SonderForschungsBereich 504
Rationalitätskonzepte,
Entscheidungsverhalten und
ökonomische Modellierung

No. 07-58

PERCEPTRON VERSUS AUTOMATON&8727;

Sylvain Béal*

August 2007

Financial support from the Deutsche Forschungsgemeinschaft, SFB 504, at the University of Mannheim, is gratefully acknowledged.

*Sonderforschungsbereich 504, email: sylvain.beal@awi.uni-heidelberg.de

Universität Mannheim
L 13,15
68131 Mannheim

PERCEPTRON VERSUS AUTOMATON*

Sylvain Beal†

August 6, 2007

Abstract

We study the finitely repeated prisoner's dilemma in which the players are restricted to choosing strategies which are implementable by a machine with a bound on its complexity. One player must use a finite automaton while the other player must use a finite perceptron. Some examples illustrate that the sets of strategies which are induced by these two types of machines are different and not ordered by set inclusion. The main result establishes that a cooperation in almost all stages of the game is an equilibrium outcome if the complexity of the machines players may use is limited enough. This result persists when there are more than T states in the player's automaton, where T is the duration of the repeated game. We further consider the finitely repeated prisoner's dilemma in which the two players are restricted to choosing strategies which are implementable by perceptrons and prove that players can cooperate in most of the stages provided that the complexity of their perceptrons is sufficiently reduced.

Keywords: prisoner's dilemma, finitely repeated games, machine games, automaton, perceptron, bounded complexity.
JEL Classification number: C72, C73.

1 Introduction

The best-understood class of dynamic games is that of repeated games for which well known results, called folk theorems, delineate the set of equilibrium outcomes. One of these results asserts that any feasible and individually rational payoff vector is obtained at some equilibrium of the infinite

*I am grateful to Richard Baron and Philippe Solal for helpful suggestions and remarks. I have also benefited from comments of seminar participants at Institut Henri Poincaré, University of Heidelberg, EURISE and at the conference Complex Behavior in Economics: Modeling, Computing and Mastering Complexity. Support by the German Science Foundation (DFG), through SFB 504, is gratefully acknowledged.

†Alfred Weber Institut, Universität Heidelberg, Grabengasse 14, 69117 Heidelberg, Germany and CREUSET, University of Saint-Etienne, France. E-mail : sylvain.beal@awi.uni-heidelberg.de

repetition of the game. Benoît, Krishna (1987) determine the conditions under which the main insights underlying this infinite-horizon folk theorem are maintained in finite time horizons. However, in some cases, the set of equilibrium outcomes of the finitely repeated game differs sharply from that of its infinitely repeated counterpart. For instance, the unique equilibrium outcome of the T-period repeated prisoner's dilemma is the one in which all players defect in all periods. This uniqueness result is disturbing in light of experiments with this game in which real players do not always choose the dominant action of defecting (See Axelrod, 1984 for references). This paper explores one attempt to resolve this finite horizon paradox in the finitely repeated prisoner's dilemma by restricting the strategies players may use.

There are numerous methods of restricting strategies. The approach we adopt in this paper consists in bounding the complexity of players' strategies. Players are restricted to choosing strategies which are implementable by a machine and the complexity of a strategy is measured by the size of the smallest machine that implements it. Two types of machines are considered : the perceptron and the finite automaton. More explicitely, player 1 is restricted to perceptrons not exceeding a certain number k^* of classifiers while player 2 is restricted to finite automata not exceeding a certain number m^* of states. In contrast to the proximate litterature (Neyman, 1985, 1998 and Neyman, Okada, 2000) where at least one player can implement any of the strategies available to his opponent, it turns out that, in our model, one cannot determine which of the players has the greatest ability to implement repeated game strategies.[1] On one side, the perceptron summarizes any history by an aggregate according to the empirical frequency of each stage-game outcome and a threshold decision rule dictates the player's decision. Therefore, the perceptron processes the entire history, but imperfectly since it cannot distinguish between two different histories with the same distribution of outcomes. On the other side, the automaton dictates the player's decision according to the last opponent's action. Therefore, an automaton does not process the entire history, but its states may perfectly keep track of some of the last opponent's moves. The present paper contains two examples which show that the set of repeated game strategies induced by perceptrons does not include, nor is included by, the set of repeated game strategies induced by finite automata. For that reason, it makes sense to ask whether equilibrium outcomes are affected by such heterogeneity. In particular, we wonder

[1]Exceptions are the situations in which $k^* \geq T$ and $m^* \geq T$. In fact, any sequence of T actions is induced by some automata with T states and we prove that this statement holds for perceptrons with T classifiers (proposition 1). Nevertheless, it does not exactly amount to say that such "complex" machines can induce any repeated game strategies.

if players with non comparable sets of strategies can achieve some mode of cooperation in the finitely repeated prisoner's dilemma.

This paper continues the line of research on machine games suggested by Aumann (1981) and initiated by Neyman (1985) for the finitely repeated prisoner's dilemma in which the two players use finite automata. Neyman (1985) proves that a cooperation in all stages is an equilibrium outcome if the players are restricted to choosing automata with less than T states and that the cooperative average payoff can be achieved in mixed strategies even when the number of states in these automata is reasonably larger than T. For instance, a pair of "grim trigger" two-state automata is an equilibrium that results in the cooperative outcome in each stage of the game. Such a cooperative equilibrium path fails to exist in our machine game due to the ability of a perceptron to defect in stage T if the cooperative outcome is induced in each of the preceding stages. It turns out that there is no equilibrium outcome in which the players cooperate in each stage, for any duration of the game and any complexity of the machine of each player (proposition 2). Neyman (1998) extends his seminal result to the class of two-player finitely repeated games played by finite automata and shows that equilibrium payoffs coincide with those of the folk theorem for infinitely repeated games provided that the sizes of the automata chosen by the players are also not too large compared to T. Our approach departs from that of Neyman (1985, 1998) in that each player in our model uses a different type of machine. In particular, the current paper is a first attempt to use perceptrons to bound the complexity of a player's strategies in a finitely repeated game since its introduction in game theory by Cho (1994).[2] The absence of equilibrium in which the cooperative outcome is induced in each stage of the game does no preclude equilibria in which players achieve some measure of cooperation. More precisely, each player can ensure a payoff arbitrarily close to the cooperative one provided that one player is restricted to perceptrons with no more than one classifier and the other player is restricted to automata with no more than two states (proposition 3). Surprisingly, the result persists when the number of states in player 2's automaton is larger than T (proposition 4). In that case, it is worth emphasizing that player 2 can behave as if he were unrestricted since he can choose an automaton which induces any sequence of T actions. Next, we consider a variant of the model in which both players are assumed to use perceptrons with limited complexity. It appears that if one player is

[2] Cho (1994) recovers the perfect folk theorem (Fudenberg, Maskin, 1986) for two-player infinitely repeated games when the players must use strategies which are implementable by perceptrons. The perceptron has also been studied in infinitely repeated games by Cho (1995, 1996a, 1996b) and Cho, Li (1999), and in economics by Rubinstein (1993).

restricted to perceptrons with at most one classifier and the other player can choose perceptrons with at most five classifiers, there is an equilibrium that results in the cooperative outcome in all but the very last stages of the game (proposition 6).

The idea to confront players with different abilities to implement strategies has been examinated in two-player repeated games in which a player restricted to choosing finite automata plays against an unrestricted player. In two-player finitely repeated games, Neyman, Okada (2000) prove a folk theorem which includes pairs of equilibrium payoffs in which the "restricted" player obtains no more than his maxmin payoff. In infinitely repeated games, Gilboa, Samet (1989) restrict one player to strategies which are implementable by finite connected automata *i.e.* automata which can be represented by a connected "transition diagram". They show that the "unrestricted" player has a dominant strategy in the repeated game but cannot prevent the "restricted" player to secure almost his best feasible and individually rational payoff. These two approaches differ from ours since the "unrestricted" player can induce any of the strategies available to the "restricted" player.

