PSEUDORANDOM PROCESSES: ENTROPY AND AUTOMATA*

Penélope Hernández and Amparo Urbano**

WP-AD 2001-22

Correspondence to Penélope Hernández: THEMA Université de Cergy-Pontoise, U.F.R. d'Economie et Gestion. 33 Boulevard du Port, 95011 Cergy-Pontoise Cedex, France. Tel. 33 1 34 25 61 70. Fax: 33 1 43 25 62 33. e-mail: penelope.hernandez@eco.u-cergy.fr .

Editor: Instituto Valenciano de Investigaciones Económicas, S.A.

First Edition July 2001.

Depósito Legal: V-3417-2001

---

* The authors thank Olivier Gossner for helpful comments and partial support by DGICYT under project PB95-1074..

** P. Hernández: THEMA, University of Cergy-Pontoise; A. Urbano: Univeristy of Valencia.

PSEUDORANDOM PROCESSES: ENTROPY AND AUTOMATA

Penélope Hernández and Amparo Urbano

A B S T R A C T

This paper studies implementation of cooperative payoffs in finitely repeated games when players implement their strategies by finite automata of big sizes. Specifically, we analyze how much we have to depart from fully rational behavior to achieve the Folk Theorem payoffs, i.e., which are the maximum bounds on automata complexity which yield cooperative behavior in long but not infinite interactions. To this end we present a new approach to the implementation of the mixed strategy equilibrium paths leading to cooperation. The novelty is to offer a new construction of the set of the pure strategies which belong to the mixed strategy equilibrium. Thus, we consider the subset of strategies which is characterized by both the complexity of the finite automata and the entropy associated to the underlying coordination process. The equilibrium play consists of both a communication phase and the play of a cycle which depends on the chosen message. The communication set is designed by tools of Information Theory. Moreover, the characterization of this set is given by the complexity of the weaker player that implements the equilibrium play. We offer a domain of definition of the smallest automaton which includes previous domains in the literature.

KEYWORDS: Complexity; Cooperation; Entropy; Automata; Repeated Games.

1 INTRODUCTION

The message of the Folk Theorem and several other results (Aumann,1960, 1981; Rubinstein, 1979, 1980) is that cooperative behavior may emerge in non-cooperative situations when the nature of interactions is long term. However, in the ¯nite repetition of most of these situations, all equilibria lead to the non-cooperative outcome of each stage. This is in clear contrast, for instance, with common observations in the experiments involving ¯nite repetitions of the prisoner's dilemma, where participants achieve some mode of cooperation. On the other hand, if players are restricted to choose automata that are too small to count the number of stages of the repeated game then both players choosing "a cooperating automata" is a Nash equilibrium. One may therefore think of "bounded rationality" or bounded ability to handle strategic complexities, as a way to solve the prisoners's dilemma paradox. It is surprising that even if the players can choose large automata, then they can get arbitrarily close to the cooperative payo®s provided that they are allowed to randomize in their choices of automata (Neyman, 1985).

A great deal of attention has been paid recently to repeated games with bounded complexity. Speci¯cally, there are several papers in the repeated games literature, which study the conditions under which the set of feasible and rational payo®s are equilibrium outcomes, when there are bounds (possibly very large) to the number of strategies that players may use. In the context of strategies implemented by ¯nite automata, these bounds are given by the complexity of the players's automata which implement the equilibrium (see Rubinstein, 1986; Abreu and Rubinstein, 1988; Neyman, 1998; Papadimitriou and Yannakakis, 1994; Neyman and Okada, 1997, among others).

The present paper studies implementation of cooperative payo®s in ¯nitely repeated games when players implement their strategies by ¯nite automata whose sizes are exogenously given; the motivation being to justify the empirical regularity of such a cooperative behavior (Axelrod, 1980). Speci¯cally, we analyze how much we have to depart from fully rational behavior to achieve the Folk Theorem payo®s, i.e., which are the maximum bounds on automata complexity which yield cooperative behavior in long but not in¯nite interactions.

Building on the work of Neyman (1998), we improve existing results in the literature (Neyman, 1998; Neyman and Okada, 1997, Zemel, 1989 and Papadimitriou and Yannakakis, 1994) by taking a di®erent approach and focusing on the complexity of mixed strategy equilibrium paths leading to the Folk Theorem payo®s. Given that in our setting players not only choose large automata but also randomize among them, the equilibrium is a mixture of such choices. Each player's pure strategy determines a possible play and the set of pure strategies which belong to the support of the mixed strategy determines the set of possible plays. Thus, the ¯rst problem to solve is to choose the subset of a player's pure strategies which generates the mixed strategy and, in turn, the set of possible plays. There are many of these subsets since the number of ¯nite automata is an exponential number of a player's complexity. Then, the second problem is to select a right subset such that the selection of the speci¯c equilibrium play satis¯es good properties of complexity and e±ciency. This implementation of a speci¯c mixed strategy equilibrium is through a coordination process which yields a payo® close enough to any of the ones belonging to the set of feasible and rational payo®s. Thus, the complexity of such a process determines that of the equilibrium path and we look for processes with satisfy both the equilibrium complexity bounds and maximal e±ciency (closer to the targeted payo®).

We characterize the above properties of a coordination scheme by its informational features. Speci¯cally, the complexity of the process is related with the associated entropy, which captures, from an Information Theory viewpoint, the cardinality of the sequences belonging to a particular set with some good properties (Typical Set). The number of equilibrium plays depends on the cardinality of the selected sequences and thus on its associated complexity. Processes with low entropy translates to small cardinalities and hence to small number of plays while processes with the maximal entropy imply a large number of sequences and then a large number of equilibrium plays. On the other hand, e±ciency of the process is translated to optimal codi¯cation schemes which produces "short" coordination processes.

Speci¯cally, to construct an equilibrium play the coordination process consists of both a communication phase and the play of a cycle, whose last part, the veri¯cation play, depends on the speci¯c chosen message. Since equilibrium plays are in a one-to-one relationship with the set of communication messages, the design of this set (with respect to the cycle's play) is crucial for the construction. Then our equilibrium conditions are determined by the inter-play communication scheme. We consider the subset of pure strategies which is characterized by both the complexity of the ¯nite automata and the entropy associated to the communication and the veri¯cation phenomena.

The novelty of the paper is to present a new approach to construct mixed strategy equilibria with ¯nite automata. This new viewpoint allows us to characterize the set of pure strategies which belong to the support of the equilibrium mixed strategies. Moreover since we o®er the less restrictive equilibrium conditions this set cannot be improved upon. The previous literature (Neyman, 1998; Neyman and Okada, 1997) give restrictions on the whole set of pure strategies. In our approach the restrictions are given on each pure strategy and thus we are able to characterize each equilibrium automaton. To impose such constraints we make use of the notion of entropy as a measure of the messages' uncertainty of our communication scheme and also as a way to measure their associated complexity. This construction also allows us to relate our communication scheme under strategic complexity (¯nite automata) with those in repeated games with communication and unbounded rationality (Lehrer, 1996; Lehrer and Sorin, 1997; Forges, 1990; Gossner, 1998; Gossner and Viellie, 1999 and Ben-Porath, 1998 among others).

A related line of research addresses the same question under speci¯c restrictions of the players' set of strategies by an exogenous bound: one of the player's strategies are restricted to those that have strategic entropy less than a prespeci¯ed bound; where a player's strategic entropy refers to the uncertainty of his mixed strategy relative to the other player's strategy (see, Neyman and Okada, 1999 and 2000).

Since punishments in the ¯nitely repeated game are in pure strategies, the main result of the paper is given in terms of the weaker player's complexity. The domain of de¯nition of this player's complexity includes all the others bounds already o®ered in the literature. This improvement is achieved by the approach that we follow: to understand the problem of constructing the set of pure strategies as a codi¯cation problem where what is being codi¯ed is the complexity of the player with the smallest automaton (the "weaker player").

Although we use the concept of entropy as a technical tool, it also gives us a much deeper understanding of the connection between communication and codi¯cation issues. The complexity costs associated to the veri¯cation play are measured in terms of the weaker player's complexity, since his automaton's capacity determines the number of plays. Moreover, since, this player's complexity bounds are related to the "-approximation to the targeted equilibrium, there are also e±ciency costs associated to the veri¯cation play. However, the communication costs are just measured in terms of the players' payo®s (in the "-approximation to the targeted equilibrium) since in our construction the weaker player's automaton need not additional states to process the information. In this framework, the entropy notion is useful to characterize both the complexity and the e±ciency costs associated to the veri¯cation play and the communication phase. On one hand, the entropy of sequences of i.i.d. random variables give us a good measure of the complexity of such sequences. On the other, the optimal (shortest) codi¯cation of the veri¯cation sequences produces the shortest communication phase, which, in turn, is bounded by the entropy of the random variable associated to the veri¯cation sequences. Thus, the entropy measures both the complexity and the efficiency costs associated to the equilibrium play.

The paper is organized as follows. Section 2 sets up the one-shot game, the finitely repeated game and the finite automata framework and some known results in play complexity are stated. Section 3 offers the main result, while section 4 presents the scheme of the play. The analysis of sequences and codification schemes is undertaken in section 5, where some tools of Information Theory are presented and a first result of our construction, stated in section 3, is proven. Section 6 is devoted to prove the main result. To this end, the constructions of (a) the set of messages, (b) the equilibrium play, and (c) the players' automata, are offered and it is checked that they satisfy the equilibrium conditions. Concluding remarks close the paper.

2 PRELIMINARIES

2.1 The one-shot game

Let $G = (\{1,2\}, (A^i)_{i \in \{1,2\}}, (r^i)_{i \in \{1,2\}})$ be a game where $\{1,2\}$ is the set of player. A^i is a finite set of actions for player i (or pure strategies of player i) and $r^i : A = A^1 \times A^2 \longrightarrow \mathbb{R}$ is the payoff function of player i.

We denote by $u_i(G)$ the individual rational payoff of player i in pure strategies, i.e., $u_i(G) = \min \max r^i(a^i, a^{-i})$ where the max ranges over all pure strategies of player i, and the min ranges over all pure strategies of player $3 - i$. For any finite set B we denote we denote by $\Delta(B)$ the set of all probability distributions on B. An equilibrium of G is a pair $\sigma = (\sigma^1, \sigma^2) \in \Delta(A^1) \times \Delta(A^2)$ such that for every i and any strategy of player i, $\tau^i \in A^i$; $r^i(\tau^i, \sigma^{-i}) \quad r^i(\sigma^1, \sigma^2)$; where $r(\sigma) = E_\sigma(\ r(a^i, a^{-i}))$ If σ is an equilibrium, the vector payoff $r(\sigma)$ is called an equilibrium payoff.

We denote by $E(G)$ the set of all equilibrium payoffs of G.

2.2 The finitely repeated game G^T

>From G we define a new game in strategic form G^T which models a sequence of T plays of G, called stages. By choosing actions at stage t, players are informed of actions chosen in previous stages of the game. Formally, let $H_t; t = 1, \ldots, T$, be the Cartesian product of A by itself $t - 1$ times, i.e.: $H_t = A^{t-1}$, with the common set theoretic identification $A^0 = \emptyset$, and let $H = \cup_{t \geq 0} H_t$. A pure strategy σ^i for player i in G^T is a mapping from H to A^i; $\sigma^i : H \longrightarrow A^i$.

Obviously, H is a disjoint union of H_t; $t = 1, \ldots, T$ and $\sigma^i_t : H_t \to A^i$ as the restriction of σ^i to H_t. We denote the set of all pure strategies of player i in G^T by $\Sigma^i(T)$. Any 2-tuple $\sigma = (\sigma^1, \sigma^2) \in \times \Sigma^i(T)$ of pure strategies induces a play $\omega(\sigma) = (\omega_1(\sigma), \ldots, \omega_T(\sigma))$ with $\omega_t(\sigma) = (\omega^1_t(\sigma), \omega^2_t(\sigma))$ defined by $\omega_1(\sigma) = (\sigma^1(\alpha), \sigma^2(\alpha)) = \sigma(\alpha)$ and by the induction relation $\omega^i_t(\sigma) = \sigma^i(\omega_1(\sigma), \ldots, \omega_{t-1}(\sigma)) = \sigma^i_t(\omega_1(\sigma), \ldots, \omega_{t-1}(\sigma))$.

Let $r_T(\sigma) = \frac{r(\omega_1(\sigma)) + \ldots + r(\omega_T(\sigma))}{T}$ be the average vector payoff during the first T stages induced by the strategy profile σ.

Two strategies σ^i and τ^i of player i in G^T are called equivalent if for every $3 - \{i\}$ tuple of pure strategies σ^{-i}; $\omega_t(\sigma^i, \sigma^{-i}) = \omega_t(\tau^i, \sigma^{-i})$ for every $1 \le t \le T$.

An equivalence class of pure strategies is called a reduced strategy.

2.3 Finitely repeated games played by finite automata

A finite automaton for player i that implements the strategy profile σ in G^T is a tuple $M^i = < Q^i, q^i_0, f^i, g^i >$, where:

- Q^i is the set of states
- q^i_0 is the initial state
- f^i is the action function, $f^i : Q^i \to A^i$
- g^i is the transition function from state to state $g^i : Q^i \times A^{-i} \to Q^i$

The size of a finite automaton is the number of its states, $|Q|$.

We define a new game in strategic form $G^T(m_1, m_2)$ which denotes the T stage repeated version of G, with the average payoff as evaluation criterion and with all the finite automata of size m_i as the pure strategies of player i, $i = 1, 2$. Let $\Sigma^i(T, m_i)$ be the set of pure strategies in $G^T(m_1, m_2)$ that are induced by an automaton of size m_i.

A finite automaton for player i can be viewed as a prescription for this player to choose his action in each stage of the repeated game. If at state q the other player chooses the action tuple a^{-i}, then the automaton's next state is $g^i(q, a^{-i})$ and the action to be taken at stage 1 is $f^i(q^i)$. The action in stage 2 is $f^i(g^i(q^i, a^{-i}_1))$ where a^{-i}_1 is the action taken by the other players in stage 1. More generally, define inductively,

$g^i(q; b_1, \ldots, b_t) = g^i(g^i(q; b_1, \ldots, b_{t-1}); b_t)$,

where $a_j^{-i} \in A^{-i}$, the action prescribed by the automaton for player i at stage j is

$f^i(g^i(q^i; a_1^{-1}, \ldots, a_{t-1}^{-i}))$.

For every automaton M for player i, define a strategy σ^i_M in G^T by

$\sigma^i_M(a_1, \ldots, a_{t-1}) = f^i(g^i(q^i; a_1^{-i}, \ldots, a_{t-1}^{-i}))$. A strategy σ^i for player i in G^T is implementable by the automaton M if σ^i is equivalent to σ^i_M i.e.: for every $\tau \in \Sigma^2(T)$, $\omega(\sigma^i; \tau) = \omega(\sigma^i_M; \tau)$.

2.4 Notation

Let $G = (\{1, 2\}; A; r)$ be the two-player game in strategic form defined in section 1.1. Denote by K twice the largest absolute value of a payoff in the game G. Thus, $|r^i(a) - r^i(b)| \leq K$ for every $a, b \in A$ and $i = 1, 2$.

Given the set X, $co(X)$ means the convex hull of X.

Recall that $u_i(G)$ is the individual rational payoff of player i in pure strategies and denote by $F(G)$ the set of feasible and rational payoffs of G i.e., the set of payoff profiles x such that $x \in co(r(A))$ and $x^i > u^i(G)$

Denote by $[x]$ the integer part of a real number x.

The number of elements of a set X is denoted by $|X|$.

