

ECONOMIC RESEARCH REPORTS

A NOTE ON FINITE AUTOMATA EQUILIBRIA
WITH DISCOUNTING

by

Michele Piccione

R.R. #89-33

November 1989

**C. V. STARR CENTER
FOR APPLIED ECONOMICS**



**NEW YORK UNIVERSITY
FACULTY OF ARTS AND SCIENCE
DEPARTMENT OF ECONOMICS
WASHINGTON SQUARE
NEW YORK, N.Y. 10003**

A Note on Finite Automata Equilibria with Discounting*

by

Michele Piccione**

October 1989

* I am extremely grateful to my adviser, Professor Charles Wilson, for his continuous help and encouragement. Without his advice this paper would have never been in its present form. I also thank Professors Ariel Rubinstein and Jean Pierre Benoit for many helpful comments. Of course, all the remaining mistakes and imperfections are mine. Technical assistance from the C.V. Starr Center for Applied Economics is gratefully acknowledged.

** New York University

ABSTRACT

In this note we will extend the work of Abreu and Rubinstein (1988) by giving necessary conditions as well as sufficient conditions for finite automata equilibria with discounting. Moreover, we will show that any automata equilibrium path is a subgame perfect equilibrium path when repeated game strategies are defined as functions of past histories.

1. Introduction

Abreu and Rubinstein (1988) derive necessary conditions for finite automata equilibria when states are costly and payoffs are evaluated according to both the limit of the means and the discounting criterion. Sufficient conditions are given for some special cases. For the discounting case the automata equilibrium path turns out to have the following features: (i) it consists of an introductory phase and a k -period cycle, (ii) in the introductory phase and the cycle all states are different, (iii) if (a,b) and (c,d) are one shot outcomes of the automata equilibrium path in a two-player game, $a = c$ if and only if $b = d$.

In the limit of the means case a further pre-cycle phase is present.

In this note we will extend their work by giving necessary conditions as well as sufficient conditions for finite automata equilibria with discounting. Moreover, we will show that any automata equilibrium path is a subgame perfect equilibrium path when repeated game strategies are defined as functions of past histories.

The first characterization of automata equilibrium paths when states are costly is to be found in Rubinstein (1986). Payoffs are evaluated according to the limit of the means criterion and a solution concept stricter than the Nash Equilibrium is used. It requires that along the equilibrium path players at no stage will prefer to switch to an automaton with a smaller number of states. The automata equilibrium path shows, after a finite time, a k -period cycle where all states are distinct. Abreu and Rubinstein (1988) extend the work of Rubinstein (1986) and give necessary conditions for finite automata (Nash) equilibria for both the limit of the means and the discounting criterion.

Banks and Sundaram (1989) introduce costs in the transition function as well. They show that the transition function will only include transitions along the equilibrium path as those that are not used will be dropped by the players. This will exclude the existence of unused threats. Consequently, the automata equilibrium path will contain only outcomes that are Nash Equilibria in the stage game.

Our approach instead will focus on the possibility of punishing a player after a deviation by using states that are used in equilibrium. This allows us to add further necessary conditions to those already in Abreu and Rubinstein (1988), which in turn lead us to show that any automata equilibrium path is subgame perfect.

As in Abreu and Rubinstein (1988) we concentrate on issues pertaining to the cost of implementing repeated game strategies leaving aside problems concerning their complexity. This has to be contrasted with the work of Kalai and Stanford (1988) where they show that for any subgame perfect equilibrium there exists a subgame perfect ϵ -equilibrium of finite complexity ϵ -close with respect to payoffs.

In section 2 the game is defined. In section 3 necessary and sufficient conditions for automata equilibrium paths are derived. Subgame perfection will be proved in section 4.

2. The game

Let $G = \{S_1, S_2, u_1, u_2\}$ be a normal form two player game where S_i is a finite set of normal form strategies for player i and $u_i : S_1 \times S_2 \rightarrow R$ his payoff function. We consider an infinite repetition of game G in which strategy space for each player is the set of finite automata. Such a game

will be denoted by G_M^∞ . A strategy for player i is then a four tuple $M_i = \{Q_i, q_i^1, \lambda_i, \mu_i\}$ where Q_i is a finite set of states, $q_i^1 \in Q_i$ is the initial state, $\lambda_i: Q_i \rightarrow S_i$ is the output function, $\mu_i: S_j \times Q_i \rightarrow Q_i$ is the transition function. The number of states in Q_i will be denoted by $\text{card}(M_i)$. Let $S = S_1 \times S_2$ and $Q = Q_1 \times Q_2$.

