Folk Theorems for the Repeated Prisoners' Dilemma with Limited Memory and Pure Strategies*

Mehmet Barlo
Sabancı University

Guilherme Carmona
Universidade Nova de Lisboa

June 15, 2007

Abstract

We establish two Folk Theorems for the repeated prisoners' dilemma with limited memory strategies and finite action spaces. The first displays that every individually rational payoff can be approximately supported by a limited memory subgame perfect strategy of finite complexity. Moreover, by adding time-dependence into the analysis, we can strengthen our first result to prove that for all payoffs bounded away from the common minmax return, the amount of memory required does not depend on how fine the approximation is. Thus, our results demonstrate that the extensive abundance of equilibrium payoffs described by the Folk Theorem is not due to players' abilities to remember actions played in distant pasts.

Journal of Economic Literature Classification Numbers: C72; C73; C79

Keywords: Repeated Prisoners' Dilemma; Memory; Bounded Rationality

Correspondent: Mehmet Barlo, FASS, Sabancı University Orhanlı; 34956 Tuzla Istanbul Turkey; Email: barlo@sabanciuniv.edu; telephone: +(90) 216 483 9284; fax: +(90) 216 483 9250.

*We are grateful to Narayana Kocherlakota, and Andy McLennan for their valuable advice and support. Special thanks are due to Efe Ok, Hamid Sabourian and an anonymous referee for very detailed comments. We thank John Huffstot for editorial assistance. Any errors that may remain are ours.

1 Introduction

Repeated games provide a framework in which long-term relationships can be analyzed. In particular, the repeated prisoners' dilemma is 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 if players are sufficiently patient, then cooperation in every period will be a (subgame perfect) equilibrium outcome. But, because players do not possess a short-term incentive to cooperate, in any equilibrium strategy that results in cooperation, current play must depend on the past.

Similar considerations apply to any (strictly) individually rational payoff vector: They can be approximated with subgame perfection due to the Folk Theorems of Fudenberg and Maskin (1986) and Aumann and Shapley (1994); and in any equilibrium strategy supporting them, players must remember and condition their play on the past. Therefore, it is reasonable to expect, as suggested by Aumann (1981), that the extensive multiplicity of equilibria described by the Folk Theorem may be reduced by restricting players to use limited memory strategies.

In contrast, this study shows that the Folk Theorem continues to hold with limited memory strategies in the repeated prisoners' dilemma with pure strategies and with both the discounting and the no-discounting payoff criteria.

Specifically, Theorem 1 establishes that any strictly individually rational payoff can be approximated with a time-independent limited memory subgame perfect strategy profile when players are sufficiently patient. It should be pointed out that the size of the time-independent memory for this result depends on how fine the approximation is. In order to make the bound on the memory independent of the degree of the approximation, we consider time-dependent strategies that allow players to condition their behavior on the time-index; and, in Theorem 2, we prove that for all $\varepsilon > 0$, there exists a natural number M_ε such that any strictly individually rational payoff bounded away by ε from the minmax return can be arbitrarily approximated employing an time-dependent M_ε-memory equilibrium strategy provided that the players are sufficiently patient.

Without any restrictions on memory the above results could be easily reached: Knowing that defection in every period is the most severe credible punishment for both of the players, we could

use an equilibrium simple strategy inducing a path with the desired payoff vector.[1] However, with limited memory such strategies may fail to be subgame perfect, and a player's deviation may fail to trigger perpetual defection by the other.

To see this, suppose that we want to implement a cycle consisting of $((C,D),(D,C))$, which yields an average payoff strictly higher than the minmax return to each player.[2] The simple strategy inducing this cycle, denoted by $\pi = \{\pi_t\}_{t=1}^{\infty}$, and involving the play of (D,D) forever for any history inconsistent with the equilibrium path, is subgame perfect with unbounded memory and with sufficiently high discount factors. However, this strategy is not subgame perfect with limited memory. This is because, if players can remember at most M periods, one player prefers to deviate at a history with its last M entries equal to $(a_1, \pi_2, \ldots, \pi_M)$ with $a_1 \neq \pi_1$ instead of playing the punishment: If $\pi_M = (D,C)$, then player 1 can play C instead of D, which will make the play return to the equilibrium outcome in the next period.[3] Hence, player 1's continuation payoff in that history strictly exceeds the payoff he would receive by not deviating.

Therefore, in contrast to the case with unbounded memory, limited memory calls for simple strategies to be carefully specified for histories when both players deviate. It should be pointed out that Sabourian (1998) also considers the Folk Theorem with limited memory and no-discounting. However, the strategy he employs is only explicitly defined for single player deviations. Consequently, it is unclear whether the conditions in his Theorem 2 are sufficient to support a cycle by a limited memory subgame perfect strategy.

The way we overcome this difficulty is by allowing the play to continue along the equilibrium path even in some histories that are inconsistent with the equilibrium path. However, this alone is not sufficient as the above example illustrates. In fact, we could modify the above strategy so that it recommends π_{M+1} at a history whose M-tail equals $(a_1, \pi_2, \ldots, \pi_M)$ with $a_1 \neq \pi_1$; and, other than this modification the strategy is unchanged. Then, player i for whom $\pi_i^M = C$, will find it profitable to deviate from D to C at a history with its M-tail equal to $(a_1, a_2, \pi_2, \ldots, \pi_{M-1})$, for any $a_2 \neq \pi_1$ and any a_1. By doing so, he produces a history with its M-tail equal to $(a_2, \pi_2, \ldots, \pi_M)$ and brings the play back to the equilibrium path. Therefore, if we continue to change the strategy by allowing the play to return to the equilibrium path at these problematic histories, an inductive argument

[1]For a definition of simple strategies, see Abreu (1988).

[2]We are thankful to an anonymous referee for providing us with a similar example.

[3]Notice that if $\pi_M = (C,D)$ by a similar argument, player 2 would deviate.

would imply that the play must be the equilibrium path after any possible history, a requirement clearly incompatible with subgame perfection. Therefore, this examples shows that the approach of changing the strategy as suggested is useful only when the equilibrium path satisfies certain properties.

