AN AXIOMATIC APPROACH OF INFINITELY REPEATED GAMES

Andrés Felipe Muñoz Gómez

Adviser: Natalia González

Master's degree in Economics
Universidad Icesi
Santiago de Cali, June 30 2016

Abstract

This document offers an axiomatization of repeated interactions between two boundedly rational players where cooperation is not achieved by a solution. In other words, I propose a collection of axioms that characterize the set of non-cooperative solutions of repeated games played by finite automata à la Rubinstein. Several axioms are presented for Individual Behaviour and Efficiency. This axioms characterize a payoff string that achieved the solution concept of Rubinstein for a serie of cooperative games.

Keywords: Axiomatization, Bounded Raionality, Complexity measures.

Abstract

Este documento ofrece una axiomatización de juegos repetidos entre dos jugadores con racionalidad acotada donde el equilibrio cooperativo no hace parte de la solución. En otras palabras, propongo una colección de axiomas que caracterizan las soluciones no-cooperativas de juegos repetidos por aumatomatas finitas á la Rubinstein. Varios axiomas son presentado para El Comportamiento Individual y Eficiencia. Los axiomas caracterizan una sucesión de pagos que alcanzan el concepto de solución de Rubinstein para una gama de juegos de cooperación.

Keywords: Axiomatización, Racionalidad acotada, medidas de complejidad.

Contents

An Axiomatic approach of Infinitely Repeated Games

1 Introduction

The choice process that is well known and studied by traditional economics is roughly the following, "*after thoughtful deliberation the decision maker must choose the best alternative according to her preferences and subject to a set of feasibility constraints*". Furthermore, the underlying assumptions implicit in the choice process of the rational decision maker are: (i) knowledge of the problem, (ii) clear preferences and (iii) the ability to optimize. That is, the decision maker is fully aware of the choice set, she has a complete ordering over the choice set and she has unlimited ability to conduct complicated computations to determine her optimal actions, respectively. Yet a reflective examination of ones choice process suggests these assumptions are rarely realistic. This is why traditional economics has argued that even if the decision maker rarely behaves as the rational procedural scheme prescribes his behaviour can be described *as if* he follows such procedure. But, in 1955, Herbert Simon presented a theory that incorporated constraints on the information-processing capacities of the decision maker; in particular, he introduced computational limitations of the actor. And yet, his contributions had limited impact on economic theory because of the difficulty to embed the procedures by which decision are made. Game theory, on the other hand, has been another source of interest when analysing strategic interactions as a way of understanding the meaning "*rational*" behaviour. For example, Aumman and Shapley 1976 [1] and Rubinstein 1979[2] proved that the individually rational outcome can support the Pareto criterion of mutual cooperation instead of mutual defection when a game is played repeatedly. However, to what extent does rationality in games depend on *irrationality*? One method which has shown to be productive in order to capture certain elements of *irrationality* i.e. *bounded rationality* is to model players as stimulus-response machines. Using these artificial models in which the stimulus of the other players' previous actions maps into a response has allowed economists to study games in which players have (i) finite memory, (ii) a finite number of input signals and (iii) a finite set of responses. Inevitably, in order to model bounded rationality one must acknowledge Herbert Simon's contributions. Nevertheless, I will not deal with models that offer *satisficing procedures* of human decisions but instead will focus on a artificial model that captures specific procedural aspects. In particular, I propose an axiomatic approach of infinitely repeated games between two boundedly rational players; that is, I offer an axiomatization that is complimentary to Rubinstein's (86), use of Finite Automata to formalize players' behaviour in repeated games.

The literature about imperfect memory has been divided into two broad fields depending on the modelling strategy. The first approach places assumptions which make implicit the memory process when players are not aware of these limitations. Among these works one can find bounded recall, where agents recall the information only for a determined number of time. Kalai & Stanford (88) and Leherer (88) are some examples of papers on multi-player games with bounded recall. Other types of imperfect memory are memory decay or restrictions holding a finite posterior history. The common aspect is that updating rules in each period are exogenous.

[1]Long term competition: A game theoretic Analysis.(Aumann & Shapley, 1976)

[2]Equilibrium in supergames.(Rubinstein, 1979)

The second approach assumes all agents are fully aware of these limitations. The problem is such that agents must decide the optimal strategy given memory restrictions, that is, now the memory rule is part of the player's strategy.

Under this approach, authors such as Futia (77), Smale (80) and Mount-Reiter (83) work the relation on bounded rationality and computational complexity. However, the notion of automaton, especially with Moore machines (56) were developed initially by Radner (80) and Aumann (81). From the latter work, several other authors such as Neyman (85), Rubinstein (85), Ben-Porath (86) and Kalai (87) have developed measures of strategic complexity in repeated games. Other authors such as Green (82), Megiddo & Wigderson (86) and Gilboa & Schmeidler (89) model bounded rationality using other types of machines (Turing) for repeated games.

We offer an axiomatic program that characterizes the set of Nash equilibria of a supergame where players simultaneously choose Moore machines to implement their strategies. Players are boundedly rational to the extent that there are limits to the complexity of response available to each player.

Following the general framework and some definitions in Section 2, I introduce the axioms in Section 3. In Section 4, I discuss the measures of complexity that characterize Rubinstein's set of Nash equilibria and in Section 5, I present the characterization. My conclusions are left for Section 6.

2 The Environment

Consider a two-person supergame with completely patient players, where a game, G, is repeated sequentially an infinite number of times. We refer to player i's opponent as player j. It is important to note that, *opponent*, does not imply that the other player is trying to "beat" player i. Rather, each player's best alternative may involve "helping" or "not-helping" the other player.

Let the *pure-action space* for each player be defined as an action vector, $A_i = (a_{i,k}, a_{i,m})$ where k and m denote each alternative. This should not be confused with a *strategy profile*, which represents a *two*-dimensional vector of individual strategies, thus for every different combination of individual actions one obtains a different strategy profile, $a = (a_{i,k}, a_{j,k})$. The *set of all strategy profiles* (or space of strategy profiles) defined on $A = A_i \times A_j$ should not to be confused with a strategy profile. Let us use the the prisoner's dilemma for a complete picture: $A_i =$ (cooperate, no cooperate), $a =$ (cooperate, cooperate) and $A = \{$(cooperate, cooperate), (cooperate, no cooperate), (no cooperate, cooperate), (no cooperate, no cooperate)$\}$ are the **pure-action space**, **strategy profile** and **set of all strategy profiles**, respectively.

At the end of each period, both players receive the *one*-period G- payoffs. Moreover, because we want to follow Rubinstein's exact specification of payoffs, we evaluate the streams of payoffs using the average reward, as opposed to future discounted criterion. Thus, the payoff function, $u_i : A \to \mathbb{R}$ gives player i's von Neuman-Morgenstern's utility for each strategy profile in the set of all strategy profiles.