A pair of automata (M_1, M_2) induces an outcome sequence $\sigma(M_1, M_2) = (s^t)_{t=1}^\infty \in S^\infty$, and a state sequence $\phi(M_1, M_2) = (q^t)_{t=1}^\infty \in Q^\infty$, in the obvious way. I will denote by $\sigma_i(M_1, M_2)$ and by $\phi_i(M_1, M_2)$ the corresponding sequences of actions and states of player i .

We assume that the players evaluate an outcome sequence by the discounted sum of utilities. For $(s^t)_{t=1}^\infty = \sigma(M_1, M_2)$, let $\pi_i(M_1, M_2) = \sum_{t=1}^\infty \delta^{t-1} u_i(s^t)$ where $0 < \delta < 1$.

Following Abreu and Rubinstein (1988) we assume the following partial preference ordering. Player i prefers strictly a pair of automata (M_i, M_j) to another pair (L_i, L_j) , which will be denoted by $(M_i, M_j) (>)_i (L_i, L_j)$ if

$$\pi_i(M_i, M_j) > \pi_i(L_i, L_j) \text{ and } \text{card}(M_i) = \text{card}(L_i)$$

or

$$\pi_i(M_i, M_j) = \pi_i(L_i, L_j) \text{ and } \text{card}(M_i) < \text{card}(L_i)^1$$

¹All the results of this article still hold if we assume a lexicographic completion of the preference ordering with $\pi_i(M_i, M_j)$ first.

3. Automata Equilibrium

In this section we will derive necessary and sufficient conditions for automata equilibria. An automata equilibrium is a Nash equilibrium for the repeated game G_M^∞ .

Definition: A pair of automata (M_1^*, M_2^*) is an **Automata Equilibrium** if there is no M_1 and no M_2 such that $(M_1, M_2^*) \succ_1 (M_1^*, M_2^*)$ or $(M_1^*, M_2) \succ_2 (M_1^*, M_2^*)$

3.1 Cycles

Given an n-tuple $s = (s^1, \dots, s^n)$ we will denote by $c(s)$ the sequence generated by repeating s an infinite number of times. Then define $\dim[c(s)] = \inf \{x : c(s) = c((s^1, \dots, s^x))\}$, x a positive integer

Lemma 1: n is a multiple of $\dim[c((s^1, \dots, s^n))]$.

Proof: Let $(d^t)_{t=1}^\infty = c(s)$ and $k = \dim[c(s)]$. Let $m = \min_\lambda (\lambda n : \lambda n = \rho k, \lambda, \rho \text{ strictly positive integers})$ denote the least common multiple of n and k . Now, $kn/m = rn + tk$ for some integers r, t (see Dudley (1978), Theorem 1 p. 24). Therefore, given any j , $d^{j + kn/m} = d^{rn + tk + j}$. By the definition of k we have that $d^{rn + tk + j} = d^j$ and therefore $d^j = d^{j + kn/m}$. Then, it follows that $\dim[c((s))] \leq kn/m \leq k$. As kn/m is the greatest common divisor of k and n , the claim follows. QED

Lemma 1 implies that to determine if there is a smaller cycle within a sequence generated by an n-tuple (s^1, \dots, s^n) , we only need to check whether the n-tuple itself can be divided into smaller identical cycles. An n-tuple

which cannot be divided into smaller identical cycles will be called **irreducible**.

Let $s = (s^1, \dots, s^n)$ be an n -tuple of outcomes of game G , and suppose that $s^i = (x, y)$ and $s^j = (z, w)$. We will say that s is **injective** if $x = z$ or $y = w$ implies that $s^i = s^j$.

