Games and Economic Behavior **29,** 191–223 (1999)
Article ID game.1998.0674, available online at http://www.idealibrary.com on IDEAL®

Strategic Entropy and Complexity in Repeated Games

Abraham Neyman*

Institute of Mathematics, The Hebrew University, 91904 Jerusalem Israel, and SUNY at Stony Brook, Stony Brook, New York 11794-4384

and

Daijiro Okada†

Center for Rationality, The Hebrew University, 91904 Jerusalem, Israel

Received April 16, 1997

We introduce the entropy-based measure of uncertainty for mixed strategies of repeated games—strategic entropy. We investigate the asymptotic behavior of the *maxmin* values of repeated two-person zero-sum games with a bound on the strategic entropy of player 1's strategies while player 2 is unrestricted, as the bound grows to infinity. We apply the results thus obtained to study the asymptotic behavior of the value of the repeated games with finite automata and bounded recall. *Journal of Economic Literature* Classification Numbers: C73, C72.

1. INTRODUCTION

Strategic complexity and bounded rationality in the context of repeated games have been extensively studied. Among the numerous directions of inquiry in this area we are interested in how equilibrium outcomes are affected when we impose an exogenous restriction on the sets of strategies available to the players by means of *finite automata* and *bounded recall*. For other complexity and rationality issues and approaches to them, the reader is referred to a survey by Kalai (1990).

Finite automata and bounded recall are possible alternatives to describe strategies in a repeated game which have "finite memories." As suggested in Aumann (1981), such formulations put bounds on the complexity of

* E-mail: aneyman@math.huji.ac.il.
† E-mail: dokada@math.huji.ac.il.

0899-8256/99 $30.00

strategies and enable an analysis in the framework of finite games. A memory may be interpreted as a ‘‘state of the player's mind.’’ Restricting strategies to those with a finite memory thus amounts to postulating limitations on the amount of information a player can process in deciding what to do in the course of a play. An example of such a constraint is the inability to count the number of stages played. Alternatively, one can think of a player using an external device, such as a computer, to carry out his plans, and this machine has some hardware constraints.

In the case of an automaton, memories are represented by its *states*. To each state is assigned an action to be taken whenever the automaton is at that state. Starting at a prespecified initial state, an automaton undergoes the transition of its states according to a rule which determines the next state as a function of the current state and the other players' actions at the last stage.[1] The more states an automaton has, the more complicated is the strategy—depending on the greater variety of past actions—that it can carry out. Thus the number of states of an automaton, called the *size* of that automaton, serves to distinguish quantitatively between ‘‘simple’’ and ‘‘complicated’’ strategies. On the other hand, bounded recall limits how far back in the past a player can remember, i.e., the length of recall. Such a strategy can be represented by a function that assigns actions only to the sequences of past actions with a limited length. In the early stages of the game, however, the sequence of actual actions may be too short for this function to be ‘‘activated’’ to play a repeated game. In order to fulfill the time span required, a surrogate memory—formally termed an *initial memory*—is introduced.[2] Again, the longer the length of recall, the more complicated the patterns of actions a player can take.

Aumann (1981) also includes a study of the undiscounted infinitely repeated prisoner's dilemma with a special case of finite memory, due to Aumann, Cave, and Kurz: action at each stage is allowed to depend only on the last action taken by the other player; i.e., the length of recall is 1. Such strategies are called ‘‘reactive’’ and can be implemented by automata with at most two states. They examine the payoff matrix of this restricted game (8×8 bimatrix) and show that the process of successive weak domination eliminates all but ‘‘Tit-for-Tat’’ strategies for both players. The

[1] This type of automaton, which we will use in this paper, is sometimes called an *exact* automaton. In contrast, a *full* automaton takes into account all players' actions. Depending on the issues and questions addressed, each of the two types has its advantage. See Kalai (1990), Section 4.1, and also Neyman (1997), Section 7. The results in Kalai (1990) are stated in terms of full automata.

[2] Again, there are two alternatives for the domain of such a function: only other players' actions or all players' actions. Unlike our choice for automata, we will take the second alternative for bounded recall. In Section 7.2 we will show that bounded recall can be viewed as a special case of automata.

outcome is cooperation at every stage. See also Kalai *et al*. (1988) for the discounted case.

Another study of repeated prisoner's dilemma with complexity bounds is found in Neyman (1985). He studies a finitely repeated (undiscounted) version in which the players are restricted to the strategies that are implementable by automata of bounded sizes. It is well known that, without any restriction on the set of strategies, every strategic equilibrium of this game leads to defections, or double-crossing, throughout the stages.[3] Neyman (1985) states that if the bounds are so small that the players are unable to count up to the last stage, then cooperation can prevail at every stage in equilibrium. Furthermore, even if players are allowed to use automata of much larger sizes, e.g., any power of the number of repetitions, one can still achieve cooperation at most of the stages provided that the game is sufficiently long. More general results on the set of equilibrium payoffs of finitely repeated two-person games with bounded automata can be found in Neyman (1998). See also Papadimitriou and Yannakakis (1994) and Zemel (1989).

In Neyman (1998), questions and results are formalized as the asymptotics of equilibrium outcomes of the restricted repeated games. He considers a sequence of finitely repeated two-person games that are parameterized by the number of repetitions and bounds on the size of automata for each player. The main theorem of this paper specifies the condition on the order of magnitude of these parameters under which the set of equilibrium payoffs converges to the set of feasible and individually rational payoffs as the parameters tend to infinity.

This approach has been useful in that it provides a framework to study a set of seemingly different questions concerning repeated games with finite automata or bounded recall. For example, Ben-Porath (1993) studies undiscounted infinitely repeated two-person zero-sum games. His investigation centers around the question regarding how much of an advantage a bigger automaton has over a smaller one. Suppose that players 1 and 2 are restricted to strategies implementable by automata of size n and $m(n)$ $(> n)$, respectively. The results are stated in terms of the asymptotics of the value of the restricted game $V_{n,m(n)}$ as $n \to \infty$. Interestingly, a player with a larger bound has an advantage only if his bound is exponentially larger than the other player's bound. Otherwise, the player with a smaller bound can asymptotically secure the value of the one-shot game. Even being polynomially bigger than the other bound is not big enough. For-

[3] Let us remark that the proof of this statement is more subtle than the simple backward induction argument which is sometimes erroneously given in the literature. One shows by induction that, at any equilibrium path, the actions in the last k stages of the game are (defect, defect) with probability 1.

mally,

$$\lim_{n \to \infty} \frac{\log m(n)}{n} = 0 \quad \Rightarrow \quad \lim_{n \to \infty} V_{n, m(n)} = v,$$

where v is the value of the stage game. Lehrer (1988) has similar results for bounded recall.

Study of the zero-sum two-person case is important for that of the non-zero-sum case; it provides insights into individually rational levels and effective punishments, and thus the set of equilibrium payoffs. See Ben-Porath (1993, Sect. 4), Lehrer (1988, Sect. 5.5), and also Lehrer (1994). More recent results and open problems in this line of research can be found in Neyman (1997).

In this paper we study repeated two-person zero-sum games in which a player, say player 1, with a restricted set of strategies plays against an unrestricted player, player 2. Concerning the values of such games, two questions arise: "What is the number of repetitions needed for player 2 to take advantage of player 1's restriction?", and "How long can player 1 protect himself against an unrestricted player 2?" The questions are again formulated as the asymptotic behavior of the value of the repeated game: "What is the relationship between the number of repetitions and the complexity bound so that, as they tend to infinity, either (a) player 2 can hold player 1's payoff down to his one-shot game maxmin value in pure actions or (b) player 1 can still secure the value of the one-shot game?" Neyman (1997) has conjectures on these questions for repeated games with finite automata and bounded recall. The present work has stemmed from an attempt to resolve the conjecture corresponding to question (a) above. Precise statements of the conjectures will be given in Section 7.

Our approach is to deduce the results concerning the value of a repeated game with finite automata and bounded recall from a result concerning the "maxmin" value of a repeated game with a restriction directly on the set of mixed strategies.

Note that finite automata and bounded recall limit the set of pure strategies in a repeated game. However, any mixture on the restricted set of pure strategies is allowed. Geometrically, this amounts to saying that the set of strategies is restricted to a *face* of the mixed strategy simplex. Since a face of a simplex is a compact convex set, the value of the repeated game with finite automata or bounded recall exists by the minimax theorem.

In contrast, we impose a restriction directly on mixed strategies. To this end we employ an information theoretic quantity called *entropy*. Entropy is a measure of uncertainty of random variables and is a functional of probability distributions. Every mixed strategy is a probability distribution over pure strategies and thus its entropy is well defined. To define *strategic entropy*, however, we will take into account the uncertainty of a mixed strategy relative to the other player's strategy. Specifically, we look at the probability distribution on the play induced by a pair of strategies and its entropy. The strategic entropy of a mixed strategy is then defined to be the maximum entropy of the play with respect to the other player's strategy. Thus it is the maximum uncertainty of the play that the other player faces against that mixed strategy.

We restrict player 1's strategies to those that have strategic entropy less than a prespecified bound. A mixed strategy with a small strategic entropy would be "not so uncertain" and hence "close" to a pure strategy, and vice versa. We argue that a strategy with relatively little strategic entropy compared to the number of repetitions eventually "reveals" its (pure) actions in the course of a play, and an unrestricted player may take advantage of it. To be more precise, our main theorem states that if the bound on strategic entropy is of a smaller order than the number of repetitions, then the *maxmin* value for player 1 in the restricted game is asymptotically his one-shot game maxmin level in pure actions.

The crucial fact in applying our result to finite automata and bounded recall is that any strategy implementable by a bounded automaton or bounded recall has also a bounded entropy. This implies that the subset of player 1's strategies with a large enough entropy bound contains a face spanned by bounded automata or bounded recall strategies. Therefore, the maxmin value for player 1 on this subset is at least the corresponding value, which is equal to *the* value, of the repeated game with finite automata or bounded recall. This enables us to deduce the conjectures mentioned above as corollaries of our theorem. For example, in Section 7, we will consider the finitely repeated game $G^n_{m(n)}$ in which player 1's strategies are restricted to those implementable by automata with $m(n)$ states, a function of the number of repetitions n, and then study the asymptotics of its value, $V^n_{m(n)}$. We will see that the number of (the equivalence classes of) such pure strategies is of the order $m(n)^{Om(n)}$ and any mixture of such strategies has entropy of the order $O(m(n) \log m(n))$. Now a main result of Section 6 implies that for every $\varepsilon > 0$ there is $\gamma > 0$ such that for any mixed strategy of player 1 whose entropy is a sufficiently small fraction of n, i.e., at most γn, player 2 has a counterstrategy which holds player 1's average payoff within ε of U_*, the maxmin payoff in pure actions of the stage game. This in turn implies that if $m(n) \log m(n) \to 0$ as $n \to \infty$, then the maxmin payoff to player 1 in $G^n_{m(n)}$ tends to U_* as $n \to \infty$. But since the mixed strategy sets in $G^n_{m(n)}$ are compact and convex, its minimax value is indeed its value $V^n_{m(n)}$.

