GROWING STRATEGY SETS IN REPEATED GAMES

DAIJIRO OKADA† AND ABRAHAM NEYMAN‡

Abstract. We consider a new type of restriction on strategy sets in repeated games, growing strategy sets. We impose a restriction on the way the set of strategies available to a player at each stage expands, possibly without bound, but not as fast as unrestricted strategy sets. In this paper growing strategy sets are defined without regard to any specific complexity measure of strategies. What is bounded is the rate of growth of the size of strategy set over time.

We then study an undiscounted infinitely repeated two-person zero-sum game in which the strategy set of player 1, the maximizer, expands "slowly" while there is no restriction on player 2's strategy space. Our main result is that, if the number of strategies available to player 1 at stage t grows subexponentially with t, then player 2 has a pure optimal strategy and the value of the game is the maxmin value of the stage game in pure actions, the lowest payoff that player 1 can guarantee for sure in one-shot game. This is a strong result in that an optimal strategy in an infinitely repeated game has, by definition, a property that, for every $\varepsilon > 0$, it holds player 1's payoff to at most the value plus ε after some stage.

We also briefly discuss how a growing strategy set may arise as a result of allowing strategic complexity, such as the size of automata, to grow with the number of repetitions. (*Journal of Economic Literature* Classification Number:)

Keywords: Growing Strategy Sets, Strategic Complexity, Repeated Games

[This Version: December 12, 2003]

---

† Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, New Jersey 08901-1248, USA.

Email: okada@econ.rutgers.edu.

‡ Institute of Mathematics, The Hebrew University of Jerusalem, Givat Ram, Jerusalem 91904, Israel.

Email: aneyman@math.huji.ac.il.

1. Introduction

Complexity of repeated games as a model of interactive decision making stems, in part, from the richness of strategies that the theory allows players to choose from. The number of theoretically possible strategies is double-exponential in the number of repetitions. (See Section 2.2.) This is due to the fact that the number of histories grows exponentially with the number of repetitions and also that we count strategies that map histories into actions in all possible ways. Some strategies are too complicated to admit a short and practically implementable description: a short description of a strategy requires an efficient encoding of histories, but some histories may have no shorter descriptions than simply writing them out in their entirety. These considerations motivate the study of how restricting strategies to simple ones may affect outcomes of repeated games.

Various methods of restricting strategies have been investigated: finite automata, bounded recall, Turing machine etc. Each method captures a particular aspect of complexity of strategies in repeated games.[1]

An advantage of restricting strategies to those of fixed finite complexity (e.g., number of states of finite automata or length of recall) is that the strategy sets become finite sets and so the existence of the value and equilibrium is guaranteed, and, in some cases one can even write down the payoff matrix.[2]

Some criticisms may be made of imposing exogenous bound of complexity. First, measures of complexity may seem ad hoc, different complexity measure leads to radically different classification of strategies (Stearns (1997)) and thus results obtained may be sensitive to particular choice of complexity measure. Second, each measure

[1]See a survey by Kalai (1990). Also see Neyman (1997) for more recent results when complexity bounds are exogenously given in contrast to the models of Rubinstein (1986) and Abreu and Rubinstein (1988) in which there are no fixed exogenous complexity bounds but the cost of complexity is considered. We do not consider the cost of complexity. Strategy sets are exogenously restricted and the restriction does not enter players' preferences. One may think of exogenous restriction as an extreme case of complexity cost: infinite cost to use a strategy outside of the restricted set. Our notions of optimal strategy, the value, and equilibrium outcomes are standard, i.e., it involves payoff comparisons only. Trade off between cost of expanding computational resource and attainable payoffs is an important topic of study.

[2]See an analysis of repeated prisoner's dilemma with "memory zero" strategies by Aumann, Cave, and Kurz as reported in Aumann (1981)

captures only a specific aspect of complexity and ignores some other important aspects of complexity such as amount of time required to determine an action at each stage, and related to this, complexity of computing best response (Stearns (1997), O'Connell and Stearns (1999), Gilboa (1988)). Papadimitriou (1992) finds an interesting trade off between easy description and implementation of strategies (bounded number of states of automata) versus the computational complexity of finding best response among the restricted strategies.

To these we add one more item which is the motivation for the research reported in this paper. Boundedly rational players are limited by the set of available (implementable) strategies, but limitation may ease over time or there may not be a drastic finite limit to the complexity for the entire horizon, possibly infinite, of the game. Computational resource may expand, e.g., by adding more memory over time or by learning. Any model that captures these intuitions would impose some restriction on the way the set of strategies available at each stage expands, possibly without bound but not as fast as unrestricted strategy sets.

In light of the discussion in the last two paragraphs, we consider growing strategy sets with an arbitrary restriction without regard to any specific complexity measure. What is bounded is the rate of growth of the strategy sets over time. To be more precise, we imagine player i with a restricted strategy set Ψ_i in a repeated game. Nature of the restriction or complexity bound that results in Ψ_i is arbitrary. In particular, Ψ_i may contain infinite number of strategies including those that cannot be implemented by any computer with finite memory. For each stage t of the repeated game, we count the number of strategies in Ψ_i that look distinct to other players up to that stage, $\psi_i(t)$. If Ψ_i is the unrestricted set itself, then, as mentioned in the beginning, $\psi_i(t)$ is double-exponential in t. Thus it is of interest to study how outcomes of repeated games are affected by the condition on the rate of growth of $\psi_i(t)$.

