Time Dependent Bounded Recall Strategies Are Enough to Play the Discounted Repeated Prisoners' Dilemma*

Mehmet Barlo
Sabancı University

Guilherme Carmona
Universidade Nova de Lisboa

May 7, 2004

Abstract

We show that for any discount factor, there is a natural number M such that all subgame perfect equilibrium outcomes of the discounted repeated prisoners' dilemma can be obtained by subgame perfect equilibrium strategies with the following property: current play depends only on the number of the time-index and on the history of the last M periods. Therefore, players who are restricted to using pure strategies, have to remember, at the most, M periods in order to play any equilibrium outcome of the discounted repeated prisoners' dilemma. This result leads us to introduce the notion of time dependent complexity, and to conclude that in the repeated prisoners' dilemma, restricting attention to finite time dependent complex strategies is enough.

*We are grateful to Narayana Kocherlakota, and Andy McLennan for their valuable advice and support. Special thanks are due to Hamid Sabourian for very detailed comments. We benefited from discussions in, and thank the participants of the Applied Microeconomics Workshop at the University of Minnesota, and the Mathematical Economics Workshop at the Bilkent University. We thank John Huffstot for editorial assistance. Any remaining errors are ours.

1 Introduction

Repeated games provide a framework in which long-term relationships can be analyzed. In particular, the repeated version of the prisoners' dilemma can be used to study whether or not it is possible for two individuals to cooperate even when they have a short-term incentive for not doing so. The well-known answer displays that the possibility of cooperation depends on how patient agents are. If the weight that players assign to future payoffs is sufficiently high, then cooperation will be a (subgame perfect) equilibrium outcome path.

Because players do not possess a short-term incentive to cooperate, in any equilibrium strategy that results in cooperation, current play must depend on past play. Therefore, players have to remember the past in order for cooperation to arise as an equilibrium phenomenon.

In this paper, we ask how much do players have to remember in order to play the discounted repeated prisoners' dilemma. In our view, this question is important insomuch as we regard strategies that depend less on memory as being more attractive to players, which will be the case whenever remembering the past has costs.

We show that time dependent bounded recall strategies are enough to play the discounted repeated prisoners' dilemma with pure strategies. More precisely, given the discount factor, we show that there is a natural number M such that all subgame perfect equilibrium outcome paths can be obtained by subgame perfect equilibrium strategies in which current play depends only on the number of the time-index and on what has transpired in the last M rounds of play. Thus, players have to remember no more than M periods in order to play any equilibrium outcome of the discounted repeated prisoners' dilemma. In particular, no equilibrium outcome requires an unbounded memory.

To see what is behind our main result, consider the cooperative outcome. This outcome can be obtained as a subgame perfect equilibrium outcome with the grim-trigger strategy, provided that the discount factor is high enough. In the grim-trigger strategy, any player in any round observes the complete history up to that round and will cooperate whenever the outcome has been cooperative in all the previous rounds; otherwise, he will not cooperate. In particular, in any round players' play will be affected by the outcome in the first round. Hence, the grim-trigger strategy requires unbounded recall. Fortunately, cooperation can also be obtained in equilibrium using an alternative strategy that requires less memory: players start by cooperating; in the following stages of the game, each player cooperates if and only if they both have cooperated in the previous stage. Clearly, this strategy requires only one memory. In effect, what we show is that if an outcome can be implemented by a subgame perfect equilibrium strategy, it can also be implemented by a bounded recall subgame perfect equilibrium strategy, where the bound does not exceed M.

Although we have been able to obtain a strong result for the repeated prisoners' dilemma, it appears that our tools are not easily extended to other games. Part of the difficulty of analyzing bounded recall strategies in the framework of this paper comes from restricting players to play pure strategies. In fact in Barlo and Carmona (2003) and Sabourian (1989) general characterizations using bounded recall strategies are obtained when players have a rich strategy space (e.g., when they can play mixed strategies or when they have a connected set of pure strategies).

Prisoners' dilemma has been the focus of analysis in the complexity literature: Aumann (1981), following Simon's bounded rationality ideas, suggested that a strategy is more intuitive if in every period the behavior depends on a finite number of states, i.e., if it is implemented by a finite automata. Aumann's suggestion was then followed by Neyman (1985) and Rubinstein (1986), who pioneered the analysis of complexity in the prisoners' dilemma.

