JOURNAL OF ECONOMIC THEORY 39, 83–96 (1986)

Finite Automata Play the Repeated Prisoner's Dilemma

ARIEL RUBINSTEIN*

Department of Economics, The Hebrew University, Jerusalem, Israel

Received December 19, 1984; revised October 23, 1985

The paper studies two-person supergames. Each player is restricted to carry out his strategies by finite automata. A player's aim is to maximize his average payoff and subject to that, to minimize the number of states of his machine. A solution is defined as a pair of machines in which the choice of machine is optimal for each player at every stage of the game. Several properties of the solution are studied and are applied to the repeated prisoner's dilemma. In particular it is shown that cooperation cannot be the outcome of a solution of the infinitely repeated prisoner's dilemma. *Journal of Economic Literature* Classification Numbers: 021, 022, 026.

1. INTRODUCTION

In the fifties Simon pointed out the importance of "bounded rationality" to economic theory (see Simon [10, 11]). Although Simon's ideas have received worldwide recognition we still find Simon pointing out "... an urgent need to expand the established body of economic analysis, which is largely concerned with substantive rationality, to encompass the procedural aspects of decision making" [11, p. 506]. The reasons why Simon's work has had a limited impact on economic theory are quite clear: it is difficult to embed the procedural aspects of decision making in formal economics models and we do not possess a unique natural theory describing these aspects. As economists we are confronted with a choice between waiting for a satisfactory description of the procedure of human decision making and analyzing somewhat artificial models capturing certain elements of "bounded rationality." I prefer the latter.

The term "bounded rationality" was used to cover a wide range of issues.

* Some of the work on this paper was done while I visited the Institute for International Economic Studies at the University of Stockholm in August 1983. I am grateful to this Institute for its hospitality. I benefited from the many discussions with Ken Binmore, Ed Green, Charles Wilson, and Asher Wolinsky, whom I would like to thank. I would also like to thank Doug McManus and Cathy Weinberger for their editorial comments.

0022-0531/86 $3.00

Therefore I would like to emphasize that I deal here only with one specific procedural aspect: rules of behavior are costly to operate and decision makers aim to minimize these costs. Thus, for example, I do not deal with the costs of computing optimal behavioral rules.

The cornerstone of the model is a two-person supergame. In the supergame, a game, G, is repeated sequentially an infinite number of times. At each repetition each player chooses a one period G-strategy, the choice of which may depend on the outcome in the previous periods. At the end of each period, the players receive the one period G-payoffs. Streams of payoffs are evaluated according to the criterion of the limit of the means. A supergame strategy is a plan how to play G at every period, conditional on every possible history.

There are three reasons for my choice of a supergame as the fundamental component of this model:

(a) The supergame has already been intensively analysed.

(b) The set of supergame strategies has a natural internal structure.

(c) We have strong intuitions about the relative complexity of supergame strategies.

This helps us to test the plausibility of the forthcoming definition of the complexity of a strategy.

Our first departure from the standard treatment of the supergame is in the strategy spaces. A player is required to play the repeated game using a kind of finite automaton called a (Moore) machine. A machine consists of a finite set of states, one of them an initial state, an output function, and a transition function. Given that the machine is at the state q^t in the tth round of the repeated game, the output function determines the one-shot game strategy that the player plays as a function of the element q^t. The transition function determines the state q^{t+1} as a function of the state q^t and of the opposing player's move at period t.

Finite automata have been used for the study of computer operation. Sometimes brain functioning is modeled as a finite automaton. However, the question of whether a finite automaton is a reasonable way to model a decision maker is certainly central to the evaluation of the current study. There is an artificial element in the description of a player's behavior as a machine. In the absence of a more established tool to model a decision maker I believe Moore's machine to be a reasonable tool for formalizing a player's behavior in a supergame.

If we just constrain the players in a supergame to choose only machines (rather than supergame strategies) we still get an extensive set of Nash equilibrium payoffs. This set includes all individually rational payoffs which are rational convex combinations of the one-shot game payoffs.