Let us make a precautionary remark: we are not suggesting that our entropy concept is a measure of strategic complexity in the way that the size of an automaton or the length of recall is intended to be. Our view in this paper is that entropy captures an abstract informational feature common in automata and bounded recall restrictions and thereby serves as a useful tool to analyze them.

A basic model of repeated games is described in the next section. In Section 3 we review some information theoretic concepts such as entropy and conditional entropy. Section 4 is devoted to the definition of strategic entropy and its variants. We discuss the relation between strategic entropy and the ordinary entropy of a strategy in Section 5. Section 6 contains the main results on repeated games with bounded strategic entropy. Section 7 is devoted to an application of the theorem to finite automata and bounded recall which gives positive resolutions to the two conjectures of Neyman (1997). Section 8 concludes the paper.

2. PRELIMINARIES

Henceforth, $\mathbb{N}$ denotes the set of all positive integers, and $\mathbb{R}$ denotes the set of all real numbers. For any $x \in \mathbb{R}$, $[x]$ is the integer part of x. For any set Θ, we denote the set of all probability distributions on it by $\Delta(\Theta)$. The n-fold Cartesian product of Θ is denoted by Θ^n. For any probability p (resp. a random variable X), the expectation operator with respect to p (resp. the distribution of X) is denoted by E_p (resp. E_X).

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 player 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. Thus $\Delta(A)$ and $\Delta(B)$ are the sets of mixed actions of the two players.

We denote the *maxmin value in pure actions* and the *value* of G by $\mathrm{U}_*(G)$ and $\mathrm{Val}(G)$, respectively. That is, $\mathrm{U}_*(G) = \max_{a \in A} \min_{b \in B} r(a, b)$, and $\mathrm{Val}(G) = \min_{\beta \in \Delta(B)} \max_{a \in A} E_\beta[r(a, b)]$. By the minimax theorem we can also write $\mathrm{Val}(G) = \max_{\alpha \in \Delta(A)} \min_{b \in B} E_\alpha[r(a, b)]$.

Given a game $G = (A, B, r)$, we next describe a new game in which G is played repeatedly. Each of the repetitions is referred to as a *stage* and G as the *stage game*. At each stage, each player independently takes a stage game action. The pair of actions is announced to both players and the transfer of the corresponding payoff from one player to the other takes place. The information structure of the repeated game is that of perfect recall. That is, at any stage each player remembers all information he received (the sequence of actions taken by both players) up to that stage.

This description is common knowledge. The formal description follows.

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 $k = 1, \dots, n$. 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., $\omega_1, \dots, \omega_n$. We denote by $\mathscr{A}_n$ the algebra on Ω_∞ generated by $\mathscr{H}_n$. Set $\mathscr{A}_0 = \{\phi, \Omega_\infty\}$. Clearly, $\mathscr{A}_{n-1} \subset \mathscr{A}_n$ for every $n \in \mathbb{N}$. 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 coordinates of $\omega = (\omega_k)_{k=1}^{\infty}$, we sometimes write $s_n(\omega_1, \dots, \omega_{n-1})$. 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. Then they are compact metrizable spaces. 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 strategy.

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 follows:

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

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 an $(A \times B)$-valued random variable. The expectation operator with respect to $P_{\sigma,\tau}$ will be denoted by $E_{\sigma,\tau}$.

For each $n \in \mathbb{N}$ we define the n-average payoff function r_n: $S \times T \to \mathbb{R}$ by $r_n(s, t) = (1/n)\sum_{k=1}^{n} r(a_k, b_k)$. Also, for each $\lambda \in [0, 1)$ we define the λ-discounted payoff function r_λ: $S \times T \to \mathbb{R}$ by $r_\lambda(s, t) = (1 - \lambda)\sum_{k=1}^{\infty} \lambda^{k-1} r(a_k, b_k)$. The bilinear extensions of r_n and r_λ to $\Delta(S) \times \Delta(T)$ are denoted by the same symbols, i.e., $r_n(\sigma, \tau) = E_{\sigma,\tau}[(1/n)\sum_{k=1}^{n} r(a_k, b_k)]$ and $r_\lambda(\sigma, \tau) = E_{\sigma,\tau}[(1 - \lambda)\sum_{k=1}^{\infty} \lambda^{k-1} r(a_k, b_k)]$. Both r_n and r_λ are continuous on $S \times T$. The continuity extends to $\Delta(S) \times \Delta(T)$ with respect to the product topology with the weak* topology on each factor.

In this paper we study three classes of repeated games differentiated by their payoff functions:

- *Finitely Repeated Game* G^n with n-average payoff r_n;
- λ-*Discounted Game* G_λ with λ-discounted payoff r_λ;
- *Undiscounted Game* G^∞, where the payoff from (s, t) is evaluated by the Cesàro limit 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 6.3.

For each $n \in \mathbb{N}$, two pure strategies 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. REVIEW OF INFORMATION THEORETIC CONCEPTS

Throughout this section we deal with finite random variables and distributions. For a more general, measure theoretic, treatment of entropy, see Smorodinsky (1971).

Let Θ be a finite set. Let X be a random variable which takes values in Θ and whose distribution is $p \in \Delta(\Theta)$, i.e., $p(\theta) = \text{Prob}(X = \theta)$ for each $\theta \in \Theta$. Note that the distribution p can be considered as a real random variable on Θ so that the expectation (with respect to p) of a measurable transformation of p itself is well defined.

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

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

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*. For example, the entropy of tossing a fair coin once is 1 bit, twice is 2 bits, and so on.

It is obvious by the definition that the entropy of a random variable depends only on its distribution and not on the particular values it takes.[4] 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. If one interprets the amount of uncertainty removed as the amount of information gained, then the entropy also measures the amount of information contained in a random variable or a probability distribution.[5]

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, if $p_0 = \frac{1}{2}$, there is no greater likelihood of either of the two outcomes and this corresponds to the highest value of the entropy, $\log 2 = 1$. Also observe that $H(p)$ is strictly concave and continuous.

It is easy to verify that the properties of entropy illustrated in this example hold in general. See 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* $\theta \in \Theta$;

[4] Furthermore, it depends only on the unordered profile of probabilities; for example, if two random variables taking values in a set of three elements have the distributions $(\frac{1}{2}, \frac{1}{3}, \frac{1}{6})$ and $(\frac{1}{3}, \frac{1}{6}, \frac{1}{2})$, then they have the same entropy.

[5] Shannon (1948) showed that, under certain axioms for a measure of uncertainty, the entropy (up to the choice of the logarithmic base) is the only one that satisfies them.

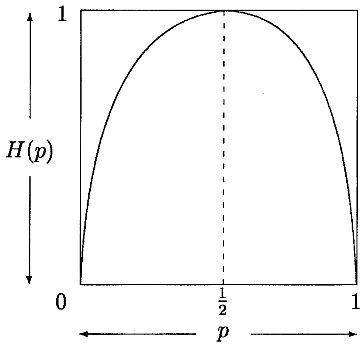

FIG. 1. The graph of entropy function.

(3) *H is continuous and strictly concave as a function on $\Delta(\Theta)$, where $\Delta(\Theta)$ is considered to be the unit simplex in the Euclidean space whose coordinates are indexed by the elements of Θ.*

The notion of entropy can be extended to an arbitrary finite dimensional vector of random variables or probability distributions. In Definition 3.1 simply replace a random variable X by a random vector $X = (X_1, \ldots, X_n)$ and the range Θ by $\times_{k=1}^{n} \Theta_k$, where each Θ_k is the range of X_k and is finite. Thus

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

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

Let (X_1, X_2) be a random vector taking values in $\Theta_1 \times \Theta_2$ with distribution $p(\theta_1, \theta_2)$. Suppose that we observe the realization of X_1. For each $X_1 = \theta_1$, we calculate the entropy of X_2 conditional on $X_1 = \theta_1$ as the entropy of the conditional distribution of X_2 given $X_1 = \theta_1$ denoted by $p(X_2 | \theta_1)$. The expected value of such entropy with respect to the marginal distribution of X_1, $p(X_1)$, would then measure the ex ante uncertainty of X_2 when we can observe X_1. Formally, let us define a function $h(X_2 | X_1)$: $\Theta_1 \to \mathbb{R}$ by

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

where $h(X_2 | \theta_1)$ is the value of $h(X_2 | X_1)$ at $\theta_1 \in \Theta_1$. We consider $h(X_2 | X_1)$ as a random variable on Θ_1 equipped with distribution $p(X_1)$.

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

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

Note that the distribution of X_2, $p(X_2)$, is equal to $E_{X_1}[p(X_2|X_1)]$ and $H(X_2) = H(p(X_2))$. Also, $h(X_2|X_1) = H(p(X_2|X_1))$. Therefore the concavity of the entropy and Jensen's inequality imply that

$$H(X_2|X_1) = E_{X_1}[H(p(X_2|X_1))] \leq H\big(E_{X_1}[p(X_2|X_1)]\big) = H(X_2).$$

That is, the conditioning reduces the entropy. The strict concavity of H implies, moreover, that the equality holds if, and only if, X_1 and X_2 are independent.

The next proposition states that the uncertainty of a pair of random variables is the uncertainty of one plus the remaining uncertainty of the other given the first. The proof is a simple marshaling of the definitions. See Cover and Thomas (1991, Chap. 2).

PROPOSITION 3.2. $H(X_1, X_2) = H(X_1) + H(X_2|X_1)$ $(= H(X_2) + H(X_1|X_2))$.

One can easily verify this in a simple example:

EXAMPLE 3.2. Consider a joint distribution of (X_1, X_2) on $\{0, 1\} \times \{0, 1\}$ given below:

$$\begin{array}{ccc} & & X_1 \\ & & \begin{array}{cc} 0 & 1 \end{array} \\ X_2 & \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & 0 \end{pmatrix} \end{array}$$

Then $p(X_1) = (\frac{1}{2}, \frac{1}{2})$, $p(X_2) = (\frac{3}{4}, \frac{1}{4})$,

$$p(X_2|X_1 = \theta_1) = \begin{cases} (\frac{1}{2}, \frac{1}{2}) & \text{if } \theta_1 = 0 \\ (1, 0) & \text{if } \theta_1 = 1 \end{cases}$$

and

$$p(X_1|X_2 = \theta_2) = \begin{cases} (\frac{1}{3}, \frac{2}{3}) & \text{if } \theta_2 = 0 \\ (1, 0) & \text{if } \theta_2 = 1. \end{cases}$$

Therefore we have

$$H(X_1, X_2) = \tfrac{3}{2}$$
$$H(X_1) = 1$$
$$H(X_2) = 2 - \tfrac{3}{4}\log 3$$
$$H(X_2 | X_1) = \tfrac{1}{2}$$
$$H(X_1 | X_2) = -\tfrac{1}{2} + \tfrac{3}{4}\log 3.$$

Note that in general $H(X_2|X_1) \neq H(X_1|X_2)$.