Let f be a real function then:

f grows polynomially is denoted by $f = O(p)$ for some polynomial p i.e.: $f = n^{O(1)}$.

f grows subexponentially is denoted by $f = o(2^{n^2})$; i.e.: $\forall \epsilon > 0$ $\frac{f}{2^{n^2}} < \epsilon$ for all sufficiently large n.

2.5 Play complexity

The main results in play complexity are those given by Kalai and Stanford (1988) and Neyman (1998). We present here the definitions of the complexity of a strategy in G^∞ and then the definitions in G^T.

First, a finite sequence of actions $(a_1, \ldots, a_t)$ is compatible with a pure strategy σ^i if for every $1 \leq s \leq t$, $\sigma^i(a_1, \ldots, a_{s-1}) = a^i_s$. Let $A^n(\sigma^i)$ be the set of all sequences of actions of length n that are compatible with σ. Consider for any sequence of actions $(a_1, \ldots, a_t)$ and a pure strategy σ^i the new strategy $(\sigma^i \mid a_1, \ldots, a_t)$ in G^∞ given by

$(\sigma^i \mid a_1, \ldots, a_s)(b_1, \ldots, b_\ell) = \sigma^i(a_1, \ldots, a_s, b_1, \ldots, b_\ell).$

The number of different reduced strategies that are induced by a given pure strategy σ^i of player i in $G^T(m_1, m_2)$ and all σ^i-compatible sequences of actions of length n, for all n, provides with a first measure of the complexity of σ, $\text{comp}_1(\sigma^i)$. This definition has a natural extension to the finitely repeated game, G^T. Let $(\sigma_t)_{t=1}^T$ where $\sigma \in \Sigma^i(T)$ and define $\text{comp}_2(\sigma) = \min\{\text{comp}_2(\tau) : \tau \in \Sigma^i \text{ and } \forall t, 1 \le t \le T, \sigma_t = \tau_t\}$.

Second, define $\text{comp}_2(\sigma^i)$ as the size of smallest automaton that implements σ^i.

The two above definitions turn out to be equivalent (Neyman, 1998, proposition 2), $\text{comp}_1(\sigma^i) = \text{comp}_2(\sigma^i)$.

We shall often need bounds on the complexity of strategies that induce a given play. Hence, for a play ω, define player i's complexity of ω, $\text{comp}^i(\omega)$, as the smallest complexity of a strategy σ^i of player i which is compatible with ω:

$\text{comp}^i(\omega) = \inf\{\text{comp}^i(\sigma) : \sigma \in \Sigma^i \text{ is compatible with } \omega\}.$

Let Q be a set of plays. A pure strategy σ^i of player i is conformable to Q if it is compatible with any $\omega \in Q$. The complexity of player i of a set of plays Q is defined as the smallest complexity of a strategy σ^i of player i that is comformable to Q.

$\text{comp}^i(Q) = \inf\{\text{comp}^i(\sigma) : \sigma \in \Sigma^i \text{ is comformable to } Q\}$

The following lemmata, proved in Neyman (1998), provide bounds of the complexity of some particular plays which will be used in the proof of the main result. The first result provides with an upper bound of the complexity of a sequence of actions of length t.

Lemma 1 Let $a = (a_1, \ldots, a_t) \in A^t$. Then $\text{comp}^i(a) \le t$.

Let $a = (a_1, \ldots, a_t) \in A^t$ and $b = (b_1, \ldots, b_s) \in A^s$, and denote by $a + b = (a_1, \ldots, a_t, b_1, \ldots, b_s) \in A^{t+s}$ the concatenation of two histories. The second lemma states the complexity bound of such a concatenation.

Lemma 2 Let $a = (a_1, \ldots, a_t) \in A^t$ and $b = (b_1, \ldots, b_s) \in A^s$. Then $\text{comp}^i(a+b) \le \max(\text{comp}^i(a), \text{com}$

For $a = (a_1, \ldots, a_t) \in A^t$ and a positive integer d, define $d * a$ by induction on d: $1 * a = a$, and $(d+1) * a = d * a + a$.

The complexity of a sequence of actions that changes in the last stage is stated next.

Lemma 3 Let $a = (a_1, \ldots, a_t) \in A^t$ with $a_1 = a_2 = \ldots = a_{t-1}$ and $a^i_{t-1} \neq a^i_t$. Then $\text{comp}^i(a) = t$.

Let $a = (a_1, \dots, a_t) \in A^t$ and $b = (b_1, \dots, b_s) \in A^s$, and s with $\min(t, s) \geq s - 1$ then define $a =_s b$ if $a_r = b_r$ for every $r < s$.

Consider two finite sequences of actions a and b such that the first action for player i in a and b is different.: $a^i_1 \neq b^i_1$. The next lemma presents a lower bound for the complexity of a play that consists of a cycle $(t \cdot a + b)$ repeated d times and there is a deviation of player i after the $t \cdot a$ action pairs on. This result is useful to measure the complexity needed to deviate from a given cycle play.

Lemma 4 Let $a = (a_1, \dots, a_k) \in A^k$ and $b = (b_1, \dots, b_n) \in A^n$ with $a^i_1 \neq b^i_1$, $t \geq 0$ and $d \geq 1$. Assume that $\omega = (\omega_1, \dots, \omega_s) \in A^s$ with $(d - 1)(tk + n) + tk + 1 < s \leq (d + 1)(tk + n)$ and $d \cdot (t \cdot a + b) =_s \omega$ and $((d + 1) \cdot (t \cdot a + b))^i_s \neq \omega^i_s$. Then $comp^i(\omega) \geq d(t + 1)$.

Let $f : A^1 \to A^2$ be a 1-1 function and let $a = (a_1, \dots, a_n) \in A^n$ be a play with $a^2_t = f(a^1_t)$ for every $1 \leq t \leq n$, then a is called a coordinated play. In case of a coordinated play, the number of equivalence classes induced by a strategy σ^i conformable with ω is exactly the length of the play. We need a complexity lower bound for a play that consists of a coordinated periodic play. This is stated next.

Lemma 5 Let $a = (a_1, \dots, a_n) \in A^n$ be a coordinated play, $b \in A$ with $b^1 \neq a^1_1$, and $d \in \mathbb{N}$. Then $comp^i(d \cdot a + b) \geq (d - 1)n + 1$.

Finally, the next result states a lower bound for a play in terms of the number of consecutive action of player i.

Lemma 6 Let $a = (a_1, \dots, a_k)$ be a play. Let $B^i \subset A^i$ be a nonempty subset of the actions of player i. Assume that $k : B^i \to \mathbb{N}$ is such that for every $b^i \in B^i$ there is $s = s(b^i) < t - k(b^i)$ with $a_{s+1} = \dots = a_{s+k(b^i)} = b^i$ and $a^i_{s+1} \neq a^i_{s+k(b^i)+1}$. Then $comp(a) \geq \sum_{a^i \in B^i} k(a^i)$.

By the definition of the complexity of a strategy, the above lemmata are proved by counting the number of different strategies obtained when all possible plays ω are induced. Each induced strategy generates an equivalence class of strategies and then the number of these equivalence classes coincides with the number of the automaton states. The overall sketch of the proofs is:

1. Let σ be a strategy compatible with ω.
2. Consider the set of strategies $\{(\sigma \mid \omega_t) \mid t \in \mathbb{N}\}$ where $(\sigma \mid \omega_t)$ denotes the strategy induced by the play ω of length t

3. For each strategy consider the number of reduced strategies with the concatenation of histories.

This last number is the cardinality of the set f(¾ j !$_t$) j t 2 Ng and thus comp(¾) is obtained.

3 MAIN RESULT

The main result establishes the existence of an equilibrium payo® of $G^T(m_1; m_2)$ which is " ¡ closed to a feasible and rational payo®. In the context of ¯nitely repeated games, deviations in the last stages could be precluded if players did not know the end of the game. This may be achieved if players implemented their strategies by playing with ¯nite automata which cannot count until the last stage of the game. On the contrary, player i will deviate if he is able to implement cycles of length at least the number of the repetitions. Hence, if players answered to di®erent plays of length smaller than the number of repetitions then they could spend their capacity and not be able to count until the end of the game. In this way, a player can ¯ll up the rival's complexity by requiring him to conform with distinct plays of su±ciently large length, i.e., approximately O("T):

To ¯ll up the complexity of the weaker player, the stronger player (the one with the biggest automaton) speci¯es the set of plays by means of a set of messages to be sent in the communication phase. The complexity of the set of plays is determined by the complexity of such a weaker player and the di®erence among the distinct plays is a small portion of each play (the veri¯cation play). Thus, what is being determined in each message is the above veri¯cation play. Hence, to design the set of plays can be understood as a codi¯cation problem where what is being coded is the weaker player's complexity.

Similarly to the existing literature (Neyman, 1998) we o®er the equilibrium conditions in terms of the complexity of the smallest automaton which implements the equilibrium play. The main di®erence is that both the upper and the lower bounds that we achieve include previous bound's domains. This is due to our optimal construction of the set of veri¯cation sequences and the associated communication scheme. We characterize the above set by selecting a subset of sequences over a ¯nite alphabet. Since messages are a codi¯cation of plays we follow the shortest codi¯cation in order to construct the communication phase[1]. We state informally

[1] This is in clear contrast with Neyman (1998) who does not construct the shortest communication phase given his set of veri¯cation sequences.

this first result which is needed to show that under our construction the sets of verification and communication sequences are the optimal sets to codify the weaker player's complexity. The formal statement of this result is presented in section 5 where we introduce the tools of Information theory which are needed to prove it. Then, Theorem 1 establishes the existence of an equilibrium payoff of $G^T(m_1; m_2)$ which is $\varepsilon -$ closed to a feasible and rational payoff under automaton bounds which are the best in the literature.

Result 1: The set of messages for the communication phase coincides with the set of sequences for the verification play, i.e. an optimal codification map is the identity. In other words, given our set of verification sequences there is not a shortest codification scheme.

The main result below presents the equilibrium conditions to reach a feasible and rational payoff in a finitely repeated game when players implement their strategies by means of finite automata.

Theorem 1 Let $G = (\{1; 2\}; A; r)$ be a two person game in strategic form. Then for every ε sufficiently small, there exist positive integers T_0 and m_0, such that if $T \geq T_0$, and $x \in co(r(A))$ with $x^i > u^i(G)$ and $m_0 \leq \min\{m_1, m_2\} \leq \exp(\varepsilon T)$ and $\max\{m_1, m_2\} > T$ then there exists $y \in E(G^T(m_1; m_2))$ with $|y^i - x^i| < \varepsilon$:

Theorem 1 will follow from conditions on: 1) a feasible payoff $x \in co(r(A))$; 2) a positive constant $\varepsilon > 0$; 3) the number of repetitions T, and 4) the bounds of the automata sizes, $m_1; m_2$, that guarantee the existence of an equilibrium payoff y of the game $G^T(m_1; m_2)$ that is ε-close to x.

To see that our bounds include previous bound's domains we include here Neyman's result:

Theorem (Neyman, 1998): Let $G = (\{1; 2\}; A; r)$ be a two person game in strategic form. Then for every ε sufficiently small, there exist positive integers T_0 and m_0, such that if $T \geq T_0$, and $x \in co(r(A))$ with $x^i > u^i(G)$ and $m_0 \leq \min\{m_1; m_2\} \leq \exp(\varepsilon^3 T)$ and $\max\{m_1, m_2\} > T$ then there exists $y \in E(G^T(m_1; m_2))$ with $|y^i - x^i| < \varepsilon$:

One of the conditions of our theorem is stated by means of the inequalities $m_i \geq m_0$ where m_0 is sufficiently large. Another condition require the bound of one or both size to be subexponential in the number of repetitions, i.e., a condition that asserts that $(\log m_i)/T$ is sufficiently small. The characterization of this condition is related with the codification schemes to be studied in Section 5.

4 THE SCHEME OF THE PLAY

In this section we present the scheme of the play to reach a feasible and rational payo® x in a ¯nitely repeated game. The plays along the equilibrium path are divided into a communication phase followed by a play phase.

Assume without loss of generality that m_1 m_2: Knowing player 1's complexity, player 2 determines a precise number of plays from which one is selected and sent to player 1 in the communication phase. This signal speci¯es one of the ¯nitely many plays of the repeated game to be played in the play phase and it uses two actions that we label 0 and 1. Player 2 plays a mixed strategy during this phase and player 1 responds properly to any message. The action of Player 1 is independent of the message (signal) sent by player 2. Since player 2 proposes the plays, messages have to be independent of the associated payo®s to each of them. We reach this independence by means of balanced sequences, i.e., sequences with the same number of zeros and ones. The speci¯cation of the set of messages and the correspondence with the set of plays is crucial in our construction, because we associate each message from the communication phase with a unique play in the play phase.

After the communication phase the equilibrium play enters into the play phase which consists of a cycle repeated along the play until T. The length of the cycle does not depend on the signal sent by player 2. Each one of the cycles has associated payo® approximately equal to the e±cient and rational payo® x. Thus, in any one of the proposed plays, player 1 has no incentive to deviate prior to the very last stages of the ¯nitely repeated game. The cycle has to parts: the veri¯cation play and the regular play. The regular play is common for every signal and it consists of a cycle of di®erent action pairs such that players reach a vector payo® " ¡ close to the e±cient and rational targeted payo® x.

Player 2 follows a veri¯cation play to check that player 1 has spent all his states following the play. It consists of a coordinated play with the identity as the function between A^1 and A^2, i.e., both players play the same actions. In words, both players follow a monitoring phase such that the sequence of actions can be understood as a coordination process which determines each pure strategy. The sequence of actions played in this phase is a sequence whose empirical distribution coincides with the uniform distribution and where the last element of the sequence is ¯xed.

The veri¯cation scheme is constructed such that it satis¯es three properties. First, it is balanced (the number of ones is equal to the number of zeros) to deter player 2's deviations by selecting the best payo® sequences. Second, this phase generates a payo® " close to x. Finally, player 2 ¯lls up player 1's capacity by generating enough pure strategies so that the number of remaining states is su±ciently small. In this way, player 1's deviations from the proposed play by counting up until the last stage of the game are avoided. For instance, player 1 could be able to select just one proposed play and deviate in the last stage of this play while repeating the cycle in all other proposed plays. Similarly, he could increase his own payo® by neglecting a subset of plays. Thus, the repetition of the cycle precludes sophisticated deviations by player 1.

There are two schemes that player 1 has to design to make a good use of his complexity. Player 1 needs all the plays in his automaton to follow the right play until T. There are many player 1's automata which could process the information sent by player 2. Given our automaton framework we minimize the information processing of player 1 by using the same states to process the signal and to follow the regular part of the di®erent cycles. However, this introduces a di±culty since these states of player 1's automaton admit both actions 0 and 1. Moreover, Player 1 uses one automaton with the minimal number of states for each play. The way to decrease this number is by reusing states for two di®erent actions. For instance, player 1 can use the same state to implement the action pairs (0; 0) and (0; 1) because for the action 0 he could accept both actions 0 and 1: This entails that there are deviations of player 2 that might be unpunished. If player 2 knew exactly the states that admit both actions, he could take advantage over them in future stages of the game. These deviations can only be undertaken by player 2 in the play phase, since the sequences from the communication phase are balanced and thus he is indi®erent among the messages. To avoid this problem player 1 uses a mixed strategy whose support consists of the minimal subset of pure strategies which are conformable with the proposed plays and such that it generates enough randomization to obscure the location of his reused states. Player 1's mixed strategy is constructed by a uniform distribution in this minimal subset.

Note that every player's behavior plays a di®erent role in the game. The signaling activity of player 2 has two purposes: how to coordinate and how to ¯ll up player 1's capacity. And these are the goals of the player 2's mixed strategy. On the contrary, player 1's role consists of supporting the \coordination" proposed by player 2 by means of a mixed strategy. To this end, player 1 builds a mechanism against player 2's undetectable deviations.

