. Mas-Colell, Eds.), NATO Advanced Science Institutes Series F: Computer and Systems Sciences, Vol. 155, pp. 233–255. Berlin: Springer–Verlag.

Neyman, A. (1999). "Finitely Repeated Games with Finite Automata," *Math. Oper. Res.* **23**, 513–552.

Neyman, A., and Okada, D. (1999). "Strategic Entropy and Complexity in Repeated Games," *Games Econom. Behav.* **29**, 191–223.

Papadimitriou, C. H., and Yannakakis, M. (1994). "On Complexity as Bounded Rationality: Extended Abstract," in *26th ACM Symposium on Theory of Computing*, Montreal, Quebec, Canada, pp. 726–733.

Papadimitriou, C. H., and Yannakakis, M. (1998). "On Bounded Rationality and Computational Complexity," mimeo.