Since no structure is imposed on the strategies that belong to Ψ_i, it would be difficult to derive results that rely on an explicit construction of strategies. For this reason, and as a first undertaking in this line of research, we will study a simplest model of repeated games with restricted strategy sets: repeated two-person zero-sum games in which the strategy set of player 1, the maximizer, is restricted in the manner mentioned above while there is no restriction on player 2's strategy set. The payoffs in the repeated games are undiscounted. Under the condition that $\psi_i(t)$, the cardinality of Ψ_i, grows subexponentially with t, we will show that player 2 has a pure optimal strategy and the value of the game is the maxmin value of the stage game in pure actions, the lowest payoff that player 1 can guarantee himself for sure in one-shot game. This is a strong result since an optimal strategy in an infinitely repeated game has, by definition, a property that, for every $\varepsilon > 0$, it holds player 1's payoff to at most the value plus ε after some stage regardless of player 1's choice of strategy. Our justification for studying the zero-sum case is a standard one: individually rational levels of payoff must be determined relative to the restricted strategy sets which will provide a useful information when one studies nonzero-sum games.

We will set the notation used throughout the paper and formalize the idea of growing strategy set in Section 2. Some examples of growing strategy sets will also be discussed in this section. Section 3 contains the main results. Some concluding remarks are made in Section 4.

2. Strategies in Repeated Games

2.1. Preliminaries. Consider a finite n-player game in strategic form,

$$G = \big(A_i, g_i\big)_{i\in N}.$$

The set of players is $N = \{1, \dots, n\}$. For each $i \in N$, A_i is the set of actions available to player i and $g_i : A_1 \times \cdots \times A_n \to \mathbb{R}$ is his payoff function. Set $A = A_1 \times \cdots \times A_n$. We call G the stage game.

In a repeated games[3], at each stage each player observes a history of actions by all players and takes an action. Thus a pure strategy for player i is a mapping from all possible histories to his actions. Formally, let $H_1 = \{\epsilon\}$ be the set of "the empty history", and, for each positive integer $t > 1$, let

$$H_t = \underbrace{A \times \cdots \times A}_{(t-1) times}.$$

[3]In this paper we talk of the most basic model of repeated games, i.e., ones with complete information, perfect monitoring and standard signaling.

Then H_t represents the information available to each player at the beginning of the t-th stage. Player i's pure strategy in a repeated game is then a mapping $\sigma_i : \cup_{t=0}^{\infty} H_t \to A_i$. Let Σ_i be the set of all pure strategies of player i.

There is an alternative way of representing a strategy in a repeated game. We define player i's strategy at the t-th stage to be a mapping $\sigma_{it} : H_t \to A_i$. Then i's strategy in the repeated game can be represented by a sequence $\sigma_{i1}, \sigma_{i2}, \ldots$. Denote by Σ_{it} the set of all σ_{it}'s. Then i's strategy set is their Cartesian product $\Sigma_{i1} \times \Sigma_{i2} \times \ldots$. We use the same symbol Σ_i to denote this product.[4]

A play of a repeated game is a sequence $(a^1, a^2, \ldots)$ where $a^t \in A$. An n-tuple of strategies $\sigma = (\sigma_1, \ldots, \sigma_n) \in \Sigma_1 \times \ldots, \Sigma_n$ recursively induces a play as follows.

$$a_1(\sigma) = (\sigma_1(\epsilon), \ldots, \sigma_n(\epsilon))$$
$$a_{t+1}(\sigma) = (\sigma_1(a_1(\sigma), \ldots, a_t(\sigma)), \ldots, \sigma_n(a_1(\sigma), \ldots, a_t(\sigma))).$$

We say that two strategies of a player i, σ_i and σ_i', are equivalent up to the t-th stage if, for every $(n-1)$-tuple of other players' strategies σ_{-i},

$$a_s(\sigma_i, \sigma_{-i}) = a_s(\sigma_i', \sigma_{-i}) \quad \text{for } s = 1, \ldots, t.$$

If two strategies are equivalent up to the t-th stage for every t, then we simply say they are equivalent.

2.2. **Restricted Strategy Sets and Their Growth.** Let us denote by m_i the number of actions available to player i, i.e., $m_i = |A_i|$, and $m = m_1 \times \cdots \times m_n = |A|$. We note first that the number of strategies available to player i in the first t stages of a repeated game is[5]

$$|\Sigma_{i1}| \times \cdots \times |\Sigma_{it}| = m_i^{m^0} \times \cdots \times m_i^{m^{t-1}} = m_i^{\frac{m^t - 1}{m-1}}.$$

This number is double exponential in t.

Suppose that player i has access to a restricted set of strategies, $\Psi_i \subset \Sigma_i$, due to limitations on some aspects of complexity of his strategies. For each positive

[4]For $\sigma_i \in \Sigma_i$, let σ_{it} be the restriction of σ_i to H_t. Then $(\sigma_{i1}, \sigma_{i2}, \ldots) \in \Sigma_{i1} \times \Sigma_{i2} \times \ldots$. Conversely, given $(\sigma_{i1}, \sigma_{i2}, \ldots) \in \Sigma_{i1} \times \Sigma_{i2} \times \ldots$, let $\sigma_i(h) = \sigma_{it}(h)$ for $h \in H_t$. Then $\sigma_i \in \Sigma_i$. So the two ways of representing strategies are equivalent.

[5]The number of reduced strategies available to player i in the first t stages is $m_i^{\frac{m_{-i}^t - 1}{m_{-i} - 1}}$ where $m_{-i} = \times_{j \neq i} m_j$.

integer t, let $\Psi_i(t)$ be formed by identifying strategies in Ψ_i that are equivalent up to the t-th stage.[6] Note that Ψ_i and $\Psi_i(t)$ may not be expressed as a Cartesian product as we did for Σ_i $(= \Sigma_{i1} \times \Sigma_{i2} \times \dots)$.

Let $\psi_i(t)$ be the number of elements in $\Psi_i(t)$. Any consideration on strategic complexity gives rise to a restricted strategy set Ψ_i and thus limitation on the rate of growth of $\psi_i(t)$. For example, if player i is restricted to those strategies that can be implemented by automata with a fixed number of states, then Ψ_i is a finite set and $\Psi_i(t) = \Psi_i$ for all sufficiently large[7] t. In this case $\psi_i(t) = O(1)$.

We illustrate the concept of growing strategy sets in a few examples. These examples are not meant to be realistic or theoretically useful, but rather to be an aid to fix the concept in readers mind more readily.