Proposition 3.2 applied 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 the "chain rule" for entropy is proved:

PROPOSITION 3.3. $H(X_1, \ldots, X_n) = H(X_1) + \sum_{k=2}^{n} H(X_k | X_1, \ldots, X_{k-1})$.

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 3.3 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 $\log|\Theta|$.

If we were to define entropy rate as the limit rather than the upper limit, it would not necessarily exist. Consider, for example, two Bernoulli random variables X and Y whose distributions are $(1, 0)$ and $(\frac{1}{2}, \frac{1}{2})$, respectively. Construct an independent sequence consisting of a string of Xs followed by an exponentially longer string of Ys followed further by a still exponentially longer string of Xs and so on, so that $(1/k)H(X_1, \ldots, X_k)$ oscillates between 0 and 1. However, it can be shown that for any stationary process the entropy rate, defined as the limit, does exist and coincide with the limit of conditional entropy, $\lim_{k\to\infty} H(X_k | X_1, \ldots, X_{k-1})$. See Cover and Thomas (1991, Chap. 4).

4. STRATEGIC ENTROPY

We start with an example which illuminates the heuristics behind the concepts we define in this section and motivates the use of entropy concepts discussed in the previous section.

EXAMPLE 4.1. Consider the twice repetition of a stage game in which player 1 has two actions, T and B, and player 2 also has two actions, L and R. The game is depicted in Figure 2.

We represent a player's pure strategy by an ordered five-tuple. For example, (T, B, T, T, B) represents player 1's strategy which assigns the action T in the first stage, i.e., at the information set I_1, B in the second stage after the history (T, L), i.e., at the information set I_2, T at I_3, and so on.

Now consider a mixed strategy σ of player 1 whose support consists of four pure strategies $s_1 = (T, B, T, T, B)$, $s_2 = (T, B, B, B, T)$, $s_3 = (T, B, T, T, T)$, and $s_4 = (B, B, B, B, B)$. Suppose that player 2 is informed of σ—its support and the probabilities it assigns—but not of its realization.

Note that s_1 and s_3 are equivalent. Hence if either one of these is selected by σ, no matter what strategy player 2 uses, he will never be able to tell which is actually selected. Now suppose that σ has selected s_2. Let player 2 play (L, L, L, L, L). (Player 2's actions in the second stage will be irrelevant in the subsequent discussion.) After the first stage player 2 observes that player 1 has taken the action T, and thus concludes that the realized strategy must be either s_1, s_2, or s_3. Will player 2 ever be able to narrow down the realization of σ further? The answer is no. All that player 2 observes at the end of the game is that player 1 has chosen T in

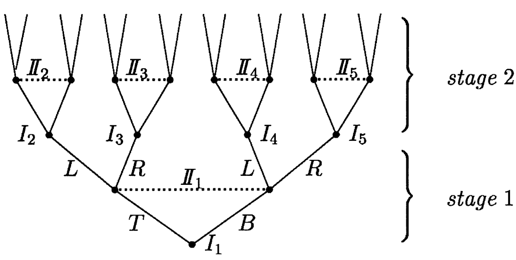

FIG. 2. The game tree of Example 4.1.

the first stage and B in the second stage after the history (T, L). Any one of s_1, s_2, and s_3 could lead to the same play of the game. Therefore uncertainty among the three strategies persists even after the game is ended provided that player 2 uses a strategy that assigns L at the first stage. Notice that s_2 is not equivalent to s_1 (and hence not to s_3, either); nevertheless, it is not distinguished if player 2 chooses L at the first stage. Had he chosen R at the first stage, however, he would have discovered that σ had indeed selected s_2 because this is the only strategy that plays B after the history (T, R).

From this example it is clear that, by using a pure strategy t, player 2 can only distinguish those pure strategies of player 1 in the support of his mixed strategy which give rise to different play of the game against t and nothing more. In other words, the amount of information (the reduction of uncertainty) about player 1's strategy σ obtained by using a pure strategy t is precisely the amount of information offered by the play induced by (σ, t). This motivates us to look at the information on the induced play measured by entropy. The following discussion focuses on player 1's strategy. It will be an easy task to extend the concepts defined below to games with an arbitrary number of players. See Section 8.

For each $n \in \mathbb{N}$, 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. The dual interpretation is that it is the amount of information on the play of the game that player 2 can obtain using t when player 1 uses σ.

Let us discuss some properties of the function $H^n(\cdot : \cdot)$. Fix $t \in T$. Each $\sigma \in \Delta(S)$ induces a probability distribution $P_{\sigma,t}$ on the finite set of n-histories. Its support is the set of n-histories induced by (s, t) for some s in the support of σ. Note that, with t fixed, the number of n-histories in the support of $P_{\sigma,t}$ is at most $|A|^n$. Therefore by Proposition 3.1(1) we see that $H^n(\cdot : \cdot)$ is bounded by $n \log|A|$. Since the probability $P_{\sigma,t}$ is continuous in (σ, t), $H^n(\cdot : \cdot)$ is continuous on $\Delta(S) \times T$.

From the above definition it is also easy to see that $H^n(\cdot : t)$ is a concave function on $\Delta(S)$ for each fixed $t \in T$. The strict concavity, however, does not hold. For example, suppose that two strategies σ and σ' differ only on those histories that would not be realized against a particular t. Let σ''

be any mixture of σ and σ'. Then the random plays induced by (σ, t), (σ', t), and (σ'', t) (or $P_{\sigma,t}$, $P_{\sigma',t}$, and $P_{\sigma'',t}$) all coincide. Consequently, the corresponding convex combination of $H^n(\sigma : t)$ and $H^n(\sigma' : t)$ is equal to $H^n(\sigma'' : t)$.

Clearly, $H^n(\cdot : t)$ depends only on the n-equivalence classes of $\Delta(S)$ for each $t \in T$, and similarly, $H^n(\sigma : \cdot)$ depends only on the n-equivalence classes of T. That is, if σ and σ' in $\Delta(S)$ are n-equivalent, then $H^n(\sigma : t) = H^n(\sigma' : t)$ for every $t \in T$, and if t and t' in T are n-equivalent, then $H^n(\sigma : t) = H^n(\sigma : t')$ for every $\sigma \in \Delta(S)$. We summarize the properties of $H^n(\cdot : \cdot)$ discussed above as a proposition.

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 3.3.

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

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

The quantity $H(X_k | X_1, \dots, X_{k-1})$ may be interpreted as the average uncertainty of actions at stage k given the $(k-1)$-history induced by (σ, t). One can think of $H(X_k | X_1, \dots, X_{k-1})$ as per-stage average reduction of uncertainty about σ (or a pure strategy selected by σ if σ is a mixed strategy) when player 2 uses t. In this context Lemma 4.1 states that the average overall reduction of uncertainty, which is the gain of information, about σ up to stage n using t is precisely $H^n(\sigma : t)$.

This lemma provides an alternative proof of Proposition 4.1(5). Given the history up to stage $k-1$, the actions at stage k, $X_k = (a_k, b_k)$, has randomness only on the part of player 1's action a_k whose entropy is at most $\log|A|$. Thus, for every $(k-1)$-history, $\omega_1, \dots, \omega_{k-1}$, $h(X_k | \omega_1, \dots, \omega_{k-1}) \le \log|A|$, and hence $H(X_k | X_1, \dots, X_{k-1}) \le \log|A|$. Therefore, by Lemma 4.1, $H^n(\sigma : t) \le n \log|A|$. Note that, unlike $H(X_k | X_1, \dots, X_{k-1})$, the unconditional entropy $H(X_k)$ of the action at stage k can be larger than $\log|A|$.

EXAMPLE 4.2. Consider the game of Example 4.1. See Fig. 2. We denote by (p, q) a mixed action of player 1 (player 2) that takes T (L) with probability p and B (R) with probability $q = 1 - p$. Define $\sigma = (\sigma_n)_{n=1}^{\infty}$

$\in \Delta(S)$ and $t \in T$ as follows. For each $\omega = (\omega_n)_{n=1}^{\infty}$, $\sigma_n(\omega) = (\frac{1}{2}, \frac{1}{2})$, and

$$t_1(\omega) = (1, 0)$$

$$t_2(\omega) = \begin{cases} (1,0) & \text{if } \omega_1 = (T, L) \\ (0,1) & \text{otherwise.} \end{cases}$$

For $n > 2$, $t_n(\omega)$ is an arbitrary pure action. It is easily verified that X_2 is uniformly distributed on $\{T, B\} \times \{R, L\}$ and hence $H(X_2) = \log 4 > \log 2 = \log|A|$.

This example also shows that even if player 1's actions are independent from stage to stage, the pair of random actions $X_n = (a_n, b_n)$ induced by (σ, t) are not necessarily so. This is because player 2's stage actions do depend on the past histories.

One may consider two alternatives for the extension of $H^n(\cdot : \cdot)$ to $\Delta(S) \times \Delta(T)$. One is the linear extension: for $\tau \in \Delta(T)$, let us define $H^n(\sigma : \tau) = E_\tau[H^n(\sigma : t)]$. This quantity may be interpreted as the average uncertainty of the play that player 2 faces when he uses τ against σ, or the average amount of information on the play that player 2 can obtain using τ against σ. The other alternative is to look directly at the random play up to stage n induced by (σ, τ) and take its entropy. We denote this by $H^n(\sigma, \tau)$. It is the uncertainty of the play up to stage n when one is informed only of the distributions σ and τ but not their realizations. Since $H^n(\sigma, t) = H^n(\sigma : t)$ for each $t \in T$, it is indeed an extension of the original function on $\Delta(S) \times T$. In general, $H^n(\sigma, \tau)$ is greater than the linear extension $H^n(\sigma : \tau)$ because of the concavity of the entropy.

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. Proposition 4.1(4) offers a more intuitive argument for the existence of the maximum: $H^n(\sigma : \cdot)$ 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. With the above discussion of the extension of $H^n(\cdot : \cdot)$ in mind, we can also write $H^n(\sigma) = \max_{\tau \in \Delta(T)} H^n(\sigma : \tau)$. Proposition 4.1(1) and (4) imply, moreover, that the n-strategic entropy $H^n(\sigma)$, considered as a function of σ, is the maximum of essentially a finite number of continuous functions of σ. Therefore, $H^n(\sigma)$ is continuous in σ. The n-strategic entropy $H^n(\sigma)$, unlike the entropy, is not a concave function in σ. This is well expected in light of Proposition 4.1(2) and (4) since they imply that $H^n(\sigma)$ is the maximum of a finite number of concave functions. See also Example 4.3(d) below. Although we will not utilize it in this paper, it is also of interest to consider the *lower n-strategic entropy* of $\sigma \in \Delta(S)$ defined by

$$\underline{H}^n(\sigma) = \min_{t \in T} H^n(\sigma : t),$$

which preserves the concavity. This is the lower bound of the uncertainty of the play up to stage n that player 2 faces no matter what pure strategy she uses against σ. We provide some examples to facilitate the understanding of the n-strategic entropy.