Consider now a pair (s, z) where $s = (s^1, \dots, s^n)$ is an n -tuple, $n \geq 0$,² and $z = (z^1, \dots, z^m)$ is an m -tuple, $m > 0$. Let $b(s, z)$ denote the sequence consisting of one repetition of s and an infinite repetition of z . Notice that if $n = 0$, then $b(s, z)$ is just $c(z)$. Also notice that there is no restriction on the set of sequences which can be generated if we restrict the second argument of $b(\cdot, \cdot)$ to irreducible m -tuples and we assume that $s^n \neq z^m$.³

Now, we have the following lemma

Lemma 2: Let $s = (s^1, \dots, s^n)$ be an n -tuple of outcomes and $z = (z^1, \dots, z^m)$ be an irreducible m -tuple of outcomes for which $s^n \neq z^m$. Suppose the $(n+m)$ -tuple $(s^1, \dots, s^n, z^1, \dots, z^m)$ is injective. If, given two automata M_1 and M_2 , $\sigma(M_1, M_2) = b(s, z)$, then $\text{card}(M_i) \geq n + m$, $i = 1, 2$.

Proof: Assume it does not hold for player i . Let $\{q_i^t\}_{t=1}^m = \phi_i(M_1, M_2)$ and $h \in \{t : q_i^t = q_i^{t+k}, t + k \leq m + n, t \geq 1\}$. By assumption this set is not empty. As $(s^1, \dots, s^n, z^1, \dots, z^m)$ is injective, at h and $h + k$ we must observe the same

²If $n = 0$, s denotes the 0-tuple

³If $s^n = z^m$, consider the sequence $b((s^1, \dots, s^{n-1}), (s^n, \dots, z^{m-1}))$ and proceed until we get an inequality (or the 0-tuple).

outcome. Consequently, $q_i^{h+1} = q_i^{h+k+1}$ and so we must observe the same outcome at $h + 1$ and $h + k + 1$. Proceeding by induction, it follows that a k -period cycle will start at h , i.e. $c(d^h, \dots, d^{h+k-1})$, where $\{d^v\}_{v=1}^\infty = b(s, z)$.

Now, if $h > n$ we have that $\dim[c(z)] \leq k < m$. By Lemma 1 it follows that z is not irreducible. A contradiction.

If $h \leq n$, we have that $s^n = d^n = d^{n+mk} = d^{n+m} = z^m$, contradicting the assumption that $s^n \neq z^m$.

QED

3.2 Best Responses

Let s and z satisfy the assumptions of Lemma 2 and let $b(s, z) = \{d^t\}_{t=1}^\infty$. For $f \in \{1, \dots, m + n\}$ define $M_{(s, z), i}^f = \{Q_i, q_i^1, \lambda_i, \mu_i\}$ to be an automaton for player i where $Q_i = (q_i^1, \dots, q_i^{n+m})$, $\lambda_i(q_i^t) = d_i^t$, $t = 1, \dots, n+m$, and

$$\mu_i(x, q_i^t) = \begin{cases} q_i^{t+1} & \text{if } x = d_j^t \text{ and } t < n + m \\ q_i^{n+1} & \text{if } x = z_j^m \text{ and } t = m + n \\ q_i^f & \text{otherwise.} \end{cases}$$

Automaton $M_{(s, z), i}^f$ conforms to the path $b(s, z)$ as long as player j does and switches to state q_i^f whenever player j deviates from the path.

Consider now the form of the responses of player j that maximize the sum of discounted utility when player i is playing the automaton $M_{(s, z), i}^f$. We will say that $\{x_j^t\}_{t=1}^\infty$ is an optimal sequence of actions for player j with respect to $M_{(s, z), i}^f$ if it maximizes $\sum_{t=1}^\infty \delta^{t-1} u_j(\lambda_i(p^t), x_j^t)$ subject to $p^{t+1} = \mu_i(x_j^t, p^t)$, $p^1 = q_i^1$. Notice that an optimal sequence for player j is the solution of a Markovian stationary problem. By Blackwell (1965), Theorem 7, a

(stationary) solution exists. If $\{x^t\}_{t=1}^{\infty}$ is such a solution and $\{q^t\}_{t=1}^{\infty}$ is the corresponding sequence of states, the value function $V_j: Q_i \rightarrow R$ satisfies $V_j(q^t) = u_j(\lambda_i(q^t), x^t) + \delta V_j(\mu_i(x^t, q^t))$.