This is why, for our first result, we restrict attention to the following kind of outcomes for the time-independent limited memory case: Cycles that have at least two consecutive (C, C)s followed by at least one (D, D); and ordered so that all (C, C)s come first, followed by all the (D, D)s, and the rest (without loss of generality) composed of either only (C, D)s or only (D, C)s.[4] Letting M equal to the size of the cycle, the time-independent M-memory strategy we specify recommends that at a history h with its M-tail either consistent with the equilibrium outcome, or given by $(a_1, \ldots, a_{M-m}, \pi_1, \ldots, \pi_m)$ for some $a_1, \ldots, a_{M-m}$, m greater or equal to the number of consecutive (C, C)s in the cycle, players should play the next equilibrium action. Moreover, for any other history, the strategy recommends to play D. As a result, if a player deviates singly, then the other player will play D, thus, players will never observe two consecutive (C, C)s. The reason why we need at least two consecutive (C, C)s followed by at least one (D, D) is illustrated in the following examples: Suppose first that the cycle we analyze has two consecutive (C, C)s, but not followed by a (D, D). Then, if the M-tail of a history is given by $(a_1, \ldots, a_{M-2}, (C, C), (C, C))$ for some $a_1, \ldots, a_{M-2}$, and (D, C) is the next action profile in the cycle, then player 1 can deviate: Play C and this deviation does not trigger player 2 to play D forever, because next period the M-tail of the resulting history would be $(a_2, \ldots, a_{M-1}, (C, C), (C, C))$. Second, suppose that the cycle consists of $((C, C), (D, D), (D, C))$. Then, if the M-tail of an history equals $(a_1, \ldots, a_{M-2}, (C, C), (D, D))$ for some $a_1, \ldots, a_{M-2}$, then the strategy recommends (D, C). However, player 1 can deviate and play C and this deviation does not trigger player 2 to play D forever; in contrast, the play continues along the equilibrium path. Note that even though these deviations in both of these examples are clearly not profitable, these examples illustrate that a single agent deviation does not necessarily trigger the other to play D forever.

Although there are still some restrictions when players can use time-dependent strategies, these

[4]When the desired strictly individually rational payoffs are the same for the two players, a cycle consisting of only (C, C)s and (D, D)s suffices. On the other hand, if the payoff of the second agent is strictly higher than that of the first, the cycle constructed will involve only (C, D)s following (C, C)s and (D, D)s. It should be noted that the reverse situation can then be handled with only (D, C)s.

are less severe. In fact, the kind of outcomes we consider are those that consist of the repetition of a cycle with at least one (C, C). In this case, the time-dependent M-memory strategy we specify only recommends π_t at a history h in period t if its M-tail is either consistent with the equilibrium outcome, or given by $(a_1, \ldots, a_{M-m}, \pi_{t-m}, \ldots, \pi_{t-1})$ for some $a_1, \ldots, a_{M-m}$, and with $\pi_{t-j} = (C, C)$ for some $1 \leq j \leq m$. Moreover, for any other history, the strategy recommends to play D. Then, with time-dependence, (C, C) in the right period preceded and/or followed by the equilibrium path signals whether or not players should play the equilibrium. As a result, if a player deviates singly, then the other player will play D, thus, players will never observe (C, C) in the periods in which such action profile is supposed to be observed. To illustrate further the differences between the time-dependent and the independent case, it should be pointed out that we no longer need two consecutive (C, C)s followed by a (D, D). This is because, when a player deviates in a cycle that involves at least one (C, C), the other player choosing D would imply that (C, C) in the period with the right time-index will never be observed with single agent deviations. That is, in contrast to the time-dependent case, in a period where the play should be (D, C), the first player deviating and choosing C no longer leads to equilibrium, but to punishment, because (C, C) happens in a period with the wrong time-index.

We employ this signalling ability of (C, C) to reduce the memory requirements in the time-dependent case. Specifically, we first divide any cycle into shorter segments with the use of (C, C)s. Secondly, for cycles approximating a strictly individually rational payoff bounded away from the minmax return, we show that the length of these shorter segments are bounded above and do not depend on how fine the approximation is. Hence, letting the size of the memory, M, be given by the length of the longest of these shorter segments, the strategy discussed in the previous paragraph achieves two goals: It implements the desired cycle and ensures that a single player deviation triggers the other to play D forever.

At the hearth of the above difficulties lies the restriction to pure strategies. In fact, when players' action spaces are connected and payoff functions continuous, players can encode the relevant information about the history at an arbitrarily low cost by using the richness of the action spaces. As a consequence, Barlo, Carmona, and Sabourian (2006) established that the Folk Theorem continues to hold even with time-independent one-memory strategies. However, this richness is lost with finite action spaces, and such encoding becomes difficult, if not impossible.

The 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. Aumann (1981) also introduced limited memory strategies as a special class of finite automata strategies, a concept first investigated by Kalai and Stanford (1988), Lehrer (1988) and Aumann and Sorin (1989). We refer the reader to Kalai (1990) for more details. In this study, Theorem 1 is established using finite automata strategies with limited memory, whereas, for Theorem 2 the strategy separates knowledge of the time of play from that regarding the past moves, a notion used in Barlo, Carmona, and Sabourian (2006) and Cole and Kocherlakota (2005).

2 Notation and Definitions

The prisoners' dilemma is described as a normal form game G with two players ($N = \{1, 2\}$), each of whom have 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	$4, 0$	$1, 1$

While this particular choice of payoffs simplifies the exposition, we note that none of our results depend on it. In fact, Theorem 1 requires only that D be a dominant strategy for both players, while Theorem 2 requires, in addition, that (C, C) be Pareto optimal. We denote player i's payoff function by $u_i : A \to \mathbb{R}$, for $i = 1, 2$, and $A = A_1 \times A_2$.

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 $a_i^t \in \{C, D\}$ 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 rth 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 in Kalai and Stanford (1988), limited memory will be modeled by restricting the strategies that players are allowed to use, and not agents' knowledge of the history of the game.

For all $i \in N$, a *strategy* for player i is a function $f_i : H \to A_i$ mapping histories into actions. 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 $f = (f_1, \ldots, f_n)$. 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 is a *time-dependent m-memory strategy*, if m is the smallest integer satisfying the property: $f_i(h) = f_i(\bar{h})$ for all $k \in \mathbb{N}$ and all $h, \bar{h} \in H_k$ with $T^m(h) = T^m(\bar{h})$. If such an m does not exist, we say that $\text{rec}(f_i) = \infty$. Similarly, f_i is a *time-independent m-memory strategy* if m is the smallest integer satisfying $f_i(h) = f_i(\bar{h})$ for all $h, \bar{h} \in H$ with $T^m(h) = T^m(\bar{h})$. If f_i is a strategy with m-memory, time-dependent or not, we write $\text{rec}(f_i) = m$.