The rest of the paper is organized as follows. Section 2 introduces the prisoner's dilemma, the finite automaton and the perceptron. Examples illustrating that the sets of strategies induced by the two types of machines are distinct and not ordered by set inclusion are given is section 3. Section 4 contains the results for the repeated prisoner's dilemma in which one player is restricted to choosing finite automata and the other player is restricted to choosing perceptrons. The variant of the model in which both players are restricted to using strategies which are implementable by perceptrons is studied in section 5. Various aspects of the model are discussed in section 6.

2 Preliminaries

2.1 Prisoner's dilemma

Let $G = (A_1, A_2, \pi_1, \pi_2)$ be a two-person prisoner's dilemma game in normal form. $A_i = \{C, D\}$ is the set of actions for player $i = 1, 2$ and $a = (a_1, a_2) \in A_1 \times A_2 =: A$ is called an action pair. For each $i = 1, 2$, $\pi_i : A \longrightarrow \mathbb{R}$ is player i's payoff function defined by table 1 where $b, c > 0$ and $c - b < 1$. Since action D is stricly dominant, (D, D) is the unique Nash equilibrium but (C, C) is the efficient outcome. Player i can secure a null payoff against any action of the opponent, *i.e.*, his minmax payoff is 0. We will refer to $G = (A_1, A_2, \pi_1, \pi_2)$ as the **stage game**. Throughout, $-i$ will denote the opponent of player i.

	C	D
C	$1, 1$	$-b, 1+c$
D	$1+c, -b$	$0, 0$

Table 1: *The prisoner's dilemma.*

2.2 Finitely repeated game

Let G^T be the supergame obtained by repeating G at stages $t = 1, 2, \ldots, T$, $T \in \mathbb{N}$. A history $h^t = (a^1, \ldots, a^{t-1})$ at the beginning of stage t records the actions taken by each player in stages $1, 2, \ldots, t-1$. Let $H^t = A^{t-1}$ be the set of all histories at stage t and $\mathcal{H} = \bigcup_{t=1}^{T} H^t$ be the set of all histories, where $H^1 = \{\emptyset\}$ is the null history at the beginning of the game.

A pure strategy s_i for player i is a sequence of functions $\{s_i^t\}_{t=1}^T$ where, for each $t \geq 1$, the function $s_i^t : H^t \longrightarrow A_i$ determines player i's action at stage t, denoted by $s_i(h^t) \in A_i$, as a function of the previous $t-1$ action pairs $h^t \in H^t$. Let S_i be the set of all strategies for player i and $S := S_1 \times S_2$.

Each pair of strategies $s = (s_1, s_2) \in S$ induces a unique sequence of action pairs $(a^t)_{t=1}^T$. The fonction $f_i : S \longrightarrow \mathbb{R}$ assigns to each pair of strategies $s \in S$ the average payoff

$$f_i(s) = \frac{1}{T} \sum_{t=1}^{T} \pi_i(a^t)$$

for player $i = 1, 2$. A pair of strategies $s \in S$ is a Nash equilibrium of the repeated game G^T if for each $i = 1, 2$ and each $s_i' \in S_i$, $f_i(s) \geq f_i(s_i', s_{-i})$. In the finitely repeated prisoner's dilemma G^T, the unique equilibrium outcome consists in T action pairs (D, D).

2.3 Machine game

We now turn to the **machine game** that consists in the repeated game G^T in which the players are restricted to choosing strategies which are implementable by a machine with bounded complexity. More precisely, player 1 must choose a perceptron with no more than $k^* \in \mathbb{N}$ classifiers and player 2 must choose a finite automaton with no more than $m^* \in \mathbb{N}$ states.

A finite **perceptron** for player 1 is a triple $\psi_1 = (Z, \mathcal{K}, R)$ where

1. For each $t \geq 2$, a summary function $Z : A \times H^t \longrightarrow \mathbb{R}$ assigns to each $(a, h^t) \in A \times H^t$ the empirical frequency

$$Z(a, h^t) = \frac{1}{t-1} \#\{\tau \leq t-1 : a^\tau = a\}$$

of a in h^t. By definition, $\sum_{a \in A} Z(a, h^t) = 1$ for each $h^t \in H^t$;

2. A finite collection $\mathcal{K}$ of k classifiers.[3] A classifier $K_l \in \mathcal{K}$ is a triple $K_l = (\alpha_l, \beta_l^1, d_l)$, where the function $\alpha_l : A \longrightarrow \mathbb{R}$ assigns a synaptic weight $\alpha_l(a)$ to each $a \in A$. The value associated to h^t by K_l is

$$\beta_l^t = \sum_{a \in A} \alpha_l(a) Z(a, h^t) + \frac{\beta_l^1}{t-1},$$

where $\beta_l^1 \in \mathbb{R}$ is the initial value of K_l. The decision unit used throughout this paper is the threshold function $d_l : \mathbb{R} \longrightarrow \{0, 1\}$ defined by

$$d_l(\beta_l^t) = \begin{cases} 1 & \text{if } \beta_l^t \geq 0, \\ 0 & \text{if } \beta_l^t < 0. \end{cases}$$

To make short, let $d_l := d_l(\beta_l^t)$.

3. A decision function $R : \{0, 1\}^k \longrightarrow A_1$ which assigns to each vector $(d_1, \ldots, d_k) \in \{0, 1\}^k$ an action $R(d_1, \ldots, d_k) \in A_1$.

In stage 1, the perceptron dictates the action to be played according to the initial values of the classifiers. Let Ψ_1 be the set of all finite perceptrons for player 1.

A **finite automaton** or **Moore machine** M_2 for player 2 is a four-tuple $(Q_2, q_2^1, \lambda_2, \mu_2)$ where

1. Q_2 is the finite set of m states in M_2 ;

2. q_2^1 is the initial state ;

3. $\lambda_2 : Q_2 \longrightarrow A_2$ is the output function which plays action $\lambda_2(q_2) \in A_2$ whenever M_2 is in state q_2 ;

[3] When $k = 0$ the perceptron has no classifier and plays the same action in all stages regardless of the opponent's strategy.

4. $\mu_2 : Q_2 \times A_1 \longrightarrow Q_2$ is the transition function. In stage t, if M_2 is in state $q_2 \in Q_2$ and player 1 chooses $a_1 \in A_1$, then the machine's next state is $\mu_2(q_2, a_1) \in Q_2$.

Let $\mathcal{M}_2$ be the set of all finite Moore machines for player 2. In the machine game G^T, a pair $(\psi_1, M_2) \in \Psi_1 \times \mathcal{M}_2$ is a Nash equilibrium is for all $\psi_1' \in \Psi_1$ and all $M_2' \in \mathcal{M}_2$, $f_1(\psi_1, M_2) \geq f_1(\psi_1', M_2)$ and $f_2(\psi_1, M_2) \geq f_2(\psi_1, M_2')$.

3 Examples

In order to illuminate how the two types of machines induce a strategy, let us construct an automaton and a perceptron which implement the GRIM TRIGGER strategy. Next we will examine a couple of examples to see that perceptrons and automata have distinct abilities to induce repeated game strategies. The GRIM TRIGGER strategy s_i for player i is defined for $t = 1$ by $s_i^1(\emptyset) = C$ and for each $t \geq 2$ by

$$s_i^t(h^t) = \begin{cases} C & \text{if } a_{-i}^\tau = C,\ \forall \tau = 1, \ldots, t-1, \\ D & \text{otherwise.} \end{cases}$$

The perceptron ψ_1^1 for player 1 with a single classifier K_1 and the two-state automaton M_2^1 for player 2 which implement this strategy are given in figure 1.[4]

$$K_1^1 : \quad \begin{aligned} &\alpha_1(C,C) = 0 \\ &\alpha_1(C,D) = -2 \\ &\alpha_1(D,C) = 0 \\ &\alpha_1(D,D) = -2 \\ &\beta_1^1 = 1 \end{aligned}$$

M_2^1 :

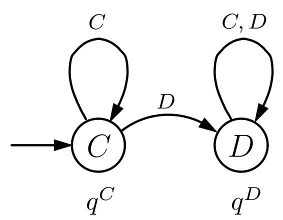

Figure 1: *Perceptron and automaton for the* GRIM TRIGGER *strategy.*

As suggested by Cho (1995, footnote page 268) for infinitely repeated games, some strategies are induced by a perceptron but not by a finite automaton. This is also the case for finitely repeated games even if the complexity of the perceptron is considerably reduced. As a first example, consider the prisoner's dilemma repeated for $T > m^*$ stages. Suppose that player 1 chooses the GRIM TRIGGER strategy and that $m_2 \leq m^*$ states are allowed in player 2's automaton M_2. Such an automaton cannot implement a strategy

[4]Observe that M_2^1 is not exactly equivalent to s_i since the automaton does not specify how to behave after its own deviation.

that consists in playing C in the first m^* stages and D thereafter. A perceptron with a unique classifier can induced such a strategy (this machine is constructed in the proof of proposition 2).