Example 1. For each t, let Ψ_{it} be a subset of Σ_{it} and $\psi_{it} = |\Psi_{it}|$. Define $\Psi_i = \Psi_{i1} \times \Psi_{i2} \times \dots$. Then $\Psi_i(t) = \Psi_{i1} \times \dots \times \Psi_{it}$ and $\psi_i(t) = \psi_{i1} \times \dots \times \psi_{it}$.

In all the examples that follow, consider a two person game in which each player has two actions: $N = \{1, 2\}$ and $A_1 = A_2 = \{0, 1\}$.

Example 2. For each positive integer k, define a strategy $\sigma_1^{(k)}$ as follows. For each history h, let $\kappa(h)$ be the number of times player 2 chose action 1.

$$\sigma_1^{(k)}(h) = \begin{cases} 1 & \text{if } \kappa(h) \geq k \\ 0 & \text{otherwise.} \end{cases}$$

Let $\Psi_1 = \{\sigma_1^{(1)}, \sigma_1^{(2)}, \dots\}$. Then $\Psi_1(t) = \{\sigma_1^{(1)}, \dots, \sigma_1^{(t)}\}$ and $\psi_i(t) = t$.

Example 3. A prefix of a history $h = (h_1, \dots, h_t)$ is any of its initial segment $h' = (h_1, \dots, h_s)$, $s \leq t$. A set of histories $L \subset \cup_{t=1}^{\infty} H_t$ is said to be prefix-free if no element of L is a prefix of another. Now, for each positive integer t, let $L(t) \subset H_1 \cup \dots \cup H_t$ be prefix-free and $L(t) \subset L(t+1)$. Write λ for the sequence $L(1), L(2), \cdots$, and define a strategy σ_1^{λ} as follows.

$$\sigma_1^{\lambda}(h_1, \dots, h_t) = \begin{cases} 1 & \text{if } (h_1, \dots, h_s) \in L(t) \text{ for some } s \leq t, \\ 0 & \text{otherwise.} \end{cases}$$

[6] If two strategies in Ψ_i are equivalent, then they are never distinguished in $\Psi_i(t)$ for any t. So the reader may consider Ψ_i to be the set of equivalence classes of strategies.

[7] In fact, this holds for all $t \geq m$ where m is the bound on the number of states of automata.

This is a generalization of the trigger strategy: σ_1^λ takes action 1 forever as soon as a history in some $L(t)$ occurs. Let $\mathcal{L}$ be the set of all increasing sequences of prefix-free sets of histories. Take a subset $\mathcal{M}$ of $\mathcal{L}$ and define Ψ_1 to be the set of player 1's strategies σ_1^λ with $\lambda \in \mathcal{M}$. Let us examine $\Psi_1(t)$ and $\psi_1(t)$.

It is easy to verify that, for any $\lambda = (L(t))_t$ and $\mu = (M(t))_t$ in $\mathcal{L}$, σ_1^λ and σ_1^μ are equivalent up to the t-th stage if, and only if, $L(t) = M(t)$. We say that λ and μ are equivalent up to the t-th stage if $L(t) = M(t)$. This is an equivalence relation on $\mathcal{L}$, and hence on $\mathcal{M}$. We denote by $\mathcal{M}(t)$ the set of the equivalence classes when this relation is taken on $\mathcal{M}$. For notational simplicity, the elements of $\mathcal{M}(t)$ will be denoted by λ, μ and so on as for the elements of $\mathcal{M}$ themselves. Then we have

$$\Psi_1(t) = \{\sigma_1^\lambda \mid \lambda \in \mathcal{M}(t)\} \quad \text{and} \quad \psi_1(t) = |\mathcal{M}(t)|.$$

Examples of $\mathcal{M}$ can be constructed as follows. Let $f : \mathbb{N} \to \mathbb{N}$ be a nondecreasing function and let $\mathcal{M} = \{(L(t))_t \in \mathcal{L} \mid |L(t)| = O(f(t))\}$. It is not difficult to construct examples of $(L(t))_t$ for which $|L(t)| = O(t)$, $O(t^p)$ for each $p > 1$, and $O(2^{\alpha t})$ for $0 < \alpha < 1$.

3. Some Results: Restricted vs. Unrestricted Players

We now derive a few consequences of restricting strategy sets in terms of the growth rate of $\psi_i(t) = |\Psi_i(t)|$, which may be interpreted as the number of strategies available to player i up to the t-th stage. We emphasize that the nature of the restricted strategy set Ψ_i is completely arbitrary. It may include infinitely many strategies and also the strategies that cannot be represented by any finite state machines or finitely bounded recall. As such, it is quite difficult, if not impossible, to obtain results on optimal strategies or equilibrium payoffs which require examination and construction of specific strategies. In what follows we study what may appear to be an extreme case of repeated games with strategy restriction: a two-person zero-sum infinitely repeated game with no discounting of payoffs and only one of the players (player 1, the maximizer) has a restricted strategy set and plays against the unrestricted player (player 2).[8] Although the repeated game we study in this paper is rather special, our results apply to any measure of strategic

[8]Although mixed strategies are not used in our results, we allow the players to use them. The choice of a strategy, according to some probability distribution, is performed before the game

complexity that gives rise to a restricted strategy set satisfying our condition on the rate of growth $\psi_1(t)$.

Let w be player 1's maxmin payoff in the stage game where max and min are taken over the pure actions: $w = \max_{a_1 \in A_1} \min_{a_2 \in A_2} g(a_1, a_2)$. This is the worst payoff that player 1 can guarantee himself for sure in the stage game. For a pair of repeated game strategies $(\sigma_1, \sigma_2) \in \Sigma_1 \times \Sigma_2$, we write $g_t(\sigma_1, \sigma_2)$ for the player 1's average payoff up to the t-th stage.