A strategy $f \in F$ induces an outcome $\pi(f)$ as follows: $\pi^1(f) = f(e)$, and $\pi^k(f) = f((\pi^t(f))_{t=1}^{k-1})$, for $k = 2, 3, \ldots$. The payoff of player i in the supergame $G(\delta)$ of G is, for all $\delta \in (0, 1]$ and $i = 1, 2$,

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

if $\delta < 1$ and by

$$U_i(f) = \liminf_{T \to \infty} \frac{1}{T} \sum_{t=k}^{T} u_i(\pi^k(f))$$

if $\delta = 1$.

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.

3 Finite Memory Folk Theorems

In this section, we present our Folk Theorems for limited memory strategies.

For all $\varepsilon \geq 0$, let $\mathcal{U}^\varepsilon = \{u \in \mathrm{co}\,(u(A)) \mid u_i > 1 + \varepsilon \text{ for all } i \in N\}$. We refer to $\mathcal{U}^\varepsilon$ as the set of ε-strictly individually rational payoffs. Moreover, $\mathcal{U} \equiv \mathcal{U}^0$, denotes the set of strictly individually rational payoffs.

Theorem 1 *For all $\varepsilon > 0$, there exists $M \in \mathbb{N}$ and $\delta^* \in (0,1)$ such that for all $u \in \mathcal{U}$ and $\delta \geq \delta^*$, there exists a time-independent, subgame perfect equilibrium strategy f of $G(\delta)$ with $\mathrm{rec}(f) \leq M$ and $\|U(f) - u\| < \varepsilon$.*

Theorem 1 implies the Folk Theorem for the prisoners' dilemma with both limited memory and finite complexity. Indeed, a time-independent M-memory strategy f can be understood as a function from A^M into A; hence, its complexity is at most $|A|^M = 2^M$. Thus, Theorem 1 is related with Theorem 4.1 of Kalai and Stanford (1988) as follows: For sufficiently high discount factors, every subgame perfect payoff can be approximated by a limited memory subgame perfect (and not merely approximate subgame perfect) strategy of finite complexity.

The proof of Theorem 1 would be standard without the requirement of limited memory: A payoff vector in $\mathcal{U}$ would be approximated with a payoff vector which is strictly above one for both players, and equals the average payoff of a finite cycle. Such a payoff could then be supported as a subgame perfect payoff of a simple strategy having that cycle as the equilibrium path and a common punishment of (D, D) forever. Then, letting L be equal to 1 plus the length of this finite cycle, the complexity of such a strategy would be at most 2^L.

As explained in the introduction, the difficulty brought by limited memory arises in the process of identifying a suitable, more constrained equilibrium path to (approximately) support a desired payoff vector. In general, we need to construct strategies that not only satisfy standard incentive requirements (as in Abreu (1988)), but also can be implemented with limited memory. Making

sure that these incentive considerations hold in our setting is not very difficult, because we modify simple strategies so that a single player deviation (even though it might not be profitable) leads to (D, D) forever. However, the equilibrium path must be chosen so that coupled with the strategy employed, together they deliver the desired properties discussed in the introduction.

Proof of Theorem 1. Let $\mathbb{Q}_n = \{k/n : k \in \mathbb{Z}\}$, and $\mathcal{U}_n$ be the set of those vectors $v \in \mathbb{R}^2$ satisfying $v > 1$ and

$$v = \lambda_1(3,3) + \lambda_2(1,1) + \lambda_3(4,0) + \lambda_4(0,4)$$

for some $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) \in \Delta$ such that there exists $0 \leq k \leq n$ with $\lambda_j \in \mathbb{Q}_k$ for all $j = 1,2,3,4$, $\lambda_3\lambda_4 = 0$ and either $\lambda_1\lambda_2\lambda_3 > 0$ or $\lambda_1\lambda_2\lambda_4 > 0$. Since $\mathcal{U}_n$ converges to $\mathcal{U}$ in the Hausdorff distance, it follows for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $\mathcal{U} \subseteq \cup_{v\in\mathcal{U}_n} B_{\varepsilon/2}(v)$ for all $n \geq N$. Define $M = 2N$. We then choose $\delta^* \in (0,1)$ such that $\delta \geq \delta^*$ implies that the average payoff of a cycle of length at most M is within $\varepsilon/2$ of its discounted sum and that the simple strategy with equilibrium path consisting of that cycle and with a common punishment path of (D, D) forever is subgame perfect.

Standard arguments show that such choices are possible. Thus, it remains to show that a payoff vectors $v \in \mathcal{U}_M$ can be supported with a time-independent, limited memory strategy with $\text{rec}(f) \leq M$.

Without loss of generality suppose that $v_1 \geq v_2$, thus

$$v = \frac{p_1}{n}(3,3) + \frac{p_2}{n}(1,1) + \frac{p_3}{n}(4,0)$$

with $p_1 + p_2 + p_3 = n \leq N$, and, consider a path consisting of repetitions of the following cycle: $p_1 > 0$ times (C, C), $p_2 > 0$ times (D, D) and $p_3 > 0$ times (D, C) (the case $p_3 = 0$ is analogous). By duplicating the size of the cycle if necessary, we may assume that $p_1 \geq 2$ (and so, the length of the cycle is at most $2N = M$). For convenience, let $\pi = \{\pi_t\}_{t=1}^{\infty}$ denote such path. Also, let $(a, b; p)$ denote a repetition of the pair $(a, b) \in A$ for p times. Then, in particular, it follows that $(\pi_1, \ldots, \pi_{p_1+p_2+p_3}) = ((C, C; p_1), (D, D; p_2), (D, C; p_3))$.

Let $M = n = p_1 + p_2 + p_3$ and define the strategy f as follows (if there is the need of duplicating the length of the cycle, then $M = 2N$). In the definition of f, and in what follows, the condition

$$T^M(h) = (a_1, \ldots, a_{k-(p_1+p_2+m)}, (C, C; p_1), (D, D; p_2), (D, C; m))$$

means that there exist $a_1, \ldots, a_{k-(p_1+p_2+m)} \in \{C, D\}^2$ such that the equality holds. The same applies to other conditions similar to that. Let $h \in H$. If the length of $T^M(h)$ is equal to $k < M$, then

$$f(h) = \begin{cases} \pi_{k+1} & \text{if } T^M(h) = (\pi_1, \ldots, \pi_k), \\ (D, C) & \text{if } T^M(h) = (a_1, \ldots, a_{k-(p_1+p_2+m)}, (C, C; p_1), (D, D; p_2), (D, C; m)) \\ & \text{for some } 0 \leq m < p_3 \\ (D, D) & \text{otherwise.} \end{cases} \tag{1}$$