The next example shows that the strategy induced by a four-state automaton cannot be induced by a perceptron even with an arbitrarily large number of classifiers.

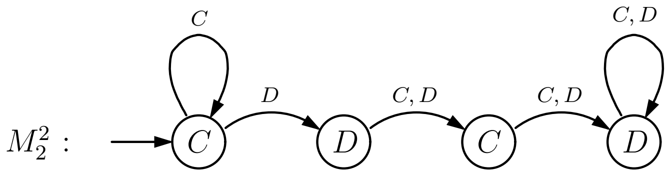

Figure 2: *Strategy which cannot be induced by a perceptron.*

Example 1 Suppose $T \geq 5$, $m^* = 4$ and $k^* \in \mathbb{N}$. Consider the strategy s_a induced by the four-state automaton represented in figure 2. A perceptron cannot induce this strategy. We argue by contradiction. Firstly, assume that player 1 can construct a perceptron ψ_1 with a unique classifier K_1 that implements s_a. Consider two sequences $h_2^5 = (C, C, D, D)$ and $h_2'^5 = (C, D, D, C)$ of four actions of player 2. According to s_a, the perceptron ψ_1 must react by playing in the first five stages the sequences of actions $h_1^6 = (C, C, C, D, C)$ and $h_1'^6 = (C, C, D, C, D)$ respectively. Observe that the two histories

$$h^5 = \begin{pmatrix} C & C & C & D \\ C & C & D & D \end{pmatrix} \quad \text{and} \quad h'^5 = \begin{pmatrix} C & C & D & C \\ C & D & D & C \end{pmatrix}$$

induced at the beginning of stage 5 contain exactly the same action pairs, even if the order in which they occur is different. As a consequence, the value β_1^5 associated to these two histories by K_1 is identical. It follows that ψ_1 must play the same action in response to h_2^5 and $h_2'^5$, which is contradictory with the two sequences h_1^6 and $h_1'^6$ specified by s_a. We conclude that s_a cannot be induced by a perceptron with at most one classifier. The same argument applies for each classifier in a perceptron with finitely many classifiers. $\square$

4 Perceptron versus automaton

In this section, we study the set of achievable pairs of equilibrium payoffs in the finitely repeated prisoner's dilemma defined in table 1. Player 1 must choose strategies which are implementable by perceptrons with $k \leq k^*$ classifiers and player 2 must choose strategies which are implementable by automata with $m \leq m^*$ states. The first result states that if there is at least

T classifiers available to player 1, then he can construct a perceptron which induces any sequence of T actions.

Proposition 1 *Assume $T \in \mathbb{N}$, $m^* \in \mathbb{N}$ and $k^* \geq T$. For any automaton M_2 of player 2 and any strategy $s_1 \in S_1$, player 1 can construct a perceptron ψ_1 which induces the sequence of T actions played by s_1 against M_2.*

Proof. Assume $T \in \mathbb{N}$, $m^* \in \mathbb{N}$ and $k^* \geq T$. Consider any automaton M_2 of player 2 and any repeated game strategy $s_1 \in S_1$. Let $(a_1^1, \ldots, a_1^T)$ be the sequence of T actions played by s_1 against M_2. We construct a perceptron ψ_1^2, with T classifiers denoted by $K_1, \ldots, K_T$, which induces $(a_1^1, \ldots, a_1^T)$. Each classifier K_l, $l = 1, \ldots, T$, is defined by :

$$K_l : \quad \alpha_l(a) = -1, \; \forall a \in A$$
$$\beta_l^1 = l - 2$$

The decision function $R : \{0,1\}^T \longrightarrow \{C, D\}$ is defined by

$$R(d_1, \ldots, d_T) = \begin{cases} a_1^t & \text{if } \displaystyle\sum_{l=1}^{T} d_l = T - t, \\ a_1 \in \{C, D\} & \text{otherwise.} \end{cases} \tag{1}$$

For each $l = 1, \ldots, T$ and each $a \in A$, we have $\alpha_l(a) = -1$. This implies that the actions chosen by the perceptron ψ_1^2 do not depend on the opponent's automaton M_2. In the first stage, $\beta_1^1 = -1$ and $\beta_l^1 \geq 0$ for each $l > 1$. Thus, $d_1 = 0$ and $d_l = 1$ for each $l = 2, \ldots, T$. It follows that

$$\sum_{l=1}^{T} d_l = T - 1,$$

which implies that $R(d_1, \ldots, d_l) = a_1^1$. In the second stage, $\beta_l^2 = l - 3$, for each $l = 1, \ldots, T$, regardless of the action played by the opponent in stage 1. Therefore, $\beta_l^2 < 0$ for $l = 1, 2$ and $\beta_l^2 \geq 0$ for each $l > 2$. This implies that $d_1 = d_2 = 0$ and $d_l = 1$ for each $l = 3, \ldots, T$. Hence, $\sum_{l=1}^{T} d_l = T - 2$ implies $R(d_1, \ldots, d_l) = a_1^2$. Then, in each stage $t = 3, \ldots, T$, it must be the case that

$$\beta_l^t = \frac{l - 1 - t}{t - 1}$$

regardless of the history h^t at the beginning of this stage, or equivalently $\beta_l^t < 0$ if $l \leq t$ and $\beta_l^t \geq 0$ if $l > t$. We deduce that $d_l = 0$ if $l \leq t$ and $d_l = 1$ if $l > t$ such that the statement $R(d_1, \ldots, d_T) = a_1^t$ follows from the equality

$$\sum_{l=1}^{T} d_l = T - t.$$

Player 1's perceptron ψ_1^2 induces the sequence $(a_1^1, \ldots, a_1^T)$ of T actions as desired. ■

Observe that the result does not depend on the assumption that player 2's strategy must be implementable by a finite automaton. It is neither related to the particular form of the prisoner's dilemma. In fact, this result is not restricted to 2×2 games, it applies for any two-player game. The next corollary follows from proposition 1.

Corollary 1 *Assume $T \in \mathbb{N}$, $k^* \geq T$ and $m^* \geq T - 1$. The only Nash equilibrium of G^T induces at each stage the Nash equilibrium (D, D) of G.*

Proof. Consider any $T \in \mathbb{N}$ and suppose that $k^* \geq T$ and $m^* \geq T - 1$. By proposition 1, player 1 can construct a perceptron which mimics the behavior of any repeated game strategy against a given automaton of the opponent. Therefore, player 1's the best reply to any automaton of player 2 induces the play of D in stage T. Next, by way of contradiction, suppose that there is a Nash equilibrium (ψ_1, M_2) that does not induce (D, D) in each stage. Since ψ_1 plays D in stage T, it is the interest of player 2 to construct a strategy that dictates D in the last two stages. Therefore, the Nash equilibrium (ψ_1, M_2) must induce (D, D) in stage T since such a strategy can be induced by an automaton with no more than $T - 1$ states. Let $t^C < T$ be the last stage in which (D, D) is not induced by (ψ_1, M_2). If ψ_1 plays C in this stage, player 1 has an incentive to construct a T-classifier perceptron ψ_1' which mimics ψ_1 up to stage $t^C - 1$ and then plays D from stage t^C. If M_2 plays C in stage $t^C \leq T - 1$, player 2 can also construct an automaton M_2' with at most $T - 1$ states which mimics M_2 during the first $t^C - 1$ stages and then plays D from stage t^C. This contradicts the initial assumption that (ψ_1, M_2) is a Nash equilibrium. We conclude that if $k^* \geq T$ and $m^* \geq T - 1$, the unique Nash equilibrium of G^T induces in each stage the Nash equilibrium (D, D). ■

The next proposition highlights that there is no Nash equilibrium that results in the cooperation in each stage if player 1 is restricted to choosing strategies which are implementable by a perceptron, whatever the duration of the game, the number of classifiers in this perceptron and the number of states in the opponent's automaton.