EXAMPLE 4.3. (a) Any pure strategy has zero n-strategic entropy for every n. Any strategy that is n-equivalent to a pure strategy has zero n-strategic entropy. This is trivial because such a strategy, by definition, induces pure actions against any pure strategy of player 2 up to stage n. Conversely, if a strategy has zero n-strategic entropy, then it is necessarily n-equivalent to a pure strategy.

(b) Suppose that $\sigma = (\sigma_k)_{k=1}^{\infty}$ is an independent sequence of mixed actions, $\sigma_k = \alpha_k \in \Delta(A)$. Take $t \in T$ and let $(X_k)_{k=1}^{\infty}$ be the random play induced by (σ, t). Then $H(X_k | X_1, \dots, X_{k-1}) = H(\alpha_k)$ for every k. Hence by Lemma 4.1, $H^n(\sigma : t) = \sum_{k=1}^{n} H(\alpha_k)$ for every $t \in T$ and thus $H^n(\sigma) = \sum_{k=1}^{n} H(\alpha_k)$. A special case of this is when $\alpha_k = \alpha$ for all k. In this case, $H^n(\sigma) = nH(\alpha)$.

(c) Consider the game of Example 4.1. Define two behavioral strategies of player 1, $\sigma = (\sigma_n)_{n=1}^{\infty}$ and $\sigma' = (\sigma_n')_{n=1}^{\infty}$, as follows. For each $\omega = (\omega_n)_{n=1}^{\infty} \in \Omega_\infty$,

$$\sigma_1(\omega) = (1, 0)$$

$$\sigma_2(\omega) = \begin{cases} \left(\frac{1}{2}, \frac{1}{2}\right) & \text{if } \omega_1 = (\cdot, L) \\ (0, 1) & \text{if } \omega_1 = (\cdot, R) \end{cases}$$

and

$$\sigma_1'(\omega) = (1, 0)$$

$$\sigma_2'(\omega) = \begin{cases} (1, 0) & \text{if } \omega_1 = (\cdot, L) \\ \left(\frac{1}{3}, \frac{2}{3}\right) & \text{If } \omega_1 = (\cdot, R). \end{cases}$$

Specification of σ_n and σ_n' for $n > 2$ is irrelevant to the subsequent discussion. The maximum of $H^2(\sigma : t)$ over t is attained by any $t = (t_n)_{n=1}^{\infty}$ with $t_1 = L$. Thus $H^2(\sigma) = H(\frac{1}{2})$, where $H(p)$ is the entropy function defined in Example 3.1. On the other hand the maximum of $H^n(\sigma' : t)$ over t is attained by any $t = (t_n)_{n=1}^{\infty}$ with $t_1 = R$ and $H^2(\sigma') = H(\frac{1}{3})$.

(d) The two strategies defined in (c) can be used to show that the n-strategic entropy $H^n(\sigma)$ is not concave in σ. Let σ'' be the mixture of σ and σ', each with equal probabilities. Then, for each $\omega = (\omega_n)_{n=1}^{\infty}$,

$$\sigma_1''(\omega) = (1, 0)$$

$$\sigma_2''(\omega) = \begin{cases} \left(\frac{3}{4}, \frac{1}{4}\right) & \text{if } \omega_1 = (\cdot, L) \\ \left(\frac{1}{6}, \frac{5}{6}\right) & \text{if } \omega_1 = (\cdot, R) \end{cases}$$

and the maximum of $H^n(\sigma'' : t)$ over t is attained by any $t = (t_n)_{n=1}^{\infty}$ with $t_1 = L$. Hence $H^2(\sigma'') = H(\frac{3}{4})$. Since $H(\frac{3}{4}) < H(\frac{2}{3}) < H(\frac{1}{2})$ (see also Fig. 1), we have

$$H^2(\sigma'') < \tfrac{1}{2}H^2(\sigma) + \tfrac{1}{2}H^2(\sigma').$$

Therefore, $H^2(\cdot)$ is not concave on $\Delta(S)$.

(e) Proposition 4.1(5) implies that $H^n(\sigma)$ is bounded by $n \log|A|$. Let $\tilde{\alpha}$ be the mixed action of player 1 which is uniformly distributed on A. From the argument of (b) above it follows that if σ plays the i.i.d. sequence of $\tilde{\alpha}$, then $H^n(\sigma) = n \log|A|$. There are other strategies that attain this bound. Again, in the game of Example 4.1, define $\sigma = (\sigma_n)_{n=1}^{\infty}$ as follows:

$$\sigma_1(\omega) = \left(\frac{1}{2}, \frac{1}{2}\right)$$

$$\sigma_2(\omega) = \begin{cases} \left(\frac{1}{2}, \frac{1}{2}\right) & \text{if } \omega_1 = (\cdot, L) \\ (0, 1) & \text{if } \omega_1 = (\cdot, R). \end{cases}$$

Then any pure strategy $t = (t_n)_{n=1}^{\infty}$ of player 2 with $t_1 = L$ would maximize $H^2(\sigma : t)$ and $H^2(\sigma) = 2 \log 2 = 2$.

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.*

It is clear that the n-strategic entropy $H^n(\sigma)$ is nondecreasing in n. We define the *total strategic entropy* of σ by the supremum of the n-strategic entropy.

DEFINITION 4.2. The total strategic entropy $H^*(\sigma)$ of $\sigma \in \Delta(S)$ is defined by

$$H^*(\sigma) = \sup_n H^n(\sigma).$$

Note that $H^*(\sigma)$ may be infinite. Also note that $H^n(\sigma)$ cannot grow faster than the linear function $f(n) = n \log|A|$.

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 the 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 4.3. 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. It is easy to construct a behavioral strategy that plays a mixed action everywhere and has zero strategic entropy rate. Since $H^n(\sigma : t) \leq H^n(\sigma)$ for every $n \in \mathbb{N}$ and $t \in T$, the strategic entropy rate of σ is at most $\tilde{H}(\sigma) = \limsup_{n \to \infty}(1/n)H^n(\sigma)$. This inequality can be strict.

EXAMPLE 4.4. Consider the following behavioral strategy σ of player 1 where the stage game is a 2×2 matrix game with $A = \{\text{Top}, \text{Bottom}\}$ and $B = \{\text{Left}, \text{Right}\}$. Play Top in the first stage and continue with Top as long as player 2 plays Left. If player 2 plays Right at stage m for the first time, play the i.i.d. mixed action $(\frac{1}{2}, \frac{1}{2})$ (Top and Bottom with equal probabilities) for the next m^2 stages and then play Top forever after that. Let us show that $H^\infty(\sigma) = 0$ while $\tilde{H}(\sigma) = 1$.

For every $t = (t_n)_{n=1}^{\infty} \in T$, the random play $(X_n)_{n=1}^{\infty}$ induced by (σ, t) is such that either $X_n = (\text{Top}, \text{Left})$ for every n or there is a unique $m < \infty$

for which

$$X_n = \begin{cases} (\text{Top}, \text{Left}) & \text{for } n = 1, \dots, m-1 \\ (\text{Top}, \text{Right}) & \text{for } n = m \\ \left(\left(\frac{1}{2}, \frac{1}{2}\right), t_n\right) & \text{for } n = m+1, \dots, m+m^2 \\ (\text{Top}, t_n) & \text{for } n \geq m+m^2+1. \end{cases} \tag{4.1}$$

In the former case $H^n(\sigma : t) = 0$ for every n. In the latter case, $H^n(\sigma : t) = m^2$ for every $n \geq m + m^2$ and hence $(1/n)H^n(\sigma : t) = m^2/n \to 0$. This shows that $H^\infty(\sigma) = 0$.

Next, it is clear that $\tilde{H}(\sigma) \leq \log|A| = \log 2 = 1$. For each $m \in \mathbb{N}$, define $t^m = (t_n^m)_{n=1}^\infty \in T$ as follows: $t_n^m =$ Left if $n \neq m$, and $t_n^m =$ Right if $n = m$. Then the random play induced by (σ, t^m) is exactly as (4.1) above with t_n replaced by Left, and $H^{m+m^2}(\sigma) \geq H^{m+m^2}(\sigma : t^m) = m^2$. Hence

$$\tilde{H}(\sigma) \geq \limsup_{m\to\infty} \frac{H^{m+m^2}(\sigma)}{m+m^2} \geq \lim_{m\to\infty} \frac{m^2}{m+m^2} = 1.$$

Therefore $\tilde{H}(\sigma) = 1$.

5. ENTROPY AND STRATEGIC ENTROPY

In the first example in the previous section we saw that if σ has two different pure strategies $s = (s_k)_{k=1}^\infty$ and $s' = (s'_k)_{k=1}^\infty$ in its support such that (s, t) and (s', t) induce n-equivalent plays, then player 2, using this t, can never tell which one was actually selected within the first n stages. In other words, the uncertainty about $(s_k)_{k=1}^n$ and $(s'_k)_{k=1}^n$ persists even at the end of stage n. This suggests that the reduction of uncertainty about (a pure strategy selected by) σ in the first n stages using t, which is $H^n(\sigma : t)$, is smaller than the uncertainty of σ as a probability on the first n stage strategies. Denote by $\hat{H}^n(\sigma)$ the entropy of σ viewed as a probability on $\times_{k=1}^n S_k$. We call $\hat{H}^n(\sigma)$ the *n-entropy* of σ. In this section we will show that the n-strategic entropy of σ is at most its n-entropy.

For this purpose it is convenient to introduce the concept of the entropy of a partition. Let $(\Phi, \mathscr{F}, \mu)$ be a probability space and let $\mathscr{P}$ be a finite partition of Φ into $\mathscr{F}$-sets. The entropy of the partition $\mathscr{P}$ with respect to μ is defined by

$$H_\mu(\mathscr{P}) = -\sum_{F\in\mathscr{P}} \mu(F)\log \mu(F).$$

The next proposition states that the coarsening of a partition cannot increase entropy. See Smorodinsky (1971).