Here comes the second departure from the standard supergame model. The players are expected to take into consideration not only the supergame payoffs, but also the complexity of the machines they use. Therefore, before proceeding, we must formalize the notion of complexity. The formalization of this notion was a central topic in the theory of automata. Several sophisticated measures have been suggested in the literature. It seems that considerable work is needed to construct complexity measures that are relevant in economic contexts. Obviously the measure of complexity has to be carefully matched to the interpretation given to the "machines."

The formalization I am using in this paper is fairly naive. What counts is the number of states in the machine. To be more precise, imagine that the players bear the cost of maintaining the states of the machine. Each period a player has to pay a "small" fee for each of the states maintained in his machine. He pays the fee for every state he chooses to keep in his machine regardless of the frequency of its usage. I will refer to these costs as the *maintenance costs* of the machine.

As to the tradeoff between maintenance costs and the supergame payoffs I have studied here a limit case: the maintenance costs are infinitesimal. The players care primarily about the average payoff and they care lexicographically only secondarily about the maintenance costs.

The complexity notion, the definition of maintenance costs, and the trade-off between the maintenance costs and the supergame payoff will induce the definition of the solution for the model. A solution (called semi-perfect-equilibrium) is a pair of machines, one for each of the players, which satisfies at every stage of the game:

(a) Neither of the players can achieve a higher average payoff by a unilateral change of his own machine.

(b) Neither of the players is able to reduce the number of the states used.

This paper concentrates on the classical repeated prisoner's dilemma. (For a reference to the importance of the game in the development of game theory, see Luce and Raiffa [6].) Figure 1 describes the one-shot prisoner's dilemma.

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

FIGURE 1

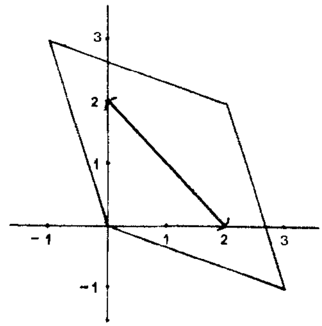

FIGURE 2

The only Nash equilibrium of the one-shot game is (D, D). In the repeated game (with the limit of the means evaluation criterion), any feasible and individually rational payoff is a repeated game payoff of a Nash equilibrium (the Folk theorem) and in fact even of a perfect equilibrium (for the study of the perfect equilibria of the repeated games see Rubinstein [9] and the references there).

In contrast, it will be shown here that a repeated game payoff of a solution must be either $(0, 0)$ or an internal point on the segment combining $(0, 2)$ and $(2, 0)$. (See Fig. 2.) The cooperative payoff vector $(2, 2)$ is not achieved by a solution.

Thus the current paper's approach is very different from that of three previous bounded rationality studies of the prisoner's dilemma in which cooperation is explained by "bounded rationality." Radner [8] applied the ε-equilibrium concept to a finitely repeated prisoner's dilemma. Under the title "Can Bounded Rationality Resolve the Prisoner's Dilemma?" he showed that the players can come "close" to the vector payoff $(2, 2)$ by using a pair of strategies which will be "almost" the best response of one against the other.

Smale [12] incorporates a bounded memory assumption. The players can retain in their memory only some kind of average of the past payoffs, and their strategies should be "good" in a well-defined way which captures a notion of bounded rationality. Again, the payoff $(2, 2)$ is established as the solution's payoff.

The closest to this paper is Green [4]. Green studies the finitely-repeated prisoner's dilemma. He assumes that the players use a restricted class of strategies. The use of a strategy is associated with a cost. Green's choice of the restricted class of strategies and of the costs is motivated by an intuitive argument in which a strategy is replaced by a Turing machine, and the machines vary in their costs.

The idea that finite automata theory may be useful for modelling bounded rationality in economic contexts is not new. Marschak and McGuire make this suggestion in unpublished notes [7]. Aumann [1] suggests the use of finite automata in the context of repeated games (see Aumann [1, p. 21].