Proposition 2 *Let $T \in \mathbb{N}$, $k^* \in \mathbb{N}$ and $m^* \in \mathbb{N}$. There is no Nash equilibrium of G^T which induces the action pair (C, C) in each stage.*

Proof. If $m^* \geq T$, the proof is immediate since the player 2 can construct an automaton with T states that plays D in stage T, regardless of the perceptron chosen by player 1. If $m^* < T$, suppose by way of contradiction that there is a Nash equilibrium (ψ_1, M_2) of G^T which induces the action pair (C, C) in each stage. We distinguish two cases according to the value of k^*.

If $k^* = 0$, the perceptron ψ_1 has no classifier. The strategy induced by ψ_1 must consist in playing the same action in each stage, here the action C by assumption, regardless of the opponent's strategy. Player 2's best reply to this perceptron consists in playing D in each stage, and this strategy is induced by a one-state automaton. In this context (ψ_1, M_2) cannot be a Nash equilibrium.

If $k^* \geq 1$, the perceptron ψ_1 is not player 1's best reply to M_2 whatever the punishment released by M_2 if player 1 does not comply with the sequence of actions $(C, \ldots, C)$. In fact, for each $k^* \geq 1$, player 1 can construct a perceptron ψ_1^3 with a unique classifier K_1 such that ψ_1^3 plays action D in stage T after the play of $T-1$ actions C. The classifier K_1 is defined by :

$$
\begin{aligned}
K_1 : \quad & \alpha_1(C, C) = -1 \\
& \alpha_1(C, D) \in \mathbb{R} \\
& \alpha_1(D, C) \in \mathbb{R} \\
& \alpha_1(D, D) \in \mathbb{R} \\
& \beta_1^1 = T - 2
\end{aligned}
$$

The decision function specifies $R(1) = C$ and $R(0) = D$. It is easy to check that $\beta_1^t \geq 0$ for each $t = 1, \ldots, T-1$ which implies that ψ_1^3 plays action C in each stage $t = 1, \ldots, T-1$ against M_2. Then, ψ_1^3 plays action D in the last stage since $\beta_1^T < 0$. Therefore, $f_1(\psi_1^3, M_2) > f_1(\psi_1, M_2)$ and the proof is complete. ■

Remark that a perceptron must have at least one classifier in order to induce a strategy which does not consist in the constant play of an action. An automaton must have at least two states to do likewise. Henceforth, let $k^* = 1$ and $m^* = 2$ be the minimal complexity in the machine of each player which allows for the play of non-constant strategies. Even if proposition 2 asserts that the players cannot cooperate in each stage of the game, the next result establishes that a cooperation in almost all stages is an equilibrium outcome provided that the machines have at most the minimal complexity which allows for the play of non-constant strategies.

Proposition 3 *Let $k^* = 1$ and $m^* = 2$. For each $\varepsilon > 0$, there exists $T_\varepsilon \in \mathbb{N}$ such that if $T \geq T_\varepsilon$, the machine game G^T has a Nash equilibrium (ψ_1, M_2) whose payoffs verify $f_i(\psi_1, M_2) \geq 1 - \varepsilon$ for each player i.*

Proof. Let $k^* = 1$ and $m^* = 2$. Assume that player 2 uses the two-state automaton M^1 represented in figure 1 which implements the GRIM TRIGGER strategy. The next step is the construction of the perceptron ψ_1^4 for player 1. Its unique classifier is defined by :

$$K_1: \begin{array}{l} \alpha_1(C,C) = -1 \\ \alpha_1(D,C) = 0 \\ \alpha_1(C,D) = -T \\ \alpha_1(D,D) = -T \\ \\ \beta_1^1 = T-2 \end{array}$$

The decision function specifies $R(1) = C$ and $R(0) = D$. The pair (ψ_1^4, M_2^1) yields the following average payoffs :

$$f_1(\psi_1^4, M_2^1) = \frac{T+c}{T}, \text{ and } f_2(\psi_1^4, M_2^1) = \frac{(T-1)-b}{T}.$$

Given any real number x, denote by $\lceil x \rceil$ the smaller integer that is larger of equal to x. For each $\varepsilon > 0$, let

$$T_\varepsilon = \left\lceil \max\left\{ \frac{1+b}{\varepsilon}, 3+c+b \right\} \right\rceil.$$

We need to check that the induced payoffs verify $f_i(\psi_1^4, M_2^1) \geq 1 - \varepsilon$ for each player $i = 1, 2$, each $\varepsilon > 0$ and each $T \geq T_\varepsilon$. Observe that $f_1(\psi_1^4, M_2^1) > f_2(\psi_1^4, M_2^1)$ for each T. Therefore, it suffices to check that $f_2(\psi_1^4, M_2^1) \geq 1 - \varepsilon \Longleftrightarrow T \geq (1+b)/\varepsilon$, which is ensured whenever $T \geq T_\varepsilon$. Next, it remains to prove that (ψ_1^4, M_2^1) is a Nash equilibrium of G^T. We proceed in two steps.

A) The perceptron ψ_1^4 is a best reply to M_2^1

Firstly, consider the possible deviations from the sequence of actions played by ψ_1^4 against M_2^1 in any stage $t \leq T-1$. Since M_2^1 releases a minmax punishment in each stage following the deviation, player 1 obtains at most 0 in each of these stages. The gain for player 1 of such a deviation increases in t such that the best deviation occurs in stage $T-1$. Thus, a deviation in one of the first $T-1$ stages yields player 1 at most the average payoff

$$\frac{T-1+c-b}{T}$$

which is strictly less than $f_1(\psi_1^4, M_2^1)$. Secondly, player 1 cannot gain by deviating in stage T since he obtains the best stage payoff. We conclude that

ψ_1^4 is a best reply to M_2^1.

B) The automaton M_2^1 is a best reply to ψ_1^4

Since $m^* = 2$, player 2's automaton can only deviate in the first two stages from the sequence of actions played by M_2^1 against ψ_1^4. In addition, the synaptic weights $\alpha_1(C, D) = -T$ and $\alpha_1(D, D) = -T$ in player 1's perceptron retaliate by a minmax punishment as soon as the opponent's deviation is observed, *i.e.* player 2 cannot obtain more than 0 in each stage that follows his deviation. The best implementable deviation by player 2 consists in playing C in stage 1 and then D in all remaining stages. In such a case, his average payoff is $(2 + c)/T$ which is less than $f_2(\psi_1^4, M_2^1)$ when $T \geq 3 + c + b$. This condition is satisfied whenever $T \geq T_\varepsilon$. The automaton M_2^1 is a best reply to ψ_1^4. We conclude that the pair (ψ_1^4, M_2^1) is a Nash equilibrium. ■

The players can cooperate almost all the time if their strategies are limited to the minimal complexity that allow for the play of non-constant strategies. The next proposition shows that this result applies even if player 2 can choose automata with an arbitrarily large number of states. Recall that when $m^* \geq T$, player 2's automaton can induce any sequence of T actions. In this context, player 2 can mimic the behavior of any repeated game strategy against a given perceptron of player 1.

Proposition 4 *Fix $k^* = 1$, $m^* \geq T$ and suppose that $b \leq \min\{1, c\}$. For each $\varepsilon > 0$, there exists $T_\varepsilon \in \mathbb{N}$ such that if $T \geq T_\varepsilon$, the machine game G^T has a Nash equilibrium (ψ_1, M_2) whose payoffs verify $f_i(\psi_1, M_2) \geq 1 - \varepsilon$ for each player i.*

The proof of this result is given in appendix. The basic idea is to generate a "complex" history up to stage $T - 1$ in the sense that the action pair (C, C) does not occur in each stage. The equilibrium automaton threatens any deviation by player 1 from this desired sequence of actions by playing D until the end of the game. Then, it is the interest of the perceptron to conform to these plays and this task mobilizes most of the computational ressources of the machine. As a consequence, although the perceptron of player 1 is sometimes able to defect in stage T (see the proof of proposition 3 for instance), the history induced by the pair (ψ_1, M_2) in the first $T - 1$ stages is "complex" enough to prevent the perceptron from reverting to action D in the last stage. The next result determine a structural property of Nash equilibria in which the equilibrium perceptron of player 1 cooperate in the last stage. For instance, the equilibrium perceptron constructed in the proof of proposition 4 belongs to this category.