If, instead, the length of $T^M(h)$ is greater or equal to M, let

$$f(h) = \begin{cases} \pi_{t+M+1} & \text{if } T^M(h) = (\pi_{t+1}, \ldots, \pi_{t+M}) \text{ for some } t \in \mathbb{N}_0, \\ (D, C) & \text{if } T^M(h) = (a_1, \ldots, a_{p_3-m}, (C, C; p_1), (D, D; p_2), (D, C; m)) \text{ for some } 0 \leq m < p_3 \\ (D, D) & \text{otherwise.} \end{cases} \tag{2}$$

Define H_E as the set of histories $h \in H$ satisfying either

$$T^M(h) = (\pi_{t+1}, \ldots, \pi_{t+k}) \tag{3}$$

for some $t \in \mathbb{N}_0$, $0 \leq k \leq M$ and $t(M - k) = 0$ or

$$T^M(h) = (a_1, \ldots, a_{k-(p_1+l+m)}, (C, C; p_1), (D, D; l), (D, C; m)) \tag{4}$$

for some $0 \leq l \leq p_2$, $0 \leq m < p_3$, $(p_2 - l)m = 0$ and $p_1 + l + m < k \leq M$. As it will be shown below, H_E is the set of histories such that the play of f leads to the equilibrium path. Also, let $H_P = H \setminus H_E$, which as it will be shown below, are the set of histories such that the play of f leads to the play of (D, D) forever.

The following claim states that if players play according to f starting from an history $h \in H_E$, then they play the equilibrium path.

Claim 1 *If $h \in H_E$, then $\pi(f|h) = \{\pi_t\}_{t=k}^{\infty}$ for some $1 \leq k \leq M$.*

Proof. Let $h \in H_E$ and assume first that $T^M(h) = (\pi_{t+1}, \ldots, \pi_{t+k})$ for some $t \in \mathbb{N}_0$, $0 \leq k \leq M$ and $t(M - k) = 0$. Then, $f(h) = \pi_{t+k+1}$ and so $T^M(h \cdot f(h)) = (\pi_1, \ldots, \pi_k, \pi_{k+1})$ if $k < M$ and $T^M(h \cdot f(h)) = (\pi_{t+2}, \ldots, \pi_{t+M}, \pi_{t+M+1})$ if $k = M$. Clearly, in both cases, $h \cdot f(h) \in H_E$. Thus, by induction, players play the equilibrium path.

Finally, suppose that $T^M(h) = (a_1, \ldots, a_{k-(p_1+l+m)}, (C,C;p_1), (D,D;l), (D,C;m))$ for some $0 \leq l \leq p_2$, $0 \leq m < p_3$, $(p_2 - l)m = 0$ and $p_1 + l + m < k \leq M$. If $l < p_2$, players play (D,D) and

$$T^M(h \cdot f(h)) = \begin{cases} (a_1, \ldots, a_{k-(p_1+l)}, (C,C;p_1), (D,D;l+1)) & \text{if } k < M \\ (a_2, \ldots, a_{M-(p_1+l)}, (C,C;p_1), (D,D;l+1)) & \text{if } k = M; \end{cases}$$

if $l = p_2$, players play (D,C) and

$$T^M(h \cdot f(h)) = \begin{cases} (a_1, \ldots, a_{k-(p_1+l+m)}, (C,C;p_1), (D,D;p_2), (D,C;m+1)) & \text{if } k < M \\ (a_2, \ldots, a_{M-(p_1+l+m)}, (C,C;p_1), (D,D;p_2), (D,C;m+1)) & \text{if } k = M. \end{cases}$$

Hence, in both cases, $h \cdot f(h) \in H_E$ and so players play the equilibrium path, i.e., they play (D,D) for $p_2 - l$ periods and then (D,C) for $p_3 - m$ periods and the M-tail of the history generated by such play is $((C,C;p_1), (D,D;p_2), (D,C;p_3))$. ■

The next claim shows that if players play according to f starting from an history $h \in H_P$, then they play (D,D) forever.

Claim 2 *If $h \in H_P$, then $\pi(f|h)$ is such that $\pi_t(f|h) = (D,D)$ for all $t \in \mathbb{N}$.*

Proof. Note that if $h \in H_P$, then $f(h) = (D,D)$. So, it is enough to show that $(h \cdot f(h)) \in H_P$.

Let $T^M(h) = (a_1, \ldots, a_k)$ for some $0 \leq k \leq M$. If $k < M$, then $T^M(h \cdot f(h)) = (a_1, \ldots, a_k, (D,D))$. Suppose, in order to reach a contradiction, that $h \cdot f(h) \in H_E$. Note that it cannot be that $T^M(h \cdot f(h)) = (\pi_1, \ldots, \pi_{k+1})$ since this would imply that $T^M(h) = (\pi_1, \ldots, \pi_k)$ and so $h \in H_E$, a contradiction. Thus, if $h \cdot f(h) \in H_E$ and $f(h) = (D,D)$, it must be that

$$T^M(h \cdot f(h)) = (a_1, \ldots, a_{k+1-(p_1+l)}, (C,C;p_1), (D,D;l))$$

for some $1 \leq l \leq p_2$ and $p_1 + l < k+1$. Thus, $T^M(h) = (a_1, \ldots, a_{k-(p_1+l-1)}, (C,C;p_1), (D,D;l-1))$ and so $h \in H_E$, a contradiction.

If $k = M$, then $T^M(h \cdot f(h)) = (a_2, \ldots, a_k, (D,D))$. Thus, if $h \cdot f(h) \in H_E$ it would follow that $T^M(h \cdot f(h)) = (a_2, \ldots, a_{M-p_1-l+1}, (C,C;p_1), (D,D;l))$ for some $1 \leq l \leq p_2$ (note that $a_2, \ldots, a_{M-p_1-l+1}$ can be consistent with the equilibrium path, i.e., it allows for $T^M(h \cdot f(h)) = ((D,D;p_2 - l), (C,D;p_3), (C,C;p_1), (D,D;l)))$. Hence,

$$T^M(h) = (a_1, \ldots, a_{M-p_1-l+1}, (C,C;p_1), (D,D;l-1))$$

and so $h \in H_E$, a contradiction. ■

The following claim shows the consequence of a single player deviation from f at an history $h \in H_P$.

Claim 3 *If $h \in H_P$, $a_i \neq f_i(h)$ and $a_{-i} = f_{-i}(h)$ for some $i \in \{1,2\}$, then $h \cdot a \in H_P$.*

Proof. Since $f(h) = (D,D)$, then $a_{-i} = D$ and $a_i = C$. Thus, the conclusion is clear if $i = 1$ since in no history in H_E has (C,D) in its 1-tail. Therefore, we may focus on $i = 2$.