5 SEQUENCES AND CODIFICATION SCHEMES

We proceed to construct the set of verification sequences and the associated communication scheme. The key points of the construction are: 1) the characterization of such sequences by both their empirical distribution and they informational properties and 2) the design of the set of communication sequences through the optimal codification of the verification set. This approach produces our result 1 and clarifies the difference between previous constructions and ours.

Notice that in order to fill up the complexity of player 1, player 2 generates sufficiently many plays which player 1 has to conform with. The difference among them is given by the sequences of action pairs for the verification play because the regular play is common. Moreover, there is a map between each play and each message related to the corresponding verification. Hence, we look for the shortest way to construct messages associated to the verification play and to be sent in the communication phase, such that this last phase is also the shortest one.

To find a solution to this problem is equivalent to solving a codification problem in Information Theory, since the verification sequences have to be coded in the communication phase. To codify means to describe a phenomenon. The realization of this phenomenon can be viewed as the representation of a random variable. Then, a codification problem is just a one-to-one mapping (the source code) from a finite set (the range of a random variable or input) to another set of sequences of finite length (output sequences). What is important here is that the length of the output sequences is the shortest one with respect to the length (or probability) of the input sequences.

In our setting the set of verification sequences is the input set and the set of messages corresponds with the output set. We start with the set of balanced sequences of length k, whose cardinality is[2] about $O(2^{k-1})$ and which are the verification sequences. Our output set consists of finite length strings from the binary alphabet with the shortest length and again with the balancedness condition.

Solving the codification problem we obtain the set of messages for the communication phase. Our codification verifies that it is the shortest one and the output sequences are balanced. By tools of Information Theory we prove our result 1, i.e., that the trivial codification (the source code is the identity) is optimal in the sense that its expected length is minimum and then there is no code with shortest expected length that the identity. This result is due to the fact that

[2]See footnote 4 below.

the set of sequences for the verification play is designed in such a way that player 1's complexity (m_1) is bounded by an integer which is the cardinal of the smallest set of balanced sequences. If the above condition is not satisfied then there will exist non-trivial optimal source codes[3].

The formal details of our construction are presented next. We consider first deterministic sequences which satisfy some properties: they are balanced and the last component of each sequence is fixed. We use the method of types and the Type set to define these sequences. In second place, we analyze the information properties of these sequences by means of concepts such as entropy and the Kullback distance. This allows us to view the Type Set as the set of random sequences of a given entropy, even without knowing the actual random variable whose distribution is emulated by the deterministic sequence. Finally, we present the minimal codification of the Type Set with this alternative approach.

5.1 Deterministic Sequences: Type Set

Let $x_1, \ldots, x_n$ be a sequence of n symbols from an alphabet $\mathcal{X} = \left\{a_1, a_2, \ldots, a_{|\mathcal{X}|}\right\}$. We will use the notation x^n and x interchangeably to denote a sequence $x_1, x_2, \ldots, x_n$.

We look for the set of sequences whose empirical distribution is close enough to a given distribution. We just consider rational distributions of a given length n.

Definition 1 The type P_x (or empirical probability distribution) of a sequence $x = x_1, x_2, \ldots, x_n$ is the relative proportion of occurrences of each symbol of $\mathcal{X}$, i.e., $P_x(a) = \frac{N(a|x)}{n}$ for all $a \in \mathcal{X}$, where $N(a \mid x)$ is the number of times that a occurs in the sequence $x \in \mathcal{X}^n$.

Definition 2 Given a length n, denote by $\mathcal{P}_n$ the set of types of sequences of length n; i.e.,

$$\mathcal{P}_n = \{P_x \mid x \in \mathcal{X}^n\}$$

For instance, if $\mathcal{X} = \{0, 1\}$, then the set of possible types for the length n is:

$$\mathcal{P}_n = \left\{(0,1), \left(\tfrac{1}{n}, \tfrac{n-1}{n}\right)\left(\tfrac{2}{n}, \tfrac{n-2}{n}\right), \ldots, \left(\tfrac{n-2}{n}, \tfrac{2}{n}\right), \left(\tfrac{n-1}{n}, \tfrac{1}{n}\right), (1,0)\right\}$$

Definition 3 If $P \in \mathcal{P}_n$, then the set of sequences of length n and type P is called the type class of P, denoted by T(P); i.e., $T(P) = \{x \in \mathcal{X}^n : P_x = P\}$.

[3]Neyman's source code (1998) is also the identity but it is not the optimal one given his set of verification sequences. In particular, he uses half of the communication sequences to specify the chosen play and the other half to balance them, in clear contrast with our codification which uses a whole sequence to determine the play. His construction produces that the set of possible plays is smaller than ours, i.e., a subset of ours, and then that our upper bound on player's 1 complexity is larger than Neyman's upper bound.

5.2 Random sequences: Typical Set

We present here some basic results from Information Theory. For a more complete treatment consult Cover and Thomas (1991).

Let X be a random variable over a finite set Θ, whose distribution is $p \in \Delta(\Theta)$; i.e., $p(\theta) = \Pr(X = \theta)$ for each $\theta \in \Theta$.

Definition 4 The entropy $H(X)$ of X is defined by $H(X) = -\sum_{\theta \in \Theta} p(\theta) \log(p(\theta)) = -E_X[\log p(X)]$, where $0 \log 0 = 0$ by convention.

Notice that the entropy of a random variable depends on the distribution and not on the values it takes and measures the amount of information contained in a random variable or in a probability distribution.

Let $X = (X_1, \ldots, X_n)$ be a vector of finite random variables over $\Theta^n_{k=1}\Theta_k$. Then by the definition of entropy,

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

Given a pair of random variables (X_1, X_2) taking values in $\Theta_1 \times \Theta_2$ with joint distribution $p(\theta_1, \theta_2)$, we denote by $p(\theta_2 \mid \theta_1)$ the conditional probability that $X_2 = \theta_2$ given that $X_1 = \theta_1$. Define $h(X_2 \mid \theta_1) = -\sum_{\theta_2 \in \Theta_2} p(\theta_2 \mid \theta_1) \log p(\theta_2 \mid \theta_1)$.

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

Definition 5 The conditional entropy $H(X_2 \mid X_1)$ of X_2 given X_1 is defined by

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

An easy computation shows that $H(X_1, X_2) = H(X_1) + H(X_2 \mid X_1)$ where $H(X_1, X_2)$ is the entropy of the variable (X_1, X_2). Then, the generalization of the above result is the next proposition.

Proposition 1 If $X = (X_1, \ldots, X_n)$ is a vector of random variables then

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

The entropy of a random variable is a measure of the uncertainty of the random variable, i.e., the amount of information required on the average to describe the random variable, while the relative entropy (or Kullback Leiber distance) gives us the distance between two distributions. It gives the level of inefficiency of assuming that the distribution is q when instead the true one is p.

Definition 6 The relative entropy of the probability mass function $p(x)$ with respect to the probability mass function $q(x)$ is defined as

$$D(p \| q) = \sum_{x \in \mathcal{X}} p(x) \log \frac{p(x)}{q(x)} = E_p \log \frac{p(X)}{q(X)}$$

Notice that the relative entropy is not a true distance since it is not symmetric and does not satisfy the triangle inequality. Nevertheless, it is often consider as a distance between distributions.

5.2.1 Typical set: Asymptotic Equipartition Property.

Consider independent, identically distributed (i.i.d) random variables $X_1, \ldots, X_n$. The law of large numbers states that $\frac{1}{n} \sum_{i=1}^{n} X_i$ is close to its expected value, EX, for large values of n. The Asymptotic Equipartition Property (AEP) is a consequence of the weak law of large numbers. If $X = X_1, \ldots, X_n$ is a vector of i.i.d random variables and $p(X_1, \ldots, X_n)$ is the probability of observing the sequence $X_1, \ldots, X_n$ then $\frac{1}{n} \log \frac{1}{p(X_1, \ldots, X_n)}$ is close to the entropy $H(X)$. The Asymptotic Equipartition Property makes it possible to divide the set of all sequences into two sets, the typical set, where the sample entropy is close to the entropy of the random variable, and the non-typical set, which contains the other sequences. Any property that is proved for the typical set will determine the behavior of a large sample. However, we might be able to predict the probability of the sequence that we actually observe. We ask for the probability $p(X_1, \ldots, X_n)$ of the outcomes $X_1, \ldots, X_n$, where $X_1, X_2, \ldots$ are i.i.d $\sim p(x)$. We are asking for the probability of an event drawn according to the same probability distribution. It turns out that $p(X_1, \ldots, X_n)$ is close to $2^{-nH(p)}$ with high probability. Almost all events are almost equally likely.

For instance consider the random variable $X \in \{0, 1\}$ with a probability mass function defined by $p(1) = p$ and $p(0) = q$. If $X_1, \ldots, X_n$ are i.i.d. according to $p(x)$. Then the probability of a sequence $x_1, x_2, \ldots, x_n$ is $\prod_{i=1}^{n} p(x_i)$. Clearly, it is not true that all 2^n sequences of length n have the same probability.

The asymptotic equipartition property is formalized in the following theorem:

Theorem 2 (AEP): If $X_1, \ldots, X_n$ are i.i.d. with common distribution $p(x)$ then

$-\frac{1}{n}\log p(X_1, \ldots, X_n) \to H(X)$ in probability.

Definition 7 The typical set $A_\delta^{(n)}$ with respect to $p(x_1, \ldots, x_n)$ is the set of sequences $(x_1, \ldots, x_n) \in \mathcal{X}^n$ such that

$2^{-n(H(X)+\delta)} \le p(x_1, \ldots, x_n) \le 2^{-n(H(X)-\delta)}$

As a consequence of the AEP, the cardinality of the set $A_\delta^{(n)}$ verifies that $(1-\delta)2^{n(H(X)+\delta)} \le \left|A_\delta^{(n)}\right| \le 2^{n(H(X)-\delta)}$; for sufficiently large n.

Thus, the typical set has probability nearly 1, all typical sequences have about the same probability $2^{-nH(X)}$ and by indexing the typical set has short descriptions of length $\approx nH$.

5.3 Information Properties of the Type Set

The essential properties of the method of types arise from the following theorem, which states that all sequences with the same type have the same probability and that the size of a type class $T(P)$ is related with the type entropy.

These expressions make it possible to compute the behavior of long sequences drawn i.i.d. according to some distribution based on the properties of the type of the sequence. Then, if $X_1, X_2, \ldots, X_n$ are drawn i.i.d. according to $q(x)$, the typical set associated with $q(x)$ can be considered as the Type Set of the empirical distribution associated with $X_1, X_2, \ldots, X_n$, where the Kullback distance between the type P and q is small.

Theorem 3 a) If $X_1, X_2, \ldots, X_n$ are i.i.d. according to q, then the probability of x depends on its type and is given by $q^n(x) = 2^{-n(H(P_x)+D(P_x\|q))}$.

b) $\frac{1}{(n+1)^{|\mathcal{X}|}}2^{nH(P)} \le |T(P)| \le 2^{nH(P)}$

For the binary case we can write a better bound of the cardinality of $T(P)$ by Stirling's formula[4]. Specifically, $|T(P)| = \alpha_n \frac{2^{2n}}{(\pi n)^{\frac{1}{2}}}$ for $P = (\frac{1}{2}, \frac{1}{2})$ and length $2n$ with $1 - \frac{1}{22n} \le \alpha_n \le 1 + \frac{1}{22n}$.

[4] Stirling's formula says that:

$$n! = n^n \exp(-n)\sqrt{2\pi n}(1 + \varepsilon_n)$$

with $1 \le \varepsilon_n \le \frac{1}{11n}$.

5.4 Codi¯cation and data compression

Given a random variable X over a ¯nite set £, we are interested in generating a one-to-one map (the source code) between the range of X and a ¯nite set with speci¯c properties. The most important property among them is that the expected length of the source code of the random variable is as short as possible. With this requirement we achieve an optimal data compression which is important to identify a variable with a lower complexity.

Our purpose is to de¯ne a code from the support of the random variable distributed uniformly over sequences of length n (where n is even) with parity of ones and zeros in each of them and with a ¯xed last component equal to one, into the sequences belonging to the minimal Type Set of length m. The input sequences are played in the veri¯cation phase and an optimal codi¯cation of these sequences is used for the communication phase.

Known results in Information Theory relate the expected length of the code with the entropy of the random variable to code. For instance, Shannon (1948) establishes that the length of the code of each element of the range of the random variable is the logarithm of the inverse of its associated probability. Then the expected length of the code is lower than the entropy of the random variable. Also, Hu®man (see Cover and Thomas, 1991) constructs an algorithm where the expected length of any source code is minimized and thus he provides with optimal coding[5]. Next we present formally the de¯nitions of codi¯cation and data compression.

De¯nition 8 A source code C from a random variable X is a mapping from £, the range of X, to D^* the set of ¯nite length strings of symbols from a D-ary alphabet. Let C(x) denote the codeword corresponding to x and let l(x) denote the length of C(x).

De¯nition 9 The expected length L(C) of a source C(x) for a random variable X with probability mass function p(x) is given by $L(C) = \sum_{x \in £} p(x) l(x)$, where l(x) is the length of the codeword associated with x.

De¯nition 10 A code is said to be non-singular if every element of the range of X maps into a di®erent string in D^*, i.e., $x_i \neq x_j \Rightarrow C(x_i) \neq C(x_j)$.

Non-singularity su±ces for an unambiguous description of a single value of X.

De¯nition 11 A codeword x is a pre¯x in a codeword y if there is a codeword z such that xz = y.

[5] Hu®man also establishes an inverse ranking between the probabilities and the length of the codes of each element. Elements with higher probability have an associated code of a shorter length and viceversa.

Definition 12 A code c is called a prefix code or an instantaneous code if no codeword is a prefix of any other codeword.

An instantaneous code can be decoded without reference to the future codewords since the end of a codeword is immediately recognizable. The above property justifies the prefix code as a good codes since there is no prefix part such that the end of each code is unique. The sufficient condition to construct instantaneous code of minimum expected length is known as the Kraft inequality. Formally:

Theorem 4 (Kraft inequality): For any instantaneous code over an alphabet of size D, the codeword lengths $l_1, l_2, \ldots, l_m$ must satisfy the inequality $\sum D^{-l_i} \le 1$.

By the above definitions we have to consider the coding of a source from a random variable such that the expected length $L(C)$ is as short as possible. This is equivalent to finding the instantaneous code with the minimum expected length, i,e., to minimize $L = \sum p_i l_i$ subject to $\sum D^{-l_i} \le 1$. By the use of the Lagrangian multipliers we get that the optimal codelengths are $l_i^* = -\log_D p_i$. Then, the expected length is $L^* = \sum p_i l_i^* = -\sum p_i \log_D p_i = H_D(X)$. Thus, $H_D(X) \le L^*$ with equality iff $D^{-l_i} = p_i$.

Remark 1 Consider now a source alphabet of size 2^k, with equidistribution. The entropy associated is $H = -\sum_{i=1}^{2^k} 2^{-k} \log 2^{-k} = k$. By the above bound on L^*, such a source is coded by all codewords with length k.