DEFINITION **1.** *Let Ω be the set of all two-person infinite games. We denote the set of payoff profiles as,*

$$\Pi(G) = \{u \in \mathbb{R}^2 | \exists a \in A, u = (u_i(a), u_j(a))\}$$

for $G \in \Omega$

Therefore, a supergame strategy is a combination described by each individuals' choice of action of how to play G at every period, conditional on every possible history.

Using a type of artificial device, a kind of finite automaton called a Moore machine implies a change in the strategy space of the game, G. Players no longer select strategies, but Moore machines. A machine consists of a finite set of internal states, Q_i, an initial state, $\bar{q}_i \in Q_i$, an action function, $\lambda_i : Q_i \to A_i$[3] and a transition

[3] Earlier we defined, A_i as the finite set of actions or choices for player i.

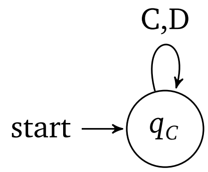

Figure 1: Moore machine: always cooperate

(or next state) function $\mu_i : Q_i \times A_j \to Q_i$. In other words, it is a finite automaton in which the player's next action (the machine's output) is contingent on the existing state of the machine, which in turn is a function of the previous state of the machine (or previous period) and the other player's previous move (the machine's input), through a transition (or next-state) function. Let (M_i, M_j) be a pair of machines for the two players, if both players in a two-person game have chosen Moore machines then the supergame is played as follows,

1. The machine of player i in the first period starts at the initial state, that is, $q_i = \bar{q}_i$.
2. Player i chooses $s_i^1 = \lambda_i(q_i^1) \in A_i$.
3. Player $i's$ machine, in the second period, moves into state, $q_i^2 = \mu_i(q_i^1, s_j^1)$.
4. The sequence is generally defined as follows: provided the machines are at period t at the states, q_i^t and q_j^t, the former chooses $s_i^t = \lambda_i(q_i^t) \in A_i$. Then, at period t+1 player i's state is $q_i^{t+1} = \mu_i(q_i^t, s_j^t)$.

In other words, the pair, (M_1, M_2) generates deterministically an infinite sequence of strategy profiles and a infinite sequence of states. In short, after each player has chosen its machine, the sequence of the game will generate actions and states as the repeated game progresses.

A Moore machine can be depicted as a *directed graph* comprised of the elements we presented above, that is, $M_i = \langle Q_i, q_i, \lambda_i, \mu_i \rangle$. The nodes of the graph correspond to the states, or elements of the finite set of internal states, Q_i and the initial state, q_i and the edges of the nodes represent the possible transition among the states. For instance, the *transition diagram* of strategies of the repeated Prisoner's Dilemma in which player i always cooperates is described in Figure 1.

The letters C and/or D above the arc show the other player's action (the input), the state, q_C, correspond to the machines action (the output) in that node. I will assume the *start* label is the initial state of the machine, $\bar{q}_i \in Q_i$. The normal form of this machine can be described as,

$$Q_i = \{q_C\}, \quad \bar{q}_i = q_C, \quad \lambda_i(q_C) = C \quad \text{and} \quad \mu_i(q_C, \cdot) = q_C$$

The following example was taken from Rubinstein (1986). Consider the case where player i plays C until the player j plays D, after which he then decides to punish the latter from there onwards. The normal form of *The Grim Moore machine* can be represented as,

$$Q_i = \{q_C, q_D\}, \quad \bar{q}_i = q_C, \quad \lambda(q_s) = s, \quad \text{and} \quad \mu(q_s, s) = q_s \quad \text{for} \quad s = C, D$$

Furthermore, the *transition diagram* is slightly different from the *always cooperate Moore machine*. Here the strategy for player i of playing C until player j plays D is depicted in Figure 2. Notice that after player j plays D, player i will play D repeatedly regardless of what the former player moves.

Now imagine that players must bear the cost of maintaining each state as part of the machine. A player must pay a fee for every state he chooses to maintain in his machine irrespective of the frequency of its usage; that is, a machine with x states is *cheaper* than one with $x + n$ for $n \in \mathbb{N}$. In this respect, players choose machines and the complexity of implementing their strategies is measured by the number of states. Among the many measures of complexity, two of which we will address section 4, most would agree that a simpler one is

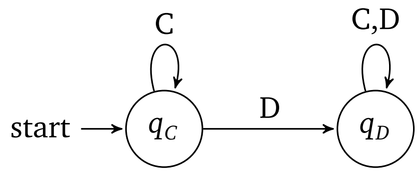

Figure 2: The Grim Moore machine

playing a shorter profile forever than playing a longer one. In other words, it is less costly and simpler for players to play a one action profile rather than alternate between two (or more). Let $M = (M_1, M_2)$ be a pair of machines; $q^t = (q_1^t, q_2^t)$ be an element of the infinite sequence of states of the machines and $s^t = (s_1^t, s_2^t)$, be an element of the infinite sequence of actual actions of the supergame. Since machines are finite, there exists a non-negative number t_2 such that at period t_2+1 the pair of states is the same as period t_1: $q^{t_1} = q^{t_2}$ for $t_2 \geq t_1$.

Definition **2.** *Let $(q^1, q^2, ..., q^{t_1-1})$ be the initial part of the play by the pair of finite automaton. We denote a* **cycle** *of the play by both machines as $(q^{t_1}, ..., q^{t_2})$ and the* **length** *of the cycle as $T = t_2 - t_1 + 1$.*

Given an infinite sequence, $(s^t)_{t=1}^{\infty}$, of strategy profiles and with players being "completely patient" we use the time-average criterion with discount factor $\delta = 1$ to define a players payoff stream.

Definition **3.** *For $s \in S$, let player i's payoff stream in the cycle be defined as,*

$$\pi_i(M_i, M_j) = \frac{1}{T}\sum_{t=1}^{T} u_i(s^t)$$

Thus, as opposed to the normal form of a standard two-person supergame, $G = (A_i, A_j, u_i, u_j)$, where players must choose from a finite set of strategies in the current supergame players must choose Moore machines. A pair of machines generates an infinite sequence of G's outcomes, $(s^t)_{t=1}^{\infty}$ and states $(q^t)_{t=1}^{\infty}$, .

In turn, the natural question that follows is when will a pair of machines be preferred to another. Let $>_L$ be the lexicographic preference defined on $\mathbb{R}^2$. Player i prefers a pair of automata, (M_i, M_j) to $(\hat{M}_i, \hat{M}_j)$ if,

$$(\pi_i(M_1, M_2), |M_i|) >_L (\pi_i(\hat{M}_1, \hat{M}_2), |\hat{M}_i|).$$

In other words, player i prefers, $(M_1, M_2) \succ_i (\hat{M}_1, \hat{M}_2)$ if he obtains a higher average payoff in the former machine, or if he gets the same average payoff by using a machine with less states.