3.1. **Finite Set of Strategies.** If the restricted strategy set Ψ_1 is a finite set, then it is obvious that the unrestricted player 2 can construct a strategy, say σ_2^*, which eventually identifies the strategy chosen by player 1 and gives him at most w at each stage thereafter.[9] Therefore $g_t(\sigma_1, \sigma_2^*)$ converges to w for every $\sigma_1 \in \Psi_1$. The first proposition provides its speed of convergence. It has appeared in Neyman and Okada (2000) in a study of nonzero-sum two person finitely repeated games with finite automata. In order to make this paper self-contained, and, since this proposition will be used in the proof of the second proposition, we will give the proof.

Proposition 1. *For every finite subset Ψ_1 of Σ_1 there exists $\sigma_2^* \in \Sigma_2$ such that for all $\sigma_1 \in \Psi_1$*

$$g_t(\sigma_1, \sigma_2^*) \leq w + \|g\| \frac{\log_2 |\Psi_1|}{t} \quad \textit{for all } t = 1, 2, \ldots$$

where $\|g\| = \max\{g(a_1, a_2) \mid a_1 \in A_1, a_2 \in A_2\}$.

Proof: For each history $h = (h_1, \ldots, h_{t-1})$, where $h_s = (a_{1s}, a_{2s})$, let Ψ_1^h be the set of strategies in Ψ_1 that are compatible with h. That is,

$$\Psi_1^h = \{\sigma_1 \in \Psi \mid \sigma_1(\epsilon) = a_{11}, \text{ and}$$
$$\sigma_1(h_1, \ldots, h_{s-1}) = a_{1s} \text{ for all } s = 2, \ldots, t-1.\}$$

For each $a_1 \in A_1$ let Ψ_1^{h,a_1} be the set of strategies in Ψ_1^h that takes the action a_1 given the history h, i.e.,

$$\Psi_1^{h,a_1} = \left\{\sigma_1 \in \Psi_1^h \mid \sigma_1(h) = a_1\right\}.$$

---

starts. Note that when strategy set is restricted, certain behavioral strategies may not be available to the player if they are not equivalent to mixtures of available pure strategies.

[9]Ben-Porath (1993)[Lemma 1, Theorem 1] provides an explicit construction of such strategy when the restriction is in terms of finite automata.

Let $a_1(h) \in A_1$ be such that $\left|\Psi_1^{h,a_1(h)}\right| \geq \left|\Psi_1^{h,a_1}\right|$ for all $a_1 \in A_1$. Now define σ_2^* by

$$\sigma_2^*(h) = \operatorname*{argmin}_{a_2 \in A_2} g(a_1(h), a_2).$$

Clearly, $\left\{\Psi_1^{h,a_1} \mid a_1 \in A_1\right\}$ is a partition of Ψ_1^h. If $a_1 \neq a_1(h)$, then $\left|\Psi_1^{h,a_1}\right|$ is at most one half of $\left|\Psi_1^h\right|$. This implies that

$$\left|\Psi_1^{(h_1,\dots,h_{t-1},h_t)}\right| \leq \frac{\left|\Psi_1^{(h_1,\dots,h_{t-1})}\right|}{2} \tag{1}$$

whenever $h_t \neq (a_1(h), \cdot)$.

Fix $\sigma_1 \in \Psi_1$ and let $(h_1, h_2, \dots)$, where $h_s = (a_{1s}, a_{2s})$, be the play generated by (σ_1, σ_2^*). If we set

$$I_s = \begin{cases} 1 & \text{if } a_{1s} \neq a_1(h_1, \dots, h_{s-1}) \\ 0 & \text{otherwise,} \end{cases}$$

then (1) implies that

$$\left|\Psi_1\right| 2^{-\sum_{s=1}^t I_s} \geq \left|\Psi_1^{(h_1,\dots,h_t)}\right| \geq 1.$$

Therefore

$$\sum_{s=1}^t I_s \leq \log_2 \left|\Psi_1\right|.$$

This means that the number of stages at which player 1's action differs from $a_1(h_1, \dots, h_{s-1})$ is at most $\log_2\left|\Psi_1\right|$. Thus

$$g_t(\sigma_1, \sigma_2^*) = \frac{1}{t}\sum_{s=1}^t g(h_s) \leq \frac{1}{t}\sum_{s=1}^t \big((1 - I_s)w + \|g\| I_s\big) \leq w + \|g\| \frac{\log_2\left|\Psi_1\right|}{t}.$$

This completes the proof. **Q.E.D.**

3.2. **Infinite Ψ_1 for which $\psi_1(t)$ grows slowly.** Next we consider an infinite strategy set Ψ_1. Recall that $\Psi_1(t)$ is formed by identifying strategies in Ψ_1 that are equivalent up to the t-th stage and $\psi_1(t) = \left|\Psi_1(t)\right|$.

Proposition 2. *Suppose that* $\log \psi_1(t)/t \to 0$ *as* $t \to \infty$. *Then there is a strategy* $\sigma_2^* \in \Sigma_2$ *such that, for every* $\sigma_1 \in \Psi_1$,

$$\limsup_{t \to \infty} g_t(\sigma_1, \sigma_2^*) \leq w.$$

Thus, if the growth rate of $\psi_1(t)$ is subexponential in t, then player 2 can guarantee not to give player 1 more than w in the long run. Note that whether player 1 can attain exactly w or not depends on what strategies are in Ψ_1. For example, if $a^* = \mathrm{argmax}_a(\min_b g(a,b))$, and a strategy that takes a^* in every stage is available, then w can be achieved by using such strategy. We first present a lemma whose proof is found in the appendix.

Lemma 1. *If $\log \psi_1(t)/t \to 0$ as $t \to \infty$, then there is an increasing sequence of positive integers $t_1, t_2, \ldots$ such that*

(A) $\dfrac{t_{k+1} - t_k}{t_k} \to 0$ *as* $k \to \infty$*, and*

(B) $\dfrac{\log \psi_1(t_{k+1})}{t_{k+1} - t_k} \to 0$ *as* $k \to \infty$*.*

Proof of Proposition 2: Let $\{t_k\}_{k=1}^{\infty}$ be a sequence satisfying the properties (A) and (B) of Lemma 1. Call a consecutive stages $t_k + 1, \ldots, t_{k+1}$ of the repeated game the k-th block.