In our problem we want to codify a subset of the Type Set[6] $T_n(\frac{1}{2}, \frac{1}{2})$ of length n such that the output of the codification verifies: 1) it consists of balanced sequences and 2) the last component of each sequence is equal to 1. Notice that the set of the verification sequences V satisfies: 1) $V \subset TP_{n-1}(\frac{n-2}{2n-2}, \frac{n}{2n-2}) \times \{1\}$ and 2) $\binom{n-3}{\frac{n-2}{2}} < |V| \le \binom{n-1}{\frac{n}{2}}$ or $\frac{2^{n-3}}{n} < |V| < \frac{2^{n-1}}{n}$ and 3) $V \subset \{0,1\}^m$ such that m is odd. Each sequence $s \in TP_{n-1}(\frac{n-2}{2n-2}, \frac{n}{2n-2}) \times \{1\}$ has an associated probability of $\frac{1}{|V|}$. The next result establishes that these sequences have optimal descriptions of length about n.

Proposition 2 Let C be a source code from $TP_{n-1}(\frac{n-2}{2n-2}, \frac{n}{2n-2}) \times \{1\}$, uniformly distributed, to the set of finite length strings of a binary alphabet. Then the expected length of C is greater than $n - 3 - \log n$ and smaller than $n - 1 - \log n$.

[6] Assume that the first component refers to the frequency of ones and then the second component to that of zeros.

Proof:

By definition $L(C) = \sum_{s \in TP_{n-1}(\frac{n-2}{2n-2}, \frac{n}{2n-2}) \times \{1\}} p(s) \log \frac{1}{p(s)} =$

$= \sum_{s \in TP_{n-1}(\frac{n-2}{2n-2}, \frac{n}{2n-2}) \times \{1\}} \frac{1}{|TP_{n-1}(\frac{n-2}{2n-2}, \frac{n}{2n-2}) \times \{1\}|} \log \left| TP_{n-1}(\frac{n-2}{2n-2}; \frac{n}{2n-2}) \times \{1\} \right| =$

$\log \left| TP_{n-1}(\frac{n-2}{2n-2}; \frac{n}{2n-2}) \times \{1\} \right|$

As $\frac{2^{n-3}}{n} < \left| TP_{n-1}(\frac{n-2}{2n-2}; \frac{n}{2n-2}) \times \{1\} \right| < \frac{2^{n-1}}{n}$ then

$n - 3 - \log n < L(C) < n - 1 - \log n$ ∎

Formal statement of Result 1: Let $V \subset TP_{n-1}(\frac{n-2}{2n-2}; \frac{n}{2n-2}) \times \{1\}$ and $Q = C(V)$, where C is a source code with minimal expected length and with a balanced output. Then $Q = V$ and C is a bijective map.

Proof of Result 1: Let Q be the set of communication messages with $Q = C(V)$; where V is the set of verification sequences. We prove here that $Q = V$.

Let m be the smallest odd integer such that $|Q| < \left| TP_m(\frac{m-1}{2m}; \frac{m+1}{2m}) \right|$: By the above theorem it is clear that $m > n - 3$: Recall that n is even and then such a smallest odd integer m is $n - 3 + 2 = n - 1$: Then the communication phase consists of sequences in $TP_{n-1}(\frac{n-2}{2n-2}; \frac{n}{2n-2}) \times \{1\}$ which already was the set V. Then the source code C is the identity[7]. ∎

The nature of the verification sequences which we want to codify is not a relevant information to find the optimal set for the communication phase. We present next an alternative approach for the construction of the set of verification sequences which allow us to relate our communication scheme under strategic complexity (finite automata) with those in repeated games with communication and full rationality (Lehrer, 1996; Lehrer and Sorin, 1997; Forges, 1990; Gossner, 1998; Gossner and Viellie, 1999 and Ben-Porath, 1998, among others).

To this end, recall that the entropy of sequences of i.i.d. random variables is a key concept to describe such sequences. Also, in the framework of finite automata, it measures how many states are needed to describe sequences and thus it is a good measure for communication schemes, since their required "good properties" (better payoffs, no deviations from the equilibrium path,

[7] Notice that the cardinality of the set of verification sequences in Neyman (1998) is $2^{n-1} < |V| \le 2^n$, and then his optimal codification corresponds with a communication set equal to $TP_{n'-1}(\frac{n'-2}{2n'-2}; \frac{n'}{2n'-2}) \times \{1\}$ where n' is the smallest even integer which satisfies $\frac{2^{n'-3}}{\sqrt{\pi n'}} > 2^n$. Then the optimal codification length of the sequences belonging to this set is about n, instead of $2n$ which is Neyman's length for the communication sequences.

etc.) are given with the minimal number of states. This minimality condition on the number of states together with that of sequence-independent payoffs drive us to choose as the set of verification sequences that of random variables with maximal entropy. Thus, we can consider the set of verification sequences as a subset of a Typical set of length n given a random variable X. A consequence of the AEP is that all sequences of the typical set of length n, $A_\delta^{(n)}$ have about the same probability $2^{-nH(X)}$ and by using the above remark 1 they have also short descriptions of length $\approx nH$. Obviously the random variable X has to be close to the empirical distribution of the chosen Type set. The next lemma establishes the condition on the random variable X such that the Typical Set of length n associated to X contains the type set $TP_{n-1}(\frac{n-2}{2n-2}; \frac{n}{2n-2})$: the distribution of X has to be close enough to the uniform distribution. This condition allow us to give an alternative proof of the result 1.

Lemma 7 Let X be a random variable with distribution q and $\left|\sum_{i \in \{0;1\}} (q(i) - \frac{1}{2}) \log q(i)\right| < \delta$ then $TP_{n-1}(\frac{n-2}{2n-2}; \frac{n}{2n-2}) \times \{1\} \subset A_\delta^{(n)}$

Proof:

Let $x = (x_1; \ldots; x_{n-1}; 1) \in TP_{n-1}(\frac{n-2}{2n-2}; \frac{n}{2n-2}) \times \{1\}$.

Let $q = \prod_{i \in \{1;\ldots;n\}} q$ be the distribution induced by a sequence of i.i.d. variables $X_1; \ldots; X_n$.

It suffices to prove that the probability of $x = (x_1; \ldots; x_{n-1}; 1)$ verifies that $2^{-n(H(X)+\delta)} \le q(x) \le 2^{-n(H(X)-\delta)}$.

By the relationship between the type of x and that of q then $q(x) = 2^{-n(H(P_x)+D(P_x \| q))}$

Notice that $(H(P_x) + D(P_x \| q)) = H(q) + H(P_x) - H(q) + D(P_x \| q) =$

$H(q) + \sum_{i \in \{0;1\}} (-\frac{1}{2} \log \frac{1}{2} + q(i) \log q(i) + \frac{1}{2} \log \frac{1/2}{q(i)}) =$

$= H(q) + \sum_{i \in \{0;1\}} (q(i) - \frac{1}{2}) \log q(i) < H(q) + \delta$

Then $q(x) = 2^{-n(H(P_x)+D(P_x \| q))} \ge 2^{-n(H(q)+\delta)}$.

Hence $2^{-n(H(X)+\delta)} \le 2^{-n(H(P_x)+D(P_x \| q))} \le 2^{-n(H(X)-\delta)}$ and we conclude that $x = (x_1; \ldots; x_{n-1}; 1) \in A_\delta^{(n)}$. ∎

Alternative proof of Result 1: By the above lemma we can consider that the set of verification sequences is a subset of the typical set $A_\delta^{(n)}$ associated to a sequence of i.i.d of random variables with common distribution q. By the AEP the probability of each sequence is about $2^{-n(H(P_x)+D(P_x \| q))}$. Then, by remark 1, the shortest description of each sequence is n

. The image of the source code corresponds with the sequences used in the communication phase. The coding map is singular and then for any sequence in the verification phase there exists a unique element (signal or message) in the type set of length n with corresponds with the communication phase. $\square$

6 PROOF OF THE MAIN RESULT

Let $G^T(m_1; m_2)$ be the finite repetition played by finite automata of the two-player game in strategic form $G = (\{1; 2\}; A; r)$ and let $x \in co(r(A))$ such that $x^i > u^i(G)$; $i = 1; 2$: Without loss of generality, x can be expressed as $x = \sum_{i=1}^{3} \lambda_i r(a_i^1; a_i^2)$ where $\sum_{i=1}^{3} \lambda_i = 1$. Consider the following three cases, according to the number of player 1's different actions to obtain x:

1. $|\{a_1^1; a_2^1; a_3^1\}| = 1$;

2. $|\{a_1^1; a_2^1; a_3^1\}| = 2$;

3. $|\{a_1^1; a_2^1; a_3^1\}| = 3$:

The proof of the main result in the first case is a subcase of the proof of the second one. A proof can be found in Neyman(1998); alternative proofs are provided by Papadimitriou and Yannakakis (1994) and by Hernandez and Urbano (2000).

We construct a mixed strategy equilibrium for the second case, i.e., when $|\{a_1^1; a_2^1; a_3^1\}| = 2$, since it is rich enough to show the main features of the more general construction of the third one yet it is easier to deal with. Thus, assume without loss of generality that $a_1^1 = a_2^1 \neq a_3^1$; and that $a_2^2 = a_3^2$ and denote a_1^1 and a_1^2 by 0 and a_2^2 and a_3^1 by 1 and assume that $x = \lambda_0 r(0; 0) + \lambda_1 r(1; 1) + \lambda_2 r(0; 1)$; with $\lambda_i > 0$, $i = 0; 1; 2$ and where $\sum_{i=0}^{2} \lambda_i = 1$: Then, either $\lambda_0 r^2(0; 0) + \lambda_2 r^2(0; 1) > (u_2(G) + 2\varepsilon)(\lambda_0 + \lambda_2)$, or $r^2(1; 1) > u_2(G) + 2\varepsilon$ and we assume this last inequality. The other subcase, i.e., when $a_1^1 = a_2^1 \neq a_3^1$ and $a_1^2 \neq a_2^2 \neq a_3^2$ and the third case, i.e., when $|\{a_1^1; a_2^1; a_3^1\}| = 3$; is analyzed at the end of the paper.

Set

$I = \left\lfloor \frac{T}{L+\delta} \right\rfloor$

$d_1 = [\lambda_1 I]$

$d = L^4$

$d_2 = \left\lfloor \frac{L^2}{d} \right\rfloor$

$d_3 = \frac{l\lambda_0(1 - \frac{1}{L})^{\rfloor}}{d}$

$d_0 = l - d_1 - dd_2 - d(d+1)/2 - dd_3$

$l_1 = d_0 + d_1$

Now, we define the play by means of a communication phase and a regular phase. This last phase consists of a cycle with two parts. The first one is a verification phase which is related with the communication phase. The second one starts with the action pairs $(0; 0)$ and also includes all the required actions pairs $(0; 1)$ to achieve the payoff x in the cycle. The third part is the remainder of the action pairs $(0; 0)$ and then all action pairs $(1; 1)$: The cycle is repeated until the end of the game.

More specifically, the number l above is the length of the cycle that both players repeat until the end of the game. The cycle consists of playing the actions pairs $(0; 0)$, $(1; 1)$, $(0; 1)$ in such a way that the payoff x is obtained, i.e., the number of times that each action pair is played is approximately $l\lambda_i$; $i = 0; 1; 2$; respectively. For every T (the length of the game), the cycle has to be repeated a large number of times, L where $\lceil \frac{3K}{\varepsilon} \rceil \quad L < \frac{1}{\varepsilon^2}$: To ensure that at the end of the repeated game player 1 is in the regular play where the action pair $(1; 1)$ is played, we choose $l = [\frac{T}{L+\delta}]$ where $\frac{1}{2} < \delta < 1$ and $(L + 1)l - T << l$ and $(L + 1)l > T$. To deter deviations it is enough to assume that $L = \lceil \frac{3K}{\varepsilon} \rceil$. The number of times that the action pair $(1; 1)$ is played is about $\lambda_1 l$ and then d_1 is the integer part. The action pairs $(0; 0)$ and $(0; 1)$ are not played consecutively. The number of times that the action pair $(0; 0)$ is played is d_0 plus dd_3 which is about $\lambda_0 l$ and that of the action pair $(0; 1)$ is dd_2: The integer number d is sufficiently large to accommodate all pair actions in such a way that the number of reused states in the player 1's automaton is relatively small.

6.1 Equilibrium play

The following is a construction of an equilibrium point $(\sigma^*; \tau^*)$ of $G^T(m_1; m_2)$ with associated equilibrium vector payoff $(y^1; y^2)$ with $|y^i - x^i| < \varepsilon$.

The mixed equilibrium strategy of player 2, τ^*, chooses randomly a pure strategy τ^{ϵ} where ϵ is an element of the message space Q. The message space Q is a set of sequences of length $2k$, where k depends on the parameters of the game, T and m_1. Moreover it verifies several conditions: every message is a sequence with the same number of ones and zeros and the last component is 1. Thus Q is a subset of $T(P)$ with $P = (\frac{1}{2}; \frac{1}{2})$ and with sequences of length $2k$.

Each pure strategy σ in the support of σ^* of player 1 and the pure strategy τ^{ϵ} of player 2 induce a play $\omega(\sigma;\tau^{\epsilon}) = (\omega_1(\sigma;\tau^{\epsilon}); :::; \omega_T(\sigma;\tau^{\epsilon}))$ that depends on ϵ, and therefore we denote it by $\omega(\epsilon) = (\omega_1(\epsilon); :::; \omega_T(\epsilon))$ and call it the proposed play. The payoff associated to $\omega(\sigma;\tau^{\epsilon})$ does not depend on the selected message ϵ.

Player 2 communicates his choice of ϵ in Q at the beginning of the play to player 1, who processes this information. The action of player 1 in the communication phase is independent of ϵ and player 2 specifies the proposed play $\omega(\epsilon)$ with his message. After the communication phase, the proposed play enters in a cycle of length l. First, players verify the proposed play by following the verification play for $2k$ stages. It consists of a coordinated play of actions pairs $(0;0)$ and $(1;1)$. Then, both players play the regular play consisting of the action pairs $(0;0)$; $(0;1)$ and $(1;1)$ for the remaining stages until l:

The strategy of player 1 will detect with positive probability any deviation of player 2. Some deviation of this player will be detected immediately with positive probability, and others will lead to a detection with positive probability in a future stage. The strategy of player 1 triggers to punishing (playing the strategy that holds player 2 down to $u_2(G)$; denoted by D^{i}) forever once he detects a deviation by player 2. We turn now to the formal construction of the proposed play and the associated equilibrium strategies.

The set of messages

We start with the construction of the set Q, and the integers k and l_1. First, let $k = k(m_1; l_1)$, be the smallest integer such that $\binom{2k}{k}\frac{l_1}{2} > m_1 - l_1$. We will see that the number of pure strategies for player 2 is at most $\binom{2k-1}{k}$ and by Lemma 6 the complexity of each pure strategy is at least l_1, filling up, in this way, player 1's complexity. It follows that $\binom{2(k-1)}{k-1}\frac{1}{2} \quad [\frac{m_1 - l_1}{l_1}] < \binom{2k}{k}\frac{1}{2}$:

Recall that l is the length of the cycle. For every T (the length of the game), the cycle has to be repeated a large number of times, L. Also, recall that $l_1 = d_0 + d_1$ where d_1 is the number of action pairs $(1;1)$ along the cycle of length l, i.e., $l\lambda_1$ and d_0 is approximately $\frac{l\lambda_1}{L}$: Then l_1 is a function of $O(\varepsilon T)$:

To build the set of messages, consider the set of equidistributed sequences of zeros and ones of length $2k$ and such that the last component of each of them is a 1. These sequences have the property that their empirical distribution correspond with the type $(\frac{1}{2};\frac{1}{2})$ of length $2k$: Recall that $T(\frac{1}{2};\frac{1}{2}) = \left\{ x \in \{0;1\}^{2k} : P_x = P \right\} = \left\{ x \in \{0;1\}^{2k} : \frac{N(a|x)}{n} = \frac{1}{2} \text{ for all } a \in \{0;1\} \right\}$: Then, the set that we consider is a subset of $T(\frac{1}{2};\frac{1}{2})$ of length $2k$ and where the last component of each sequence is a 1 to mark the end of both the communication phase and the verification play.