Now, let $H_j^k = \sum_{t=k}^{\infty} \delta^{t-k} u_j(d^t)$ denote player j's payoff at k if he conforms to the outcome path $b(s, z)$. Let $v_j^k = \max_a \{u_j(d_i^k, a) : a \in S_j/d_j^k\}$ be the maximum one period utility player j can attain by deviating when the output of player i's machine is d_i^k . Then let $D_j^k(f) = v_j^k + \delta H_j^f$ denote the payoff player j gets at k if he deviates at k and conforms to the punishment path thereafter. The following lemma tells us that to see whether $b(s, z)$ is an optimal sequence of actions for player j with respect to $M_{(s, z), i}^f$, we only need to check one-shot deviations.

Lemma 3: a) $V_j(q_i^1) = H_j^1$ if and only if $H_j^k \geq D_j^k(f)$ for any $k = 1, \dots, n + m$.
 b) Suppose $H_j^k > D_j^k(f)$ for any $k = 1, \dots, n + m$. Then, given a sequence of actions $\{a^t\}_{t=1}^{\infty}$ for player j and the corresponding sequence of states $\{p^t\}_{t=1}^{\infty}$ for player i, $\sum_{t=1}^{\infty} \delta^{t-1} u_j(\lambda_i(p^t), a^t) = V_j(q_i^1)$ only if $\{a^t\}_{t=1}^{\infty} = b(s, z)$.

Proof: a) Necessity: By definition, for $k = 1, \dots, m + n$

$$V_j(q_i^k) \geq v_j^k + \delta H_j^f = D_j^k(f).$$

Therefore, if $V_j(q_i^1) = H_j^1$, then $H_j^k \geq D_j^k(f)$ for any $k = 1, \dots, m + n$.

Sufficiency: Suppose that $H_j^k \geq D_j^k(f)$, $k = 1, \dots, n + m$. We establish first that $V_j(q_i^f) = H_j^f$. Suppose not. Then $V_j(q_i^f) > H_j^f$ and there is a $k \geq f$ such that

$$V_j(q_i^f) = \sum_{t=f}^{k-1} u_j(d^t) \delta^{t-f} + \delta^{k-f} v_j^k + V_j(q_i^f) \delta^{k-f+1}.$$

By definition,

$$H_j^f = \sum_{t=f}^{k-1} u_j(d^t) \delta^{t-f} + H_j^k \delta^{k-f}$$

Therefore,

$$H_j^f \geq \sum_{t=f}^{k-1} u_j(d^t) \delta^{t-f} + D_j^k(f) \delta^{k-f} = \sum_{t=f}^{k-1} u_j(d^t) \delta^{t-f} + \delta^{k-f} v_j^k + H_j^f \delta^{k-f+1}$$

$V_j(q_i^f) > H_j^f$ then implies

$$V_j(q_i^f) > \sum_{t=f}^{k-1} u_j(d^t) \delta^{t-f} + \delta^{k-f} v_j^k + V_j(q_i^f) \delta^{k-f+1}.$$

A contradiction.

Now suppose that $V_j(q_i^t) > H_j^t$ for some $q_i^t \neq q_i^f$. Then there is a $k > t$ such that

$$V_j(q_i^t) = \sum_{\tau=t}^{k-1} u_j(d^\tau) \delta^{\tau-t} + \delta^{k-t} v_j^k + V_j(q_i^f) \delta^{k-t+1}.$$

Then,

$$V_j(q_i^t) = \sum_{\tau=t}^{k-1} u_j(d^\tau) \delta^{\tau-t} + \delta^{k-t} D_j^k(f) \leq H_j^t$$

A contradiction.

b) Assume it is not unique. Let $\{x^t\}_{t=1}^{\infty}$ be an optimal sequence for player j and let $k = \inf\{t : x^t \neq d_j^t\}$. Then,

$$V_j(q_i^1) = \sum_{\tau=1}^{k-1} u_j(d^\tau) \delta^{\tau-1} + \delta^{k-1} u_j(d_i^k, x^k) + \delta^k V_j(q_i^f) \leq \sum_{\tau=1}^{k-1} u_j(d^\tau) \delta^{\tau-1} + \delta^{k-1} D_j^k(f) < H_j^1.$$

A contradiction. QED

We will now construct optimal punishment paths. These are optimal in the sense that they minimize for each player the sum of future discounted utilities along the path $b(s, z)$. Let $f_i \in \operatorname{argmin}_j \{ \sum_{t=j}^{\infty} \delta^{t-j} u_i(d^t) : 1 \leq j \leq n + m \}$ be the time at which the sum of future discounted utilities for player i is at its minimum.