Definition **4.** *A pair of automata (M_1^*, M_2^*) is a semi-perfect-equilibrium if there does not exists a M_1, M_2 and t such that,*

$$(M_1, M_2^*(q_2^t)) \succ_1 (M_1^*(q_1^t), M_2^*(q_2^t)) \tag{1}$$

or

$$(M_1^*(q_1^t), M_2) \succ_2 (M_1^*(q_1^t), M_2^*(q_2^t)) \tag{2}$$

Correspondingly, q_1^t and q_2^t are the states of M_1^* and M_2^* in period t when the game is played by the pair, (M_1^*, M_2^*). Conversely, a pair of automata, (M_1^*, M_2^*) is a Nash equilibrium if there does not exist a M_1 or M_2 such that $(M_1, M_2^*) >_1 (M_1^*, M_2^*)$ or $(M_1^*, M_2) >_2 (M_1^*, M_2^*)$. Notice that former solution concept is some what different from the latter in the following way: *M_i^* must be optimal, not only at the beginning of the game but also at the start of each repetition*

After defining the two-person supergame, in the following section we present the entire set of axioms we will use to characterize Rubinstein's *semi-perfect*-equilibrium.

3 The Axioms

3.1 Individual Behaviour

The first set of axioms aim to ensure the existence of a utility function that represents each player's choice of machine. that is, for each machine's infinite sequence of strategies. A solution $S(G)$ assigns a set of s to each stage game $G \in \Omega$, that is, the set $S(G)$ contains all the sequences that are expected to occur in the repeated stage game, G. In other words, a sequence of play written as $s = (s^1 s^2 ...)$ where $s^t = (a_1^t, a_2^t)$ and $s \in S(G)$ defined as the solution.

Let a supergame strategy, $s \in S(G)$ be sequence of functions, $\lambda^t = (\lambda_i^t, \lambda_j^t)$ where λ_i^t determines player i's strategy at period t as a function of i's state, that is, $s_i^t = \lambda_i^t(q_i^t)$.

The existence of the utility function allows us to represent the payoff of any sequence and also compare the payoffs of different string lengths. Furthermore, let a player's minimum security payoff in the stage game (whatever the other player's choice is) be,

$$\underline{u}_i(G) = \max_{\psi_i \in \Psi_i} \min_{a_j \in A_j} u_i(\psi_i, \underline{a}_j)$$

Where Ψ_i denote player i's set of mixed actions with elements ψ_i.

Axiom 1. *Security Payoff: for all $s' \notin S(G)$ such that a player i ensures his security payoff it must be true that,*

$$\underline{u}_i(G) > u_i(s') \longleftrightarrow \underline{s} \succ s' \quad \forall i$$

With axiom 1 every average payoff of an infinite sequence of strategies, s', which is strictly less than his maxmin level will not belong to $S(G)$. Moreover, provided $\underline{s} \succ s'$ the set $u_i(G)$ is bounded below by $\underline{u}_i(G)$ (given the opponent's decision). Every player will guarantee that his payoff will be as large as his cutoff point. The proposition allows any player to change at any time from his maxmin level to his maxmin mixed action provided his continuation payoff is weakly greater.

Let us consider the Samaritan's dilemma where player (country) 1's action set is to transfer resources, or not, and player (recipient) 2's is to work or not. Of course, player 1's donation is important to both players' welfare, but if 1 donates, 2 prefers not working. Hence, ex-post payments could set ex-ante incentives that crowd out the willingness of political agents to implement preventive action resulting in non-cooperative behaviour. In the payoff matrix, (T, W), represents the cooperation strategy of the game. However, in figure 3 this strategy is not part of the thick dotted area, which represents the minimal payoff of each player. To understand why, take the pair of Nash equilibria of the game, Notice the pair of maxmin average payoffs is $(3, 2)$ and the cooperative strategy payoff is $(2, 3)$, but the latter is clearly not part of the boundary.

	W	$\bar{W}$
$\bar{T}$	4,2	1,1
T	2,3	3,4

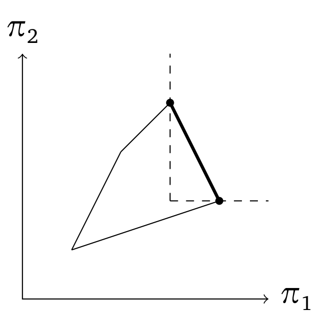

Figure 3: Samaritan's dilemma.

Before explaining the dark thick segment which connects two Nash equilibria of the game let us firs present our next axiom.

Axiom 2. *Convexity: If $\pi(s) \geq \pi(s')$, then $\pi(ts + (1-t)s') \succeq \pi(s')$ for all $t \in [0,1]$.*

Consider the prisoner's dilemma (Figure 4). Let $a = (C, D)$ and $b = (D, C)$ and fix $s = (a...)$ and $s' = (b...)$ so that in the first supergame strategy both play (C, D) and in the second they play (D, C) indefinitely. In the case they play a sequence in which both strategies appear they create a mixed action profile from the both s and s'. A simple example could be $\dot{s} = (aab...)$ or $\ddot{s} = (abb...)$. In the figure the thick solid line represents the infinite set of linear combinations from the original pair of strings.

	C	D
C	2,2	-1,3
D	3,-1	0,0

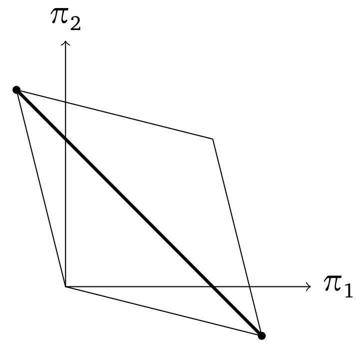

Figure 4: Prisoner's dilemma under axiom 4

Consider a situation where a player believes she can do better by playing some action, as opposed to following string s, then she will break the arrangement. Therefore, it is necessary to define the set of feasible payoffs above certain individually rational level. Let the action and payoff profiles that appear in string s be defined as,

$$\begin{aligned} P(s) &= \{a \in A : s^t = a\} \quad \text{for some t} \\ U(s) &= \{(u_1(a), u_2(s)) : a \in P(s)\} \end{aligned}$$

where $a \in A$ denotes an action profile; hence $A_j = \{a_j \in A_j : \exists a_i \in A_i \text{ for which } a \in P(s)\}$. The following proposition says that if a player can do better by playing a mixed action instead of following string s, given every action played by the opponent, then she might probably follow the mixed action.