Thus, Q is a subset of $T(\frac{k-1}{2k-1};\frac{k}{2k-1})\times\{1\}\subset T(\frac{1}{2};\frac{1}{2})$ and with cardinality $\left|T(\frac{k-1}{2k-1};\frac{k}{2k-1})\times\{1\}\right| = \frac{|T(\frac{1}{2};\frac{1}{2})|}{2} = \binom{2k}{k}\frac{1}{2}$.

The associated play to a given message

For every ε we define the associated play $\omega(\varepsilon)$ of G^T, i.e., a sequence $\omega(\varepsilon) = (\omega_1(\varepsilon); \ldots; \omega_T(\varepsilon))$ with $\omega_t(\varepsilon) = (\omega_t^1(\varepsilon); \omega_t^2(\varepsilon))$ in A. As noted above, the play consists of a communication phase followed by a play phase. We set $\theta(\varepsilon)$ as the communication phase. The play phase, denoted by $c(\varepsilon)$, is a cycle which is repeated until the end of the game except for the last stage T: This phase consists of the verification play $\theta^*(\varepsilon)$ and the regular play $\mathbf{e}$.

The purpose of the regular play $\mathbf{e}$ is twofold: to achieve the payoff x and with the lowest complexity[8]. Since $x = \lambda_0 r(0;0) + \lambda_1 r(1;1) + \lambda_2 r(0;1)$, an easy way to reach x would be to play $l\lambda_0$ times $(0;0)$, followed by $l\lambda_2$ times $(0;1)$ and by $l\lambda_1$ times $(1;1)$, with an associated complexity for player 1 of l. By lemma 6, x could even be achieved with a complexity of $l(\lambda_0 + \lambda_1)$: However, it is possible to reduce the above complexity by repeating the action pairs in a different way while keeping the same proportion than above. For instance, the action pair $(0;0)$ could be played a number of times and then introduce subplays of appropriated length of the other action pairs $(0;1)$ and $(1;1)$. The connection among different subplays is marked by the action pair $(0;1)$. Specifically, the play of the action pair $(0;0)$ consists of its "shortest" repetition such that player 1 can safely accept the remaining action pairs $(0;0)$ and $(0;1)$ (by using his reused states). To this end $\mathbf{e}$ is composed of three different parts: The play c^*, plus the play of $d_0 - k$ times of $(0;0)'$s and the play of $d_1 - k$ times of $(1;1)'$s. In this way, player 1 can insert the $l\lambda_2$ repetitions of $(0;1)'$s in the states with a 0 as the action function and thus the play c^* consists of action pairs $(0;0)$ and $(0;1)$, while the second play is just $(0;0)$ action pairs, and the third one represents about $l\lambda_1$ times $(1;1)$ action pairs. The regular play is common for every signal.

Let

$$\theta(\varepsilon) = ((0;\varepsilon_1);\ldots;(0;\varepsilon_{2k}))$$

$$\theta^*(\varepsilon) = ((\varepsilon_1;\varepsilon_1);\ldots;(\varepsilon_{2k};\varepsilon_{2k}))$$

[8] Notice that we look for a construction which can be implemented by the player with the lowest complexity. In this way we achieve the less restrictive equilibrium conditions.

The construction of the cycle is as follows. Let

$$\mathbf{e} = c^* + (d_0 - k) \cdot (0;0) + (d_1 - k) \cdot (1;1)$$

Define the play c^* by,

$$c^* = \sum_{i=1} (d_3 \cdot (0;0) + (d_2 - 1) \cdot (0;1) + (i - 1) \cdot (0;0) + (0;1))$$

Notice that $|(d_3 \cdot (0;0) + d_2 \cdot (0;1) + (i - 1) \cdot (0;0) + (0;1))| = d_3 + d_2 + i$; which does not follow a cyclical pattern. Also, observe that the difference between the payoff of a run of c^* and that of the corresponding part of $(0;0)'$s and $(0;1)'$s of x is sufficiently small, i.e.,

$$\left\| R(d_3 \cdot (0;0) + d_2 \cdot (0;1) + (i - 1) \cdot (0;0) + (0;1)) - \frac{\lambda_0 r(0;0) + \lambda_2 r(0;1)}{\lambda_0 + \lambda_2} \right\| < O(\tfrac{1}{L}).$$

The play c^* is designed such that the action pair $(0;1)$ is played about $l\lambda_2$ times and the complexity of the regular play is minimized. To this end, player 1 uses the same action pair $(0;1)$ as a signal or marker to change from a subplay to another in each run of c^*. In this way, the complexity of $\mathbf{e}$ decreases from $l(\lambda_0 + \lambda_1)$ to $d_0 + d_1$ (see lemma 10). Notice that the above upper bound is the number of action pairs $(0;0)$ and $(1;1)$ which are needed (to reach the payoff x) in a cycle of length l, where the pair $(0;1)$ is used as a signal for player 1. The last i-$(0;0)$ action pairs are used as a counting device to assure that the number of runs is exactly d. Notice that the regular play is designed to compress the actions pairs $(0;0)$ by means of the action pairs $(0;1)$ included in c^*.

Recall that the verification play $\theta^*(\epsilon)$ and the regular play $\mathbf{e}$ form the cycle $c(\epsilon)$ that is repeated until the end of the game except the last stage T: Then, define this cycle $c = c(\epsilon)$ of length l by:

$c = c(\epsilon) = \theta^*(\epsilon) + \mathbf{e} = \theta^*(\epsilon) + c^* + (d_0 - k) \cdot (0;0) + (d_1 - k) \cdot (1;1)$.

Also, recall that $comp^1(\mathbf{e}) \geq (d_0 - k) + (d_1 - k)$ and then $comp^1(c) = comp^1(\theta^*(\epsilon) + (d_0 - k) \cdot (0;0) + (d_1 - k) \cdot (1;1)) = d_0 + d_1 = l_1$. The play c^* allows player 1 to reduce his complexity of $\mathbf{e}$ and then the complexity of $c(\epsilon)$.

In the last stage of the game player 2 plays the best response to the action 1 of player 1, denoted by b^2. Then $\omega_T(\epsilon) = (1;b^2)$.

The associated play to a given ϵ in Q is given by:

$$
\begin{aligned}
\omega(\epsilon) = {} & \theta(\epsilon) + L(\theta^*(\epsilon) + c^* + (d_0 - k) * (0,0) + (d_1 - k) * (1,1)) + \\
& \theta^*(\epsilon) + c^* + (d_0 - k) * (0,0) + \\
& (T - 2k - lL - (l - (d_1 - k)) - 1) * (1,1) + (1, b^2).
\end{aligned}
$$

To summarize, a play $\omega(\epsilon) = (\omega_1(\epsilon), \ldots, \omega_T(\epsilon))$ with $\omega_t(\epsilon) = (\omega_t^1(\epsilon), \omega_t^2(\epsilon))$ in A is as follows:

$$
\omega_t(\epsilon) = \left\{
\begin{array}{lll}
(0, \epsilon_t) & \text{if} & 0 \le t \le 2k \\
\theta^*(\epsilon) & \text{if} & 2k < t \bmod l \le 4k \\
c^* & \text{if} & 4k < t \bmod l \le 4k + l - d_1 - d_0 \\
(0,0) & \text{if} & 4k + l - d_1 - d_0 < t \bmod l \le 3k + l - d_1 \\
(1,1) & \text{if} & 3k + l - d_1 < t \bmod l \le 2k + l
\end{array}
\right\},
$$

$\omega_T(\epsilon) = (1, b^2)$

The first row corresponds with the communication phase where player 2 sends the message ϵ and player 1 plays 0. The verification phase is represented by the second row. The third, fourth and fifth rows coincide with the rest of the cycle of length l. The cycle is repeated until the end of the game.

Properties of the associated play

In this section we study first how close to x is the payoff induced by the cycle $c(\epsilon)$ and by its associated play $\omega(\epsilon)$; and second, the complexity of player 1 associated to both the play $\omega(\epsilon)$ and the set of plays Q. The first two lemmae assert that for T sufficiently large, the payoff induced by $c(\epsilon)$ and by the proposed play $\omega(\epsilon)$ is ε-close to the equilibrium payoff x and it is independent of the signal. The last lemma of this section establishes a lower bound for the different plays to measure player 1's complexity on the set of plays Q:

Lemma 8 The vector payoff $R(c(\epsilon))$ is independent of ϵ, and for sufficiently large values of T,

$$
|R^i(c(\epsilon)) - x^i| < \frac{\varepsilon}{2}.
$$

Proof:

The number of action pairs $(0;0)$, $(1;1)$ and $(0;1)$ has to be approximately $l\lambda_0$, $l\lambda_1$ and $l\lambda_2$ (respectively). The number of times of $(0;0)'s$; $(1;1)'s$ and $(0;1)'s$ in the play c is $k + dd_3 + d(d-1)/2 + d_0 - k$, $k + d_1 - k$, and $dd_2 + d$ respectively.

Since $d_0 = l - d_1 - dd_2 - d(d+1)/2 - dd_3$ then

$k + dd_3 + d(d-1)/2 + d_0 - k = l - dd_2 - d - d_1$.

Notice that $|l\lambda_0 - k + dd_3 + d(d-1)/2 + d_0 - k| = |l\lambda_0 - l - dd_2 - d - d_1| =$
$= |l\lambda_1 - d_1 + l\lambda_2 - dd_2 - d| \leq |l\lambda_1 - d_1| + |l\lambda_2 - dd_2 - d| < 1 + d$

Then for sufficiently large values of T, $|R^i(c(\epsilon)) - x^i| < \frac{\varepsilon}{2}$. ∎

Lemma 9 The vector payoff $\sum_{t=1}^{T} r(\omega_t(\epsilon))$ is independent of ϵ, and for sufficiently large values of T,

$$|R^i(\omega(\epsilon)) - x^i| < \varepsilon.$$

Proof:

Clearly $\sum_{t=1}^{T} r(\omega_t(\epsilon))$ is independent of ϵ because the communication and the verification plays consist of balanced sequences. Then, both phases are independent of the chosen sequence.

Notice that $|R^i(\omega_t(\epsilon)) - R^i(c(\epsilon))| < \frac{K}{L}$.

By the above lemma $|R^i(\omega(\epsilon)) - x^i| = |R^i(\omega(\epsilon)) - R^i(c(\epsilon)) + R^i(c(\epsilon)) - x^i|$
$\leq |R^i(\omega(\epsilon)) - R^i(c(\epsilon))| + |R^i(c(\epsilon)) - x^i| < \frac{K}{L} + \frac{\varepsilon}{2} = \frac{\varepsilon}{3} + \frac{\varepsilon}{2} < \varepsilon$
∎

Both players' complexity give us the equilibrium conditions on the automaton sizes. Player 2's complexity on a given play $\omega(\epsilon)$; i.e., $comp^2(\omega(\epsilon))$; is equal to $T+1$. To find out a lower bound of player 1's complexity, we study his play complexity associated to ϵ, i.e., $comp^1(\omega(\epsilon))$. Player 1 has to respond correctly to each signal and thus we compute his complexity on the set of plays $\omega(\epsilon)$ for $\epsilon \in Q$, $comp^1(Q)$; where Q is the set of plays. Recall that a player's complexity of a set of plays Q is defined as the smallest complexity of a strategy σ which is conformable to Q.

To compute $comp^1(Q)$, we have to consider the coordinated and the non-coordinated plays. The coordinated plays consist of the play of both the verification phase of length $2k$ and the last action pairs $(d_1 - k)(1;1)$. Hence, a lower bound of player 1's complexity is the number of different coordinated plays in the play phase. Their complexity is exactly their length which coincides with the number of the action pairs in the verification play plus the number of $(1;1)'$s after c^*. Notice that the play of c^*, i.e., $(\omega_{d_3++2k+1}(\varepsilon);\ldots\omega_{l-d_1-d_0+k}(\varepsilon))$, is not a coordinated play: its play complexity is obtained by lemma 6. Then, to bound player 1's complexity on the set of plays $\omega(\varepsilon);\ \varepsilon\in Q$, we find lower bounds of both the two coordinated plays in $\omega(\varepsilon)$ and the non-coordinated part of $\omega(\varepsilon)$. With them, it is shown that a lower bound of player 1's complexity, $comp^1(Q)$; is $|Q|l_1$:

Lemma 10 1)For every $(\varepsilon;t);(\varepsilon';t')\in Q\times(\{1;\ldots;2k+d_3\}\cup\{l-d_1+k+1;\ldots;l-1\})$ with $(\varepsilon;t)\neq(\varepsilon';t')$

$(\omega_t(\varepsilon);\ldots\omega_{t+l_1-1}(\varepsilon))\neq(\omega_{t'}(\varepsilon');\ldots;\omega_{t'+l_1-1}(\varepsilon'))$

2) Let $\omega=(\omega_{d_0+1}(\varepsilon);\ldots\omega_{l-d_1+k}(\varepsilon))$, a lower bound of player 1's complexity of ω is $comp^1(\omega)\geq d_0$

3) By 1) and 2) $comp^1(Q)\geq|Q|l_1$

Proof:

1) To bound player 1's complexity on $\omega(\varepsilon);\ \varepsilon\in Q$, we find first lower bounds of both the two coordinated plays in $\omega(\varepsilon)$ and the non-coordinated part of it.

After the communication phase for $2k<t \bmod l\leq 4k+d_3$ and $l-d_1+k<t \bmod l\leq l$ both players follow a coordinated play. We have to prove that for every $(\varepsilon;t);(\varepsilon';t')\in Q\times(\{1;\ldots;2k+d_3\}\cup\{l-d_1+k;\ldots;l\})$ with $(\varepsilon;t)\neq(\varepsilon';t')$, then

$(\omega_t(\varepsilon);\ldots\omega_{t+l_1-1}(\varepsilon))\neq(\omega_{t'}(\varepsilon');\ldots;\omega_{t'+l_1-1}(\varepsilon'))$.

It suffices to show that for any pair $(\varepsilon;t)\neq(\varepsilon';t)$ and $2k<t \bmod l\leq 4k+d_3$ and $l-d_1+k<t \bmod l\leq l$ either there exists $0\leq s\leq l$ with $(\omega_t(\varepsilon);\ldots;\omega_{t+s}(\varepsilon))\neq(\omega_{t'}(\varepsilon');\ldots;\omega_{t'+s}(\varepsilon'))$; or there exists $0\leq s\leq l$ with $\omega_{t+s}(\varepsilon)\neq\omega_{t'+s}(\varepsilon')$:

Suppose that $t=t'$ and thus $\varepsilon\neq\varepsilon'$. Therefore there exists $0\leq s'<2k$ with $\varepsilon_{s'}\neq\varepsilon'_{s'}$. Let $s=l-t+s'$ such that $0\leq s\leq l$: We conclude that $\omega_{t+s}(\varepsilon)\neq\omega_{t+s}(\varepsilon')$:

Next, suppose that $t\neq t'$. We can always choose one s such that the $\omega_{t+s}(\varepsilon)$ is in the regular part and $\omega_{t'+s}=\omega_{l+2k}$. With that we conclude that $\omega_{t+s}(\varepsilon)=(0;0)$ and $\omega_{t'+2^k}(\varepsilon')=(1;1)$: More specifically, suppose that $t<t'$. If $t'-t>l-d_1-2k$, and $t0+2k+1<l$ setting $s=2k+l-t+1;\ t'+s=l+2k+1+t'-t$ then $\omega_{t'+s}(\varepsilon)=(1;1)$ and $\omega_{t+s}(\varepsilon)=(0;0)$: If $t'-t\leq d_1-k$ as $d_1>2k$ setting $s=2k+l-t'+1$; then $\omega_{t+s}(\varepsilon)=(1;1)$ and $\omega_{t'+s}(\varepsilon')=(0;0)$. Note that this choice is independent of $\varepsilon;\varepsilon'\in Q$.