Proposition 5 *Let $k^* \in \mathbb{N}$. Consider a Nash equilibrium (ψ_1, M_2) of G^T with ψ_1 playing C in stage T. Then, ψ_1 has exactly $k = k^*$ classifiers.*

Proof. Let $k^* \in \mathbb{N}$ and consider a Nash equilibrium (ψ_1, M_2) of G^T in which ψ_1 dictates C in the last stage.[5] Suppose by way of contradiction that ψ_1 dictates C in stage T and has $k < k^*$ classifiers. Then, player 1 can construct a perceptron ψ_1' with the same k classifiers than ψ_1 and a $(k+1)$th classifier K' defined by a synaptic weight $\alpha'(a) = -1$ for each $a \in A$ and the initial value $\beta'^1 = T - 2$. The decision function R' of ψ_1' is given by

$$R'(d_1, \ldots, d_k, d') = \begin{cases} R(d_1, \ldots, d_k) & \text{if } d' = 1, \\ D & \text{if } d' = 0. \end{cases}$$

For all the strategies available to the opponent, the value of K' is positive in all stages except in the last one. Therefore, by definition of R', perceptron ψ_1' mimics in each of the $T-1$ first stages the action played by ψ_1 against M_2. In stage T, the decision function R' specifies to play action D since $\beta'^T < 0$. This contradicts the fact that (ψ_1, M_2) is a Nash equilibrium of G^T. We conclude that each Nash equilibrium of G^T for which the perceptron plays C in stage t has exactly $k = k^*$ classifiers. ■

This result may be generalized to any two-player finitely repeated games in the following way. If there is a Nash equilibrium in which the perceptron does not play in stage T a best reply to the action chosen by the opponent, then it has exactly $k = k^*$ classifiers. Remark also that a Nash equilibrium (ψ_1, M_2) for which ψ_1 plays C in stage T is not a Nash equilibrium if one adds or removes a classifier from the number k^* of classifiers allowed in player 1's perceptron.

5 Perceptron versus perceptron

In this section, we consider a machine game in which both players are restricted to choosing strategies which are implementable by perceptrons. The next proposition shows that if players 1 and 2 are restricted to using perceptrons with at most one and five classifiers respectively, there is a Nash equilibrium in which they can approximate the cooperative average payoff. Denote by k_i^* the maximal number of classifiers in player i's perceptron.

[5] Note that it must be the case that $k^* < T$ otherwise by proposition 1, player 1 can construct a perceptron with T classifiers that mimics ψ_1 against M_2 up to stage $T-1$ and defects in the last stage.

Proposition 6 *Fix $k_1^* = 1$, $k_2^* = 5$ and suppose that $b \leq \min\{1, c\}$. For each $\varepsilon > 0$, there exists $T_\varepsilon \in \mathbb{N}$ such that if $T \geq T_\varepsilon$, the machine game G^T has a Nash equilibrium (ψ_1, ψ_2) whose payoffs verify $f_i(\psi_1, \psi_2) \geq 1 - \varepsilon$ for each player i.*

The proof of this proposition is also given in appendix since it rests on the proof of proposition 4. It illustrates the difficulty of dealing with perceptrons with many classifiers. The computation process which leads to the action specified by the decision function of the perceptron becomes heavier as its number of classifiers increases. In fact, it requires to handle with many (simple) calculations which seem less practical than the task of transiting from one state to other in an automaton with many states.

6 Conclusion

The most evident open question in the current setting is whether there exist bounds on the complexity of the machines for which the set of achievable payoffs coincides with those of the folk theorem for infinitely repeated games. In this paper, we have answered the question of whether cooperation is a possible equilibrium outcome in the finitely repeated prisoner's dilemma in which the players have non comparable bounds on the complexity of the strategies they may choose. Each existence result has been proved using trigger strategies. Folk theorems also often employ such "vengeful" strategies. However, in our model, the equilibrium histories are rather simple since the cooperative outcome is induced consecutively in most of the stages. Constructing perceptrons which retaliate by a definitive minmax punishment to sustain more elaborated histories raises additional difficulties. In particular, it points to the need for constructing perceptrons with many classifiers and yet little is known about the exact computational abilities of such machines. We hope to provide findings about this question in a future work on finitely repeated zero-sum machine games.

The preceding discussion addresses the question of extending the analysis to games with more than two actions available to the players and/or more than two players. The former extension has been examined for machine games played by automata (Neyman, 1998, and Neyman, Okada, 2000) and for machine games played by perceptrons (Cho, 1996a). The latter extension has been studied by Neyman (1998) for machine games played by automata but only for results regarding the complexity of various plays even if the extension to n players of several theorems on the set of equilibrium payoffs needs only minor modifications. The difficulty noted earlier however reappears for n-player machine games played by perceptrons : equilibrium perceptrons should incorporate many classifiers in order to punish deviations by the opponents.

We do not consider mixed strategy equilibria because of the lack of a modelisation of mixed strategies in machine games played by perceptrons. Mixed strategies induced by automata are constructed in Neyman (1985, 1998). In these two papers, each player randomizes among his available automata. Unfortunately, mixed strategies induced by perceptrons has not been considered so far. Players may randomize among their available perceptrons but one might also consider a perceptron which receives the mixed history as input and which dictates the mixed action to be played as output. Nevertheless, we do not now whether the results established in this paper persist when the players are allowed to use such mixed strategies.

Appendix

Proof. [proposition 4] Fix $k^* = 1$, $m^* \geq T$ and suppose that $b \leq \min\{1, c\}$. The proof is divided in three parts. Firstly, for each $\varepsilon > 0$, define $T_\varepsilon \in \mathbb{N}$ by

$$T_\varepsilon = \max\left\{7, \left\lceil \frac{2+2b-c}{\varepsilon} \right\rceil \right\}.$$

Secondly, consider the following pair (ψ_1^5, M_2^3) of machines. The perceptron ψ_1^5 of player 1 has a single classifier K_1 defined by

$$\begin{aligned} K_1 : \quad & \alpha_1(C,C) = -1 \\ & \alpha_1(D,C) = 2 \\ & \alpha_1(C,D) = -3T \\ & \alpha_1(D,D) = -3T \\ & \beta_1^1 = T - 6 \end{aligned}$$

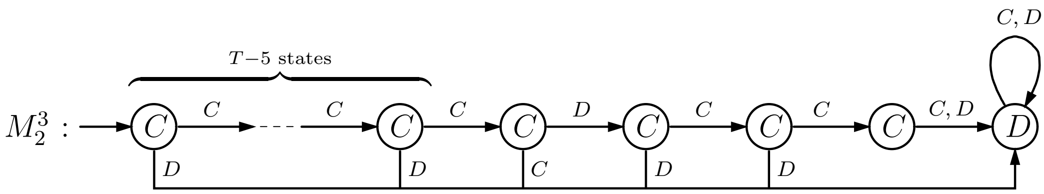

Figure 3: *The automaton M^3 of player 2.*

The decision function specifies once again $R(1) = C$ and $R(0) = D$. The automaton M_2^3 of player 2 has $T \geq T_\varepsilon \geq 7$ states and is represented in figure 3. The history induced by the pair (ψ_1^5, M_2^3), the current value of the classifier and the stage payoffs obtained by both players are indexed in table 2. Players 1 and 2 obtain the following average payoffs :

$$f_1(\psi_1^5, M_2^3) = \frac{T + 2c - b - 1}{T}, \text{ and}$$
$$f_2(\psi_1^5, M_2^3) = \frac{T + c - 2b - 2}{T}.$$

t	1	2	$\dots$	$T-6$	$T-5$	$T-4$	$T-3$	$T-2$	$T-1$	T
β_1^t	$T-6$	$T-7$	$\dots$	$\frac{1}{T-7}$	0	$\frac{-1}{T-5}$	$\frac{1}{T-4}$	0	$\frac{-1}{T-2}$	$\frac{1}{T-1}$
a_1^t	C	C	$\dots$	C	C	D	C	C	D	C
a_2^t	C	C	$\dots$	C	C	C	C	C	C	D
$\pi_1(a^t)$	1	1	$\dots$	1	1	$1+c$	1	1	$1+c$	$-b$
$\pi_2(a^t)$	1	1	$\dots$	1	1	$-b$	1	1	$-b$	$1+c$

Table 2: *History, stage payoffs and values of the classifier.*

Observe that for all $T \geq T_\varepsilon$, $f_1(\psi_i^5, M_2^3) > f_2(\psi_i^5, M_2^3)$ since $1 + c > -b$. Therefore, it is easy to check that for each $\varepsilon > 0$ and each $T \geq T_\varepsilon$, the pair of machines (ψ_1^5, M_2^3) yields each player $i = 1, 2$ an average payoff of $f_i(\psi_i^5, M_2^3) \geq 1 - \varepsilon$. Thirdly, it remains to prove that (ψ_1^5, M_2^3) is a Nash equilibrium of G^T. As in the proof of proposition 3, we proceed in two steps.