Now we have the following theorem

Theorem 1: Let $s = (s^1, \dots, s^n)$ be an n -tuple of outcomes and $z = (z^1, \dots, z^m)$ be an irreducible m -tuple of outcomes for which $s^n \neq z^m$. Then $b(s, z)$ is an automata equilibrium outcome path only if

- (1) $(s^1, \dots, s^n, z^1, \dots, z^m)$ is injective, and
- (2) $H_i^k \geq D_i^k(f_i)$, $i = 1, 2$, $k = 1, \dots, m + n$.

Moreover, if $H_i^k > D_i^k(f_i)$, $i = 1, 2$, $k = 1, \dots, m + n$, then $b(s, z)$ is an automata equilibrium outcome path.

Proof: Necessity of (1) follows from Abreu and Rubinstein, (1988), Theorem 1. Assume now that $b(s,z)$ can be generated as an automata equilibrium by the pair (M_1^*, M_2^*) . As s and z satisfy the hypotheses of Lemma 2, it follows by the definition of equilibrium that M_1^* and M_2^* have exactly $n + m$ states. Assume now that (2) does not hold for player i for $k = \tau$.

We will now construct an automaton that yields player i a payoff greater than $\pi_i(M_1^*, M_2^*)$ and uses at most $n + m$ states. This automaton will always deviate from the path $b(s,z)$ when player j 's automaton is at state q_j^τ and will conform to it otherwise. Let $r_j : S \rightarrow S_j$ be a best deviation mapping of player j in game G , i.e. $r_j(s_i, s_j) \in \operatorname{argmax}_a \{u_j(s_i, a) : a \in S_j/s_j\}$ and let $(g_\alpha^t)_{t=1}^\infty = \phi_\alpha(M_1^*, M_2^*)$, $\alpha = i, j$. Let γ be the smallest x such that $\mu_j(r_i(d^\tau), g_j^\tau) = g_j^x$. q_j^τ is the state in which the automaton of player j is at $\tau + 1$ if player i plays $r_i(d^\tau)$ at τ . Define an automaton $M_i = (Q_i, q_i^1, \lambda_i, \mu_i)$ where $Q_i = (q_i^1, \dots, q_i^{n+m})$

$$\lambda_i(q_i^t) = \begin{cases} d_i^t & \text{for } t \neq \tau \\ r_i(d^\tau) & \text{for } t = \tau \end{cases}$$

$$\mu_i(x, q_i^t) = \begin{cases} q_i^{t+1} & \text{for } t \neq \tau \text{ and } t \leq n + m - 1 \\ q_i^\tau & \text{for } t = \tau \\ q_i^{n+1} & \text{for } t = n + m \text{ and } t \neq \tau. \end{cases}$$

If (2) does not hold, by simple calculations one can show that

$$(3) \quad v_i^\tau + \delta H_i^\tau > H_i^\tau$$

If $\gamma \geq \tau$ and $\tau \leq n$ or if $\tau = \gamma$, then from (3) $\pi_i(M_i, M_j^*) > \pi_i(M_i^*, M_j^*)$, a contradiction. Otherwise, (3) implies that

$$(4) \quad v_i^r + \delta(1 - \delta^{\xi-\gamma+1})^{-1} \left(\sum_{t=\tau}^{\xi-1} \delta^{t-\tau} u_i(d^t) + \delta^{\xi-\tau} v_i^r \right) > H_i^r$$

where $\xi = \tau + (\gamma - \tau)^{-1} \max[\gamma - \tau, 0]m$.

The LHS of inequality (4) is the discounted payoff player i gets at τ , depending on the positions of τ and γ , by playing M_i at $t = 1$. Thus, we have that $\pi_i(M_i, M_j^*) > \pi_i(M_i^*, M_j^*)$, a contradiction.

Sufficiency follows as, by Lemma 2 and Lemma 3, $(M_{(s,z),1}^{f_2}, M_{(s,z),2}^{f_1})$ is an automata equilibrium. QED

Notice that (1) and (2) are not sufficient for automata equilibria. In fact, if $H_i^k = D_i^k(f_i)$, although player i cannot increase repeated game payoffs, by deviating he may be able to decrease the number of states and leave repeated game payoffs unchanged.