PROPOSITION 5.1. *If a partition $\mathscr{P}$ is a coarsening of a partition $\mathscr{Q}$, i.e., every element of $\mathscr{P}$ is a union of elements of $\mathscr{Q}$, then*

$$H_{\mu}(\mathscr{P}) \leq H_{\mu}(\mathscr{Q}).$$

Recall that $S = \times_{k \geq 1} S_k$, where S_k is a finite set of stage k strategies. Let $\mathscr{P}_n$ be the finite partition of S induced[6] by $\times_{k=1}^{n} S_n$. The algebra on S generated by $\mathscr{P}_n$ is denoted by $\mathscr{S}_n$. We denote by $\mathscr{S}$ the σ-algebra on S generated by $\bigcup_{n \geq 1} \mathscr{S}_n$. A mixed strategy is thus a probability on $(S, \mathscr{S})$. Since $\mathscr{P}_n$ is a finite partition of S into $\mathscr{S}$-sets, its entropy $H_{\sigma}(\mathscr{P}_n)$ is defined as above. It is clear from the definition of $\mathscr{P}_n$ that $\hat{H}^n(\sigma) = H_{\sigma}(\mathscr{P}_n)$.

Given player 2's pure strategy t, we say that s and s' in S are n-equivalent *with respect to* t, or (n,t)-equivalent for short, if (s,t) and (s',t) induce n-equivalent plays. Since (n,t)-equivalence is an equivalence relation on S, it induces a partition of S denoted by $\mathscr{P}_n(t)$. Thus with each $\mathscr{D} \in \mathscr{P}_n(t)$ is associated an n-history $C = \omega(D) \in \mathscr{H}_n$ in one-to-one and onto manner. Therefore,

$$\begin{aligned} H^n(\sigma : t) &= -\sum_{C \in \mathscr{H}_n} P_{\sigma,t}(C) \log P_{\sigma,t}(C) \\ &= -\sum_{D \in \mathscr{P}_n(t)} \sigma(D) \log \sigma(D) \\ &= H_{\sigma}\big(\mathscr{P}_n(t)\big). \end{aligned}$$

Clearly, $\mathscr{P}_n(t)$ is a coarsening of $\mathscr{P}_n$. Indeed, $\mathscr{P}_n(t)$ is a coarsening of the partition of S into the n-equivalence classes which is in turn a coarsening of $\mathscr{P}_n$. Therefore it follows from Proposition 5.1 that $H^n(\sigma : t) \leq \hat{H}^n(\sigma)$ for every $t \in T$. By taking the maximum over t we see that the n-strategic entropy of σ does not exceed its n-entropy: $H^n(\sigma) \leq \hat{H}^n(\sigma)$. Consequently, $\tilde{H}(\sigma)$ is at most $\limsup_{n \to \infty} (1/n) \hat{H}^n(\sigma)$, and hence so is the strategic entropy rate $H^{\infty}(\sigma)$.

6. THE REPEATED GAMES WITH ENTROPY BOUND

In this section we study the three classes of repeated games each with a restriction on player 1's strategies in terms of one of the three strategic entropy concepts formulated in Section 4. All the results in this section are based on the main lemma which we present first.

[6] That is, the partition induced by the natural projection of S onto the first n factors.

Observe that if $\sigma \in \Delta(S)$ has a very small n-strategic entropy, then Lemma 4.1 implies that the entropy of player 1's stage action must necessarily be small at most of the first n stages, i.e., it is "close to" a pure action. So if player 2 plays a one-shot best response at every stage, the n-average payoff for player 1 will be close to his maxmin value in pure actions of the stage game. This is the heuristics for the lemma below. Note that if $(X_k)_{k=1}^{\infty}$ is the random play induced by a pair of strategies, then $h(X_k | X_1, \ldots, X_{k-1})$ is $\mathscr{A}_{k-1}$-measurable. By a slight abuse of notation we denote $h(X_k | X_1, \ldots, X_{k-1})$ by $h(X_k | \mathscr{A}_{k-1})$ and $H(X_k | X_1, \ldots, X_{k-1})$ by $H(X_k | \mathscr{A}_{k-1})$.

LEMMA 6.1. $\forall \varepsilon > 0$, $\exists \gamma(\varepsilon) > 0$ *such that* $\forall \sigma \in \Delta(S)$, $\exists t \in T$ *such that*

$$r_n(\sigma, t) \leq \mathrm{U}_*(G) + \varepsilon$$

whenever $n \in \mathbb{N}$ *and* $H^n(\sigma) \leq \gamma(\varepsilon) n$.

Proof. Without loss of generality we assume that $0 \leq r \leq 1$. We shall identify $\Delta(A)$ with the unit simplex in $\mathbb{R}^{|A|}$ while A itself is identified with the extreme points of this simplex. For each $\alpha \in \Delta(A)$ let $d(\alpha, A) = \min_{a \in A} \|\alpha - a\|$, where $\|\cdot\|$ is the L_1-norm.

Let $\varepsilon > 0$ be given. For every pure action $b \in B$, the stage game payoff $E_\alpha[r(a,b)]$ is continuous in $\alpha \in \Delta(A)$. The minimum of finitely many continuous functions is continuous, and thus $\min_{b \in B} E_\alpha[r(a,b)]$ is continuous in $\alpha \in \Delta(A)$. For every $a \in A$, $\min_{b \in B} r(a,b) \leq \mathrm{U}_*(G)$. Therefore there is $\delta(\varepsilon) > 0$ so that, for every $\alpha \in \Delta(A)$ with $d(\alpha, A) < \delta(\varepsilon)$, we have $\min_{b \in B} E_\alpha[r(a,b)] < \mathrm{U}_*(G) + \varepsilon/2$.

Take $\sigma \in \Delta(S)$. Without loss of generality suppose that σ is a behavioral strategy, $\sigma = (\sigma_k)_{k=1}^{\infty}$. Define player 2's strategy $t_\sigma = (t_{\sigma,k})_{k=1}^{\infty}$ in such a way that, for every k, $t_{\sigma,k}(\omega)$ minimizes $E_{\sigma_k(\omega)}[r(a,b)]$ over $b \in B$. That is, $t_{\sigma,k}$ takes a best response action against σ_k at every k-history. Let $(X_k)_{k=1}^{\infty}$ be the random play induced by (σ, t_σ).

The second and third properties of entropy in Proposition 3.1 imply that there is $\kappa(\varepsilon) > 0$ such that $d(\alpha, A) \geq \delta(\varepsilon)$ implies $H(\alpha) \geq \kappa(\varepsilon)$. Thus $I(d(\sigma_k, A) \geq \delta(\varepsilon)) \leq h(X_k | \mathscr{A}_{k-1}) / \kappa(\varepsilon)$, where $I(\cdot)$ is the indicator function. Notice that both sides of this inequality are measurable with respect to $\mathscr{A}_{k-1}$ and recall our assumption $0 \leq r \leq 1$. Then at every stage k we have

$$\begin{aligned} &E_{\sigma, t_\sigma}\left[r(X_k) | \mathscr{A}_{k-1}\right] \\ &\quad \leq \left(\mathrm{U}_*(G) + \frac{\varepsilon}{2}\right) I\big(d(\sigma_k, A) < \delta(\varepsilon)\big) + I\big(d(\sigma_k, A) \geq \delta(\varepsilon)\big) \\ &\quad \leq \mathrm{U}_*(G) + \frac{\varepsilon}{2} + \frac{h\big(X_k | \mathscr{A}_{k-1}\big)}{\kappa(\varepsilon)}. \end{aligned}$$

Set $\gamma(\varepsilon) = \varepsilon\kappa(\varepsilon)/2$. Assume that n is such that $H^n(\sigma) \le \gamma(\varepsilon)n$. Using the above inequality together with Lemma 4.1 and the definition of $H^n(\sigma)$, we obtain the following:

$$\begin{aligned}
E_{\sigma, t_\sigma}\left[\frac{1}{n}\sum_{k=1}^{n} r(X_k)\right] &= \frac{1}{n}\sum_{k=1}^{n} E_{\sigma, t_\sigma}\left[E_{\sigma, t_\sigma}\left[r(X_k)|\mathscr{A}_{k-1}\right]\right] \\
&\le \mathrm{U}_*(G) + \frac{\varepsilon}{2} + \frac{1}{\kappa(\varepsilon)n}\sum_{k=1}^{n} E_{\sigma, t_\sigma}\left[h(X_k|\mathscr{A}_{k-1})\right] \\
&= \mathrm{U}_*(G) + \frac{\varepsilon}{2} + \frac{1}{\kappa(\varepsilon)n}\sum_{k=1}^{n} H(X_k|\mathscr{A}_{k-1}) \\
&= \mathrm{U}_*(G) + \frac{\varepsilon}{2} + \frac{H^n(\sigma : t_\sigma)}{\kappa(\varepsilon)n} \\
&\le \mathrm{U}_*(G) + \frac{\varepsilon}{2} + \frac{H^n(\sigma)}{\kappa(\varepsilon)n} \\
&\le \mathrm{U}_*(G) + \frac{\varepsilon}{2} + \frac{\gamma(\varepsilon)}{\kappa(\varepsilon)} \\
&= \mathrm{U}_*(G) + \varepsilon.
\end{aligned}$$

This completes the proof. Q.E.D.

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

We will modify the finitely repeated game G^n by restricting the set of player 1's strategies. This is done by imposing an exogenous bound on the n-strategic entropy of his strategies. Player 2's strategy set remains intact.

For $\eta > 0$, let $\Sigma^n(\eta)$ be the set of player 1's strategies whose n-strategic entropy is at most η. That is, $\Sigma^n(\eta) = \{\sigma \in \Delta(S) | H^n(\sigma) \le \eta\}$. Let $G^n(\eta)$ be the game $(\Sigma^n(\eta), \Delta(T), r_n)$. Since $H^n(\sigma)$ is continuous in σ (Proposition 4.2(1)), $\Sigma^n(\eta)$ is a compact subset of $\Delta(S)$. 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. As $\Sigma^n(\eta)$ is compact, the maxmin value of $G^n(\eta)$,

$$\mathrm{W}^n(\eta) = \max_{\sigma \in \Sigma^n(\eta)} \min_{t \in T} r_n(\sigma, t),$$

is well-defined.

As illustrated in the next example, the set $\Sigma^n(\eta)$ is in general *not* convex. Therefore, the minimax theorem does not apply to the game $G^n(\eta)$. Consequently, the value of $G^n(\eta)$ need not exist. Since pure strategies have zero n-strategic entropy, $S \subset \Sigma^n(\eta)$ for every $\eta \geq 0$. So the minimax value of G^n, $\overline{\mathrm{V}}^n = \min_{\tau \in \Delta(T)} \max_{s \in S} r_n(s, \tau)$, also well-defined by an argument similar to the above, remains the minimax value of $G^n(\eta)$. But the maxmin value $\mathrm{W}^n(\eta)$ is in general smaller than $\overline{\mathrm{V}}^n$.