The construction of player 2's strategy σ_2^* is similar to the one in the proof of Proposition 1. Given a history $h = (h_1, \ldots, h_{t-1})$, there is a unique k with $t_k \le t < t_{k+1}$. Let $\Psi_1^h(t_{k+1})$ be the set of player 1's strategies in $\Psi_1(t_{k+1})$ that are compatible with h and, for each $a_1 \in A_1$, set

$$\Psi_1^{h,a_1}(t_{k+1}) = \{\sigma_1 \in \Psi_1^h(t_{k+1}) \mid \sigma_1(h) = a_1\}.$$

Let $a_1(h) = \mathrm{argmax}_{a_1 \in A_1} \left|\Psi_1^{h,a_1}(t_{k+1})\right|$ and define

$$\sigma_2^*(h) = \operatorname*{argmin}_{a_2 \in A^2} g(a_1(h), a_2).$$

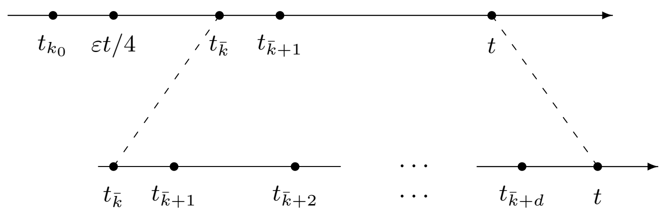
In short, σ_2^* plays the strategy constructed in the proof of Proposition 1 against $\Psi_1^h(t_{k+1})$ during the k-th block.

Fix $\varepsilon > 0$ and let k_0 be such that, for all $k \ge k_0$,

$$\frac{t_{k+1} - t_k}{t_k} < \frac{\varepsilon}{4}. \tag{2}$$

$$\frac{\log_2 \psi_1(t_{k+1})}{t_{k+1} - t_k} < \frac{\varepsilon}{4}. \tag{3}$$

Take $t > 4t_{k_0}/\varepsilon$ and let $\bar{k}$ be the smallest index k for which $t_k > \varepsilon t/4$. Then, $\bar{k} > k_0$ and $t_{\bar{k}-1} < \varepsilon t/4 < t_{\bar{k}}$. See the figure above.



Fix $\sigma_1 \in \Psi_1$. Let $h = (h_1, h_2, \dots)$ be the play induced by (σ_1, σ_2^*). Note that $\sigma_1 \in \Psi_1^{(h_1,\dots,h_{t-1})}(t_{k+1}) \subset \Psi_1(t_{k+1})$ whenever $t_k + 1 \le t \le t_{k+1}$. The average payoff to player 1 up to stage t is

$$g_t(\sigma_1, \sigma_2^*) = \frac{1}{t}\Bigg(\sum_{s=1}^{t_{\bar{k}}} g(h_s) + \sum_{s=t_{\bar{k}}+1}^{t} g(h_s)\Bigg).$$

W.l.o.g, assume $\|g\| \le 1$. First, note that, by (2),

$$\frac{1}{t}\sum_{s=1}^{t_{\bar{k}}} g(h_s) \le \frac{t_{\bar{k}}}{t} = \frac{t_{\bar{k}} - (\varepsilon t/4) + (\varepsilon t/4)}{t} \le \frac{t_{\bar{k}} - t_{\bar{k}-1}}{t_{\bar{k}-1}} + \frac{\varepsilon}{4} < \frac{\varepsilon}{2}.$$

Next suppose that there are d blocks between $t_{\bar{k}} + 1$ and t. Then,

$$\frac{1}{t}\sum_{s=t_{\bar{k}}+1}^{t} g(h_s) = \frac{1}{t}\sum_{j=1}^{d}\sum_{s=t_{\bar{k}+j-1}+1}^{t_{\bar{k}+j}} g(h_s) + \frac{1}{t}\sum_{s=t_{\bar{k}+d}+1}^{t} g(h_s).$$

The definition of σ_2^* and Proposition 1, together with (3) above, imply

$$\begin{aligned}\frac{1}{t}\sum_{s=t_{\bar{k}+j-1}+1}^{t_{\bar{k}+j}} g(h_s) &= \frac{t_{\bar{k}+j} - t_{\bar{k}+j-1}}{t} \cdot \frac{1}{t_{\bar{k}+j} - t_{\bar{k}+j-1}} \sum_{s=t_{\bar{k}+j-1}+1}^{t_{\bar{k}+j}} g(h_s) \\ &\le \frac{t_{\bar{k}+j} - t_{\bar{k}+j-1}}{t}\Big(w + \frac{\varepsilon}{4}\Big),\end{aligned}$$

and (2) implies that

$$\frac{1}{t}\sum_{s=t_{\bar{k}+d}+1}^{t} g(h_s) \le \frac{t_{\bar{k}+d+1} - t_{\bar{k}+d}}{t_{\bar{k}+d}} < \frac{\varepsilon}{4}.$$

Hence

$$\begin{aligned}\frac{1}{t}\sum_{s=t_{\bar{k}}+1}^{t} g(h_s) &< \frac{1}{t}\Big(w + \frac{\varepsilon}{4}\Big)\sum_{j=1}^{d}(t_{\bar{k}+j} - t_{\bar{k}+j-1}) + \frac{\varepsilon}{4} \\ &= \Big(w + \frac{\varepsilon}{4}\Big)\frac{t_{\bar{k}+d} - t_{\bar{k}}}{t} + \frac{\varepsilon}{4} \\ &< w + \frac{\varepsilon}{2}.\end{aligned}$$