Let $T^M(h) = (a_1, \ldots, a_k)$ for some $0 \leq k \leq M$. If $k < M$, then $T^M(h \cdot f(h)) = (a_1, \ldots, a_k, (D,C))$. Thus, $h \cdot f(h) \in H_E$ only if $T^M(h \cdot a) = (a_1, \ldots, a_{k+1-(p_1+p_2+m)}(C,C;p_1),(D,D;p_2),(D,C;m))$ with $1 \leq m \leq p_3$ and $k \geq p_1 + p_2 + m - 1$ and so

$$h = T^M(h) = (a_1, \ldots, a_{k-(p_1+p_2+m-1)}(C,C;p_1),(D,D;p_2),(D,C;m-1)) \in H_E,$$

a contradiction.

If $k = M$, then $T^M(h \cdot f(h)) = (a_2, \ldots, a_k, (D,C))$. If $h \cdot f(h) \in H_E$, then it would follow that $T^M(h \cdot f(h)) = (a_2, \ldots, a_{M-(p_1+p_2+m)+1}, (C,C;p_1),(D,D;p_2),(D,C;m))$ for some $1 \leq m \leq p_3$. Hence, $T^M(h) = (a_1, \ldots, a_{M-(p_1+p_2+m-1)}, (C,C;p_1),(D,D;p_2),(D,C;m-1))$ and so $h \in H_E$, a contradiction. ■

It follows from Claims 2 and 3 that if $h \in H_P$ and $\{a_{-i}^t\}_{t=1}^{\infty}$ is the sequence of player $-i$'s actions induced by his strategy $f_{-i}|h$ and arbitrarily player i's actions $\{a_i^t\}_{t=1}^{\infty}$, then $a_{-i}^t = D$ for all $t \in \mathbb{N}$.

Finally, we consider the effect of a single player deviation from f at an history $h \in H_E$.

Claim 4 *If $h \in H_E$, $a_i \neq f_i(h)$ and $a_{-i} = f_{-i}(h)$ for some $i \in \{1,2\}$, then $h \cdot a \in H_P$.*

Proof. Let k be the length of $T^M(h)$. Suppose, in order to reach a contradiction, that $h \cdot a \in H_E$. If $k < M$ and $T^M(h \cdot a) = (\pi_1, \ldots, \pi_{k+1})$, then $T^M(h) = (\pi_1, \ldots, \pi_k)$ and $a = \pi_{k+1} = f(h)$, a contradiction.

If $k = M$ and $T^M(h \cdot a) = (\pi_{t+1}, \ldots, \pi_{t+M})$, then either

$$T^M(h \cdot a) = ((C,C;l),(D,D;p_2),(D,C;p_3),(C,C;p_1 - l)) \tag{5}$$

for some $0 \leq l \leq p_1$,

$$T^M(h \cdot a) = ((D,D;l),(D,C;p_3),(C,C;p_1),(D,D;p_2 - l)) \tag{6}$$

for some $0 \leq l \leq p_2$, or

$$T^M(h \cdot a) = ((D,C;l),(C,C;p_1),(D,D;p_2),(D,C;p_3-l)) \tag{7}$$

for some $0 \leq l \leq p_3$.

We consider several possible cases. If (5) holds and $l = p_1$, then $a = (D,C)$, $h = (a_1,(C,C;p_1),(D,D;p_2),(D,C;p_3-1))$ and $f(h) = (D,C) = a$, a contradiction. If (5) holds and $l = p_1 - 1$, then $a = (C,C)$ and $h = (a_1,(C,C;p_1-1),(D,D;p_2),(D,C;p_3))$. The value of f depends on a_1: If $a_1 = (C,C)$, then $f(h) = (C,C) = a$, a contradiction; if $a_1 \neq (C,C)$, then $f(h) = (D,D)$ and so $a_{-i} = C \neq D = f_{-i}(h)$, a contradiction. Finally, if (5) holds and $l < p_1 - 1$, then $a = (C,C)$, $h = (a_1,(C,C;l),(D,D;p_2),(D,C;p_3),(C,C;p_1-l-1))$ and $f(h) = (D,D)$. Hence, $a_{-i} = C \neq D = f_{-i}(h)$, a contradiction.

If (6) holds and $l = p_2$, then $a = (C,C)$, $h = (a_1,(D,D;p_2),(D,C;p_3),(C,C;p_1-1))$ and the value of f depends on a_1: If $a_1 = (C,C)$, then $f(h) = (C,C) = a$, a contradiction; if $a_1 \neq (C,C)$, then $f(h) = (D,D)$ and so $a_{-i} = C \neq D = f_{-i}(h)$, a contradiction. Finally, if (6) holds and $l < p_2$, then $a = (D,D)$, $h = (a_1,(D,D;l),(D,C;p_3),(C,C;p_1),(D,D;p_2-l-1))$ and $f(h) = (D,D) = a$, a contradiction.

If (7) holds and $l = p_3$, then $a = (D,D)$, $h = (a_1,(D,C;p_3),(C,C;p_1),(D,D;p_2-1))$ and $f(h) = (D,D) = a$, a contradiction. If (5) holds and $l < p_3$, then $a = (D,C)$, $h = (a_1,(D,C;l),(C,C;p_1),(D,D;p_2),(D,C;p_3-l-1))$ and $f(h) = (D,C) = a$, a contradiction.

Thus, we have shown so far that it cannot be that $T^M(h \cdot a) = (\pi_{t+1},\ldots,\pi_{t+k})$ for some $t \in \mathbb{N}_0$ and $0 \leq k \leq M$ such that $t(M-k) = 0$. So, if $T^M(h \cdot a) \in H_E$, it must be that $T^M(h \cdot a) = (a_1,\ldots,a_{k+1-(p_1+l+m)},(C,C;p_1),(D,D;l),(D,C;m))$ for some $0 \leq l \leq p_2$, $0 \leq m < p_3$, $(p_2-l)m = 0$ and $p_1 + l + m < k + 1 \leq M$.

If $k < M$, then $T^M(h \cdot a) = (a_1,\ldots,a_{k+1-j},\pi_1,\ldots,\pi_j)$ with $p_1 \leq j < k+1$. Hence, $h = (a_1,\ldots,a_{k-j},\pi_1,\ldots,\pi_{j-1})$. Since $h \in H_E$ and $p_1 \geq 2$, then it cannot be that $j = p_1$ and $a_{k+1-j} \neq (C,C)$. So, either $j > p_1$ or $a_{k-j} = (C,C)$. In both cases, $f(h) = \pi_j = a$, a contradiction.