In contrast, we do not employ finite automata strategies to model limited memory. Instead time dependent bounded recall separates knowledge of the time of play from that regarding the past moves, a notion used in Cole and Kocherlakota (2000) and Barlo and Carmona (2003). Consequently, even a time dependent zero-recall strategy might be of infinite complexity, which is due to the action plan varying across time (see Barlo and Carmona (2003) for a more elaborate discussion).

Our formulation is also distinguished from the one used by Aumann (1981) and others in the following important aspect. In the papers mentioned above, each player's automaton is such that its input consists only of the actions of his opponent. In particular, each player will choose the same action in two different histories provided that his opponent has chosen the same actions; this is so even if his own actions are distinct. In contrast, the strategies that we use do not have this property; in fact, players have to condition their behavior in their own past actions in order for our result to hold. Thus, we follow the formulation given in Kalai and Stanford (1988), and hence, work with unmodified notions of strategies and equilibria.

Another significance of this research involves the introduction of the notion of *time dependent complexity*, henceforth *tdc*. We follow standard interpretation of strategies in extensive form games, therefore, we regard a strategy to be a history contingent plan that is written before the game starts. Therefore, we can imagine each player having an almanac that has countably many pages, each page corresponding to a time period. Consequently, in this interpretation each agent writes his plan of action into his almanac before the game starts. Given a natural number K, a strategy profile is of at most $K - tdc$, if each agent cannot write more than K entries on any one of the pages of his almanac.

Our main result can then be restated to show that in the prisoners' dilemma (pure strategies), restricting attention to at most $K - tdc$ strategies is enough, that is, any subgame equilibrium outcome path can be obtained with an (at most) $K - tdc$ subgame perfect equilibrium strategy.

2 Notation and Definitions

The prisoners' dilemma is described as a normal form game G with two players ($N = \{1, 2\}$), each of which having two actions: $A_i = \{C, D\}$ for $i = 1, 2$. Players' payoff functions are described by the following table:

$1\backslash 2$	C	D
C	$3, 3$	$0, 4$
D	$4, 0$	$1, 1$

We denote player i's payoff function by $u_i : A \to \mathbb{R}$, for $i = 1, 2$ and $A = A_1 \times A_2$.

Let $\delta \in (0, 1)$ be the discount factor, common to both players. The discount factor will be fixed throughout the analysis. The *supergame of* G consists of an infinite sequence of repetitions of G taking place in periods $t = 1, 2, 3, \ldots$. In period t the players make simultaneous moves denoted by $a_i^t \in A_i$ and then each player learns his opponent's move.

We assume that players have complete information. For $k \geq 1$, a $k-$*stage history* is a $k-$length sequence $h_k = (a_1, \ldots, a_k)$, where, for all $1 \leq t \leq k$, $a_t \in A$; the space of all $k-$stage histories is H_k, i.e., $H_k = A^k$ (the $k-$fold Cartesian product of A). We use e for the unique 0–stage history — it is a 0–length history that represents the beginning of the supergame. The set of all histories is defined by $H = \bigcup_{n=0}^{\infty} H_n$.

For every $h \in H$, define $h^r \in A$ to be the projection of h onto its r^{th} coordinate. For every $h \in H$ we let $\ell(h)$ denote the *length of* h. For two positive length histories h and $\bar{h}$ in H we define the *concatenation of* h *and* $\bar{h}$, in that order, to be the history $(h \cdot \bar{h})$ of length $\ell(h) + \ell(\bar{h})$: $(h \cdot \bar{h}) = (h^1, h^2, \ldots, h^{\ell(h)}, \bar{h}^1, \bar{h}^2, \ldots, \bar{h}^{\ell(\bar{h})})$. We follow the convention that $e \cdot h = h \cdot e = h$ for every $h \in H$.

For a history $h \in H$ and an integer $0 \leq m \leq \ell(h) - 1$, the *m-stage tail of* h is denoted by $T^m(h) \in H : T^0(h) = e$ and $(T^m(h))^j = h^{\ell(h)-(m+1)+j}$ for $j = 1, 2, \ldots, m$ and $1 \leq m \leq \ell(h) - 1$. We also follow the convention that $T^m(h) = h$, for all $m \geq \ell(h)$.

It is assumed that at stage k each player knows h_k, that is, each player knows the actions that were played in all previous stages. As is common in this area of research, e.g. Kalai and Stanford (1988), limited memory will be modelled by restricting the strategies that players are allowed to use, and not agents' knowledge of the history of the game.