EXAMPLE 6.1. Consider the matching pennies:

1	-1
-1	1

In the one-shot game, $H^1(\sigma)$ is simply the entropy of the first stage action. So for all $\eta \geq 1$ $(= \log 2)$, $\Sigma^1(\eta)$ coincides with the whole set of mixed strategies. Fix a positive number $\eta < 1$. Let p be the probability of choosing the top row. For every $0 \leq \gamma < 1$ there is $\eta = \eta(\gamma) < \frac{1}{2}$ such that $\{\eta, 1 - \eta\} = H^{-1}(\gamma)$ (see also Fig. 1). Thus $\Sigma^1(\eta)$ is the union of two disjoint intervals: $\Sigma^1(\eta) = \{(p, 1-p) | 0 \leq p \leq \eta\} \cup \{(p, 1-p) | 1 - \eta \leq p \leq 1\}$. It is easy to see that $\mathrm{W}^1(\eta) = -1 + 2\eta < 0 = \overline{\mathrm{V}}^1$ $(= \mathrm{Val}(G))$.

Consider the bound on the n-strategic entropy to be a function of the number of repetitions, i.e., $\eta: \mathbb{N} \to \mathbb{R}$. We are interested in the asymptotics of $\mathrm{W}^n(\eta(n))$. The next theorem asserts that if $\eta(n)$ grows more slowly than any linear function of n, then the unrestricted player 2 can asymptotically hold player 1's payoff down to his maxmin value in pure actions of the stage game.

THEOREM 6.1. *If* $\lim_{n \to \infty} \eta(n)/n = 0$, *then* $\lim_{n \to \infty} \mathrm{W}^n(\eta(n)) = \mathrm{U}_*(G)$.

Proof. Player 1 can guarantee himself at least $\mathrm{U}_*(G)$ at every stage using a pure action, and hence $\mathrm{W}^n(\eta(n)) \geq \mathrm{U}_*(G)$ for every n. It follows that $\liminf_{n \to \infty} \mathrm{W}^n(\eta(n)) \geq \mathrm{U}_*(G)$.

Next, take $\varepsilon > 0$ arbitrarily. Let $\gamma(\varepsilon) > 0$ be as specified in Lemma 6.1. By the condition of the theorem, there is $n(\varepsilon)$ such that for, every $n \geq n(\varepsilon)$, we have $\eta(n) \leq \gamma(\varepsilon) n$. Then by Lemma 6.1, for every $n \geq n(\varepsilon)$ and every $\sigma \in \Sigma^n(\eta(n))$, $r_n(\sigma, t_\sigma) \leq \mathrm{U}_*(G) + \varepsilon$, where t_σ is as defined in the proof of the lemma. This shows that $\limsup_{n \to \infty} \mathrm{W}^n(\eta(n)) \leq \mathrm{U}_*(G) + \varepsilon$. As this inequality holds for every $\varepsilon > 0$, we conclude that $\limsup_{n \to \infty} \mathrm{W}^n(\eta(n)) \leq \mathrm{U}_*(G)$, which completes the proof. Q.E.D.

It is easy to see that if $\eta(n)/n$ is bounded away from zero, then $\mathrm{W}^n(\eta(n))$ is bounded away from $\mathrm{U}_*(G)$ provided, of course, that there is a feasible payoff in G for player 1 which is strictly greater than $\mathrm{U}_*(G)$. Consider, for example, the matching pennies of Example 6.1. Suppose that $\eta(n) = \kappa n$ for every n for some $0 < \kappa \leq 1$. Let $p \in (0, \frac{1}{2}]$ be such that $H(p) = \kappa$ and define $\sigma = (\sigma_k)_{k=1}^\infty$ by $\sigma_k(\cdot) = (p, 1-p)$. Then $H^n(\sigma) = \kappa n$ and $\min_{b \in B} E_\sigma[r(a_k, b_k)] = -1 + 2p$ for every k. Therefore $\min_{t \in T} r_n(\sigma, t) = -1 + 2p > -1$ and thus $\mathrm{W}^n(\eta(n)) > -1$.

6.2. *The Discounted Game* $G_\lambda^*(\eta)$

In this section we consider the λ-discounted game. In contrast to the previous section, we will impose a bound on the total strategic entropy of player 1's strategies. Again, player 2's strategy set remains intact. We will also show that bounding the strategic entropy rate is not an essential restriction in the discounted game. Let us start with a lemma which is an analogue of Lemma 6.1.

LEMMA 6.2. $\forall \varepsilon > 0$, $\exists \theta(\varepsilon) > 0$ *such that* $\forall \sigma \in \Delta(S)$, $\exists t \in T$ *such that*

$$r_\lambda(\sigma, t) \leq \mathrm{U}_*(G) + \varepsilon$$

whenever $\lambda \in [0,1)$ *and* $H^*(\sigma) \leq \theta(\varepsilon)/(1-\lambda)$.

Proof. As a preparatory step, we express the λ-discounted payoff as an average of n-average payoffs.[7] Fix $(\sigma, \tau) \in \Delta(S) \times \Delta(T)$. First,

$$\begin{aligned} E_{\sigma,\tau}[r(a_n, b_n)] &= \sum_{k=1}^{n} E_{\sigma,\tau}[r(a_k, b_k)] - \sum_{k=1}^{n-1} E_{\sigma,\tau}[r(a_k, b_k)] \\ &= nr_n(\sigma,\tau) - (n-1)r_{n-1}(\sigma,\tau), \end{aligned}$$

where we set $r_0(\sigma, \tau) = 0$. From this it follows that, for each $m \in \mathbb{N}$,

$$\begin{aligned} &E_{\sigma,\tau}\left[\sum_{n=1}^{m}(1-\lambda)\lambda^{n-1}r(a_n,b_n)\right] \\ &\quad= \sum_{n=1}^{m}(1-\lambda)\lambda^{n-1}E_{\sigma,\tau}[r(a_n,b_n)] \\ &\quad= \sum_{n=1}^{m}(1-\lambda)\lambda^{n-1}\{nr_n(\sigma,\tau) - (n-1)r_{n-1}(\sigma,\tau)\} \\ &\quad= \sum_{n=1}^{m}(1-\lambda)\lambda^{n-1}nr_n(\sigma,\tau) - \sum_{n=1}^{m}(1-\lambda)\lambda^{n}nr_n(\sigma,\tau) \\ &\qquad+(1-\lambda)\lambda^m mr_m(\sigma,\tau) \\ &\quad= \sum_{n=1}^{m}(1-\lambda)(\lambda^{n-1}-\lambda^n)nr_n(\sigma,\tau) + (1-\lambda)\lambda^m mr_m(\sigma,\tau). \end{aligned}$$

Since $0 \leq \lambda < 1$ and $r(a_n, b_n) \leq K = \max_{a,b}|r(a,b)|$ for every $n \in \mathbb{N}$, the left-hand side of the above equality converges to $r_\lambda(\sigma, \tau)$ as m tends to

[7] Although it is a well-known formula of Abel's partial summation, we present it here explicitly with our payoff function to facilitate the reading of the proof.

infinity. On the other hand, $r_n(\sigma, \tau) \leq K$ for every $n \in \mathbb{N}$ and $\lim_{m \to \infty} \lambda^m m = 0$. Therefore we have obtained the following equality:

$$r_\lambda(\sigma, \tau) = \sum_{n=1}^{\infty} (1 - \lambda)(\lambda^{n-1} - \lambda^n) n r_n(\sigma, \tau). \tag{6.1}$$

Assume, again, that $0 \leq r \leq 1$. Let $\varepsilon > 0$ be given. Lemma 6.1 implies that there is $\gamma(\varepsilon) > 0$ such that for every n and $\sigma \in \Delta(S)$

$$r_n(\sigma, t_\sigma) \leq \mathrm{U}_*(G) + \frac{\varepsilon}{2} + I(H^n(\sigma) \geq \gamma(\varepsilon)n). \tag{6.2}$$

Using (6.1) and (6.2), and as $\sum_{n=1}^{\infty}(1 - \lambda)(\lambda^{n-1} - \lambda^n)n = 1$, we have

$$\begin{aligned} r_\lambda(\sigma, t_\sigma) &\leq \sum_{n=1}^{\infty} (1 - \lambda)(\lambda^{n-1} - \lambda^n) \\ &\qquad \times n\left(\mathrm{U}_*(G) + \frac{\varepsilon}{2} + I(H^n(\sigma) \geq \gamma(\varepsilon)n)\right) \\ &= \mathrm{U}_*(G) + \frac{\varepsilon}{2} + \sum_{n=1}^{\infty} (1 - \lambda)(\lambda^{n-1} - \lambda^n) \\ &\qquad \times nI(H^n(\sigma) \geq \gamma(\varepsilon)n). \end{aligned} \tag{6.3}$$

If σ is such that $H^*(\sigma) \leq \gamma(\varepsilon)\varepsilon/2(1 - \lambda)$, then for every $n > \varepsilon/2(1 - \lambda)$ we have

$$H^n(\sigma) \leq H^*(\sigma) \leq \frac{\gamma(\varepsilon)\varepsilon}{2(1 - \lambda)} < \gamma(\varepsilon)n,$$

that is, $I(H^n(\sigma) \geq \gamma(\varepsilon)n) = 0$. Write $n(\varepsilon, \lambda)$ for $[\varepsilon/2(1 - \lambda)]$ and $\theta(\varepsilon)$ for $\gamma(\varepsilon)\varepsilon/2$. Then, for every σ with $H^*(\sigma) \leq \theta(\varepsilon)/(1 - \lambda)$, the last term in (6.3) is bounded by

$$\sum_{n=1}^{n(\varepsilon, \lambda)} (1 - \lambda)(\lambda^{n-1} - \lambda^n)n \leq (1 - \lambda)n(\varepsilon, \lambda) \leq \frac{\varepsilon}{2},$$

which, together with (6.3), completes the proof. Q.E.D.

For $\eta \geq 0$ let $G_\lambda^*(\eta)$ be the λ-discounted game in which player 1's strategies are restricted to those with total strategic entropy at most η. Set $\Sigma^*(\eta) = \{\sigma \in \Delta(S) | H^*(\sigma) \leq \eta\}$ and $G_\lambda^*(\eta) = (\Sigma^*(\eta), \Delta(T), r_\lambda)$. Notice that $\Sigma^*(\eta) = \bigcap_n \Sigma^n(\eta)$. As each $\Sigma^n(\eta)$ is compact, $\Sigma^*(\eta)$ is also compact. By an argument analogous to the one in the previous section, one can show that the maxmin value of $G_\lambda^*(\eta)$,

$$\mathrm{W}_\lambda^*(\eta) = \max_{\sigma \in \Sigma^*(\eta)} \min_{t \in T} r_\lambda(\sigma, t)$$

is well defined. We will consider the bound on the total strategic entropy η to be a function of the discount factor, η: $[0,1) \to \mathbb{R}_+$, and study the asymptotics of $W_\lambda^*(\eta(\lambda))$ as $\lambda \to 1$.

Note that a λ-discounted game can be viewed as a finitely repeated game whose number of repetitions is a geometrically distributed random variable with mean $1/(1-\lambda)$. With this view, the next theorem is comparable to Theorem 6.1. It asserts that if $\eta(\lambda)$ grows more slowly than any linear function of $1/(1-\lambda)$, then $W_\lambda^*(\eta(\lambda))$ converges to $U_*(G)$ as λ tends to 1.