Proposition **1.** *Given a sequence s' if some player i has a mixed action ψ_i that satisfies $\pi(s') < u_i(\psi_i, a_j), \ \forall a_j \in A_j \Rightarrow s' \notin S(G)$*

Consider the example of the prisoner's dilemma (See table 1). Let the following two pair of strategies (C, C) and (D, C) be tagged as a and b respectively. From the two strategy pairs, a and b we consider the next two strings for $s = (ab...)$, $s' = (b...)$. In the first string player 2 never punishes her opponent. By proposition **??** player 1 will immediately discard sequence s as $\pi_1(s) = 2.5 < 3 = \pi_1(s')$. The idea here is that, unless player 1 has a reason to believe her behaviour might be punished, she may behave myopically, thus she might try defecting more often.

Proposition 1 is satisfied any time a player can switch to her maxmin (possible mixed) action and guarantee that her continuation payoff will be weakly greater than her security payoff. Notice that when one combines axiom 4 and proposition 1 we can no longer consider the set of strategies that do cannot ensure each players reservation payoff, that is:

$$\forall s \in S : \alpha(D, C) + (1-\alpha)(C, D) < \underline{u} \Rightarrow s \notin S(G)$$

Consider the entire set of convex combinations of strategies, (C, D) and (D, C). There are a subset of linear combinations of the latter and former strategies that does not ensure each players minimum cutoff payoffs,

	C	D
C	2,2	-1,3
D	3,-1	0,0

Table 1: Normal game for prisoner's dilemma

for this reason we call them *Unranked Pareto Points*. Moreover, we can see this if we compare figures 4 and 5.

For the following two axioms we introduce some additional notation. Let $(q^{t_1}, ..., q^{t_2})$ be a cycle for a any pair of machines and k_1, k_2 be two numbers between t_1 and t_2.

AXIOM 3. *For both players, i and j, and all* $t_1 \leq k_1 < k_2 \leq t_2$, $q_i^{k_1} \neq q_i^{k_2}$ *and* $q_j^{k_1} \neq q_j^{k_2}$.

AXIOM 4. *There is no* $t_1 \leq k$ *such that* $\lambda_1^*(q_1^k) = \lambda_1^*(q_1^{k+1})$ *and* $\lambda_2^*(q_2^k) \neq \lambda_2^*(q_2^{k+1})$

3.2 Efficiency

With the following set of axioms we intend to eliminate the sequences of play that generate complex machines. Previously, we referred to, $|M_i|$, as the size of a machine, in other words, as the number of states it has. Furthermore, the complexity of a strategy is determined by the size of the machine that can implement it. From the transition diagram 1 it appears that the Always cooperate Moore machine represents a strategy of lower complexity than the trigger strategy of diagram 2. In the former diagram, the number of different states induced by the initial state in all subgames is nil, whereas in the latter case it is one.[4]

AXIOM 5. *Weak efficiency: If s contains a unique string of actual plays, a and both players prefer it, as opposed to a sequence* $s' = (ab)$ *that is,* $u(a) \gg \pi(s')$ *then* $s' \notin S(G)$

Players will play a single strategy profile forever rather than play this profile as part of a more elaborate supergame strategy if the latter decreases their payoffs. In other words, strings of higher complexity obtained from a simpler sequence need not be beneficial for both. In order to ensure a collective improvement we introduce our next axiom.

AXIOM 6. *Strong efficiency: If* s' *has a strictly Pareto improving string such that for both* i *and* j,

$$\pi(s) \gg \pi(s')$$

then, $s' \notin S(G)$

One remark about this axiom is that it allows for inefficiencies, which can be the case of the "always defect" strategy as it is considered to be a simpler sequence than one in which mutual harmful and harmless behaviour among players co-exist indefinitely. The axiom, therefore excludes these latter types of strings from $S(G)$.

PROPOSITION **2.** *Axiom 6 implies Axiom 5.*

When $l(s) \leq l(s')$ this means that the cycle in s^t is shorter. Thus, the idea is to play a simpler sequence and therefore a shorter one in general.

[4]Of course, this measure does not take account of the complexity of the action function and the transition function which Radner (1986) identified. Gottinger (1983) referred to this as the tradeoff between structural complexity and computational complexity.

Axiom 7. *Lexicographic efficiency: Let s^t be a strictly Pareto sequence compared to s'^t. If both players primarily obtain a higher average payoff in s^t than in s'^t, $\pi(s^t) \geq \pi(s'^t)$ and only secondarily s^t is simpler than s'^t, then $s'^t \notin S(G)$*

Let us consider, once again, the prisoner's dilemma. Let $a = (C,C)$, $b = (D,C)$and $c = (C,D)$ be three strategy profiles. Consider sequence $s = (babc...)$ whose expected payoff for each player is equal to $\pi_1(s) = \frac{3+2+3-1}{4} = 7/4$ and $\pi_2(s) = \frac{-1+2-1+3}{4} = 3/4$ respectively. However, notice that the sequence $s' = (baa...)$ represents a strictly Pareto-improvement for both players as $\pi_1(s') = \frac{2+2+3}{3} = 7/3$ and $\pi_2(s') = \frac{2+2-1}{3} = 1$.

3.3 Examples

If we consider the prisoner's dilemma, the unique Nash equilibrium occurs when both players defect. When the game is repeated a finite number of times the unique subgame-perfect equilibrium will still be defect in every period, by an iterated dominance argument: *finking in the last period dominates cooperating; iterating once, both players fink in the second period, and so on*.[5] If the game is repeated infinitely many times and players' preferences are represented by the discounted sum of the stage game payoff, we now see that cooperation by both players in every period is an equilibrium as long as players are sufficiently patient, see left Figure 5. Hence, cooperation can be considered an equilibrium by having both players adopt grim or trigger strategies, in other words, by having each player condition her actions on earlier actions of her opponents.

	C	D
C	2,2	-1,3
D	3,-1	0,0

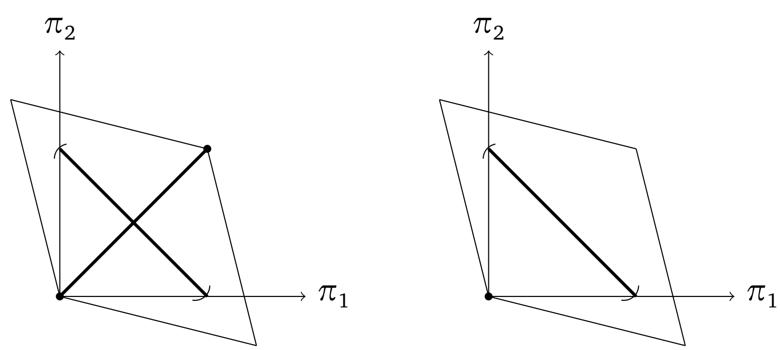

Figure 5: Solution Payoff under Individual rationality and Efficient (thick dots, thick solid lines). Left: Original game equilibrium ; Right: Rubinstein (1986)'s equilibrium payoff-Axiomatization.