Finally, consider the case of $k = M$, i.e., $T^M(h \cdot a) = (a_1,\ldots,a_{M-(p_1+l+m)},(C,C;p_1),(D,D;l),(D,C;m))$ for some $0 \leq l \leq p_2$, $0 \leq m < p_3$ and $(p_2-l)m = 0$. If $l > 0$, then we can write $T^M(h \cdot a) = (a_1,\ldots,a_{M-(p_1+l+m)},\pi_1,\ldots,\pi_{p_1+l+m})$, $a = \pi_{p_1+l+m}$,

$$T^M(h) = (a_0,a_1,\ldots,a_{M-(p_1+l+m)},\pi_1,\ldots,\pi_{p_1+l+m-1})$$

and, since $p_1 + l + m - 1 \geq p_1$, $f(h) = \pi_{p_1+l+m} = a$, a contradiction. If $l = 0$, then $T^M(h \cdot a) = (a_1, \ldots, a_{M-p_1}, (C, C; p_1))$, $a = (C, C)$ and $T^M(h) = (a_0, a_1, \ldots, a_{M-p_1}, (C, C; p_1 - 1))$. Since $h \in H_E$ and $p_1 \geq 2$, either $a_{M-p_1} = (C, C)$ or $T^M(h) = ((C, C; 1), (D, D; p_2), (D, C; p_3), (C, C; p_1 - 1))$. In the first case, $f(h) = (D, D)$ and so $a_{-i} \neq f_{-i}(h)$, a contradiction; in the second, $f(h) = (C, C) = a$, a contradiction. From this last contradiction, we conclude that $h \cdot a \in H_P$. ■

It follows from Claim 4 that if $\{a_1^t\}_{t=1}^{\infty}$ is the sequence of player 1's actions induced by his strategy $f_1|h$ and arbitrarily player 2's actions $\{a_2^t\}_{t=1}^{\infty}$, then $a_1^t = D$ for all $t > 1$.

We have shown that a single player deviation at any history leads the other to play D forever. This shows that f is a subgame perfect equilibrium for sufficiently large δ. ■

We note that the size of the memory and the complexity of the strategy used in the proof of Theorem 1 depends on how fine the approximation is. In general, as the approximation gets finer, the size of the time-independent memory explodes to infinity.

Thus, in order to obtain a Folk-Theorem-like result where the memory requirement is independent from how fine the approximation is, we consider time-dependent limited memory strategies.

Theorem 2 *For all $\varepsilon > 0$, there exists $M_\varepsilon \in \mathbb{N}$ such that for all $u \in \mathcal{U}^\varepsilon$, and all $\zeta > 0$, there is $\delta^* \in (0, 1)$ with the following property: For all $\delta > \delta^*$, there exists an equilibrium strategy f with $\mathrm{rec}(f) \leq M_\varepsilon$, and $\|U(f) - u\| < \zeta$.*

It needs to be pointed out that the only addition to the standard requirements is that the strictly individually rational payoffs under consideration must be bounded away from the common minmax payoff. Theorem 2 displays that identifying the amount with which strictly individually rational payoffs under analysis are bounded away from the minmax return, suffices to manufacture an upper bound on time-dependent memory. Moreover, clearly any payoff not bounded away from the minmax return (by $\varepsilon > 0$) can be ($\varepsilon-$)approximated in equilibrium with a payoff in $\mathcal{U}^\varepsilon$.

We note that the complexity of the strategies used to establish Theorem 2 might be large due to time-dependence. However, we may still regard them as not complex by decomposing their dependence on the past into two components. 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 $i = 1, 2$: 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 $n \in \mathbb{N}$. 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 $|\Omega_i| \leq K$.[5] In this representation, one can imagine that each player possesses an almanac that (1) has countably many pages, each page 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. More precisely, we could define the complexity $\text{comp}(f_i)$ of a strategy f_i by $\text{comp}(f_i) = \inf\{K \in \mathbb{N} : f_i \text{ is of at most } K - \text{time-dependent complexity}\}$. It is clear that for any time-dependent M-memory strategy, we can let $\Omega_i = A^M$ and $T_i = T^M$. Thus, when we restrict attention to strictly individually rational payoffs that are bounded away from the common minmax payoff, Theorem 2 shows that the Folk Theorem for the prisoners' dilemma holds with uniformly bounded time-dependent complexity.

Proof of Theorem 2. Let $\varepsilon > 0$, $\gamma = \varepsilon/2$, $\zeta > 0$ and $u \in \mathcal{U}^\varepsilon$. Then, for all $w \in \mathcal{U}^\gamma$ there exist $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4) \in \Delta$ such that $w = \lambda_1(3,3) + \lambda_2(1,1) + \lambda_3(4,0) + \lambda_4(0,4)$ and $(\lambda_2 + \lambda_3 + \lambda_4)/\lambda_1 < (3-\gamma)/\gamma$. In fact, if $w_1 = w_2$, let $\lambda_3 = \lambda_4 = 0$, and so, $w_1 > 1 + \gamma$ implies that $\lambda_1 > \gamma/2$. Hence, $(1-\lambda_1)/\lambda_1 < (2-\gamma)/\gamma < (3-\gamma)/\gamma$. If, instead, $w_2 > w_1$, then let $\lambda_3 = 0$ and it follows from $w_1 > 1 + \gamma$ that $\lambda_1 > \gamma/3$. This implies that $(1-\lambda_1)/\lambda_1 < (3-\gamma)/\gamma$. Since the case $w_1 > w_2$ is just symmetric, the proof of the above claim is completed. Because that $(3-\gamma)/\gamma = (6-\varepsilon)/\varepsilon$, let $\bar{M}_\varepsilon = \max\{1, (6-\varepsilon)/\varepsilon\}$. We may assume, without loss of generality, that $\zeta < 2\gamma$. Let v be such that $\|v - u\| < \zeta/2$. Then, it follows that $v \in \mathcal{U}^\gamma$. Moreover, v can be chosen so that $v = \lambda_1(3,3) + \lambda_2(1,1) + \lambda_3(4,0) + \lambda_4(0,4)$, where λ_j is a rational number for all $j = 1,2,3,4$ and $(1-\lambda_1)/\lambda_1 < \bar{M}_\varepsilon$.