A) The perceptron ψ_1^5 is a best reply to M_2^3

We consider successively each possible deviation by player 1 from the sequence of actions played by ψ_1^5 against M_2^3 :

• Deviations in a stage $t \leq T - 5$. Player 1 deviates in stage t by playing action D, which implies that M_2^3 reacts by moving to the absorbing state that plays D from stage $t+1$. Thus, player 1 cannot gain more than 0 in each stage $t+1, \dots, T$. The average payoff he obtains from this deviation increases with t. Deviating yields at most $(T+c-5)/T$ which is no more than $f_1(\psi_1^5, M_2^3)$ if $c+4 \geq b$. This condition is guaranteed by the assumption that $b \leq \min\{1, c\}$.

• Deviations in stage $T-4$. Player 1's optimal deviation consists in playing C in stage $T-4$ and then D in the remaining stages since M_2^3 retaliates by a definitive minmax punishment in response to the deviation. Acting likewise, player 1 obtains the average payoff $(T-4)/T$ which is no more than

$f_1(\psi_1^5, M_2^3)$ if $2c + 3 \geq b$. Once again, the assumption $b \leq \min\{1, c\}$ ensures that it is not the interest of player 1 to deviate.

• Deviations in a stage $t \in \{T-3, T-2\}$. Player 1 deviates in stage t by playing action D and his interest is to play D in each remaining stage because of M_2^3's punishment. This deviation yields player 1 at most the average payoff $(T + 2c - 2)/T$ which is no more than $f_1(\psi_1^5, M_2^3)$ if $b \leq 1$, a condition guaranteed by the assumption $b \leq \min\{1, c\}$.

• Deviations in stage $T-1$. Player 1's deviation consists in playing C in stage $T-1$ and M_2^3 punishes this move by the play of D in the last stage. Player 1 cannot do best than playing D in stage T. Deviating yields him the average payoff of $(T - 1 + c)/T$ which is no more than $f_1(\psi_1^5, M_2^3)$ since $b \leq \min\{1, c\}$ by assumption.

• Deviation in stage T. Of course, it is optimal for an "unrestricted" player to implement the sequence of actions played by ψ_1^5 against M_2^3 in the first $T-1$ stages and then to play D in stage T. We show than such a strategy cannot be induced by a perceptron with at most one classifier. By way of contradiction, assume that there is a perceptron ψ_1^6 with a single classifier K_2 which implements the optimal sequence of actions $(C, \ldots, C, D, C, C, D, D)$ against M_2^3.[6] We describe the necessary conditions on the initial value β_2^1 and on the synaptic weights of classifier K_2 in ψ_1^6. Remark that ψ_1^6 cannot induce a sequence of actions that contains actions C and D if $R(1) = R(0)$. Thus, it must be the case that $R(1) \neq R(0)$ and without any loss of generality, we can suppose that $R(1) = C$ and $R(0) = D$. It follows that player 1 has to fix $\beta_2^1 \geq 0$ since ψ_1^6 must play C in the first stage against M_2^3. By definition of the value β_2^t of classifier K_2 in each stage $t = 1, \ldots, T$, the system of inequalities $(L_1), \ldots, (L_T)$ below must have a solution to construct the desired perceptron ψ_1^6 with a unique classifier.

[6] Perceptron ψ_1^6 must have at least one classifier since the desired sequence of T actions contains both actions C and D.

$$
\begin{cases}
\beta_2^1 \geq 0 & (L_1) \\
\beta_2^1 + \alpha_2(C,C) \geq 0 & (L_2) \\
\beta_2^1 + 2\alpha_2(C,C) \geq 0 & (L_3) \\
\qquad\vdots & \vdots \\
\beta_2^1 + (T-6)\alpha_2(C,C) \geq 0 & (L_{T-5}) \\
\beta_2^1 + (T-5)\alpha_2(C,C) < 0 & (L_{T-4}) \\
\beta_2^1 + (T-5)\alpha_2(C,C) + \alpha_2(D,C) \geq 0 & (L_{T-3}) \\
\beta_2^1 + (T-4)\alpha_2(C,C) + \alpha_2(D,C) \geq 0 & (L_{T-2}) \\
\beta_2^1 + (T-3)\alpha_2(C,C) + \alpha_2(D,C) < 0 & (L_{T-1}) \\
\beta_2^1 + (T-3)\alpha_2(C,C) + 2\alpha_2(D,C) < 0 & (L_T)
\end{cases}
$$

Since $T_\varepsilon \geq 7$, it is guaranteed that the history induced by ψ_1^6 against M_2^3 begins with two action pairs (C,C). Therefore, we deduce from lines (L_1) to (L_{T-4}) that $\beta_2^1 > 0$ and $\alpha_2(C,C) < 0$. Then, it follows from lines (L_{T-3}) and (L_{T-3}) that $\alpha_2(D,C) > 0$. Next, from lines (L_{T-2}) and (L_T) we can write the restriction

$$-\beta_2^1 - (T-4)\alpha_2(C,C) \leq \alpha_2(D,C) < -\frac{\beta_2^1}{2} - \frac{(T-3)}{2}\alpha_2(C,C).$$

As a consequence, it must be the case that

$$
\begin{aligned}
& -\beta_2^1 - (T-4)\alpha_2(C,C) < -\frac{\beta_2^1}{2} - \frac{(T-3)}{2}\alpha_2(C,C) \\
\iff \quad & \beta_2^1 + (T-5)\alpha_2(C,C) > 0. \qquad (2)
\end{aligned}
$$

However, equation (2) contradicts line (L_{T-4}). The system of inequalities $(L_1), \ldots, (L_T)$ has no solution. In other words, there is no perceptron with at most one classifier which implements the optimal sequence of T actions $(C, \ldots, C, D, C, C, D, D)$ against M_2^3. We conclude that the perceptron ψ_1^5 is a best reply to M_2^3.

B) The automaton M_2^3 is a best reply to ψ_1^5

Firstly, since M_2^3 plays C in each stage $t = 1, \ldots, T-1$, any deviation by player 2 in stage t induces the play of D. By construction, the synaptic weights $\alpha_1(C,D) = \alpha_1(D,D) = -3T$ of classifier K_1 of perceptron ψ_1^5 are such that any play of D by player 2 in any stage $t < T$ implies $\beta_1^\tau < 0$ for each $\tau = t+1, \ldots, T$, regardless of the action pairs played after the deviation. In other words, ψ_1^5 retaliates by a definitive minmax punishment in reaction to any deviation prior to stage T. Secondly, we consider each possible deviation by player 2.

• Deviations in a stage $t \leq T-5$. Deviating yields player 2 an average payoff bounded above by $(T+c-5)/T$ which is no more than $f_1(\psi_1^5, M_2^3)$ if $b \leq 3/2$. This condition is fullfilled by the assumption $b \leq \min\{1, c\}$. Player 2 won't deviate in a stage $t \leq T-5$.