Therefore, $g_t(\sigma_1, \sigma_2^*) < w + \varepsilon$. **Q.E.D.**

When Ψ_i is a Cartesian product, as in Example 1, the construction of σ_2^* is easy. Let $\Psi_i = \Psi_{i1} \times \Psi_{i2} \times \ldots$. So $\Psi_i(t) = \Psi_{i1} \times \cdots \times \Psi_{it}$. Hence $\psi_i(t) = \left|\Psi_{i1}\right| \times \cdots \times \left|\Psi_{it}\right|$. For each t, let $n(t) = \left|\{s \mid s \leq t, \left|\Psi_{is}\right| \geq 2\}\right|$. Then, $\psi_i(t) \geq 2^{n(t)}$. Thus $\log_2 \psi_i(t)/t \to 0$ (as $t \to \infty$) implies that $n(t)/t \to 0$ (as $t \to \infty$). For each history $h = (h_1, \ldots, h_{t-1})$, define $\sigma_2^*(h)$ by

$$\sigma_2^*(h) = \begin{cases} \operatorname{argmin}_{a_2 \in A_2} g(\sigma_{1t}(h), a_2) & \text{if } \Psi_{it} = \{\sigma_{it}\} \\ \textit{arbitrary action} & \text{if } \left|\Psi_{is}\right| \geq 2 \end{cases}$$

Then

$$g_t(\sigma_1, \sigma_2^*) \leq \frac{t - n(t)}{t} w + \frac{n(t)}{t} \|g\| \to w \quad \text{as} \quad t \to \infty.$$

3.3. **Entropy and Growing Strategy Sets.** In this section we prove a generalization of Proposition 2 for the case when $\log \psi_1(t)/t$ converges to an arbitrary positive number. To do this we will use the concept of entropy and its properties which we will now introduce.

Let X be a random variable that takes values in a finite set Ω and let $p(x)$ denote the probability that $X = x$ for each $x \in \Omega$. Then the entropy of X is defined as the negative of the expected values of the logarithm of p, that is,

$$H(X) = -\sum_{x \in \Omega} p(x) \log p(x).$$

The entropy of a vector of random variables, $H(X_1, \cdots, X_n)$, is similarly defined.

The conditional entropy of a random variable X given another random variable Y is defined as follows. Given the event $Y = y$, let $H(X|y)$ be the entropy of X with respect to the conditional distribution of X given y, that is,

$$H(X|y) = -\sum_{x} p(x|y) \log p(x|y).$$

Then the conditional entropy of X given Y is the expected value of $H(X|y)$ with respect to the (marginal) distribution of Y:

$$H(X|Y) = E_Y[H(X|y)] = \sum_{y} p(y) H(X|y).$$

The following "chain rule" for entropy, which we will use in the proof of the next theorem, is easy to verify. See Cover and Thomas (1991).

Lemma 2. $H(X_1, \cdots, X_T) = H(X_1) + \sum_{t=2}^{T} H(X_t|X_1, \cdots, X_{t-1})$.

Let $(\Omega, \mathcal{F}, \mu)$ be a probability space and let $\mathcal{P}$ be a finite partition of Ω into sets in $\mathcal{F}$. Then the entropy of the partition $\mathcal{P}$, with respect to μ is defined by

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

It is easy to see that if $\mathcal{Q}$ is a refinement of $\mathcal{P}$, then $H_\mu(\mathcal{P}) \leq H_\mu(\mathcal{Q})$.

Given a restricted strategy set of player 1, $\Psi_1 \subset \Sigma_1$, we have defined, for each t, the set $\Psi_1(t)$ to be the partition of Ψ_1 induced by an equivalence of pure strategies. That is, we define an equivalence relation $\underset{t}{\sim}$ by

$$\sigma \underset{t}{\sim} \sigma' \iff \forall \tau \in \Sigma_2,\ a_s(\sigma, \tau) = a_s(\sigma, \tau) \quad \text{for } s = 1, \ldots, t.$$

Then $\Psi_1(t) = \Psi_1/\underset{t}{\sim}$.

Now fix player 2's strategy τ. Define an equivalence relation $\underset{t,\tau}{\sim}$ by

$$\sigma \underset{t,\tau}{\sim} \sigma' \iff a_s(\sigma, \tau) = a_s(\sigma, \tau) \quad \text{for } s = 1, \ldots, t.$$

and let $\Psi_1(t, \tau) = \Psi_1/\underset{t,\tau}{\sim}$. Clearly $\Psi_1(t, \tau)$ is a finite partition of Ψ_1 and $\Psi_1(t)$ is its refinement. Hence, by the property of the entropy of partitions mentioned above,

$$H_\sigma(\Psi_1(t, \tau)) \leq H_\sigma(\Psi_1(t)) \leq \log|\Psi_1(t)| = \log \psi(t). \tag{4}$$

By the definition of the equivalence relation defining $\Psi_1(t, \tau)$, each equivalence class $S \in \Psi_1(t, \tau)$ is associated with a history of length t, $h(S) \in H_t$. More precisely, $h(S)$ is the history of length t which results when the strategy profile (s, τ) is played, for any $s \in S$. Conversely, for any history $h \in H_t$, there is an equivalence class $S \in \Psi_1(t, \tau)$ such that $h = h(S)$. Clearly, this correspondence between $\Psi_1(t, \tau)$ and H_t is one-to-one. Furthermore, the event "a strategy $s \in S \subset \Psi_1(t, \tau)$ is selected by σ" is equivalent to the event "the history $h(S)$ occurs when (σ, τ) is played". Therefore,

$$\sigma(S) = P_{\sigma,\tau}(h(S)).$$

Let us write $X_1, \ldots, X_t$ for the sequence of action profiles up to stage t when (σ, τ) is played. So it is a random vector with distribution $P_{\sigma,\tau}$. Then the observation in this paragraph implies that

$$\begin{aligned} H_\sigma(\Psi_1(t,\tau)) &= -\sum_{S\in\Psi_1(t,\tau)} \sigma(S)\log\sigma(S) \\ &= -\sum_{h\in H_t} P_{\sigma,\tau}(h)\log P_{\sigma,\tau}(h) \\ &= H(X_1,\ldots,X_t). \end{aligned}$$