THEOREM 6.2. *If* $\lim_{\lambda\to 1}(1-\lambda)\eta(\lambda) = 0$, *then* $\lim_{\lambda\to 1} W_\lambda^*(\eta(\lambda)) = U_*(G)$.

Proof. Since player 1 can guarantee $U_*(G)$ by a pure strategy and $S \subset \Sigma^*(\eta(\lambda))$ for every $\lambda \in [0,1)$, we have $\liminf_{\lambda\to 1} W_\lambda^*(\eta(\lambda)) \geq U_*(G)$. Next we show that $\limsup_{\lambda\to 1} W_\lambda^*(\eta(\lambda)) \leq U_*(G)$.

Let $\varepsilon > 0$ be given. By the condition of the theorem there is $\lambda(\varepsilon) \in [0,1)$ such that, for every $\lambda \in [\lambda(\varepsilon),1)$, we have $\eta(\lambda) \leq \theta(\varepsilon)/(1-\lambda)$, where $\theta(\varepsilon) > 0$ is as specified in Lemma 6.2. Then, for every $\lambda \in [\lambda(\varepsilon),1)$ and $\sigma \in \Sigma^*(\eta(\lambda))$, we have $r_\lambda(\sigma, t_\sigma) \leq U_*(G) + \varepsilon$. This shows that $\limsup_{\lambda\to 1} W_\lambda^*(\eta(\lambda)) \leq U_*(G) + \varepsilon$. As $\varepsilon > 0$ was taken arbitrarily, we have the desired result. Q.E.D.

To conclude this subsection we demonstrate that in the λ-discounted game there is an ε-optimal strategy with zero strategic entropy rate for every $\varepsilon > 0$. Therefore, putting a bound on the strategic entropy rate is not a restriction in the discounted games.

OBSERVATION. $\forall\lambda \in [0,1)$, $\forall\varepsilon > 0$, $\exists\sigma \in \Delta(S)$ *such that* (i) $H^\infty(\sigma) = 0$ *and* (ii) $\forall t \in T$, $r_\lambda(\sigma, t) \geq \mathrm{Val}(G) - \varepsilon$.

Proof. Without loss of generality, assume that $0 \leq r \leq 1$. Given $\lambda \in [0,1)$ and $\varepsilon > 0$, choose m large enough so that $(1-\lambda)\sum_{k=m+1}^\infty \lambda^{k-1} < \varepsilon$. Let α^* be an optimal (mixed) action of player 1 in the stage game G, i.e., $E_{\alpha^*}[r(a,b)] \geq \mathrm{Val}(G)$ for every $b \in B$. Define $\sigma = \{\sigma_n\}_{n=1}^\infty$ as follows. For $n = 1,\ldots,m$, $\sigma_n = \alpha^*$ regardless of the history, and for $n > m$, σ_n is an arbitrary pure action.

Observe that, for any $t \in T$, if $(X_n)_{n=1}^\infty$ is the random play induced by (σ, t), then $(1/n)H^n(\sigma : t) = (1/n)H(X_1,\ldots,X_n) = (m/n)H(\alpha^*) \to 0$ as $n \to \infty$. Hence (i). By playing σ, regardless of player 2's strategy, player 1 receives an expected payoff at least $\mathrm{Val}(G)$ at every stage up to stage m and at least 0 thereafter. Therefore, by the choice of m, player 1's discounted payoff is at least $(1-\varepsilon)\mathrm{Val}(G)$, which in turn is at least $\mathrm{Val}(G) - \varepsilon$ under the assumption $r \leq 1$. This proves (ii). Q.E.D.

The above phenomenon is due to the facts that (a) the definition of strategic entropy rate does not impose any restriction on the entropy of the induced play in finite stages and (b) with discounting, payoffs in the distant future have little effect on the present value. We will see in the next section that in the undiscounted games, where payoffs in a finite number of stages have no effect on the overall payoff, a bound on the strategic entropy rate has an appreciable effect on what the restricted player can guarantee in the long run.

6.3. *The Undiscounted Game* $G^{\infty}(\gamma)$

For each $\gamma \geq 0$ the game $G^{\infty}(\gamma)$ is the infinite repetitions of G in which player 1's strategies are restricted to those that have strategic entropy rate at most γ. Define $\Sigma^{\infty}(\gamma) = \{\sigma \in \Delta(S) | H^{\infty}(\sigma) \leq \gamma\}$. The set of strategies available to player 2 remains $\Delta(T)$.

For a payoff x to be a maxmin value of an infinitely repeated game, it is natural to require two things. First, player 1 should be able to "guarantee" himself x in the sense that he has a strategy that yields something very close to x in a long run regardless of player 2's strategy. Such a strategy may depend on how close the payoff should be to x (ε-optimality). Second, player 2 should be able to "defend" x in the sense that, for every strategy of player 1, player 2 has a counterstrategy that prevents player 1 from gaining substantially more than x in a long run. Again such a counter-strategy may depend not only on player 1's strategy but also on how close the resulting payoff should be to x. The next theorem asserts in particular that the maxmin value of $G^{\infty}(0)$ equals $\mathrm{U}_*(G)$. It asserts, moreover, that the payoff $\mathrm{U}_*(G)$ has stronger properties than the two requirements defining the maxmin value.

THEOREM 6.3. (i) $\exists \sigma \in \Sigma^{\infty}(0)$ *such that* $\forall t \in T$, $\forall n$, $r_n(\sigma, t) \geq \mathrm{U}_*(G)$.

(ii) $\forall \sigma \in \Sigma^{\infty}(0), \exists t \in T$ *such that* $\forall \varepsilon > 0$ $\exists n(\varepsilon)$ *such that* $\forall n > n(\varepsilon)$, $r_n(\sigma, t) \leq \mathrm{U}_*(G) + \varepsilon$.

Proof. Player 1 can obtain at least $\mathrm{U}_*(G)$ at every stage using a pure action. This proves (i). To prove (ii), take σ with $H^{\infty}(\sigma) = 0$. Let $t_\sigma \in T$ be as in Lemma 6.1, and let $\varepsilon > 0$ be given. Then by the definition of $H^{\infty}(\sigma)$ there is $n(\varepsilon)$ such that, for every $n \geq n(\varepsilon)$, $H^n(\sigma : t_\sigma) \leq \gamma(\varepsilon) n$, where $\gamma(\varepsilon)$ is as specified in Lemma 6.1. Therefore, $r_n(\sigma, t_\sigma) \leq \mathrm{U}_*(G) + \varepsilon$. This completes the proof. Q.E.D.

7. BOUNDED COMPLEXITY AND STRATEGIC ENTROPY

In this section we will apply Theorem 1 to repeated games with finite automata and bounded recall. The fact that any probability distribution over k points has entropy at most $\log k$ plays a crucial role in the subsequent discussion.

7.1. *Finite Automata*

Given a stage game $G = (A, B, r)$, an *automaton* of player 1 is a four-tuple $M = \langle Q, q_1, f, g \rangle$, where Q is a set of *states*, $q_1 \in Q$ is an *initial state*, f: $Q \to A$ is an *action function*, and g: $Q \times B \to Q$ is a *transition function*. (Player 2's automaton is defined by substituting A with B.) By the *size* of an automaton we mean the cardinality of the set of its states $|Q|$, which may be infinite. A *finite automaton* is an automaton of a finite size.

An automaton M plays a repeated game as follows. At each stage n it takes an action prescribed by f for the current state, say q_n, i.e., $f(q_n)$; it is set for q_1 at the first stage. Then it changes its state to q_{n+1} specified by g as a function of the current state q_n and player 2's action b_n, that is, $q_{n+1} = g(q_n, b_n)$.

Every automaton M *induces* a pure strategy $s = (s_n)_{n=1}^{\infty}$ of player 1 in the repeated game in the following manner. First, for any sequence of player 2's actions $b_1, \dots, b_n$ define an extension of the transition function inductively by

$$g(q, b_1, \dots, b_n) = g\big(g(q, b_1, \dots, b_{n-1}), b_n\big).$$

Then $s_1 = f(q_1)$ and, for $n > 1$, for each $\omega = ((a_k, b_k))_{k=1}^{\infty} \in \Omega_\infty$,

$$s_n(\omega) = f\big(g(q_1, b_1, \dots, b_{n-1})\big).$$

Conversely, every pure strategy of player 1 has a representation by an automaton, possibly of an infinite size. If a pure strategy s is equivalent to a pure strategy induced by an automaton, we say that s is *implementable* by that automaton. Here we set a bound on the size of automata that player 1 may use so that not all pure strategies are available to him.

For each $m \in \mathbb{N}$ let FA_m be the set of all pure strategies implementable by automata of size m. Denote by G_m^n the game (FA_m, T, r_n) which is the n-fold repetition of G in which player 1 is restricted to FA_m (and mixtures over it) while there is no restriction on player 2's strategies. The payoff function r_n is restricted accordingly.

Since $\Delta(FA_m)$ is a compact convex set, the value of G_m^n, denoted by V_m^n exists, i.e.,

$$V_m^n = \min_{\tau \in \Delta(T)} \max_{s \in FA_m} r_n(s, \tau) = \max_{\sigma \in \Delta(FA_m)} \min_{t \in T} r_n(\sigma, t).$$

Thus it is obvious, nonetheless important to realize, that in studying the value of a game, if it exists, we can focus our attention on either the minimax value or the maxmin value. It is the latter that we utilize in the subsequent discussion.

We consider the bound on the size of automata to be a function of the number of repetitions, m: $\mathbb{N} \to \mathbb{N}$, and study the asymptotics of $V^n_{m(n)}$. It is easily seen that $|FA_m| \le m^{Cm}$ ($m \ge 2$) for some positive constant C, e.g., $C = |B| + \log|A| + 1$. Let $\sigma \in \Delta(FA_m)$ and $H^n(\sigma) = H^n(\sigma : t^*)$. Let C^* be the number of the (n, t^*)-equivalence classes of FA_m. Obviously $C^* \le |FA_m|$. Then from the discussion of Section 5 and Proposition 3.1(1) it follows that $H^n(\sigma) \le \log C^* \le Cm \log m$. The following theorem is an immediate consequence of this observation and Theorem 6.1.

COROLLARY 7.1. *If* $\lim_{n\to\infty} m(n) \log m(n)/n = 0$, *then* $\lim_{n\to\infty} \mathrm{V}^n_{m(n)} = \mathrm{U}_*(G)$.