After a player defects, her opponent switches to the open-loop punishment phase of "fink" every period which makes the deviant do no better than to accept nil in every remaining period, that is, coordinate (0,0) in figure 5). The prospective deviant must ask herself which is more desirable: receiving two today and zero forever or conforming and receiving one today and forever. This will depend on her discount factor δ; if she is very impatient, δ is close to zero and the single-period gain in the defecting period will outweigh the loss in later periods. Conversely, receiving one rather than zero in every subsequent period outweighs the temptation of obtaining two rather than one, provided she is more patient. Hence, we see in Figure 1 that cooperation in every period was a Nash equilibrium as long as each player's discount factor exceeded one-half.

Yet, by both imposing additional costs in order to capture complexity of the automata seems to be a valid way to constrain agent's rationality which induces the semi-perfect solution concept. A solution in the machines game will eliminate strategy profiles that were previously Nash equilibria. The trigger and *tit-for-tat* strategies for the iterated prisoner's dilemma will no longer be achieved by Rubinstein's solution

[5] Nonetheless, considerable amount of experimental evidence shows that subjects do tend to cooperate on many, if not most, periods.

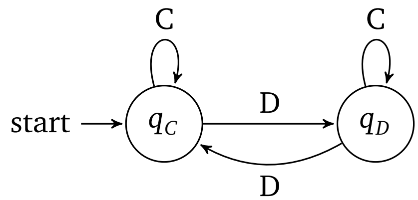

Figure 6: The Tit-for-Tat Moore machine

concept. Since each agent has "punishment" states in his/her machine which are never used, each player is better off by removing them; see right side of figure 5. Moreover, because the semi-perfect concept requires that a pair (M_1^*, M_2^*) includes only states which are used infinitely often, if both players use the machine in figure 6 the pair of machines that is Nash equilibrium is not semi-perfect. A solution à la Rubinstein must be optimal at the introductory part of the play, as well as the beginning of each repetition. If players use the *Tit-for-Tat* machine, the punishment phase will only be used temporarily before both players reach the cycle making both players drop the initial part of the play. This does not imply that semi-perfect equilibrium excludes players threatening each other, that is, if threats aren't exercised and they are costly, then players shouldn't maintain states if they were used only in the past.

Considering this example, but now from the viewpoint of the axioms listed above we need to discuss why some of kind of strings belong to the equilibrium path according to the semi-perfect-equilibrium. While the first axioms refer to the relationship of social preference, a sequence that satisfies both the axiom 1 and axiom 2 is ensuring that a space defined choice between two sequences to compare. For example (see Figure 5) if we take as a reference any sequence belonging to the inside of the figure linear combination on is it could be concluded that the sequence of the single strategy (D, D) is preferred, however succession of (C , C) it being less preferred.

The idea of the axiom 3 was explained before with proposition 1, since it is necessary for both players define a sequence of a single strategy as a lower bound. For axiom 6 however they are comparing the total payoff of the string and making a better choice for both. But with axiom 7 with the axiom 7 tries to refine this idea including the extent of complexity as the second item for successions.

	C	D
C	3,3	1,4
D	4,1	0,0

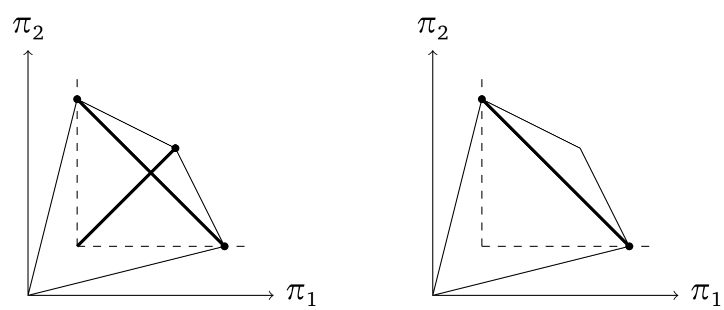

Figure 7: Solution Payoff under Individual rationality and Efficient (thick dots, thick solid lines). Left: Original game equilibrium ; Right: Rubinstein (1986)'s equilibrium payoff-Axiomatization.

Another game where we can reject the cooperative equilibria is the chicken game. Similar to the previous example, the trigger and the *tit-fot-tat* strategies support cooperative strategies in a infinitely repeated game. But, as we mentioned earlier, in a semi-perfect equilibrium if both players use, for instance, *tit-for-tat* machines, players will drop the punishment, or initial part of the play. Hence, cooperation is no longer an equilibrium as one can see in the right side of Figure 7.

Let us return to the samaritan's dilemma and compare the analytical results with the two previous games.

Both in the chicken game and in the prisoners' dilemma players have not only an incentive to cooperate but the set of Nash equilibria of the game support cooperation when the game is played infinitely many times. Consider the following two payoff matrices which correspond to the "passive" and "active" good Samaritan presented below.

	W	$\bar{W}$
$\bar{T}$	4,2	1,1
T	2,3	3,4

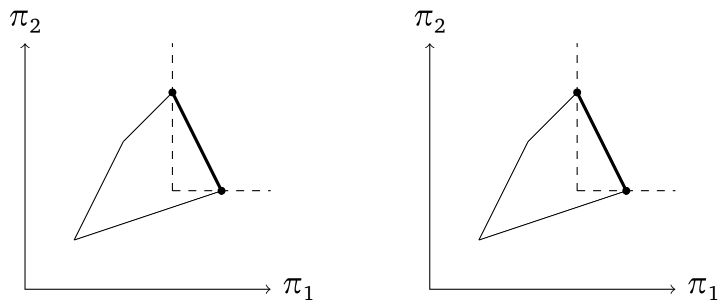

Figure 8: Solution Payoff under Individual rationality and Efficient (thick dots, thick solid lines). Left: Original game equilibrium ; Right: Rubinstein (1986)'s equilibrium payoff-Axiomatization.

We name tagged the second type of Samaritan, "active", because the donor country has a dominant strategy which is to provide assistance to the recipient, as opposed to the "passive" Samaritan. Notice that the pair (T, W) is not an equilibrium of the game, regardless of the type of Samaritan. For instance, in Figure 8 the payoff vector, $(2, 3)$ that supports cooperative behaviour is not part of the feasible area. Conversely, cooperation is supported by the payoff vector $(4, 3)$ which is not a Nash equilibrium of the infinitely repeated game.