Combining this equality with (4) we have

Lemma 3. *Let $\sigma\in\Delta(\Psi_1)$ and $\tau\in\Sigma_2$ and $(X_1,\ldots,X_t)$ be the random play up to stage t induced by (σ,τ). Then, for every t,*

$$H(X_1,\ldots,X_t)\le\log\psi_1(t).$$

For each mixed action μ of player 1, let $H(\mu)$ be its entropy, i.e.,

$$H(\mu) = -\sum_{a\in A}\mu(a)\log\mu(a).$$

Define a function $U:\mathbb{R}_+\to\mathbb{R}_+$ by

$$U(x) = \max_{\substack{\mu\in\Delta(A)\\ H(\mu)\le x}}\min_{b\in B} r(\mu,b).$$

Thus $U(x)$ is what player 1 can secure in the stage game G using a mixed action of entropy at most x. Clearly, $U(0)=w$, the maximin value in pure actions. Let cav U be the concavification of U, i.e., the smallest concave function which is at least as large as U at every point of its domain.

Theorem 1. *Suppose that* $\limsup_{t\to\infty}\dfrac{\log\psi(t)}{t} = x$. *Then, for every $\sigma\in\Delta(\Psi_1)$, there is $\tau\in\Sigma_2$ such that*

$$\limsup_{t\to\infty} g_t(\sigma,\tau)\le(\text{cav }U)(x).$$

Proof: Fix player 1's strategy $\sigma\in\Delta(\Psi_1)$. For the purpose of the payoff calculation, identify σ with its equivalent behavioral strategy. Define player 2's strategy as follows. At each stage t, and at each history $h\in H_{t-1}$, $\tau(h)$ minimizes player 1's stage payoff, that is,

$$E_{\sigma,\tau}\big[g(a_t)\big|h\big] = \min_{b\in B} E_{\sigma(h)}\big[g(a,b)\big].$$

Let $X_1, X_2, \cdots$ be the sequence of random actions induced by (σ, τ). Let $H(X_t|h)$ be the entropy of X_t given that a history h is realized. Note that, conditional on the history h, the entropy of player 1's mixed action at stage t is $H(X_t|h)$. Hence, by the definitions of U, cav U, and τ, we have

$$E_{\sigma,\tau}[g(X_t)|h] \leq U(H(X_t|h)) \leq (\text{cav } U)(H(X_t|h)).$$

Taking the expectation, we have

$$E_{\sigma,\tau}[g(X_t)] \leq E_{\sigma,\tau}[(\text{cav } U)(H(X_t|h))] \leq (\text{cav } U)(E_{\sigma,\tau}[h(X_t|h)])$$

where the second inequality follows from the concavity of cav U and Jensen's inequality. Summing over $t = 1, \cdots, T$ we have

$$\begin{aligned} g_T(\sigma, \tau) &= \frac{1}{T}\sum_{t=1}^{T} E_{\sigma,\tau}[g(X_t)] \\ &\leq \frac{1}{T}\sum_{t=1}^{T}(\text{cav } U)(E_{\sigma,\tau}[H(X_t|h)]) \\ &\leq (\text{cav } U)\left(\frac{1}{T}\sum_{t=1}^{T} E_{\sigma,\tau}[H(X_t|h)]\right) \\ &= (\text{cav } U)\left(\frac{1}{T}\sum_{t=1}^{T} H(X_t|X_1, \cdots, X_{t-1})\right) \\ &= (\text{cav } U)\left(\frac{1}{T}H(X_1, \cdots, X_T)\right) \\ &\leq (\text{cav } U)\left(\frac{\psi_1(T)}{T}\right). \end{aligned}$$

The second inequality follows from Jensen's inequality. The second and the third equalities follow from the definition of conditional entropy and Lemma 2, respectively. The last inequality follows from Lemma 3. Since $\limsup_{T\to\infty}(\psi_1(T)/T) = x$, the last term in the above inequality is at most x for all sufficiently large T. This completes the proof. **Q.E.D.**

As in Proposition 2, whether player 1 can achieve $(\text{cav } U)(x)$ or not depends on what strategies are available to him.

4. Concluding Remarks

In this section we briefly discuss growing strategy sets arising from growing complexity bounds. Let $m_1, m_2, \ldots$ be a nondecreasing sequence of positive integers. Let Ψ_i be the set of player i's strategies such that the strategies belonging to the set $\Psi_i(t)$, as defined in Section 2, are all implementable by automata with at most m_t states. It is easily seen that, if $m_{t'} = m_{t'+1}$ for some t', then $\Psi_i(t) = \Psi_i(t')$ for all $t > t'$. Thus in order to obtain a growing strategy set with this method, one must allow $m_t \geq t$ for all t. That is, the number of states must grow at least as fast as the number of repetitions. A similar conclusion applies to bounded recall strategies since any such strategy can be represented by finite automata. See Neyman (1997).

Perhaps closest in spirit to this paper is O'Connell and Stearns (1999). In their model, player 1 chooses a set of K pure strategies which can be randomized. The number K is exogenously given. There is no other restriction on player 1's strategy set, or, choice of strategy set. Player 2, whose strategies are not restricted, is informed of the set chosen by player 1. Thereafter they start playing an undiscounted finitely repeated game. This is a fairly general set up for studying strategic complexity in repeated games. For whenever one limits a complexity of strategy such as the number of states of finite automata or the length of recall, one puts a bound on the number of possible pure strategies that conform to the restriction. They address important issues mentioned in the third paragraph in Section 1: the amount of memory and time needed to implement/execute strategies. They explicitly present an algorithm to compute an optimal set of K pure strategies for player 1 and show that the algorithm runs in a polynomial time in K. They also show that, for $2 \times k$ games, one can encode the optimal set almost optimally using approximately $\log_2 K$ bits, and, one can execute each pure strategy in an optimal set by spending a fixed amount of time at each stage and a fixed amount of memory during the entire game which is independent of K and of the number of repetitions. They also provide upper and lower bounds, with respect to K, on the value of the repeated game.