Players choose behavioral strategies, that is, in each stage k, they select a function from H_{k-1} to A_i denoted f_i^k, for player $i \in N$. The set of player i's strategies is denoted by F_i, and $F = \prod_{i \in N} F_i$ is the joint strategy space. Finally, a strategy vector is $f = \left(\left\{f_i^k\right\}_{k=1}^{\infty}\right)_{i \in N}$.

Given an individual strategy $f_i \in F_i$ and a history $h \in H$ we denote the *individual strategy induced at* h by $f_i|h$. This strategy is defined pointwise on H: $(f_i|h)(\bar{h}) = f_i(h \cdot \bar{h})$, for every $\bar{h} \in H$. We will use $(f|h)$ to denote $(f_1|h, \ldots, f_n|h)$ for every $f \in S$ and $h \in H$. We let $F_i(f_i) = \{f_i|h : h \in H\}$ and $F(f) = \{f|h : h \in H\}$.

Given a strategy of player i, $f_i \in F_i$, we say that f_i has *time dependent recall of order* m, $\text{rec}(f_i) = m$, if m is the smallest integer satisfying the property: $f_i(h) = f_i(\bar{h})$ for all $k \in \mathbb{N}$, and all $h, \bar{h} \in H_k$ satisfying $T^m(h) = T^m(\bar{h})$. If such an m does not exist, we say that $\text{rec}(f_i) = \infty$. We let F_i^m be the set of all player i's strategies with recall of order at most m, and $F^m = \prod_{i \in N} F_i^m$.

Any strategy $f \in F$ induces an outcome $\pi(f)$ as follows:

$$\pi^1(f) = f(e) \quad \pi^k(f) = f(\pi^1(f), \ldots, \pi^{k-1}(f)),$$

for $k \in \mathbb{N}$. Thus, we have defined a function $\pi : F \to A^\infty$, where $A^\infty = A \times A \times \cdots$.

The payoff in the supergame of G is, for $\delta \in (0,1)$, the discounted sum of stage game payoffs:

$$U_i(f) = (1-\delta)\sum_{k=1}^{\infty} \delta^{k-1} u_i(\pi^k(f)).$$

Let $V_i : A^\infty \to \mathbb{R}$ be defined by $V_i(a^\infty) = (1-\delta)\sum_{k=1}^{\infty} \delta^{k-1} u_i(a^k)$; then $U_i = V_i \circ \pi$.

A strategy vector $f \in F$ is a *Nash equilibrium* of the supergame of G if for all $i \in N$, $U_i(f) \geq U_i(\hat{f}_i, f_{-i})$ for all $\hat{f}_i \in F_i$. A strategy vector $f \in F$ is a *subgame perfect equilibrium* of the supergame of G if every $\bar{f} \in F(f)$ is a Nash equilibrium. We let SPE denote the set of subgame perfect equilibria and $E = \pi(SPE)$ denote the *set of subgame perfect equilibrium outcomes*. Also, we let $E^m = \pi(SPE \cap F^m)$ be the set of subgame perfect equilibrium outcomes that can be obtained by subgame perfect equilibrium strategies with recall no greater than m.

3 Subgame Perfect Equilibria

Our main result is:

Theorem 1 *There exists $M \in \mathbb{N}$ such that $E = E^M$.*

The main idea of the proof is as follows: Consider any subgame perfect equilibrium outcome π. The subgame perfect M-recall strategy we use to obtain the given subgame perfect equilibrium outcome π, is as follows: If the last M periods of a given history agree with the given outcome π, then player i chooses as specified by the outcome. Otherwise, player i chooses D.

This construction requires a player to distinguish, by remembering in a finite number of periods, if there were any deviations from the given equilibrium outcome π. Hence, first we have to analyze subgame perfect outcomes which involve unbounded sequences of consecutive (D,D)s, because clearly such a path cannot be distinguished from a punishment path using finite recall.

Lemma 1 displays that if from some period on, (D,D) will be played for the remainder of the game, then that period must be the first period of the game. This outcome consists of repeating the Nash equilibrium of the stage game, and thus, no memory is needed to implement it. Then, in Lemma 2 we show that unless defection starts in the first period, no equilibrium path will entail unbounded sequences of consecutive (D, D)s.