4. Subgame Perfection and Automata Equilibria.

Let us consider an infinite repetition of game G . Let S^n be the n -fold cartesian product of S . A repeated-game strategy for player i is a sequence $\rho_i = \{\rho_i^t\}_{t=1}^{\infty}$, where $\rho_i^1 \in S_i$, $\rho_i^t: S^{t-1} \rightarrow S_i$ if $t > 1$. A pair (ρ_1, ρ_2) induces a sequence $\eta(\rho_1, \rho_2) = \{s^t\}_{t=1}^{\infty}$, $s^t \in S$, in the obvious way. Define the payoff of player i by $w_i(\rho_1, \rho_2) = \sum_{t=1}^{\infty} \delta^{t-1} u_i(s^t)$. This game will be denoted by G^{∞} .

Theorem 1 allows us to prove the following result.

Theorem 2: Any automata equilibrium path is a subgame perfect equilibrium path of G^∞ .

Proof: Let $Z_0 = \{d^t\}_{t=1}^\infty$ be an automata equilibrium outcome path. Then, by Abreu and Rubinstein (1988), Theorem 1, there exist n and m such that $b(s, z) = Z_0$ where $s = (d^1, \dots, d^n)$ and $z = (d^{n+1}, \dots, d^{n+m})$ satisfy the assumptions of Theorem 1.

Let $Z_i = \{d_t\}_{t=f_i}^\infty$. Let us define the simple strategy profile $T(Z_0, Z_1, Z_2)$ which specifies (Abreu (1988)):

- 1) play Z_0 until some player singly deviates from Z_0
- 2) play Z_i if and only if player i deviates singly from Z_j , $j = 0, 1, 2$

As Z_0 can be sustained as an automata equilibrium, (2) must hold for both players for any $1 \leq k \leq n + m$ and therefore for any t . Then by Abreu (1988) Proposition 1, $T(Z_0, Z_1, Z_2)$ is a perfect equilibrium.

QED

Remark: An automaton M_i naturally induces the following repeated-game strategy (Kalai and Stanford (1988)): $\rho_i^1 = \lambda_i(q_1^1)$ and for any $s \in S^m$, $s = (s^1, \dots, s^m)$, $\rho_i^{m+1}(s) = \lambda_i(g_i^m(s))$ where $g_i^{t+1}(s) = \mu_i(s_j^t, g_i^t(s))$ and $g_i^1(s) = q_1^1$. Notice that Theorem 2 does not establish that the repeated game strategy profile which is induced by the automata equilibrium is subgame perfect. This is because a repeated game strategy for player i induced by an automaton like the one we

defined does not distinguish among histories which differ only by the actions played by i .

5. Conclusion

In this note we prove that any automata equilibrium path is a perfect equilibrium path for a repeated game where strategies are functions of past histories. The intuition behind this result is quite simple. As in equilibrium all the states of an automaton are used, any punishment path will be part of the equilibrium play and thus it is an equilibrium path itself.

Of course, we are far from saying that any automata equilibrium is subgame perfect. Apart from the remark after Theorem 2, there are many difficulties inherent in the usual way of making behavior off the equilibrium path matter when strategies are automata with costly states. As Abreu and Rubinstein (1988) point out, a player would not use automata of bigger sizes when reacting optimally to mistakes of arbitrary small probability. On the contrary, our results follow from the trivial observation that in an automata equilibrium we do not need extra states to deter a player from deviating.

References

- Abreu D. (1988), On the Theory of Infinitely Repeated Games with Discounting, **Econometrica**, vol. 56, 383 - 396.
- Abreu D. and Rubinstein A. (1988), The Structure of Nash Equilibrium in Repeated Games with Finite Automata, **Econometrica**, vol. 56, 1259 - 1281.
- Banks J. and Sundaram R. (1989), Repeated Games, Finite Automata and Complexity, **University of Rochester**, Working Paper No. 183.
- Blackwell D. (1965), Discounted Dynamic Programming, **Annals of Mathematical Statistics**, vol. 36, 226 - 235.
- Dudley U. (1978), Elementary Number Theory, W.H. Freeman, San Francisco.
- Kalai E. and Stanford W. (1988), Finite Rationality and Interpersonal Complexity in Repeated Games, **Econometrica**, vol. 56, 397 - 410.
- Rubinstein A. (1986), Finite Automata Play and the Prisoner's Dilemma, **Journal of Economic Theory**, vol. 39, 1 - 9.