2) To bound the complexity of the non-coordinated part, i.e., $\omega=(\omega_{d_3+2k+1}(\varepsilon);\ldots\omega_{l-d_1+k}(\varepsilon))$ we use lemma 6 where $B^1=\{0\}$ and $k(0)=d_0$. Then $comp^1(\omega)\geq d_0$:

3) By adding the above complexity bounds then $comp^1(Q) \geq |Q| l_1$. $\square$

6.2 Construction of the equilibrium strategy of player 2

We now describe player2's equilibrium strategy. It consists of a mixed strategy supported by $|Q|$ pure strategies. For every $\varepsilon \in Q$, a proposed play $\omega(\varepsilon)$ is associated to a pure strategy in the support of τ^*, the equilibrium mixed strategy. Player 2 follows the proposed play and punishes forever as soon as he detects a deviation. Thus, for any $\varepsilon \in Q$, $\tau^\varepsilon = (\tau_t^\varepsilon)_{t=1}^T$ is the pure strategy of player 2 defined by,

$$\tau_t^\varepsilon(s_1, \ldots, s_{t-1}) = \begin{cases} \omega_t^2(\varepsilon) & \text{if } (s_1, \ldots, s_{t-1}) = (\omega_1(\varepsilon), \ldots, \omega_{t-1}(\varepsilon)); \\ D^2 & \text{otherwise} \end{cases}$$

The pure strategy $\tau^\varepsilon \in \Sigma^2(T, T+1)$, i.e., τ^ε is implemented by an automaton $< \{1, \ldots, T, T+1\}, 1, f_\varepsilon^2, g_\varepsilon^2 >$ of size $T+1$ where:

- $\{1, \ldots, T, T+1\}$ is the set of states.
- 1 is the initial state.
- The action function f_ε^2 defined by $f_\varepsilon^2(t) = \omega_t^2(\varepsilon)$ if $t \leq T$; $f_\varepsilon^2(T+1) = D^2$.
- The transition function g_ε^2, defined by $g_\varepsilon^2(t, a) = t+1$ if $a = \omega_t^1(\varepsilon)$ and $t \leq T$, and $g_\varepsilon^2(t, a) = T+1$ otherwise, i.e., if $a \neq \omega_t^1(\varepsilon)$, or if $t = T+1$.

6.3 Construction of the equilibrium strategy of player 1

Player 1's equilibrium strategy is a mixed strategy. Player 1 has to answer correctly to any signal sent by player 2. Hence, each pure strategy that belongs to the mixed equilibrium strategy must be conformable with the set of plays $\{\omega(\varepsilon) : \varepsilon \in Q\}$. In the communication phase player 1 has to process the information sent by player 2 and he does it by using the same states than those for the regular play. The verification play consists of a coordinated play where both players play the same action at the same time. The regular play is composed of two different parts. In the first one c^* is played. The second part consists of a coordinated play with $d_0 - k$ action pairs of $(0,0)$'s followed by $d_1 - k$ action pairs of $(1,1)$'s. To reduce the associated complexity player 1 reuses states with action function 0 to implement both action pairs $(0,0)$ and $(0,1)$:

Recall that in these states player 1 cannot punish deviations since both action are admitted and thus he construct an equilibrium mixed strategy that conceals the disposition of his reused states in the regular play. The difference among player 1's pure strategies in the support of the equilibrium strategy is the location of these states for the communication phase and for the play of c^* in the regular play. Player 1's mixed strategy is a uniform distribution over the minimal subset of pure strategy $\sigma \in \Sigma^1(m_1)$ where σ is conformable with $\{\omega(\varepsilon) : \varepsilon \in Q\}$. The minimal set is understood as the minimal set with enough uncertainly about the true locations of his reused states.

6.3.1 The Automaton of player 1

The mixed equilibrium strategy of player 1, $\sigma^* \in \Delta(\Sigma(m_1; T))$; is a mixture of pure strategies, each one being implemented by an automaton conformable with Q. Each automaton has to implement the communication phase and the play phase. We define first, the state space and the action function which implement $\omega(\varepsilon)$; for all $\varepsilon \in Q$. Second, we present the transition function for the play phase, i.e., the verification play and the regular play. Finally, we construct the transition function for the communication phase which determines the initial state.

The state space is

$$M^1 = \{\alpha\} \cup Q \times \{1; \ldots; l_1\}$$

The action function of the automaton is given by,

$$f^1(\alpha) = D^1;$$

and

$$f^i(\varepsilon; j) = \begin{cases} \theta_j(\varepsilon) & \text{if } 1 \quad j \quad 2k \\ 0 & \text{if } 2k < j \quad d_0 + k \\ 1 & \text{if } d_0 + k < j \quad d_0 + d_1 = l_1 \end{cases}$$

The play phase: The play phase is a cycle which is composed of the verification play and the regular play. The first one is a coordinated play of length 2k and it is independent of the pure strategy selected by player 1. The regular play consists of a play which is independent, or deterministic part, and another play which depends on the pure strategy selected by player 1. We start with the description of the deterministic part which is quite similar to that of Neyman (1998).

We visualize the states of the automaton of the form $(²; j)$ as arranged in a rectangular array with jQj rows and l_1 columns. Recall that $l_1 = 2k + d_0$ ¡ $k + d_1$ ¡ k: The rows are indexed by the di®erent elements ² in Q and the columns are indexed by $1; :::; l_1$: We may think that every row corresponds to a pure strategy of player 2. Given ² in Q, in the ¯rst 2k states the action function assigns an action $²_j$ if $1 \quad j \quad 2k$ which depends on the row (veri¯cation phase). Then, the action is 0 in each state whose column is between $2k + 1$ to l_1 ¡ d_1 ¡ k. For the last d_1 ¡ k columns, the action function assigns the action 1. The number of columns coincides with the complexity of player 1's cycle.

Figure 1 illustrates the automaton of player 1 for $k = 3$; and $j\,Q\,j =$ ¡$^{2k ¡ 1}_{k}$¢ = ¡$^{5}_{3}$¢ $= 10$: Suppose that the regular play has 67 columns, where $d_0 = 33$ and $d_1 = 40$ and that the veri¯cation play has 6 associated columns. The ¯lled disks (²) represent states of the automaton whose action function is a 0, when player 2 plays a 0 as well. The small disks (±) represent states that play the action 1 when player 2 plays a 1. The big disks (°) mean the ¯nal states of the regular play where both players have to play 1 at the same time. The transition function in these last states goes to the ¯rst state in the same row. The horizontal arrows indicate the transition of the automaton when player 1 follows a coordinated play.

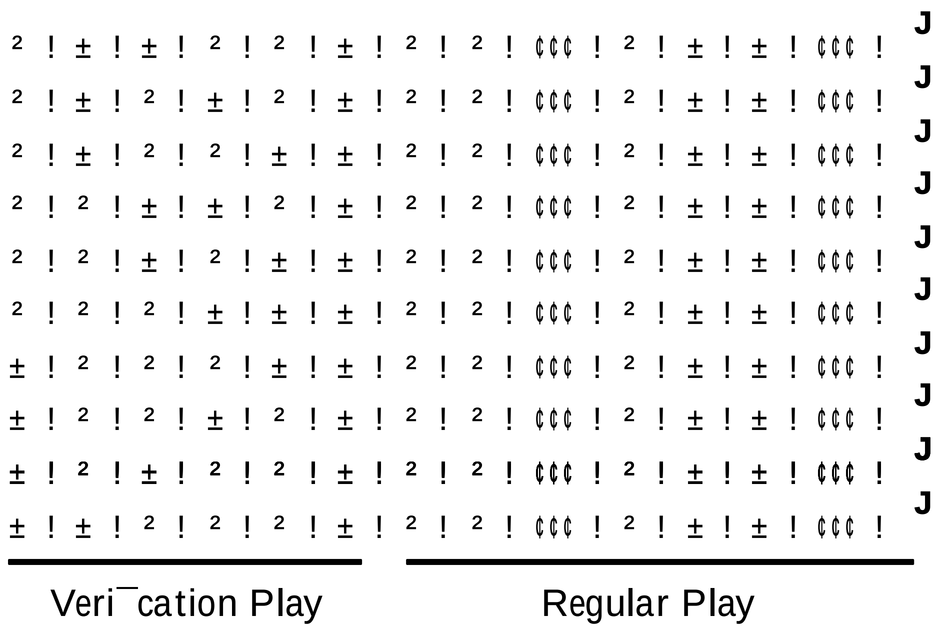


Figure 1.

Next we define the transition function for the play phase. According to the different nature of the plays in this phase (deterministic and random) the transition function is designed such that it allows both punishing deviations immediately in the deterministic part and precluding deviations in the random one.

The transition of the automaton is defined such that for each fixed $\epsilon \in Q$, player 1 remains in the same row and goes to the next column in case player 2 plays correctly in the verification phase and for states $(\epsilon; j)$ with $2k < j \le d_0 + k$ if player 2 plays a 0 and for states $d_0 + k \le j < l_1$ when player 1 plays a 1. For the state $(\epsilon; l_1)$; if player 2 plays 1 then the transition function goes to the first column in this row, i.e., player 1 starts another repetition of the cycle if player 2 plays a 1 in this stage. This leads to the following transitions:

$$g^1((\epsilon; j); 0) = \begin{cases} (\epsilon; j+1) & \text{if} \quad 1 \le j < 2k \quad \text{and } \epsilon_j = 0 \\ (\epsilon; j+1) & \text{if} \quad 2k < j \le l_1 - d_1 + k \end{cases}$$

$$g^1((\epsilon; j); 1) = \begin{cases} (\epsilon; j+1) & \text{if} \quad 1 \le j < 2k \quad \text{and } \epsilon_j = 1 \\ (\epsilon; j+1) & \text{if} \quad l_1 - d_1 + k \le j < l_1 \\ (\epsilon; 1) & \text{if} \quad j = l_1 \end{cases}$$

The states of the automaton of the form $(\epsilon; j)$ such that $1 \le j \le 2k$ or $d_0 + k \le j \le l_1$ implement a coordinated play. Any deviation from this play at these states results in punishing forever.

$g^1((\epsilon; j); e) = \alpha$ if $1 \le j \le 2k$ and $\epsilon_j \ne e$

$g^1((\epsilon; j); 1) = \alpha$ if $d_0 + k \le j \le l_1$

The state α is an absorbing state and then player 1 punishes forever after the first deviation is detected. The transition function is as follows:

$g^1(\{\alpha\}; *) = \alpha.$

Up to now, we have defined the deterministic part of the regular phase. To reduce the complexity of the cycle, player 1 reuses states whose action function is a 0 and he uses the action 1 of player 2 as a signal to start another run of c^*. These states are of the form $(\epsilon; j)$ with $2k < j \le d_0 - 5k$ with no reused state following c^* and processing the signal in the communication phase. There are dd_2 states that tolerate both actions 0 and 1: To conceal the location of these states we add a random procedure to implement the action pairs $(0;1)$ which are played in the play c^*. This random procedure is defined by the following random integers:

Let z be an integer number such that $1 \le z \le 2$ and $\bar{L} = 2d_2 + d_3 + d$. Set a random increasing function $\rho : \{1;\dots;L\} \to \{2k + d_3 + 1;\dots;d_0 - 4k - \bar{L}\}$ with $\rho(i+1) > \rho(i) + Ld_2 + d_3 + d$, and consider a random sequence of elements $i_1;\dots;i_d$ of $\{1;\dots;L\}$.

Recall that $c^* = \sum_{i=1}^{d}(d_3 * (0;0) + (d_2 - 1) * (0;1) + (i - 1) * (0;0) + (0;1))$. Now, we can define the transition function of player 1's automaton implementing c^*:

We start with the definition of the transition function of the state $(\varepsilon;2k)$, i.e., when the verification play finishes. Player 2 has to play the action 1 and then player 1 jumps to the column $\rho(i_1) - d_3$ which it is unknown to player 2. In this way player 2 is uncertain about the first reused states in c^*. The transition function is defined by:

$g^1((\varepsilon;2k);1) = (\varepsilon;\rho(i_1) - d_3)$.

For every $1 \le t \le d$ we define the transition function for the states whose action function is a 0 but accept the action 1 of player 2, i.e., these states implement the action pairs $(0;1)$ in c^* as:

$g^1((\varepsilon;\rho(i_t) + zs);1) = (\varepsilon;\rho(i_t) + zs + s)$ if $0 \le s < d_2$

and

$$g^1((\varepsilon;\rho(i_t) + s);1) = \begin{cases} (\varepsilon;\rho(i_{t+1}) - d_3) & \text{if } s = \rho(i_t) + zd_2 + t \text{ and } t < d \\ (\varepsilon;2k+1) & \text{if } s = \rho(i_t) + zd_2 + t \text{ and } t = d \end{cases}$$

The first row is the transition function for every d_2 stages of $(0;1)$ in c^* given i_t, for $0 < t \le d$. The second one defines the transition function for the last $(0;1)$; for every repetition $t < d$. Notice that the assumptions on the random sequence $i_1;\dots;i_d$, imply that for $1 \le t < t'$, $\rho(i_t) + t \ne \rho(i_{t'}) + t'$. Finally, the last row is the transition function for the last $(0;1)$ for the last repetition of c^*. The states that admit both actions are properly located in the first d_0 states.

The next figure illustrates the transition function in the regular phase implementing c^*. We consider two cases: 1) assume that $L = 2$ then $d = 2^4 = 16$ and $d_3 = 3 = d_2$. Let $i_1;\dots;i_{16} = 1;2;1;\dots$ and $\rho(1) = 14$ and $\rho(2) = 58$; 2) assume now that $L = 2$ then $d = 2^4 = 16$ and $d_3 = 3 = d_2$. Let $i_1;\dots;i_{16} = 2;1;1;\dots$ and $\rho(1) = 14$ and $\rho(2) = 58$.

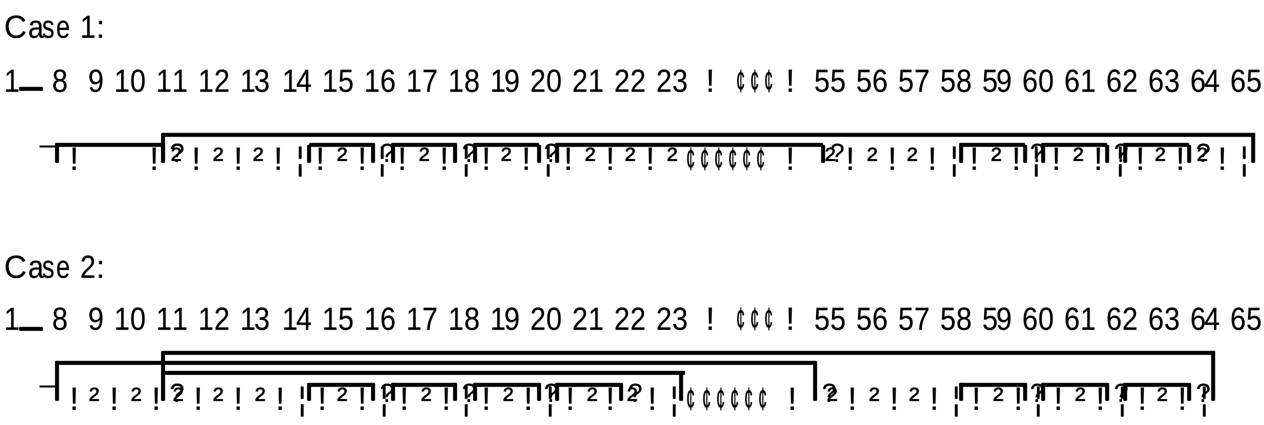


Figure 2.