Let $M_\varepsilon = \max\{2, \bar{M}_\varepsilon\}$. The above shows that for all $u \in \mathcal{U}^\varepsilon$ and for all $\zeta > 0$ there is a payoff vector v with the above properties and $\zeta/2$ close to u. Since the coefficients are all rational, such payoff can be obtained as the average payoff of some cycle with length K, say. We then choose $\delta^* \in (0,1)$ such that $\delta \geq \delta^*$ implies that the average payoff of a cycle of length K is within $\zeta/2$ of its discounted sum and that the simple strategy with equilibrium path consisting of that cycle and with a common punishment path of (D,D) forever is subgame perfect.

We proceed to the definition of a path π that consists of a repetition of a cycle and yields an average payoff of v. Let q and $(p_1, \ldots, p_4)$ be such that $p_j/q = \lambda_j$ for all $j = 1,2,3,4$. Then, the loop has length of q and involves p_1 occurrences of (C,C), p_2 of (D,D), p_3 of (D,C), and p_4 of

[5] As suggested by an anonymous referee, this notion is a particular case of a timed automata. Please refer to **?**.

(C, D). Let $\theta = (p_2 + p_3 + p_4)/p_1 = (\lambda_2 + \lambda_3 + \lambda_4)/\lambda_1$. The path π is defined as follows.

If $\theta < 1$, let $\alpha = 1/\theta$, let m be the greatest integer smaller or equal to α. Then, define the loop in the following way: Let $t_1 = 1$ and $t_j = t_{j-1} + m + 1$ for all $j = 1, \ldots, p_2 + p_3 + p_4$ and let $\pi^t = (C, C)$ for all $t \in \{1, \ldots, q\} \setminus \{t_1, \ldots, t_{p_2+p_3+p_4}\}$. Finally, insert the p_2 (D, D)s, the p_3 (D, C)s and the p_4 (C, D)s arbitrarily in periods $t \in \{t_1, \ldots, t_{p_2+p_3+p_4}\}$. In this case, define $M = 2$. Thus, it follows from the definition of π, that for all $k \geq 2$ there are at least one (C, C) in $T^M(\pi_1, \ldots, \pi_k) = (\pi_{k-1}, \pi_k)$.

If $\theta \geq 1$, let m be the smallest integer greater or equal to θ. Then, define the loop in the following way: Let $t_1 = 1$ and $t_j = t_{j-1} + m + 1$ for all $j = 1, \ldots, p_1$ and let $\pi^t = (C, C)$ for all $t \in \{t_1, \ldots, t_{p_1}\}$. Finally, insert the p_2 (D, D)s, the p_3 (D, C)s and the p_4 (C, D)s arbitrarily in periods $t \in \{1, \ldots, q\} \setminus \{t_1, \ldots, t_{p_1}\}$. Clearly, $m \leq \bar{M}_\varepsilon$ since $\theta \leq M_\varepsilon$. In this case, define $M = m + 1 \leq \bar{M}_\varepsilon + 1$. Thus, it follows from the definition of π, that for all $k \in \mathbb{N}$ there are at least one (C, C) in $T^M(\pi_1, \ldots, \pi_k) = (\pi_{k-M+1}, \ldots, \pi_k)$.

The strategy that implements π is defined as follows: When $h \in H_k$ with $k < M$, then $f(h) = \pi_{k+1}$ if

$$T^M(h) = (a_1, \ldots, a_l, \pi_{l+1}, \ldots, \pi_k)$$

with either $l = 0$ or with $0 < l \leq k - 1$ if there exists $l + 1 \leq j \leq k$ such that $\pi_j = (C, C)$; otherwise, $f(h) = (D, D)$. When $h \in H_k$ with $k \geq M$, then $f(h) = \pi_{k+1}$ if

$$T^M(h) = (a_1, \ldots, a_l, \pi_{l+k-M+1}, \ldots, \pi_k)$$

for some $0 \leq l \leq M - 1$ and there exists $l + k - M + 1 \leq j \leq k$ such that $\pi_j = (C, C)$; otherwise, $f(h) = (D, D)$.

Define $H_E = \cup_k H_{E,k}$ and, for all $k \geq$, $H_{E,k}$ as the set of histories $h \in H_k$ such

$$T^M(h) = (a_1, \ldots, a_l, \pi_m, \ldots, \pi_k)$$

satisfying $\pi_j = (C, C)$ for some $m \leq j \leq k - 1$, $l + k - m + 1 = k$ and $0 \leq l \leq k - 2$ if $k < M$ and $l + k - m + 1 = M$ and $0 \leq l \leq M - 2$ if $k \geq M$. Also, let $H_{E,0}$ be equal to the initial history and $H_{E,1} = \{\pi_1\}$. Finally, let $H_P = H \setminus H_E$.

The following claim states that if players play according to f starting from an history $h \in H_E$, then they play the equilibrium path.

Claim 5 *If $h \in H_E$, then $\pi(f|h) = \{\pi_t\}_{t=k}^{\infty}$ for some $1 \leq k \leq M$.*

Proof. Let $h \in H_{E,k}$, $k \in \mathbb{N}$, and note that $f(h) = \pi_{k+1}$. Assume first that $k < M$. Then, $h = (a_1, \ldots, a_l, \pi_{l+1}, \ldots, \pi_k)$ with $l = 0$ or with $\pi_j = (C, C)$ for some $l+1 \leq j \leq k$ and $0 < l \leq k-1$. Then, $h \cdot f(h) \in H_{E,k+1}$ since $T^M(h \cdot f(h)) = (a_1, \ldots, a_l, \pi_{l+1}, \ldots, \pi_k, \pi_{k+1})$.

Finally, consider $k \geq M$. Since $h \in H_{E,k}$, it follows that $T^M(h) = (a_1, \ldots, a_l, \pi_{l+k-M+1}, \ldots, \pi_k)$ with $\pi_j = (C, C)$ for some $l + k - M + 1 \leq j \leq k$ and $0 \leq l \leq M - 1$. Then,

$$T^M(h \cdot f(h)) = \begin{cases} (a_2, \ldots, a_l, \pi_{l+k-M+1}, \ldots, \pi_{k+1}) & \text{if } l \geq 1 \\ (\pi_{l+k-M+2}, \ldots, \pi_{k+1}) & \text{if } l = 0; \end{cases}$$

Hence, in both cases, $h \cdot f(h) \in H_{E,k+1}$. It follows by induction that players play the equilibrium path. ■

The next claim shows that if one player plays according to f starting from an history $h \in H_P$, then he plays D forever.

Claim 6 *If $h \in H_P$, $a_i = f_i(h)$ for some $i \in \{1, 2\}$, then $h \cdot a \in H_P$ for all $a_{-i} \in \{C, D\}$.*

It follows from this claim that if $\{a_i^t\}_{t=1}^{\infty}$ is the sequence of player i's actions induced by his strategy $f_i|h$ and arbitrarily player $-i$'s actions $\{a_{-i}^t\}_{t=1}^{\infty}$, then $a_i^t = D$ for all $t \in \mathbb{N}$.