	W	$\bar{W}$
$\bar{T}$	2,2	1,1
T	4,3	3,4

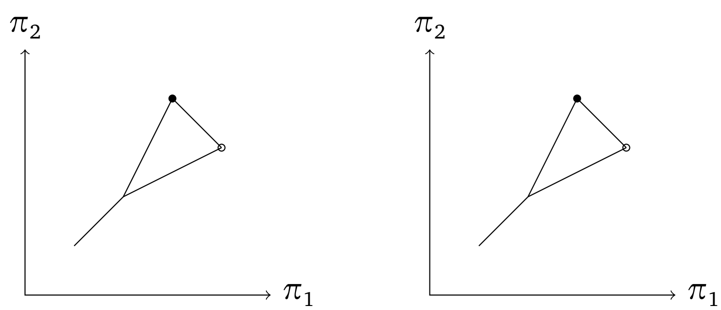

Figure 9: Solution Payoff under Individual rationality and Efficient (thick dots, thick solid lines). Left: Original game equilibrium ; Right: Rubinstein (1986)'s equilibrium payoff-Axiomatization.

Lastly, consider the battle of the sexes and the stag hunt games. Both examples share the same results: the folk theorem and our axiomatization identifies the same set of payoffs vectors which can be supported by Nash equilibria and semi-perfect equilibria, respectively. Cooperation, in both games is achieved.

4 Measures of complexity

Regardless of the notion of complexity one can argue that simpler strings are less complex. We have used Rubinstein's generalization of length of a cycle as our measure. In this section we present two other notions of complexity, namely, Kolmogorov complexity and Shannon entropy and provide a justification for Axiom 6. With the help of a couple of examples we stress the differences between the notion of complexity we use and these two others.

	C	D
C	4,2	1,1
D	1,1	2,4

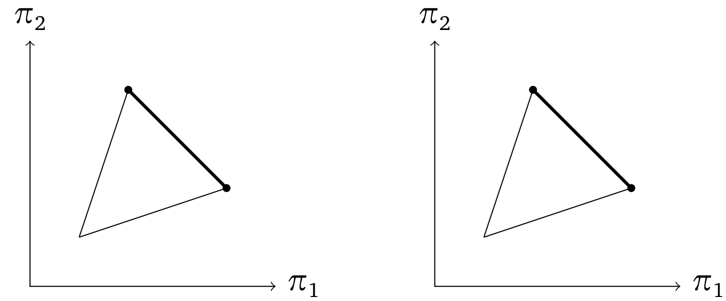

Figure 10: Solution Payoff under Individual rationality and Efficient (thick dots, thick solid lines). Left: Original game equilibrium ; Right: Rubinstein (1986)'s equilibrium payoff-Axiomatization.

	S	H
S	3,3	0,2
H	2,0	1,1

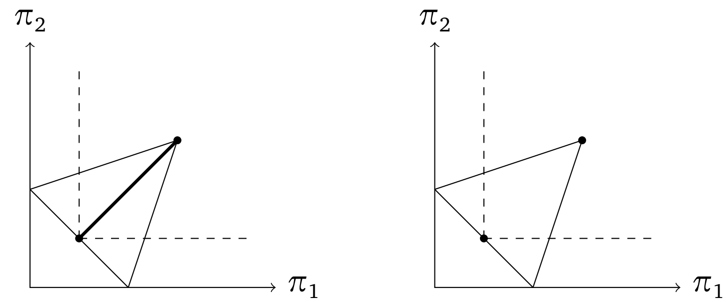

Figure 11: Solution Payoff under Individual rationality and Efficient (thick dots, thick solid lines). Left: Original game equilibrium ; Right: Rubinstein (1986)'s equilibrium payoff-Axiomatization.

In algorithmic information theory, a measure of complexity based on data compression is Kolmogorov complexity. In the same way that the measure of complexity we use is defined by the length of a cycle the Kolmogorov complexity of an object (string) is defined by the minimum number of information that produces the object as output.

Conversely, Shannon entropy ignores the object itself and considers only the characteristics of the random source from which the object generates one of the possible outcomes. Consider two machines, M^1 and M^2, each one generates symbols, a, b, c and d, according to the following probabilities: $p_a^1, p_b^1, p_c^1, p_d^1) = (0.25, 0.25, 0.25, 0.25)$ and $(p_a^2, p_b^2, p_c^2, p_d^2) = (0.5, 0.125, 0.125, 0.25)$, respectively. Rather than asking which machine produces more information, Shannon posed the question quite differently. *If one had to predict the next symbol in each machine what is the minimum amount of* **yes** *or* **no** *you would be expected to ask*. The most efficient way would be to pose questions which divide the possibilities in half. In the case of machine 1, one would begin by asking: *is it a or b?*. Since there is a 50% chance of it being a or b and a 50% chance of it being c or d. The next question, is it a, will have correctly identified that symbol. Hence, the uncertainty of machine 1 is two questions per symbol.

If we move to machine 2, because the probabilities are different we can pose questions differently. In this case, symbol a has a 50% chance and $b + c + d$ have a 50%, thus we can ask, is it a. If the answer is yes, we only needed to pose one question for this symbol, otherwise, we are left with two equal outcomes: d or b and c. We then ask is it d, and if the answer is positive we are done with the number of questions for this symbol. If not, we have a third question to ask in order to identify which of the remaining two symbols it is. Shannon proposed the following question, on average how many questions is one expected to ask to determine a symbol from machine 2? By calculating a weighted average of the number of questions per symbol, $\#q_s$

$$\begin{aligned} E_{M^2} &= p_a * \#q_a + p_b * \#q_b + p_c * \#q_c + p_d * \#q_d \\ &= 0.5 + 0.125 * 3 + 0.125 * 3 + 0.25 * 2 \\ &= 1.75 \end{aligned}$$

he concluded that machine 1 requires asking 2 questions on average per symbol, whereas ma chine 2 requires only 1.75 per symbol. This implies that the latter machine is producing less information because there is less uncertainty or surprise about its output. This average measure is what Shannon called, entropy. The unit of measure he used is based on the uncertainty of a fair coin toss which he called a bit, that is, a fair result.

DEFINITION **5.** *Let X be a finite or countable set, $x \in X$ be a random variable with probability distribution $P(X)$. Then the Shannon entropy is given by*

$$H(X) = \sum_{x \in X} p_x log_2 \left(\frac{1}{p_x} \right)$$

The definition above is simply the summation for each symbol of the probability of that symbol times the number of fair results, where the latter is defined by,

$$log_2 \left(\frac{1}{p_x} \right)$$

Entropy will be maximum when all outcomes are equally likely, but anytime we move away from equally likely outcomes entropy falls. In other words, with the "bit", each time we introduce predictability one can ask fewer questions to guess the outcome.

Kolmogorov complexity, on the other hand, considers only the object itself to determine the minimum number of "bits" required to produce the output. Let p be a string with length, $l(p)$, X the set of all strings, A the computable functions from X and $n \in \mathbb{N}$.