Proof. Obviously $\mathrm{V}^n_{m(n)} \ge \mathrm{U}_*(G)$ for every n and $m(n)$. As noted above, $\Delta(FA_{m(n)}) \subset \Sigma^n(Cm(n)\log m(n))$, and since the maxmin value is nondecreasing in the range over which max is taken,

$$\mathrm{V}^n_{m(n)} = \max_{\sigma\in\Delta(FA_{m(n)})} \min_{t\in T} r_n(\sigma, t) \le \mathrm{W}^n(Cm(n)\log m(n)).$$

Assume that $m(n)\log m(n)/n \to 0$ as $n \to \infty$. By Theorem 6.1 the right-hand side of the above inequality converges to $\mathrm{U}_*(G)$ as $n \to \infty$. Thus $\mathrm{V}^n_{m(n)}$ converges to $\mathrm{U}_*(G)$ as $n \to \infty$. Q.E.D.

In Neyman (1997) it is conjectured that if $m(n) = o(n/\log n)$, then $\mathrm{V}^n_{m(n)}$ converges to $\mathrm{U}_*(G)$ as n tends to infinity. Since $m(n) = o(n/\log n)$ is equivalent to $m(n)\log m(n) = o(n)$, Corollary 7.1 affirms the truth of the conjecture. Note that Corollary 7.1 gives only a sufficient condition for $\mathrm{V}^n_{m(n)}$ to converge to $\mathrm{U}_*(G)$. It is also conjectured in Neyman (1997) that if $\lim_{n\to\infty} n/m(n)\log n = 0$, then $\mathrm{V}^n_{m(n)}$ converges to $\mathrm{Val}(G)$.

Next we consider the game $G_{\lambda, m(\lambda)} = (FA_{m(\lambda)}, T, r_\lambda)$, the λ-discounted game in which player 1 is restricted to finite automata of size $m(\lambda)$ (and mixtures of them) which is a function of λ. Denote by $\mathrm{V}_{\lambda, m(\lambda)}$ the value of $G_{\lambda, m(\lambda)}$. Since $H^n(\sigma) \le Cm \log m$ for every $\sigma \in \Delta(FA_m)$, by taking the supremum over n, we have $H^*(\sigma) \le Cm \log m$. That is, $\Delta(FA_m) \subset \Sigma^*(Cm \log m)$. The next corollary follows from this and Theorem 6.2. The proof is analogous to that of Corollary 7.1 and is omitted.

COROLLARY 7.2. *If* $\lim_{\lambda\to 1}(1 - \lambda)m(\lambda)\log m(\lambda) = 0$, *then* $\lim_{\lambda\to 1} \mathrm{V}_{\lambda, m(\lambda)} = \mathrm{U}_*(G)$.

7.2. *Bounded Recall*

A pure strategy $s = (s_n)_{n=1}^{\infty}$ of player 1 in a repeated game is said to be of *bounded recall of size* l, or simply l*-recall*, if its choice of action at each stage depends only on actions taken by both players in the last l stages. Formally, such a strategy is represented by a function $z: (A \times B)^l \to A$ and an *initial memory* $e = (e_1, \dots, e_l) \in (A \times B)^l$, where, for each $\omega = (\omega_k)_{k=1}^{\infty} \in \Omega_\infty$,

$$s_n(\omega) = \begin{cases} z(e_n, \dots, e_l, \omega_1, \dots, \omega_{n-1}) & \text{if } n < l \\ z(\omega_{n-l}, \dots, \omega_{n-1}) & \text{if } n \geq l. \end{cases}$$

We will write $s = (e, z)$ and denote the set of all pure strategies of l-recall by BR_l.

Denote by $G^{n,l}$ the game (BR_l, T, r_n) which is the n-fold repetition of G in which player 1 is restricted to BR_l (and mixtures over it) while there is no restriction on player 2's strategies. The payoff matrix r_n is restricted accordingly. As in G_m^n, the value of $G^{n,l}$ exists:

$$V^{n,l} = \min_{\tau \in \Delta(T)} \max_{s \in BR_l} r_n(s, \tau) = \max_{\sigma \in \Delta(BR_l)} \min_{t \in T} r_n(\sigma, t).$$

Given $s = (e, z) \in BR_l$, we can construct an automaton $M = \langle Q, q, f, g \rangle$ of size $m(l) = |A \times B|^l$ that implements s. Set $Q = (A \times B)^l$, $q_1 = e$, and $f = z$. For $\omega = (\omega_1, \dots, \omega_l) \in Q$ and $b \in B$, define $g(\omega, b) = (\omega_2, \dots, \omega_l, (f(\omega), b))$. Therefore, by identifying strategies with their equivalence classes, we have $BR_l \subset FA_{m(l)}$. The next corollary, stated in Neyman (1997) as a conjecture, now follows from this and Corollary 7.1. We omit an easy proof.

COROLLARY 7.3. *There is a positive constant K such that if n: $\mathbb{N} \to \mathbb{N}$ satisfies $n(l) > \exp(Kl)$, then* $\lim_{l \to \infty} \mathrm{V}^{n(l),l} = \mathrm{U}_*(G)$.

Take, for example, $K = \log|A \times B| + 1$.

8. CONCLUDING REMARKS

We have defined strategic entropy for two-person games. The definition can be readily extended to n-person games. Given a vector of mixed or behavioral strategies $\sigma = (\sigma^1, \dots, \sigma^n)$, let $H(\sigma^{-i} : s^i)$ be the entropy of the play up to stage n induced by $(\sigma^{-i} : s^i)$ in which player i plays a pure strategy s_i and player $j \neq i$ plays σ^j. For a mixed strategy τ^i of player i, define $H^n(\sigma^{-i} : \tau^i)$ to be the average of $H^n(\sigma^{-i} : s^i)$ with respect to τ^i. This is the average amount of information on the other players' strategies

σ^{-i} that player i can obtain in the first n stages using τ^i. One then defines the n-strategic entropy of σ^{-i} by the maximum of $H(\sigma^{-i} : s^i)$ over all pure strategies of player i. The total strategic entropy and the strategic entropy rate of σ for player i are defined as in Section 4.

It is also of interest to extend the strategic entropy concept so as to measure the uncertainty of the play that a group of players collectively faces against (possibly correlated) strategies used by another group of players, or the amount of information that a group of players can obtain about strategies used by another group of players. One such extension is as follows.

Given a subset J of the n players, let σ^J be a correlated strategy of J and let σ^{-J} be a correlated strategy of the players not in J. For each pure strategy vector $s^J = (s^i)_{i \in J}$ define $H^n(\sigma^{-J} : s^J)$ to be the entropy of the play up to stage n induced by (σ^{-J}, s^J), and let $H^n(\sigma^{-J} : \sigma^J)$ be the expectation of $H^n(\sigma^{-J} : s^J)$ with respect to σ^J. The latter is the average amount of information on σ^{-J} that the coalition J can collectively obtain using σ^J. Then the n-strategic entropy of σ^{-J} is defined to be the maximum of $H^n(\sigma^{-J} : s^J)$ over all s^J. The study of these entropy-based quantities may be useful in the study of n-person games with bounded complexity, in particular, when some mode of correlation is possible through signaling of actions among the members of a coalition even if the rule of the game does not allow correlation in the stage game. For example, consider a three-player game in which there is a bound on the size of automata each player may use. Then the player with the smallest bound may be able to send a signal by means of a long sequence of actions that can be deciphered by the player with the largest bound but not by the one with the middle bound, thereby generating a correlation. The presence of such a phenomenon clearly affects the individually rational level of each player and hence equilibrium payoffs. One such example is given in Neyman (1997).

As an application of the results obtained in this paper for the non-zero-sum case, there are folk theorem type results for finitely repeated and λ-discounted two-person games with finite automata. Given a finite two-person game in strategic form $G = (A_1, A_2, r_1, r_2)$, let w_1 be the maxmin value for player 1 in pure actions, i.e., $\mathrm{w}_1 = \max_{a \in A_1} \min_{b \in A_2} r_1(a, b)$, and let $\bar{\mathrm{u}}_2$ be the minimax value for player 2 also in pure actions, i.e., $\bar{\mathrm{u}}_2 = \min_{a \in A_1} \max_{b \in A_2} r_2(a, b)$. Let $\tilde{F}$ be the set of feasible payoff vectors in which player 1 receives at least w_1 and player 2 receives at least $\bar{\mathrm{u}}_2$. Denote by $E^n(m(n))$ and $E_\lambda(m(\lambda))$ the set of equilibrium payoffs of the n-time repeated game and the λ-discounted game, with G as the stage game, in which player 1 is restricted to automata of size $m(n)$ and $m(\lambda)$, respectively. In Neyman and Okada (1998) it is shown that if there is a payoff vector (x, y) in $\tilde{F}$ such that $x > w_1$, then (1) $E^n(m(n))$ converges in the Hausdorff topology to $\tilde{F}$ as $n \to \infty$ under the condition that $\lim_{n \to \infty} m(n) \log m(n)/n = 0$ and (2) $E_\lambda(m(\lambda))$ converges also to $\tilde{F}$ as $\lambda \to 1$ under the condition that $\lim_{\lambda \to 1}(1 - \lambda) m(\lambda) \log m(\lambda) = 0$. As usual, part of the proof involves constructing an equilibrium path and punishment strategies for a deviation from it. Theorems 6.1 and 6.2 ensure the existence of effective punishment strategies of player 2.

REFERENCES

Aumann, R. J. (1981). "Survey of Repeated Games," in *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern*, pp. 11–42, Zurich: Bibliographisches Institut.

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

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

Kalai, E. (1990). "Bounded Rationality and Complexity in Repeated Games," in *Game Theory and Applications* (T. Ichiishi, A. Neyman, and Y. Tauman, Eds.). San Diego: Academic Press.

Kalai, E., Samet, D., and Stanford, W. (1988). "A Note on Reactive Equilibria in the Discounted Prisoner's Dilemma and Associated Games," *Int. J. Game Theory* **3**, 117–186.

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

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

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

Neyman, A. (1997). "Cooperation, Repetition, and Automata," in *Cooperation: Game-Theoretic Approaches* (S. Hart and A. Mas-Colell, Eds.), NATO ASI—Series F, Vol. 155, pp. 233–255. Berlin/New York: Springer-Verlag.

Neyman, A. (1998). "Finitely Repeated Games with Finite Automata," *Math. Oper. Res.*, to appear.

Neyman, A. and Okada, D. (1998). "Two-Person Repeated Games with Finite Automata," Discussion Paper 173, Center for Rationality and Interactive Decision Theory, Hebrew Univ. of Jerusalem, Israel.

Papadimitriou, C. H. and Yannakakis, M. (1994). "On Complexity as Bounded Rationality: Extended Abstract," in *STOC 94*, pp. 726–733.

Shannon, C. E. (1948). "A Mathematical Theory of Communication," *Bell System Tech. J.* **27**, 379–423, 623–656.

Smorodinsky, M. (1971). *Ergodic Theory, Entropy*, Lecture Notes in Mathematics. Berlin/New York: Springer-Verlag.

Zemel, E. (1989). "Small Talk and Cooperation: A Note on Bounded Rationality," *J. Econ. Theory* **49**, 1–9.