These properties are necessary for our result, since they guarantee that players can then conclude as to whether they are in the equilibrium path or in the punishment path. But they are not enough. The reason is that our construction critically depends on the following property: after a deviation from the given equilibrium outcome π, the behavior induced by the strategy we employ should result in a defection path for the remainder of the game. The example below is designed to demonstrate this point:

Assume that the discount factor is sufficiently high, so that the following path is a subgame perfect outcome:

$$\pi = ((C, D), (D, D), (D, D), (C, D), (C, D), (C, C), (C, C), \ldots, (C, C), \ldots) .$$

In light of Lemma 2, for the given discount factor suppose that the longest consecutive sequence of (D, D)s is 2, as in the above given outcome. Suppose that the memory figure to be used is 4. Assume that player 1 plays D in the first period. Hence, we wish the resulting outcome to be given by repeating (D, D) forever. But now consider player 1 deviating in the punishment phase and playing C in periods 4 and 5. Then the outcome of the first 5 periods is $((D, D), (D, D), (D, D), (C, D), (C, D))$. Since the history of the last 4 periods coincides with the equilibrium path, players will play (C, C) forever. Hence, a unilateral deviation does not necessarily lead to (D, D) forever.

This problem is easily solved. The length of the longest consecutive sequence of Ds played by any of the players in this outcome is 5. Then, one easily shows that players need to remember at most 6 periods. In general, if the length of the longest consecutive sequence of Ds played by any of the players in an outcome is finite, then this outcome can be implemented with a time dependent bounded recall equilibrium.

Our result then follows from our Lemma 3 below, which proves that unless defection starts in the first period, no equilibrium path will entail unbounded sequences of consecutive Ds for any of the players. Thus, confusing instances such as the one above are eliminated by setting the memory size high enough to ensure that a deviating player cannot "trick" the other.

Lemma 1 *Let $\pi \in E$ be such that*

$$\pi_n = (D, D) \quad \textit{for all } n \geq k, \tag{1}$$

for some $k \in \mathbb{N}$. Then, $k = 1$.

Remark 1 *Note that such a subgame perfect outcome is easily obtained as a result of a 0-recall strategy: players play defect in every period independent from the history.*

Proof of Lemma 1. Let π be a subgame perfect outcome satisfying condition (1). We may assume that k is the minimum $j \in \mathbb{N}$ such that $n \geq j$ implies $\pi^n = (D, D)$. Suppose, in order to reach a contradiction, that $k > 1$. Then, $k - 1 \in \mathbb{N}$ and the minimality of k implies that $\pi^{k-1} \neq (D, D)$.

Note that in period $k - 1$ the outcome cannot be D for player i and C for player j, for $i \neq j$, otherwise player j would clearly deviate. Hence, $\pi^{k-1} = (C, C)$.

Thus, in period $k - 1$ player 1's payoff is $3 + (1 - \delta) \sum_{t=1}^{\infty} \delta^t \cdot 1 = 3 + \delta$. However, by deviating and choosing D in period $k-1$ player 1 would receive a payoff of $4 + (1 - \delta) \sum_{t=1}^{\infty} \delta^t \cdot 1 = 4 + \delta$. This contradicts the assumption that π is a subgame perfect outcome. ∎

Let $\tilde{E} \subsetneq E$ be the set of equilibrium outcomes that do not satisfy (1). For $\pi \in \tilde{E}$, let k_1 be the first period in which players play (D, D); if it does not exist, we let $k_1 = 0$. Let K_1 be the maximal $k \in \mathbb{N}$ satisfying $\pi^j = (D, D)$ for all $k_1 \leq j \leq k$, unless $k_1 = 0$, in which case we let $K_1 = 0$. We refer to the (finite) sequence $\{\pi^{k_1}, \ldots, \pi^{K_1}\}$ as the *first sequence of consecutive* (D, D)*s under* π. Assuming that $k_1, k_2, \ldots, k_{n-1}$ and $K_1, K_2, \ldots, K_{n-1}$ have been defined, we let k_n be the first period after K_{n-1} in which players play (D, D); again, if it does not exist, we let $k_n = 0$. In this latter case let $K_n = 0$, while if $k_n > 0$ we let K_n be the maximal $k \in \mathbb{N}$ satisfying $\pi^j = (D, D)$ for all $k_n \leq j \leq k$. The sequence $\{\pi^{k_n}, \ldots, \pi^{K_n}\}$ is called the n^{th} *sequence of consecutive* (D, D)*s under* π. Finally, let $M_n = M_n(\pi)$ be the *length of* n^{th} *sequence of consecutive* (D, D)*s under* π, i.e., $M_n = K_n - k_n + 1$.

Lemma 2 *There exists* $B \in \mathbb{N}$ *such that* $\sup_{n \in \mathbb{N}} M_n(\pi) \leq B$ *for all* $\pi \in \tilde{E}$.