DEFINITION **6.** *The Kolmogorov complexity (K-complexity) of x^n with respect to A is given by*

$$K_A(x^n) = \min_{A(p)=x^n} l(p)$$

if there exists a string p such that $A(p) = x^n$. Otherwise,

$$K_A(x^n) = \infty$$

With the following two examples we show that both measures fail to satisfy Rubinstein's measure of complexity. Let $s = (b, b, b, a, a, ...)$ and $s' = (b, b, b, b, a, ...)$ be any pair of strings of a game. By using Rubinstein's measure of complexity, that is the length of each string, both sequences are equally complex. However, with *k-complexity* string s' is simpler than s. Intuitively, string s' requires a smaller (positive) integer to describe itself, *4b's and 1a* as opposed to string s, *3b's and 2a's*.

Consider strings, $s = (7a, 2b)$ and $s' = (21a, b, c, d)$ the positive integer multiplied by each category is the number of times it is repeated. Using definition 5 we calculate the entropy for both strings, and obtain $H(s) = 0.7642$ and $H(s') = 0.7416$, respectively. Again, this result contradicts the results of using Rubinstein's notion.

5 The main result

The following theorem limits the set of equilibria to those that satisfy individual rationality and axioms 6-7 which in turn are generated by automata. In other words, the set of axioms imply that players must play the same static semi-perfect equilibrium.

THEOREM **1.** *If exists a solution **S** in the set $S(G)$ that satisfies axioms 1-7, it must be true that there exists a pair of machines (M_1, M_2) that corresponds to a semi-perfect-equilibrium (SPE)*

6 Conclusions

This document has proposed an axiomatic approach to study repeated interactions between two boundedly rational players. Our axioms were designed to define a set of equilibria that includes all individually rational payoffs that are convex combinations of the one-shot game.

In order to capture the aspect of bounded rationality à la Rubinstein, that is by fixing agents to pay a cost per state within each players machine, we introduced axiom 6 which allowed players to play simpler strings, as opposed to more complex ones. In addition, we discussed the limitations of using two other sophisticated measures of complexity and showed why both measures fail to satisfy Rubinstein's measure of complexity. Moreover, this allowed us to eliminate the tit-for-tat and trigger strategies from the set of Nash equilibria in repeated games.

7 Appendix

7.1 Axioms for Lexicographic Order

Let $\succ$ represent a binary relation on $X = \times_{n\in N} X_n$.

Axiom 1. Asymmetric weak order: a player's preference relation is defined on X.

a. Asymmetric: $\forall x, y \in X, \quad x \succ y$ implies $\neg(y \succ x)$ hence, $x \sim y$ if and only if $\neg(x \succ y)$ and $\neg(y \succ x)$.

b. Negatively transitive: for all x, y and $z \in X$ if $x \succ y$ then either $x \succ z$ or $z \succ y$ or both.

Axiom 2. Independence: Let α_i and $\beta_i \in X$ for all $i \in \{1, 2, ..., n\}$ $(x_1, x_2, ..., \alpha_i, ..., x_{n-1}, x_n) \succ (x_1, x_2, ..., \beta_i, ..., x_{n-1}, x_n)$ if and only if $(y_1, y_2, ..., \alpha_i, ..., y_{n-1}, y_n) \succ (y_1, y_2, ..., \beta_i, ..., y_{n-1}, y_n)$ and all four n-tuples are in X.

Let $P(x, y) = \{n \in N : x_n \succ_n y_n\}$ represent the set of indices to which $x_n \succ y_n$.

Axiom 3. Decisiveness: for all x and $y \in X$, and a linear order[2] $<_o$ on $\{1, 2, ...n\}$ if $P(y, x) \neq \emptyset$ and $\min\{P(x, y)\} <_o \min\{P(y, x)\}$ then, $x \succ y$.

Example: Let $N = 4$ and $P(x, y) = \{1, 3, 4\}$ and $P(y, x) = \{2\}$. Because the most smallest index, that is, the most important element, is in the set $P(x, y)$ one can conclude $x \succ y$.

Axiom 4. Noncompensation: If, for all $x, y, w, z \in X$,

$$[\min\{P(x, y)\} <_o \min\{P(y, x)\} \iff \min\{P(z, w)\} <_o \min\{P(w, z)\}]$$

and

$$[\min\{P(y, x)\} <_o \min\{P(x, y)\} \iff \min\{P(w, z)\} <_o \min\{P(z, w)\}]$$

[2]The binary relation $<$ must be irreflexive, asymmetric, transitive and complete

then

$$(x \succ y \quad iff \quad z \succ w) \wedge (y \succ x \quad iff \quad w \succ z)$$

PROPOSITION **3** (1.). *Suppose Axiom 1. holds, then* $\forall x, y$ *and* $z \in X$ *the following must hold,*

a. $x \succeq y \iff x \succ y$ *or* $x \sim y$.
b. $x \succ y$ *and* $y \succeq z \Rightarrow x \succ z$.
c. $x \succeq y$ *and* $y \succ z \Rightarrow x \succ z$.

Proof. a. Suppose by equivalence that $x \succeq y \equiv \neg(y \succ x)$. So, either $x \succ y$ or $\neg(x \succ y)$ and $\neg(y \succ x)$, which implies $x \sim y$.

By contrast, suppose $x \succ y$ or $x \sim y$. If $x \succ y$, by Axiom 1.a. we must have $\neg(y \succ x)$, which by equivalence is $x \succeq y$. If $x \sim y$ then $\neg(y \succ x)$ is equivalent to $x \succeq y$.

b. By Axiom 1.b. $x \succ y$ implies either $x \succ z$ or $z \succ y$. Because $y \succeq x$ by equivalence we have $\neg(z \succ y)$. Since $z \succ y$ is impossible, hence $x \succ z$ must hold.

c. By Axiom 1.b. $y \succ z$ implies either $y \succ x$ or $x \succ z$. By equivalence, $x \succeq y \equiv \neg(y \succ x)$, thus $y \succ x$ cannot hold, hence, $x \succ z$ must be true.

□

PROPOSITION **4** (2.). *Supppose Axiom 1. and Axiom 2. hold. The independence axiom allows us to define a preference relation* $\succ_i$ *on* X_i *such that,*

a. *If* $\alpha_i \succ_i \beta_i$ *then*
$$(x_1, x_2, ..., \alpha_i, ..., x_{n-1}, x_n) \succ (x_1, x_2, ..., \beta_i, ..., x_{n-1}, x_n)$$
for all x_j *where* $j \neq i$.
b. *The preference relation* $\succ_i$ *is asymmetric and negatively transitive.*

Proof. a. This is straight forward as, Axiom 2. implies that $(y_1, y_2, ..., \alpha_i, ..., y_{n-1}, y_n) \succ (y_1, y_2, ..., \beta_i, ..., y_{n-1}, y_n)$ for any $(y_j)_{j \neq i}$.

b. From part a. there does not exist some $(x_j)_{j \neq i}$ such that $(x_1, x_2, ..., \beta_i, ..., x_{n-1}, x_n) \succ (x_1, x_2, ..., \alpha_i, ..., x_{n-1}, x_n)$ and $\neg(\beta_i \succ_i \alpha_i)$ hold; hence, $\succ_i$ is asymmetric.