• Deviation in stage $T-4$. Acting for the best within such a deviation, player 2 obtains the average payoff $(T-5)/T$ which is no more than $f_1(\psi_1^5, M_2^3)$ if $c+3 \geq 2b$. By assumption $c > 0 \Longrightarrow c+3 > 2$ and $b \leq \min\{1, c\}$ such that it follows that $c+3 > 2 \geq 2b$ as desired.

• Deviations in a stage $t \in \{T-3, T-2\}$. Player 2 obtains at most

$$\frac{(T-4)a+b+c+2d}{T}$$

from deviating in one of these stages. Not deviating yields a greater average payoff whenever $b \leq 1$, a condition satisfied by assumption.

• Deviations in stage $T-1$. Deviating yields player 2 an average payoff bounded above by

$$\frac{(T-3)a+b+2d}{T}$$

which is no more than $f_1(\psi_1^5, M_2^3)$ if $c+1 \geq b$. This condition is guaranteed by the assumption $b \leq \min\{1, c\}$.

• Deviation in stage T. In this stage, player 2 obtains the best stage payoff from using M_2^3. Therefore, he cannot gain from switching towards action C.

Thus, M_2^3 is a best reply to ψ_1^5 which implies that (ψ_1^5, M_2^3) is a Nash equilibrium of G^T. This completes the proof. ■

Proof. [proposition 6] Assume $k_1^* = 1$, $k_2^* = 5$ and $b \leq \min\{1, c\}$. For each $\varepsilon > 0$, define $T_\varepsilon \in \mathbb{N}$ by

$$T_\varepsilon = \max\left\{7, \left\lceil \frac{2+2b-c}{\varepsilon} \right\rceil \right\}.$$

Let player 1 use the perceptron ψ_1^5 with a single classifier constructed page 16. Player 2's perceptron, denoted by ψ_2^7, has 5 classifiers K_2, K_3, K_4, K_5 and K_6 defined by :

$$
\begin{array}{llll}
K_2: & \alpha_2(C,C)=1 & K_3: & \alpha_3(C,C)=0 \\
 & \alpha_2(D,C)=1 & & \alpha_3(D,C)=1 \\
 & \alpha_2(C,D)=1 & & \alpha_3(C,D)=1 \\
 & \alpha_2(D,D)=1 & & \alpha_3(D,D)=1 \\
 & \beta_2^1=-T+1 & & \beta_3^1=-1
\end{array}
$$

$$
\begin{array}{llll}
K_4: & \alpha_4(C,C)=-1 & K_5: & \alpha_5(C,C)=0 \\
 & \alpha_4(D,C)=-1 & & \alpha_5(D,C)=1 \\
 & \alpha_4(C,D)=T & & \alpha_5(C,D)=1 \\
 & \alpha_4(D,D)=T & & \alpha_5(D,D)=1 \\
 & \beta_4^1=T-5 & & \beta_5^1=-2
\end{array}
$$

$$
\begin{array}{ll}
K_6: & \alpha_6(C,C)=-1 \\
 & \alpha_6(D,C)=-1 \\
 & \alpha_6(C,D)=T \\
 & \alpha_6(D,D)=T \\
 & \beta_6^1=T-2
\end{array}
$$

The decision function $R:\{0,1\}^5 \longrightarrow \{C,D\}$ of perceptron ψ_2^7 is given by

$$
R(d_2,d_3,d_4,d_5,d_6)=\begin{cases} C & \text{si} \displaystyle\sum_{l=2}^{6} d_l = 2 \\ D & \text{sinon} \end{cases}
$$

Table 3 indexes the dynamics induced by the pair of perceptrons (ψ_1^5, ψ_2^7). By proposition 4, each player obtains at least $1-\varepsilon$ from (ψ_1^5, ψ_2^7) for each $\varepsilon > 0$ and each $T \geq T_\varepsilon$. It remains to show that (ψ_1^5, ψ_2^7) is a Nash equilibrium. Also, by proposition 4, it suffices to prove that ψ_2^7 induces the same strategy than M_2^3. Moreover, since one can easily verify from table 3 that ψ_2^7 generates the same T-period history against ψ_1^5 than M_2^3, it is enough to check that ψ_2^7 releases a definitive and minmax punishment in case of a deviation by player 1 from the play induced by M_2^3. We consider successively each possible deviation by player 1.

• Deviation in a period $t \in \{1,\ldots,T-5\}$. The deviation from the sequence of actions played by ψ_2^7 against ψ_1^5 consists in playing action D in stage t. The action pair (D,C) is induced in stage t such that at the beginning of stage $t+1$, β_4^{t+1} and β_6^{t+1} are still positive and β_3^{t+1} becomes positive. It follows that $\sum_{l=2}^{6} d_l \geq 3$ with the consequence that ψ_2^7 reacts to the deviation by playing action D in stage $t+1$. Next, consider the situation in stage $t+2$. The action pair $(\cdot, D)$ played in stage $t+1$ implies that β_5^{t+2} becomes positive. This value remains positive until the end of the game since this

t	1	2	...	$T-6$	$T-5$	$T-4$	$T-3$	$T-2$	$T-1$	T
β_1^t	$T-6$	$T-7$	...	$\frac{1}{T-7}$	0	$\frac{-1}{T-5}$	$\frac{1}{T-4}$	0	$\frac{-1}{T-2}$	$\frac{1}{T-1}$
β_2^t	$-T+1$	$-T+2$	...	$\frac{-6}{T-7}$	$\frac{-5}{T-6}$	$\frac{-4}{T-5}$	$\frac{-3}{T-4}$	$\frac{-2}{T-3}$	$\frac{-1}{T-2}$	0
β_3^t	-1	-1	...	$\frac{-1}{T-7}$	$\frac{-1}{T-6}$	$\frac{-1}{T-5}$	0	0	0	$\frac{1}{T-1}$
β_4^t	$T-5$	$T-6$	...	$\frac{2}{T-7}$	$\frac{1}{T-6}$	0	$\frac{-1}{T-4}$	$\frac{-2}{T-2}$	$\frac{-3}{T-2}$	$\frac{-4}{T-1}$
β_5^t	-2	-2	...	$\frac{-2}{T-7}$	$\frac{-2}{T-6}$	$\frac{-2}{T-5}$	$\frac{-1}{T-4}$	$\frac{-1}{T-2}$	$\frac{-1}{T-2}$	0
β_6^t	$T-2$	$T-3$	...	$\frac{5}{T-7}$	$\frac{4}{T-6}$	$\frac{3}{T-5}$	$\frac{2}{T-4}$	$\frac{1}{T-2}$	0	$\frac{-1}{T-1}$
$\sum_{l=2}^{6} d_l$	2	2	...	2	2	2	2	2	2	3
a_1^t	C	C	...	C	C	D	C	C	D	C
a_2^t	C	C	...	C	C	C	C	C	C	D
$\pi_1(a^t)$	1	1	...	1	1	$1+c$	1	1	$1+c$	$-b$
$\pi_2(a^t)$	1	1	...	1	1	$-b$	1	1	$-b$	$1+c$

Table 3: *History induced by* (ψ_1^5, ψ_2^7).

classifier has no negative synaptic weight associated to an action pair of the stage game. It is also the case of the value associated by K_3 to any history induced by player 1's deviation. Furthermore, the synaptic weights associated to the action pairs $(\cdot, D)$ by classifiers K_4 and K_6 are large enough to ensure that no sequence of action pairs could imply $\beta_4^\tau < 0$ or $\beta_6^\tau < 0$ for any $\tau \in \{t+2, \ldots, T\}$. Therefore, it must be the case that $\sum_{l=2}^{6} d_l \geq 4$ in stages $t+2, \ldots, T$ which means that ψ_2^7 plays action D until the end of the game. We conclude that a deviation by player 1 in stage t releases a definitive and minmax punishment.

• Deviations in stage $T-4$. Player 1's deviation consists in playing C in stage $T-4$. Since (C, C) is induced in stage $T-4$, only the value associated to the history at stage $T-4$ by K_6 is positive at the beginning of stage $T-3$. According to the decision function R, ψ_2^7 plays D in stage $T-3$. The action pair $(\cdot, D)$ which is induced at stage $T-3$ implies that $\beta_3^{T-2} \geq 0$, $\beta_4^{T-2} \geq 0$ and $\beta_6^{T-2} \geq 0$. The play of D by ψ_2^7 in stage $T-2$ follows from the fact that $\sum_{l=2}^{6} d_l \geq 3$. Exactly as in the previous case, these three values remain positive until the end of the machine game, *i.e.* ψ_2^7 punishes player 1's deviation by the play of action D until the end of the game.