Proof. Since $f(h) = (D, D)$, then $a_i = D$. If $h{\cdot}a \in H_{E,k}$, $k \leq M$ and $h{\cdot}a = (a_1, \ldots, a_l, \pi_{l+1}, \ldots, \pi_k)$ with $\pi_j = (C, C)$ for some $l + 1 \leq j \leq k$ and $0 < l \leq k - 1$, then it must be that $j \leq k - 1$ and $l \leq k - 2$. Indeed, $\pi_k = a \neq (C, C)$ and so the (C, C) appear on $(\pi_{l+1}, \ldots, \pi_{k-1})$. Thus, $h = (a_1, \ldots, a_l, \pi_{l+1}, \ldots, \pi_{k-1}) \in H_{E,k-1}$, a contradiction. If, instead, we have that $l = 0$, then $h \cdot a = (\pi_1, \ldots, \pi_k)$ and so $h = (\pi_1, \ldots, \pi_{k-1}) \in H_{E,k-1}$, a contradiction.

Finally, consider $k > M$. If $h \cdot a \in H_{E,k}$, then $T^M(h \cdot a) = (a_1, \ldots, a_l, \pi_{l+k-M+1}, \ldots, \pi_k)$ with $\pi_j = (C, C)$ for some $l + k - M + 1 \leq j \leq k$ and $0 \leq l \leq M - 1$. Since $\pi_k = a \neq (C, C)$ then the (C, C) appear on $(\pi_{l+k-M+1}, \ldots, \pi_{k-1})$ and so it follows that $j \leq k - 1$ and $l \leq M - 2$. Since $T^M(h) = (a_0, a_1, \ldots, a_l, \pi_{l+k-M+1}, \ldots, \pi_{k-1})$ for some a_0, it follows that $h \in H_{E,k-1}$, a contradiction. ■

Claim 7 *If $h \in H_E$, $a_i = f_i(h)$ and $a_{-i} \neq f_{-i}(h)$ for some $i \in \{1, 2\}$, then $h \cdot a \in H_P$.*

Thus, it follows that if $\{a_i^t\}_{t=1}^{\infty}$ is the sequence of player i's actions induced by his strategy $f_i|h$ and arbitrarily player $-i$'s actions $\{a_{-i}^t\}_{t=1}^{\infty}$, then $a_i^t = D$ for all $t > 1$.

Proof. Let $k \in \mathbb{N}$ be such that $h \in H_{E,k-1}$. Then, $f(h) = \pi_k$. Thus,

$$T^M(h) \neq (a_1, \ldots, a_l, \pi_m, \ldots, \pi_k)$$

for all j, l, m satisfying $m \leq j \leq k-1$, $l + k - m + 1 = k$ and $0 \leq l \leq k-2$ if $k < M$ or $l + k - m + 1 = M$ and $0 \leq l \leq M-2$ if $k \geq M$. Thus, $h \cdot a \in H_P$. ■

We have shown that a single player deviation by any player lead the other to play D forever. This shows that f is a subgame perfect equilibrium for sufficiently large δ. ■

4 Concluding Remarks

Implementing equilibrium outcomes of any discounted repeated game with limited memory strategies involves the following essential aspects: First, the equilibrium outcome and its associated punishment paths have to be such that players can distinguish them with limited memory. Second, the paths must obey the property that players can precisely identify who has deviated. And third, at any history (including off the equilibrium and off the punishment paths) any player who deviates singly must be punished, in particular, such deviations should not be leading the play to the equilibrium or the punishment paths.

In the case with time-independent one-memory, these properties were fully characterized by Barlo, Carmona, and Sabourian (2006) using the notion of confusion-proofness of a simple strategy. It is intuitively clear that with M-memory (and possibly with time-dependence) we have more flexibility, but the characterization of M-memory confusion-proof simple strategies is still not known to us. In this paper, with the help of the structure of the prisoners' dilemma, instead of characterizing M-memory simple strategies, we identify certain outcomes and strategies that are sufficient for our results. However, such a characterization appears to be necessary in order to extend our analysis to more general games and low discounting.

With low discounting and finite actions, the construction of an M-memory subgame perfect simple strategy becomes quite difficult. This is because, when designing an outcome path approximating a desired payoff with a M-memory simple strategy, the encoding ability given by placing (C, C)s in the right periods and players' incentive constraints are not necessarily compatible. In contrast, this difficulty can be eliminated with rich action spaces: Barlo, Carmona, and Sabourian (2006) shows that every subgame perfect outcome path can be approximated with a time-independent

one-memory equilibrium strategy even for low discount factors. Moreover, as our results show, this difficulty can also be eliminated when the discount factor is high even when action spaces are finite.

It turns out that the nature of this difficulty is rather subtle. Indeed, one can show that given a discount factor strictly less than 1, the number of consecutive Ds played by a player along any subgame perfect path is bounded. Using this property, one can show that any subgame perfect outcome can be supported by a time-dependent limited memory Nash equilibrium strategy. The strategy used in this construction would have a memory size equal to the maximal number of consecutive Ds played by a player on the equilibrium outcome, would recommend continuing with the equilibrium outcome in any history whose M-tail is consistent with the equilibrium outcome, and the play of (D, D) otherwise. However, as the first example in the introduction illustrates, such a strategy may fail to be a Nash equilibrium in subgames reached by several deviations of both players.

References

Abreu, D. (1988): "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, 56, 383–396.

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

Aumann, R., and L. Shapley (1994): "Long-Term Competition – A Game-Theoretic Analysis," in *Essays in Game Theory in Honor of Michael Maschler*, ed. by N. Megiddo. Springer-Verlag, New York.

Aumann, R., and S. Sorin (1989): "Cooperation and Bounded Recall," *Games and Economic Behavior*, 1, 5–39.

Barlo, M., G. Carmona, and H. Sabourian (2006): "Limited Memory in Repeated Games," unpublished.

Cole, H., and N. Kocherlakota (2005): "Finite Memory and Imperfect Monitoring," *Games and Economic Behavior*, 53, 59–72.

Fudenberg, D., and E. Maskin (1986): "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54, 533–554.

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

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

Lehrer, E. (1988): "Repeated Games with Stationary Bounded Recall Strategies," *Journal of Economic Theory*, 46, 130–144.

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. (1998): "Repeated Games with M-period Bounded Memory (Pure Strategies)," *Journal of Mathematical Economics*, 30, 1–35.