The communication phase: In the communication phase player 1 has to process the information sent by player 2. He uses the same states to be used in the regular play. We design the transition function for the first 2k stages such that player 1 follows a specific play after the communication phase and he conceals his reused states by changing their locations in his pure strategies. In other words, each pure strategy in the support of player 1's mixed strategy is designed such that it selects the right row along the communication phase and it does not reveal which states admit both actions.

The transition function of player 1's automaton in this phase depends on the pure strategy selected. Each pure strategy is given by two random numbers p and n. The first of them determines the initial state of the automaton. We denote this initial state by $(1; p)$. Thus, p is the column where player 1 processes the signal sent by player 2 and it verifies that $d_0 - 5k \le p \le d_0 - 3k$.

Given $\epsilon = (\epsilon_1; \ldots; \epsilon_i; \ldots; \epsilon_{2k-1}; 1) \in Q$ let k_ϵ be the smallest integer such that $\sum_{i=1}^{k_\epsilon} \epsilon_i = k$ or $k_\epsilon - \sum_{i=1}^{k_\epsilon} \epsilon_i = k$. The random integer $n \in \{1; 2\}$ determines the jumps in the columns (along the same row) that player 1 follows in the communication phase when player 2 sends a 1 after k_ϵ stages.

The transition function of the communication phase consists of three parts: the first one corresponds to the first stages until k_ϵ; the second to k_ϵ until $2k - 1$, and finally the third part refers to the last stage of the communication.

Thus, to select the right row during the first stages, the transition function jumps among the different rows guarantying that when the number of either ones or zeros is greater or equal than k the state of the automaton is in the row that corresponds to player 2's sequence of actions in the first k_ε stages of the game. This row is the one where the first components are the corresponding to the signal sent by player 2, followed by the maximal number of zeros. Thus, player 2's signals are ranked in this way. This is achieved through the following partial transition function:

$$\text{If } \varepsilon = (\varepsilon_1, \ldots, \varepsilon_{2k-1}, 1) \in Q \text{ and } \sum_{i=1}^{j-p} \varepsilon_i < k \text{ and } (j-p) - \sum_{i=1}^{j-p} \varepsilon_i < k \text{ and}$$
$$h = k - (j-p) - \sum_{i=1}^{j-p} \varepsilon_i \text{ then}$$
$$g^1((\varepsilon, j), 1) = ((\varepsilon_1, \ldots, \varepsilon_{j-p}, 1, 0, \ldots, 0, \varepsilon_{j-p+1+h}, \ldots, 1), j+1) \quad \text{if } p \leq j \leq p+2k$$
$$\text{If } \varepsilon = (\varepsilon_1, \ldots, \varepsilon_{2k-1}, 1) \in Q \text{ and } \sum_{i=1}^{j-p} \varepsilon_i = k \text{ or } (j-p) - \sum_{i=1}^{j-p} \varepsilon_i = k \text{ and}$$
$$h = k - (j-p) - \sum_{i=1}^{j-p} \varepsilon_i \text{ then } g^1((\varepsilon, j), 1) = (\varepsilon, j+n) \quad \text{if } p \leq j \leq p+2k$$
$$g^1((\varepsilon, j), 0) = (\varepsilon, j+1) \quad \text{if } p \leq j \leq p+2k$$

In second place, we design the transition function[9] when player 2 is sending the last part of the signal except for the last stage, i.e., for $t : k_\varepsilon > t > 2k$. Here, the randomness of the jumps, n, allows player 1 to hide his reused states. Recall that $n \in \{1, 2\}$, then:

If $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_{2k-1}, 1) \in Q$ and $\sum_{i=1}^{j-p} \varepsilon_i = k$ or $(j-p) - \sum_{i=1}^{j-p} \varepsilon_i = k$ and $n = 1$; then

$$g^1((\varepsilon, j), 1) = (\varepsilon, j+1) \quad \text{if } p + k_\varepsilon \leq j < p + 2k$$

If $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_{2k-1}, 1) \in Q$ and $\sum_{i=1}^{j-p} \varepsilon_i = k$ or $(j-p) - \sum_{i=1}^{j-p} \varepsilon_i = k$ and $\overline{k_\varepsilon} = 2 \times (k - \sum_{i=1}^{k_\varepsilon} \varepsilon_i - 1) + (k - k_\varepsilon + \sum_{i=1}^{k_\varepsilon} \varepsilon_i)$ and $n = 2$, then

$$g^1((\varepsilon, j), 1) = (\varepsilon, j+2) \quad \text{if } p + k_\varepsilon \leq j < p + \overline{k_\varepsilon}$$

Finally, the last state in the communication phase is not in the same column for every row. It depends on ε, n, p, i.e., on where the communication starts, on the distribution of ones in ε and on the number of jumps.

Let $\mathbf{e}$ be a function

$\mathbf{e}$:

$$Q \longrightarrow [p, \ldots, p+3k]$$
$$\varepsilon \longrightarrow \mathbf{e}(\varepsilon) = p + 3k - \sum_{i=1}^{k} \varepsilon_i - 2$$

[9] Notice that we do not use a distribution over transition functions, but we produce enough uncertainty on the final states of the transition function to deter deviations.

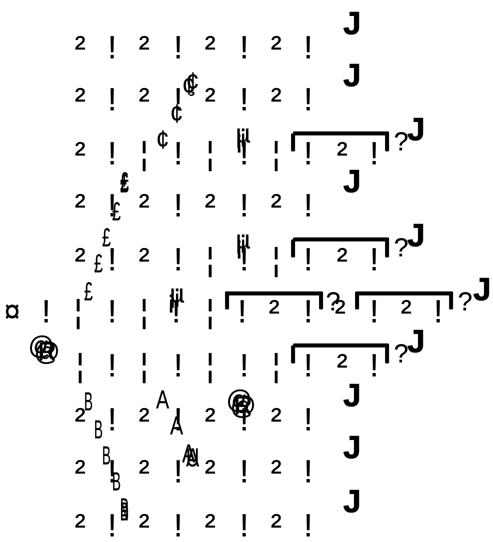
Notice that the max $\hat{e}(\varepsilon)$ is when $\varepsilon = (0; \ldots; 0; 1; \ldots; 1) \in Q$ and then $\hat{e}(\varepsilon) = p + 3k - 2$.

Now it is possible to define the final state's transition function for every row: $g^1((\varepsilon; \hat{e}(\varepsilon)); 1) = (\varepsilon; 1)$:

This is equivalent to: $g^1((1; p); \varepsilon) = (\varepsilon; 1)$:

In all other cases the value of g^1 equals α.

Figure 3 illustrates the communication phase associated to the verification play in the above example for $k = 3$ and $n = 2$. The star ($\star$) is the initial state. The diamonds ($\diamond$) represent those states in the regular play that are used to process the information sent by player 2 in the communication phase, and thus admit both actions 0 and 1 from player 2. The big states with a dot are the states in the regular play that player 1 uses to determine the end of the communication phase. These states also admit both actions, 0 and 1.



0 1 1 0 0 1
0 1 0 1 0 1
0 1 0 0 1 1
0 0 1 1 0 1
0 0 1 0 1 1
0 0 0 1 1 1
1 0 0 0 1 1
1 0 0 1 0 1
1 0 1 0 0 1
1 1 0 0 0 1

Figure 3.

As noted above, player 1's automaton is a matrix with l columns and $\binom{2k-1}{k}$ rows. Thus, the communication phase starts in the p column that player 1 has chosen randomly. Hence, the states used to process the signal are located in a submatrix with $2k$ rows and a number of columns which depends on n and ε.

Finally, we note that the conditions to find out player 1's bounds come from

$$\varepsilon \quad |R^i(\omega(\varepsilon)) - x^i| < \varepsilon:$$

- The relationship between the number of reused states and the number of states with action function 0 is approximately $\frac{1}{L}$.
- $\binom{2(k-1)}{k-1}^{\frac{1}{2}} \leq \left[\frac{m_1 - l_1}{l_1}\right] < \binom{2k}{k}^{\frac{1}{2}}$

With the first and the second condition we obtain a bound on k with respect to T and ε by counting the number of action pairs played when the game is repeated until T and the maximal number of reused states. Then, with the last condition we obtain the upper bound of player 1's complexity.

6.4 Equilibrium conditions:

We check here that the constructed strategies are indeed an equilibrium. We show first that any profitable deviation by player 1 cannot be implemented by a finite automata of complexity m_1. We study the complexity of a strategy of player 1 which yields a higher payoff when playing against τ^*, i.e. $comp^1(\sigma)$ where $r^1_T(\sigma, \tau^*) \geq \sum_{t=1}^{T} \frac{r^1(\omega(\epsilon))}{T}$. Secondly, we show that with a probability close to 1 there is no profitable deviation from player 2.

Let σ be a strategy of player 1 and let $\epsilon \in Q$, with $r^1_T(\sigma, \tau^\epsilon) \geq \sum_{t=1}^{T} \frac{r^1(\omega(\epsilon))}{T}$. Then, $\omega_t(\sigma, \tau^\epsilon) = \omega(\epsilon)$ for any $t \leq \frac{T}{z}$ where z is a fixed number that depends on the action pair (1,1), with payoffs x, and on the other payoffs of the stage game G. Therefore, for any strategy σ of player 1, $r^1_T(\sigma, \tau^\epsilon) \leq \sum_{t=1}^{T} \frac{r^1(\omega(\epsilon))}{T} + \frac{C}{T}$ where C depends on the game G.

Let σ be a pure strategy of player 1 with $r^1_T(\sigma, \tau^*) \geq \sum_{t=1}^{T} \frac{r^1(\omega(\epsilon))}{T}$ and such that σ is implemented by an automaton of size m_1.

In order to characterize the size of the automaton which implements a profitable deviation, consider the following partition of the set of messages.

Let

$$Q(1;\sigma) = \left\{ \epsilon \in Q \text{ such that } r^1_T(\sigma, \tau^\epsilon) > \sum_{t=1}^{T} \frac{r^1(\omega(\epsilon))}{T} \right\}$$
$$Q(2;\sigma) = \left\{ \epsilon \in Q \text{ such that } r^1_T(\sigma, \tau^\epsilon) = \sum_{t=1}^{T} \frac{r^1(\omega(\epsilon))}{T} \right\}$$
$$Q(3;\sigma) = \left\{ \epsilon \in Q \text{ such that } r^1_T(\sigma, \tau^\epsilon) < \sum_{t=1}^{T} \frac{r^1(\omega(\epsilon))}{T} \right\}$$

To study the complexity of σ we must know the one of $\omega(\epsilon)$ for every $\epsilon \in Q$, hence we analyze the complexity of every set of the partition of Q. Define $Q_1 = \{\omega(\sigma, \tau^\epsilon) : \epsilon \in Q(1;\sigma)\}$;

Q_2 = f ! (¾; ¿ ²) : ² 2 Q(2; ¾)g ; and Q_3 = f ! (¾; ¿ ²) : ² 2 Q(3; ¾)g : Notice that $\text{comp}^1(Q_2)$ ¸ l_1 jQ(2; ¾)j by lemma 10.

As ¾ veri¯es that r^1_T(¾; ¿ ¤) ¸ $\mathbf{P}^T_{t=1} \frac{r^1(!(²))}{T}$ then r^1_T(¾; ¿ ¤) ¸ $\frac{1}{jQj}$ $\mathbf{P}_{²2Q}$ $\mathbf{P}^T_{t=1} \frac{r^1(!(²))}{T}$: Hence,

r^1_T(¾; ¿ ¤) = $\mathbf{P}_{²2Q} \frac{1}{jQj}$ r^1_T(¾; ¿ ²) =

$\frac{1}{jQj}$ £ §$_{²2Q(1;¾)}$ r^1_T(¾; ¿ ²) + §$_{²2Q(2;¾)}$ r^1_T(¾; ¿ ²) + §$_{²2Q(3;¾)}$ r^1_T(¾; ¿ ²) ¤

Now, since any strategy ¾ of player 1, r^1_T(¾; ¿ ²) $\mathbf{P}^T_{t=1} \frac{r^1(!(²))}{T}$ + $\frac{C}{T}$ and by the de¯nition of Q(3; ¾); then

jQ(1; ¾)j ³ $\mathbf{P}^T_{t=1} \frac{r^1(!(²))}{T}$ + $\frac{C}{T}$ ´ + jQ(3; ¾)j ³ $\mathbf{P}^T_{t=1} \frac{r^1(!(²))}{T}$ ´ ¸

jQ(1; ¾) + Q(3; ¾)j ³ $\mathbf{P}^T_{t=1} \frac{r^1(!(²))}{T}$

Thus $\frac{C}{T}$ jQ(1; ¾)j ¸ jQ(3; ¾)j and for T large enough jQ(1; ¾)j ¸ 2jQ(3; ¾)j

In the next lemma we study the least complexity of a strategy of player 1 which can give him more that $\mathbf{P}^T_{t=1} \frac{r^1(!_t(²))}{T}$.

Lemma 11 The complexity of ¾ such that r^1_T(¾; ¿ ¤) ¸ $\mathbf{P}^T_{t=1} \frac{r^1(!(²))}{T}$ is
comp^1(¾) ¸ (L ¡ 1)l_1 jQ(1; ¾)j + l_1 jQ(2; ¾)j

Proof:

By the de¯nition of complexity, comp^1(¾) = comp^1 f ! (¾; ¿ ²) : ² 2 Qg ¸
comp^1 f ! (¾; ¿ ²) : ² 2 Q(1; ¾) [Q(2; ¾)g = $\text{comp}^1(Q_1)$ + $\text{comp}^1(Q_2)$:

Notice that $\text{comp}^1(Q_2)$ ¸ l_1 jQ(2; ¾)j by lemma 10. Let us bound the complexity of Q_1:

By the de¯nition of Q(1; ¾), for every ² 2 Q(1; ¾), r^1_T(¾; ¿ ²) > R^1(! (²)): Therefore there exists a deviation from the proposed play at the end of the game i.e., for every t 4k + Ll; $!_t$(¾; ¿ ²) = $!_t$(²): Now by lemma 4, a deviation takes place after 4k + Ll. By the de¯nition of complexity with ¯nite automata it su±ces to prove that for every pair (²; t); (²^0; t^0) with (²; t) 6= (²^0; t^0) and t ¸ t^0 in

Q(1; ¾) £ f4k + jc¤j ; :::; 4k + l ¡ 1; 4k + jc¤j + l ¡ 1; :::; 4k + 2l ¡ 1; 4k + jc¤j + 2l; :::; 4k + (L ¡ 1)l ¡ 1g

there exists s < T ¡ t such that

$$(!^2_t(²); :::; !^2_{t+s}(²)) = (!^2_{t^0}(²^0); :::; !^2_{t^0+s}(²^0))$$

and

$$\sigma(\omega_1(\varepsilon),\dots,\omega_{t+s}(\varepsilon)) \neq \sigma(\omega_1(\varepsilon'),\dots,\omega_{t'+s}(\varepsilon'))$$

First, we study the coordinated plays. The play $(\omega_{4k+1+|c^*|}(\varepsilon),\dots,\omega_{4k+l+d_3}(\varepsilon))$ is a coordinated play with the first $d_0 - k$ and the last d_3 actions pairs being $(0,0)$ and $\omega_{0k+l}(\varepsilon) = (1,1)$. As $d_3 > 2k$, the string $(1,1) + d_3 * (0,0)$ only appears at the end of the play and then if $4k + |c^*| \le t' < t < 4k + l$,

$$(\omega_{t+1}(\varepsilon),\dots,\omega_{4k+l}(\varepsilon)) \neq (\omega_1(\varepsilon'),\dots,\omega_{t'+4k+l-t'}(\varepsilon'))$$

and

$(\sigma \mid \omega_{t+1}(\varepsilon),\dots,\omega_{4k+l}(\varepsilon)) \neq (\sigma \mid \omega_1(\varepsilon'),\dots,\omega_{t'+s}(\varepsilon'))$ because each one of these two plays is a coordinated play.