• Deviations in a stage $t \in \{T-3, T-2\}$. Player 1's deviation is similar to that considered in the first case. The consequences on the values associated to the history by the six classifiers of ψ_2^7 are identical. Therefore, the punishments induced by this deviation consist in playing D in each of the remaining stages as desired.

• Deviations in stage $T-1$. Player 1's deviation is similar to that considered in the second case such that we can conclude that ψ_2^7 reacts by playing D in stage T.

• Deviation in stage T. From proposition 4, player 1 cannot construct a perceptron with one classifier which mimics the behavior of ψ_1^5 up to stage $T-1$ and then defects in stage T.

We conclude that for each $\varepsilon > 0$ and each $T \geq T_\varepsilon$, the pair of perceptrons (ψ_1^5, ψ_2^7) defines a Nash equilibrium whose payoffs satisfy $f_i(\psi_1, \psi_2) \geq 1 - \varepsilon$ for each player $i = 1, 2$. ■

References

[1] Aumann Robert J., "Survey of Repeated Games", Essays in Game Theory and Mathematical Economics in Honour of Oskar Morgenstern (Aumann R.J., éd.), Bibliographische Institut, Manheim, 1981.

[2] Axelrod Robert, *The Evolution of Cooperation*, Basic Books, New York, 1984.

[3] Benoît Jean-Pierre, Krishna Vijay, "Nash Equilibria of Finitely Repeated Games", *International Journal of Game Theory* (1987), **16**:197–204.

[4] Cho In-Koo, "Bounded Rationality, Neural Network and Folk Theorem in Repeated Games with Discounting", *Economic Theory* (1994), **4**:935–957.

[5] Cho In-Koo, "Perceptrons Play the Repeated Prisoner's Dilemma", *Journal of Economic Theory* (1995), **67**:267–284.

[6] Cho In-Koo, "On the Complexity of Repeated Principal Agent Games", *Economic Theory* (1996a), **7**:1–17.

[7] Cho In-Koo, “Perceptrons Play Repeated Games with Imperfect Monitoring”, *Games and Economic Behavior* (1996b), **16**:22–53.

[8] Cho In-Koo, Li Hao, “How Complex are Networks Playing Repeated Games”, *Economic Theory* (1999), **13**:93–123.

[9] Fudenberg Drew, Maskin Eric, “The Folk Theorem in Repeated Games with Discounting or with Incomplete Information”, *Econometrica* (1986), **54**:533–556.

[10] Gilboa Itzhak, Samet Dov, “Bounded versus Unbounded Rationality: The Tyranny of the Weak”, *Games and Economic Behavior* (1989), **1**:213–221.

[11] Neyman Abraham, “Bounded Complexity Justifies the Cooperation in the Finitely Repeated Prisoner’s Dilemma”, *Economic Letters* (1985), **19**:227–229.

[12] Neyman Abraham, “Finitely Repeated Games with Finite Automata”, *Mathematics of Operations Research* (1998), **23**:513–552.

[13] Neyman Abraham, Okada Daijiro, “Two-Person Repeated Games with Finite Automata”, *International Journal of Game Theory* (2000), **29**:309–325.

[14] Rubinstein Ariel, “On Price Recognition and Computational Complexity in a Monopolistic Model”, *Journal of Political Economy* (1993), **101**:473–484.

Nr.	Author	Title
07-58	Sylvain Béal	PERCEPTRON VERSUS AUTOMATON&8727;
07-57	Sylvain Béal Jacques Durieu Philippe Solal	Farsighted Coalitional Stability in TU-games
07-56	Alen Nosic Martin Weber	Determinants of Risk Taking Behavior: The role of Risk Attitudes, Risk Perceptions and Beliefs
07-55	Michael F. Meffert Thomas Gschwend	Strategic Voting under Proportional Representation and Coalition Governments: A Simulation and Laboratory Experiment
07-54	Klaus Fiedler	Pseudocontingencies - A key paradigm for understanding adaptive cognition
07-53	Florian Kutzner Peter Freytag Tobias Vogel Klaus Fiedler	Base-rate neglect based on base-rates in experience-based contingency learning
07-52	Klaus Fiedler Yaakov Kareev	Implications and Ramifications of a Sample-Size Approach to Intuition
07-51	Klaus Fiedler	The Ultimate Sampling Dilemma in Experience-Based Decision Making
07-50	Jürgen Eichberger David Kelsey	Ambiguity
07-49	Tri Vi Dang	Information Acquisition in Double Auctions
07-48	Clemens Kroneberg	Wertrationalität und das Modell der Frame-Selektion
07-47	Dirk Simons Nicole Zein	Audit market segmentation and audit quality
07-46	Sina Borgsen Martin Weber	False Consensus and the Role of Ambiguity in Predictions of Othersí Risky Preferences
07-45	Martin Weber Frank Welfens	An Individual Level Analysis of the Disposition Effect: Empirical and Experimental Evidence

Nr.	Author	Title
07-44	Martin Weber Frank Welfens	The Repurchase Behavior of Individual Investors: An Experimental Investigation
07-43	Manel Baucells Martin Weber Frank Welfens	Reference Point Formation Over Time: A Weighting Function Approach
07-42	Martin Weber Frank Welfens	How do Markets React to Fundamental Shocks? An Experimental Analysis on Underreaction and Momentum
07-41	Ernst Maug Ingolf Dittmann	Lower Salaries and No Options: The Optimal Structure of Executive Pay
07-40	Ernst Maug Ingolf Dittmann Christoph Schneider	Bankers and the Performance of German Firms
07-39	Michael Ebert Nicole Zein	Wertorientierte Vergütung des Aufsichtsrats - Auswirkungen auf den Unternehmenswert
07-38	Ingolf Dittmann Ernst Maug Christoph Schneider	How Preussag became TUI: Kissing too Many Toads Can Make You a Toad
07-37	Ingolf Dittmann Ernst Maug	Valuation Biases, Error Measures, and the Conglomerate Discount
07-36	Ingolf Dittmann Ernst Maug Oliver Spalt	Executive Stock Options when Managers are Loss-Averse
07-35	Ernst Maug Kristian Rydqvist	Do Shareholders Vote Strategically? Voting Behavior, Proposal Screening, and Majority Rules
07-34	Ernst Maug Abraham Ackerman	Insider Trading Legislation and Acquisition Announcements: Do Laws Matter?
07-33	Dirk Simons	Independence, low balling and learning effects
07-32	Rainer Greifeneder Herbert Bless	Relying on accessible content versus accessibility experiences: The case of processing capacity

Nr.	Author	Title
07-31	Rainer Greifeneder Herbert Bless	Depression and reliance on ease-of-retrieval experiences
07-30	Florian Heiss Axel Börsch-Supan Michael Hurd David Wise	Pathways to Disability: Predicting Health Trajectories
07-29	Axel Börsch-Supan Alexander Ludwig Mathias Sommer	Aging and Asset Prices
07-28	Axel Börsch-Supan	GLOBAL AGING - Issues, Answers, More Questions
07-27	Axel Börsch-Supan	MIND THE GAP: THE EFFECTIVENESS OF INCENTIVES TO BOOST RETIREMENT SAVING IN EUROPE
07-26	Axel Börsch-Supan	Labor market effects of population aging
07-25	Axel Börsch-Supan	Rational Pension Reform
07-24	Axel Börsch-Supan	European welfare state regimes and their generosity towards the elderly
07-23	Axel Börsch-Supan	Work Disability, Health, and Incentive Effects
07-22	Tobias Greitemeyer Rainer Greifeneder	Why the Euro looked like a price booster: Differential perception of increasing versus decreasing prices
07-21	Patrick A. Müller Rainer Greifeneder Dagmar Stahlberg Herbert Bless	Relying on accessibility experiences in procedural fairness judgments
07-20	Volker Stocké	The Motive for Status Maintenance and Inequality in Educational Decisions. Which of the Parents Defines the Reference Point?