Appendix

We rephrase the lemma by setting $\xi_t = \log_2 \psi_1(t)$.

Lemma 1 *Let $\{\xi_t\}_{t=1}^{\infty}$ be an increasing sequence of positive integers with the property $\xi_t/t \to 0$ as $t \to \infty$. Then there is an increasing sequence of positive integers $\{t_k\}_{k=0}^{\infty}$ such that*

(A) $\dfrac{t_{k+1} - t_k}{t_k} \to 0$ *as $k \to \infty$, and*

(B) $\dfrac{\xi_{t_{k+1}}}{t_{k+1} - t_k} \to 0$ *as $k \to \infty$.*

Proof: Fix an $\alpha \in (0,1)$. Let $f : \mathbb{N} \to \mathbb{R}_{++}$ be the smallest nondecreasing function such that

$$f(t) \geq \left(\frac{t}{\xi_t}\right)^{\alpha}.$$

More precisely, $f(1) = \left(\frac{1}{\xi_1}\right)^{\alpha}$ and $f(t) = \max\left\{f(t-1), \left(\frac{t}{\xi_t}\right)^{\alpha}\right\}$. It is easily verified that, under the hypothesis of this lemma, this function has the following properties. (In fact, (c) implies (b).)

(a) $f(t) \to \infty$ as $t \to \infty$,

(b) $\dfrac{t}{f(t)} \to \infty$ as $t \to \infty$,

(c) $\dfrac{f(t)\xi_t}{t} \to 0$ as $t \to \infty$.

Choose a $\beta \in (0,1)$. Set $t_1 > 1$ arbitrarily and define by induction

$$t_{k+1} = t_k + \left[\frac{t_k}{f(t_k)^{\beta}}\right]$$

where $[x]$ denotes the integer part of x. Property (b) of f ensures that $\{t_k\}_{k=0}^{\infty}$ is an increasing sequence. Property (a) implies (A) because

$$\frac{t_{k+1} - t_k}{t_k} \leq \frac{1}{f(t_k)^{\beta}} \to 0 \quad \text{as } k \to \infty.$$

To show that the sequence $\{t_k\}_{k=0}^{\infty}$ satisfies (B), first note that, by (c),

$$\xi_{t_{k+1}} < \frac{t_{k+1}}{f(t_{k+1})}$$

holds for all sufficiently large k. Since $t_{k+1} - t_k = \left[\frac{t_k}{f(t_k)^{\beta}}\right] > \frac{t_k}{f(t_k)^{\beta}} - 1$, we have

$$\frac{\xi_{t_{k+1}}}{t_{k+1} - t_k} < \frac{\dfrac{t_{k+1}}{f(t_{k+1})}}{\dfrac{t_k}{f(t_k)^{\beta}} - 1} = \frac{t_{k+1}}{t_k - f(t_k)^{\beta}} \cdot \frac{f(t_k)^{\beta}}{f(t_{k+1})}$$

holds for all sufficiently large k. Note that

$$\begin{aligned}\frac{t_{k+1}}{t_k - f(t_k)^\beta} &< \frac{t_k + \dfrac{t_k}{f(t_k)^\beta}}{t_k - f(t_k)^\beta} \\ &= \frac{t_k}{t_k - f(t_k)^\beta}\left(1 + \frac{1}{f(t_k)^\beta}\right) \\ &= \frac{\dfrac{t_k}{f(t_k)^\beta}}{\dfrac{t_k}{f(t_k)^\beta} - 1}\left(1 + \frac{1}{f(t_k)^\beta}\right) \to 1 \quad \text{as } k \to \infty,\end{aligned}$$

by (a) and (b), and

$$\frac{f(t_k)^\beta}{f(t_{k+1})} \leq \frac{f(t_k)^\beta}{f(t_k)} = \frac{1}{f(t_k)^{1-\beta}} \to 0 \quad \text{as } k \to \infty$$

by (a) and since $f(t_{k+1}) \geq f(t_k)$. This completes the proof. **Q.E.D.**

References

Abreu, D., and A. Rubinstein (1988): "The Structure of Nash Equilibrium in Repeated Games with Finite Automata," *Econometrica*, 56(6), 1259–1281.

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

Ben-Porath, E. (1993): "Repeated Games with Finite Automata," *Journal of Economic Theory*, 59(1), 17–32.

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

Gilboa, I. (1988): "The Complexity of Computing Best-Response Automata in Repeated Games," *Journal of Economic Theory*, 45(2), 342–352.

Kalai, E. (1990): "Bounded Rationality and Complexity in Repeated Games," in *Game Theory and Applications*, ed. by A. N. Tatsuro Ichiishi, and Y. Tauman, pp. 131 – 157. Academic Press.

Neyman, A. (1997): "Cooperation, Repetition, and Automata," in *Cooperation: Game-Theoretic Approaches*, ed. by S. H. Mas-Colell, and A., vol. 155 of *NATO ASI-Seies F*, pp. 233–255. Springer-Verlag, New York.

Neyman, A., and D. Okada (2000): "Two-Person Repeated Games with Finite Automata," *International Journal of Game Theory*, 29(3), 309 – 325.

O'Connell, T. C., and R. E. Stearns (1999): "On Finite Strategy Sets for Finitely Repeated Zero-sum Games," Technical report, Department of Computer Science, SUNY ay Albany.

Papadimitriou, C. H. (1992): "On Players with a Bounded Number of States," *Games and Economic Behavior*, 4(1), 122–131.

Rubinstein, A. (1986): "Finite Automata Play the Repeated Prisoner's Dilemma," *Journal of Economic Theory*, 39(1), 83–96.

Stearns, R. E. (1997): "Memory-bounded Game-playing Computing Devices," .