Proof. Suppose not. Then there exists $\pi \in \tilde{E}$, and $n \in \mathbb{N}$ such that $1 - \delta > 4\delta^{M_n(\pi)+1}$ and $k_n > 1$. By the definition of K_{n-1} and k_n, it follows that in period $k_n - 1$ the outcome is different from (D, D). Let $i \in N$ be such that $\pi_i^{k_n - 1} = C$. Suppose that player i deviates and plays D from period $k_n - 1$ onwards, giving rise to an outcome $\tilde{\pi}$. Note that we have $u_i(\tilde{\pi}^{k_n-1}) - u_i(\pi^{k_n-1}) = 1$, $u_i(\tilde{\pi}^k) - u_i(\pi^k) \geq 0$ for $k_n \leq k \leq K_n$ and $u_i(\tilde{\pi}^k) - u_i(\pi^k) \geq -4$ for $k > K_n$. Hence,

$$V_i(\tilde{\pi}) - V_i(\pi) \geq \delta^{k_n - 2} \left[(1 - \delta) - 4\delta^{M_n + 1} \right] > 0. \tag{2}$$

Thus, player i has an incentive to deviate in period $k_n - 1$ from π, which contradicts the assumption that π is a subgame perfect equilibrium outcome. ■

Let $i \in \{1, 2\}$ and $\pi \in \tilde{E}$; we will define the length of the n^{th} sequence of consecutive Ds played by player i under π. Let k_1^i be the first period in which player i plays D; if it does not exist, we let $k_1^i = 0$. Let K_1^i be the maximal $k \in \mathbb{N}$ satisfying $\pi_i^j = D$ for all $k_1^i \leq j \leq k$, unless $k_1^i = 0$, in which case we let $K_1^i = 0$. Assuming that $k_1^i, k_2^i, \ldots, k_{n-1}^i$ and $K_1^i, K_2^i, \ldots, K_{n-1}^i$ have been defined, we let k_n^i be the first period after K_{n-1}^i in which player i plays D; again, if it does not exist, we let $k_n^i = 0$. In this latter case let $K_n^i = 0$, while if $k_n^i > 0$ we let K_n^i be the maximal $k \in \mathbb{N}$ satisfying $\pi_i^j = D$ for all $k_n^i \leq j \leq k$. Finally, let $M_n^i = M_n^i(\pi)$ be the *length of n^{th} sequence of consecutive Ds played by player i under* π, i.e., $M_n^i = K_n^i - k_n^i + 1$.

Lemma 3 *For every $i = 1, 2$, there exists $B^i \in \mathbb{N}$ such that* $\sup_{n \in \mathbb{N}} M_n^i(\pi) \leq B^i$ *for all $\pi \in \tilde{E}$.*

Proof. Suppose not. Let B be given by Lemma 2. Then there exists $i \in \{1, 2\}$, $\pi \in \tilde{E}$, and $n \in \mathbb{N}$ such that $1 - \delta > 4\delta^{M_n^i(\pi) - B}$ and $M_n^i(\pi) > B$. This last condition implies that there exists $k_n^i \leq k \leq K_n^i$ such that $\pi^k \neq (D, D)$ and so, $\pi_{-i}^k = C$. In fact, $k \leq k_n^i + B$.

Suppose that player $-i$ deviates and plays D from period k onwards, giving rise to an outcome $\tilde{\pi}$. Note that we have $u_i(\tilde{\pi}^k) - u_i(\pi^k) = 1$, $u_i(\tilde{\pi}^l) - u_i(\pi^l) \geq 0$ for $k \leq l \leq K_n^i$ and $u_i(\tilde{\pi}^l) - u_i(\pi^l) \geq -4$ for $l > K_n^i$. Hence,

$$V_i(\tilde{\pi}) - V_i(\pi) \geq \delta^{k-1} \left[(1 - \delta) - 4\delta^{K_n^i - k + 1} \right] \geq \delta^{k-1} \left[(1 - \delta) - 4\delta^{M_n^i - B} \right] > 0. \tag{3}$$

Thus, player $-i$ has an incentive to deviate in period k from π, which contradicts the assumption that π is a subgame perfect equilibrium outcome. ■

We can now prove Theorem 1.

Proof of Theorem 1. Let $\pi \in E$. By Remark 1, we may assume that $\pi \in \tilde{E}$. Let B, B^1, B^2 be as in Lemmas 2 and 3, and let $B^* = \max\{B, B^1, B^2\}$. Thus, we have that M_n, M_n^1 and M_n^2 are all less than or equal to B^* for all $n \in \mathbb{N}$.