To prove negative transitivity, let $\alpha_i \succ_i \beta_i$. We must check the following cases: (i.) if $\gamma_i \succ_i \beta_i$ and (ii.) $\neg(\gamma_i \succ_i \beta_i)$. In the first case, negative transitivity holds. As for case (ii.), $(x_1, x_2, ..., \beta_i, ..., x_{n-1}, x_n) \succeq (x_1, x_2, ..., \gamma_i, ..., x_{n-1}, x_n)$ for all for all x_j where $j \neq i$. Consequently, by part a. and asymmetry $(x_1, x_2, ..., \alpha_i, ..., x_{n-1}, x_n) \succ (x_1, x_2, ..., \gamma_i, ..., x_{n-1}, x_n)$. Thus, $\alpha_i \succ_i \gamma_i$ and negative transitivity is satisfied.

□

PROPOSITION **5** (3.). *Supppose Axiom 3. and asymmetry of* $>_o$ *hold. The decisiveness axiom allows us to define a preference relation* $\succ_i$ *on* X_i *such that,*

If $P(x,y) = \emptyset = P(y,x)$ *then (i.)* $\neg(x_i \succeq_i y_i) \quad \forall i$ *and* $\nexists i^* : x_{i^*} \succ_i y_{i^*}$, *(ii.)* $\neg(y_i \succeq x_i) \forall i$ *and* $\nexists i^* : y_{i^*} \succ_i x_{i^*}$ *and (iii.)* $P(x,y) = \emptyset = P(y,x)$ *then (i.) (ii.) and (iii.)* $\forall i (x_i \sim_i y_i) \Rightarrow x \sim y$.

THEOREM **2** (The Lexicographic machine). *Let* $\succ$ *be a binary relation on* $X = \times_{n \in N} X_n$. *The preference relation is lexicographic if and only if it satisfies asymmetric of weak order (A1), indepedence (A2), decisiveness (A3) and noncompensantion (A4).*

7.2 Theorem of characterization

Proof. **Theorem 1.** Suppose that solution $\pi(G^*)$ satisfies the axioms.

Take any sequence $s \in S(G)$ this can be either (i) a constant action sequence or (ii) a mixed strategy sequence, thus, players switch between two Pareto unranked action profiles. In (i) if s a constant string let be $s = (a, ...)$ by proposition 1 it follows that $u(s) = u(a) \geq \underline{u}$ are the best response for both players and it must be a Nash Equilibrium in one stage game. In case (ii) suppose if $s = (s_1, s_2)$ and $s' = (s'_1, s'_2)$ are two Pareto unranked, if $u(s) > \underline{u}$ but $u(s') < \underline{u}$ (the argument is similar in the other case) then any combination from both sequences is SPE but it violates proposition 2 because for $\alpha(s_1, s_2) + (1-\alpha)(s'_1, s'_2) < \underline{u}$ for $\alpha = 0$. Take s'' as a mixed strategy sequence form sequence s and s' (both are a constant action sequence) by contradiction, none of these are Nash equilibrium in one stage game, that is $\alpha(s_1, s_2) + (1 - \alpha)(s'_1, s'_2) = s''$ is neither a Nash equilibrium, it must be true that either s or s' is Nash equilibrium. Lexicographic efficiency follows inmediately since it represent the case of constant action sequence. □

References

Abreu, D., & Rubinstein, A. (1988). The structure of nash equilibrium in repeated games with finite automata. *Econometrica*, *56*(6), 1259-1281.

Agliardi, E. (2004). Axiomatization and economic theories: Some remarks. *Revue économique*, *55*(1), 123-129.

Aumann, R. J., & Shapley, L. S. (1976). Long-term competition: A game-theoretic analysis. *Essays in Game Theory*, 1-15.

Blonski, M., Peter, O., & Spagnolo, G. (2011). Equilibrium selection in the repeated prisoner's dilemma: Axiomatic approach and experimental evidence. *American Economic Journal: Microeconomics*, *3*(3), 164-192.

Fryer, R., & Jackson, M. (2003). Categorical cognition: A psychological model of categories and identi cation in decision making. *NBER*, *9579*.

Fudenber, D., & Levine, D. K. (1998). *The theory of learning in games*. MIT press.

Gibbons, R. (1992). *Game theory for applied economists*. Princeton University Press.

Green, E. (1982). Internal costs and equilibritum: The case of repeated prisoner's dilemma. *mimeo*.

Huck, S., & Sarin, R. (2004). Players with limited memory. *Contributions to Theoretical Economics*, *4*, 1109-1109.

Kalai, E., & Stanford, W. (1988). Finite rationality and interpersonal complexity in repeated games. *Econometrica*, *56*(2), 387-410.

Lehrer, E. (1988). Repeated games with stationary bounded recall strategies,. *Journal of Economic Thery*, *46*, 130-144.

Marschak, T., & McGuire, C. B. (1971). *Economic models for organization design.* (unpublished)

Mathevet, L. (2012). *An axiomatic approach to bounded rationality in repeated interactions: theory and experiments.* (Department of economics, University of Texas)

Monte, D. (2007). *Reputation and bounded memory in repeated games with imcomplete information* (Unpublished doctoral dissertation). Yale University.

Moore, E. (1956). Gedanken-experiments on sequential machines. In C. Shannon & J. McCarthy (Eds.), (p. 129-154). Princeton University Press.

Mullainathan, S. (2001). *Thinking through categories.* (workingpaper, MIT)

Neyman, A. (1985). Bounded complexity justi es cooperation in the finitely repeated prisoners dilemma. *Economic letters*, *19*, 227-229.

Radner, R. (1978). Can bounded rationaloty resolve the prisoner's dilemma? *mimeo*.

Rubinstein, A. (1979). Equilibrium in supergames with the overtaking criterion. *Journal of Economic Theory*, *21*, 1-9.

Rubinstein, A. (1986). Finite automata play the repeated prisioner's dilemma. *Journal of Economic Thery*, *39*, 83-96.

Rubinstein, A. (1998). *Modeling bounded rationality*. The MIT press.

Smale, S. (1980). The prisoner's dilemma and dynamical systems associated to non-cooperative games. *Econometrica*, *48*, 1617-1634.

Thomson, W. (2001). On the axiomatic method and its recent applications to game theory and resource allocation. *Social Choice and Welfare*, *18*(2), 327-386.