We just consider the case where $t \neq t' (\mathrm{mod}\, l)$. Notice that the play c^* is independent of the signal ε. Moreover $(\omega_{4k+|c^*|+l+1}(\varepsilon),\dots,\omega_{4k+2l}(\varepsilon))$ is a coordinated play. Then, if $t = t' \,\mathrm{mod}(l)$ and $\varepsilon \neq \varepsilon'$

$$(\omega_t(\varepsilon),\dots,\omega_{t+l}(\varepsilon)) \neq (\omega_{t'}(\varepsilon'),\dots,\omega_{t'+l}(\varepsilon'))$$

Let s be the largest positive integer such that

$$(\omega_t(\varepsilon),\dots,\omega_{t+s}(\varepsilon)) \neq (\omega_{t'}(\varepsilon'),\dots,\omega_{t'+s}(\varepsilon'))$$

then, it follows that $\omega^1_{t+s}(\varepsilon) \neq \omega^1_{t'+s}(\varepsilon')$.

Suppose now that $t > t'$, $t = t' \bmod (l)$ and $\varepsilon \in Q(1,\sigma)$.

Let s be the largest positive integer such that $\omega^1_{t+s}(\varepsilon) \neq \omega^1_{t'+s}(\sigma,\tau^{\varepsilon})$.

As $r^1_T(\sigma,\tau^{\varepsilon}) \geq \sum_{t=1}^{T} \frac{r^1(\omega(\varepsilon))}{T} \Rightarrow s < T - t$ and $\sigma(\omega_t(\varepsilon),\dots,\omega_{t+s}(\varepsilon)) \neq \sigma(\omega_{t'}(\varepsilon'),\dots,\omega_{t'+s}(\varepsilon'))$.

∎

Lemma 12 For any strategy $\sigma \in \Sigma^1(m_1)$

$$r^1_T(\sigma,\tau^*) \geq \frac{1}{|Q|}\sum_{Q} R^1(\omega(\varepsilon))$$

Proof:

Suppose that $r^1_T(\sigma,\tau^*) \ge \sum_{t=1}^{T} \frac{r^1(\omega(\varepsilon))}{T}$.

Consider the partition of $Q = Q(1,\sigma) \cup Q(2,\sigma) \cup Q(3,\sigma)$.

First, if $|Q(3,\sigma)| = \emptyset$ then $|Q| = |Q(1,\sigma)| + |Q(2,\sigma)|$. By the above lemma the complexity of σ is greater than or equal to $3l_1|Q(1,\sigma)| + l|Q(2,\sigma)|$.

As $m_1 \ge (L-1)l_1|Q(1,\sigma)| + l_1|Q(2,\sigma)| = l_1|Q| + (L-2)l_1|Q(1,\sigma)|$ and since $\binom{2(k-1)}{k-1}\left[\frac{m_1-l}{l}\right] < \binom{2k}{k}$; then

$m_1 \ge m_1 - 2l_1 + (L-2)l_1|Q(1,\sigma)| \Leftrightarrow |Q(1,\sigma)| = \emptyset$

We conclude that $r^1_T(\sigma,\tau^*) \le \sum_{t=1}^{T} \frac{r^1(\omega(\varepsilon))}{T}$.

Next, if $|Q(3,\sigma)| \neq \emptyset$, as already noted, we can assume that for T large enough $|Q(1,\sigma)| \ge 2|Q(3,\sigma)|$. Then,

$m_1 \ge (L-1)l_1|Q(1,\sigma)| + l_1|Q(2,\sigma)| = \frac{l_1}{2}|Q(1,\sigma)| + \frac{(2L-1)l_1}{2}|Q(1,\sigma)| + l_1|Q(2,\sigma)| > l_1|Q| + \frac{(2L-3)l_1}{2} l_1|Q(1,\sigma)| > m_1$, which is a contradiction. $*$

Lemma 13 For any strategy $\tau \in \Sigma^2$ and every $\varepsilon \in Q$

$$r^2_T(\sigma^*,\tau) \le r^2_T(\sigma^*,\tau^{\varepsilon}).$$

Proof:

Let τ be a pure strategy of player 2 such that for some $\varepsilon \in Q$, $\omega_t(\sigma^*,\tau) = \omega_t(\varepsilon)$ for every $1 \le t \le 2k$ and $r^2(\sigma^*,\tau) \ge r^2(\sigma,\tau^{\varepsilon})$.

Let s' be the smallest integer such that $2k < s' \le T$ with $\omega_s(\sigma^*,\tau) \neq \omega_{s'}(\varepsilon)$ and $\omega_t(\sigma^*,\tau) = \omega_t(\varepsilon)$ for $1 < t < s'$.

If $\omega_t(\varepsilon) = (1,1)$, player 1 punishes immediately forever, since when player 1 plays the action 1 he uses states which do not tolerate both actions. Recall that $r^2(1,1) \ge u^2(G) + 2\varepsilon$. Then player 2 will lose about $2\varepsilon(\lambda_1 - \varepsilon^2)l$. Then $r^2_T(\sigma^*,\tau) \le r^2_T(\sigma^*,\tau^*)$.

If $\omega_t(\varepsilon) = (1,1)$ then $t \le T - \lambda_1 l/3$ and with a probability close to one player1 punishes in the next $d_0 - 2k - 1$ stages. Then, $r^2_T(\sigma^*,\tau) \le r^2_T(\sigma^*,\tau^*)$.

If $\omega_t(\varepsilon) = (0,1)$ then player 2 deviates in c^* and with a probability close to one player 1 punishes in the next $d_0 - 2k - 1$ stages. Then $r^2_T(\sigma^*,\tau) \le r^2_T(\sigma^*,\tau^*)$.

Finally if player 2 deviates in the communication phase, i.e.: if $(\omega_1^2(\sigma^*;\tau);\dots;\omega_{2k}^2(\sigma^*;\tau))$ is not in Q, then with a probability of at least $\frac{1}{2}$ player 1 will detect the deviation in one of the next $5k$ stages.

Therefore τ^* is a best reply against σ^*. $\square$

We finish by giving some details of the above equilibrium construction for the remaining cases: when the payoff x is obtained by three different actions of player 2 and two of player 1 (the other subcase of case 2) and when it is obtained by three different actions of both players (case 3).

Subcase 2.2: Assume that $a_1^1 = a_2^1 \neq a_3^1$, and that $a_1^2 \neq a_2^2 \neq a_3^2$ and denote a_1^1, a_2^1 and a_1^2 by 0; a_3^1 and a_2^2 by 1 and a_3^2 by 2;and assume that $x = \lambda_0 r(0;0) + \lambda_1 r(0;1) + \lambda_2 r(1;2)$, with $\lambda_i > 0$, $i = 0;1;2$ and where $\sum_{i=0}^{2} \lambda_i = 1$:

Here the communication phase entails using the action pairs $(0;0)$ and $(0;2)$; while those of the verification play are $(0;0)$ and $(1;2)$, where the first one is played whenever player 2 sends a 0 in the communication phase and the second whenever he sends the action 2. By the definition of x, the regular play consists of the three pair of actions $(0;0)$, $(0;1)$ and $(1;2)$:

Case 3: $|\{a_1^1;a_2^1;a_3^1\}| = 3$

Subcase 3.1: Assume that $a_1^1 \neq a_2^1 \neq a_3^1$, and that $a_1^2 = a_2^2 \neq a_3^2$ and without loss of generality denote a_1^1, a_1^2 and a_3^2 by 0; a_2^1 and a_2^2 by 1 and a_3^1 by 2; and assume that $x = \lambda_0 r(0;0) + \lambda_1 r(1;1) + \lambda_2 r(2;0)$, with $\lambda_i > 0$, $i = 0;1;2$ and where $\sum_{i=0}^{2} \lambda_i = 1$:

Now the communication phase consists of the action pairs $(0;0)$ and $(0;1)$; while the pairs $(0;0)$ and $(1;1)$ are for the verification play, where the first one is played whenever player 2 sends a 0 in the communication phase and the second whenever he sends the action 1.

Subcase 3.2: Finally, assume that $a_1^1 \neq a_2^1 \neq a_3^1$, and that $a_1^2 = a_2^2 = a_3^2$ and denote a_1^1 and a_1^2 and by 0; a_2^1 and a_2^2 by 1 and a_3^1 and a_3^2 by 2; and assume that $x = \lambda_0 r(0;0) + \lambda_1 r(1;1) + \lambda_2 r(2;2)$, with $\lambda_i > 0$, $i = 0;1;2$ and where $\sum_{i=0}^{2} \lambda_i = 1$:

The communication phase consists now of the action pairs $(0;0)$, $(0;1)$ and $(0;2)$ while the verification play of the pairs $(0;0)$, $(1;1)$ and $(2;2)$. Here the verification set is bigger since the cardinality of the verification sequences' alphabet is three.

7 CONCLUDING REMARKS

We conclude by summarizing the main features of our construction. Let $G^T(m_1; m_2)$ be the finite repetition played by finite automata of the two-player game in strategic form $G = (\{1; 2\}; A; r)$ and let $x \in co(r(A))$ such that $x^i > u^i(G)$ with $x = \sum_{a \in A} \lambda_a r(a_i^1; a_i^2)$, and $a \in A$, where $\sum_{a \in A} \lambda_a = 1$; and $\lambda_a > 0$.

The equilibrium play to achieve x as the equilibrium outcome follows a communication phase and a specific cycle of action pairs play which depends on this communication phase, and whose frequencies are approximately λ_a. The cycle play consists of two parts. One is independent of the communication, the regular play where the payoff x is obtained, while the other, the verification play, is uniquely determined by the message sent in the communication phase. Each part of the cycle play is codified taking into account that the action pairs in the regular play have increasing payoffs for the stronger player, which precludes his deviations as the cycle goes on. In order to keep the distortion from x as small as possible, i.e. ε small, the action pairs used in the verification play should be also used in the regular play (although, this is not always possible). Finally the communication scheme is designed such that the sender player uses different actions to this end while the receiver uses just one action.

The above features establish the codification alphabet for the equilibrium play. Also the communication and verification sequences satisfy an entropy condition to ensure a fixed complexity. In particular, efficient verification to fill up the weaker player's complexity, translates to sequences of maximal entropy, since the number of verification sequences determines this player's complexity.

The construction of the equilibrium play can be understood as a codification problem where what is being codified is the game parameters: the complexity of the weaker player and the targeted payoff x. The inter-play communication phenomenon allows to connect the notion of automaton complexity with that of communication entropy.

Finally, notice that when the players' automata have the same number of states, i.e. $m_1 = m_2$, the above construction remains the same: players could flip up a coin to decide the one who undertakes the communication. Alternatively, other constructions with the flavor of the one presented above could be designed. For instance, players could both send a message in the communication phase, follow a regular play and then verify through the following construction. Let Q and Q' be the communication set of messages of players 1 and 2, respectively. Recall that Q and Q' are subsets of $TP_{k-1}(\frac{k-2}{2k-2}, \frac{k}{2k-2}) \times \{1\}$: Let k' be the smallest even integer such that $\binom{k'}{\frac{k'}{2}-1} > 2^k$ and let c be an element of the following typical set $TP_{k'-1}(\frac{k'-2}{2k'-2}, \frac{k'}{2k'-2}) \times \{1\}$. Let $a \in Q$ be a message of Player 1 and consider a biyective map $a \times TP_{k-1} \to TP_{k'-1}$. The verification consists of a subset of $TP_{k'-1}$ via the above biyective mapping, denoted by (δ), such that each sequence $c = a \delta b$, for a given message b of Player 2. Notice that both players' signals are balanced and the sequence used in the verification phase is balanced as well. The length of the communication is two times the one in the asymmetric case, while that of the verification play is about the same than in the previous case. Nevertheless, the number of possible plays does not vary. With this new construction the rate of distortion, ε, is about the same than above. Notice that here, the number of players' messages has to be the same to fill up the their automata capacity, and that each player's pure strategy consists of a part related with it signal (sent in the communication phase) and of a second part related with all possible messages of the other player.

BIBLIOGRAPHY

Abreu, D. and Rubinstein, A. (1988): "The Structure of Nash Equilibrium in Repeated Games with Finite Automata", Econometrica, 56, 1259-1282.

Aumann, R., J. (1960): "Acceptable points in games of perfect information", Paci¯c Journal of Mathematics, 10, 381-417.

Aumann, R., J. (1981): "Survey of Repeated Games", in Essays in Game Theory and Mathematical Economics in honor of Oskar Morgenstern, 11-42, Mannheim/Vienna/Zurich: Bibliographisches Institute.

Axelrod, R. (1980): "E®ective choice in the prisoner's dilemma", Journal of Con°ict Resolution 24, 3-25.

Ben-Porath, E. (1998): "Correlation without mediation: expanding the set of equilibrium outcomes by cheap pre-play procedures", Journal of Economic Theory, 80, 108-122.

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

Forges, F. (1990): "Universal Mechanisms". Econometrica, 58, 1341-1364.

Gossner, O. and Viellie, N. (1999): "How to play with a biased Coin", Discussion Paper, 99.31. THEMA.

Gossner, O. (1998): "Secure Protocols or how Communication generates Correlation", Journal of Economic Theory, 83, 69-89.

Kalai, E, and Standford,W (1988): "Finite Rationality and Interpersonal Complexity in Repeated Games", Econometrica 56, 2, 397-410.

Lehrer, E. (1996): "Mediated Talk", International Journal of Game Theory, 25, 177-188.

Lehrer, E. and Sorin, S. (1997): "One shot Public Mediated Talk", Games and Economic Behavior, 20, 132-148.

Neyman, A. (1985): "Bounded complexity justi¯es cooperation in the ¯nitely repeated prisoner's dilemma", Economics Letters 19, 227-229.

| | | | (1997): "Cooperation, repetition, and automata", S. Hart, A. Mas Colell, eds., Cooperation: Game-Theoretic Approaches, NATO ASI Series F, 155. Springer-Verlag, 233-255.

———— (1998): "Finitely Repeated Games with Finite Automata", Mathematics of Operations Research 23, 3, 513-552.

Neyman, A. and Okada, D. (1997): "Two-Person Repeated Games with Finite Automata". Mimeo.

Neyman, A. and Okada, D. (1999): "Strategic Entropy and Complexity in Repeated Games", Games and Economic Behavior, 29, 191-223.

Neyman, A. and Okada, D. (2000): "Repeated Games with Bounded Entropy", Games and Economic Behavior, 30, 228-247.

Papadrimitriou, C.H. and Yannakakis, M. (1994): "On complexity as bounded rationality (extended abstract). STOC-94, 726-733.

Rubinstein, A (1979): "Equilibrium in Supergames with the Overtaking Criterion", Journal of Economic Theory 21, 1-9.

Rubinstein, A (1980): "Strong Perfect Equilibrium in Supergames", International Journal of Game Theory 9, 1-12.

Rubinstein, A (1986): "Finite Automata play the repeated prisoners' dilemma",.Journal of Economic Theory 39, 83-96.

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

Zemel, E. (1989): "Small talk and cooperation: A note on bounded rationality", Journal of Economic Theory 49, 1, 1-9.