Define f_i as follows: start by playing according to the path π; i.e., $f_i(e) = \pi^1$. For an arbitrary $h \in H_k$, $k \in \mathbb{N}$, let:

$$f_i(h) = \begin{cases} \pi_i^{k+1} & \text{if } T^{B^*+1}(h) = (\pi_{k-B}, \ldots, \pi_k) \\ D & \text{otherwise.} \end{cases}$$

This definition implies that if player i deviates unilaterally from π in period k, then (D, D) will be played in all periods $t \geq k+1$. This follows because player $-i$ will be able to observe the deviation on the first $B^* + 1$ periods after it has occurred, which will lead to a sequence in which he plays D for $B^* + 1$ periods; after that, since along the equilibrium path player $-i$ plays at most B^* consecutive Ds, it follows that the $(B^* + 1)-$stage tail of the history following the deviation can never be consistent with equilibrium play; thus, he keeps playing D forever.

Obviously, $\text{rec}(f) = B^* + 1$ and $\pi(f) = \pi$. Furthermore, since $\pi \in E$, then f is a subgame perfect equilibrium. So, set $M = B^* + 1$. ■

4 Concluding Remarks

This paper studies a notion of limited memory and complexity in the repeated discounted prisoners' dilemma with the additional feature that agents may use time in their behavior.

Implementing equilibrium outcomes of any discounted repeated game by finite recall strategies involves the following essential aspects: First, the equilibrium outcome and its associated punishment paths have to be such that players can distinguish them by employing finite recall strategies. Second, those paths must obey the property that players can precisely identify who has deviated.

In the case of the prisoners' dilemma, the first difficulty was solved by Lemmas 1 and 2, whereas Lemma 3 deals with the second. The fact that in the prisoners' dilemma the most severe punishment path is the same for both players and has a simple structure[1] enables us to overcome the above difficulties.

Our approach allows us to introduce a new notion of complexity modelled as follows: Note that for any strategy f_i, there exists a function $g_i : \Omega_i \times \mathbb{N} \to A_i$ and a function $T_i : H \to \Omega_i$ such that $f_i(h) = g_i(T_i(h), n)$ for all $h \in H_n$, all $n \in \mathbb{N}$ and all $i = 1, 2$.[2] We then say that a strategy f_i is of *at most $K-$time dependent complexity* if it can be represented as above by (Ω_i, T_i, g_i) with $|\Omega_i| \leq K$. In this representation, one can imagine that each player possesses an almanac that (1) has countably many pages, each page

[1] That is, it consists of repetitions of the unique Nash equilibrium of the stage game.

[2] Simply let $\Omega_i = H$, T_i be the identity function and $g_i(h, n) = f_i(h)$ for all $h \in H$ and $n \in \mathbb{N}$.

corresponding to a time period, and (2) cannot contain more than K entries on any page. The number K then provides us partial information about the complexity of the almanac.[3]

For the strategies establishing Theorem 1, we can let $\Omega_i = A^M$ and $T_i = T^M$. In particular, Ω_i is finite; thus, the following corollary is then just a reformulation of our main result:

Corollary 1 *There is $M \in \mathbb{N}$ such that for any $\pi \in E$, there is a subgame perfect equilibrium strategy f with at most M-time dependent complexity such that $\pi(f) = \pi$.*

References

Aumann, R. (1981): "Survey of Repeated Games," in *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern.* Bibliographisches Institut, Mannheim.

Barlo, M., and G. Carmona (2003): "Time Dependent, Bounded Recall Equilibria in Discounted Repeated Games," *Sabancı University and Universidade Nova de Lisboa.*

Cole, H., and N. Kocherlakota (2000): "Finite Memory and Imperfect Monitoring," *Federal Reserve Bank of Minneapolis Working Paper*, (604).

Kalai, E., and W. Stanford (1988): "Finite Rationality and Interpersonal Complexity in Repeated Games," *Econometrica*, 56, 397–410.

Neyman, A. (1985): "Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoner's Dilemma," *Economic Letters*, 9, 227–229.

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

Sabourian, H. (1989): "The Folk Theorem of Repeated Games with Bounded Recall (One-Period) Memory," *Economic Theory Discussion Paper 143, University of Cambridge.*

[3]More precisely, we could define the complexity $\text{comp}(f_i)$ of a strategy f_i by $\text{comp}(f_i) = \inf\{K \in \mathbb{N} : f_i \text{ is of at most } K - \text{time dependent complexity}\}$.