I am aware of some studies of the complexity of multi-stage decision processes. Futia [2] concentrates on applications of an algebraic approach for complexity due to Rhodes. (See Futia [2] and the reference there.) This measure is quite sophisticated but complicated. It enables Futia to draw a few conclusions on the complexity of some stopping rules for sequential search problems. Varian, in an unpublished work [13], applied the algebraic automata theory approach to social decision theory. He also uses the number of states as a measure of complexity of a social welfare rule. For other references see Gottinger [3].

Following presentation of the model, I define the solution (Sect. 2). I then present some results regarding the structure of a solution for a general supergame (Sect. 3) and apply the results to the special case of the repeated prisoner's dilemma (Sect. 4). I refrain from any final conclusions. By all accounts this work should be considered only a small step forward on a very long path. I have my own doubts about some of the details of the model. Even so I hope that the paper will serve as a demonstration of the scope of the formal approach for the study of "bounded rationality" elements in the framework of "game theory."

2. The Model

2.1. *The Supergame*

Let $G = \langle S_1, S_2, u_1, u_2 \rangle$ be a two person game in normal form, where S_i is a finite set of strategies for player i and u_i: $S_1 \times S_2 \to R$ is player i's payoff function.

The supergame of G consists of an infinite sequence of repetitions of G taking place at periods $t = 1, 2, 3,\ldots.$ At period t the players make simultaneous moves denoted by $s_i^t \in S_i$ and then each player learns his opponent's move. A supergame strategy is a sequence of functions $\{f^t\}_{t=1}^{\infty}$, where f^t determines a player's choice of action at period t as a function of the previous $t-1$ outcomes.

In the standard supergame a player has to choose a supergame strategy. In the current model a player must choose a Moore machine.

2.2. *Moore Machines*

A Moore machine for player i, denoted M_i, is a four-tuple $\langle Q_i, q_i^0, \lambda_i, \mu_i \rangle$, where Q_i is a finite set, $q_i^0 \in Q_i$, $\lambda_i: Q_i \to S_i$, and $\mu_i: Q_i \times S_j \to Q_i$ $(j \neq i)$. The set Q_i is the set of the internal states of the machine M_i. The state q_i^0 is the initial state. The element $\lambda_i(q_i)$ is a strategy in the game G that player i chooses whenever his machine is at state q_i. The function μ_i is called a transition function. If the machine is at state q_i and the other player chooses $s_j \in S_j$ then the machine's next internal state is $\mu_i(q_i, s_j)$.

Given the players choose machines M_1 and M_2 the supergame is played as follows: At the first period the machine of player i starts at the state $q_i^1 = q_i^0$. Player i chooses $s_i^1 = \lambda_i(q_i^1) \in S_i$. In the second period the machine of player i moves into the state $q_i^2 = \mu_i(q_i^1, s_j^1)$ $(j \neq i)$. In general, given that the machines are at period t at the states q_1^t and q_2^t, player i chooses $s_i^t = \lambda_i(q_i^t) \in S_i$ and at period $t+1$ his state is $q_i^{t+1} = \mu_i(q_i^t, s_j^t)$. Thus a pair of machines generates deterministically a sequence of G's outcomes, $(s_1^t, s_2^t)_{t=1}^{\infty}$, and a seuquence of pairs of states, $(q_1^t, q_2^t)_{t=1}^{\infty}$.

2.3. *Examples*

The following is a list of examples of machines which carry out familiar strategies in the repeated prisoner's dilemma. Diagrams, called transition diagrams, are used to illustrate the machines. The vertices of the diagram correspond to the states. One of the states is indicated to be the starting point. A letter C (or D) on an arch connecting the state q to the state q' means that given that the machine is at state q and the other player's move C (or D) is observed, the machine's state is changed to q'. The letter below the circle of a state q indicates the value of the λ function of the machine at the state q.

(a) The one-state machine diagrammed in Fig. 3 plays C constantly:

$$Q = \{q^*\}, \qquad q^0 = q^*, \qquad \lambda(q^*) = C, \qquad \text{and} \qquad \mu(q^*, \cdot) \equiv q^*.$$

(b) The two-states machine represented in Fig. 4 carries out the "tit for tat":

$$Q = \{q_C, q_D\}, \quad q^0 = q_C, \quad \lambda(q_s) = s, \quad \text{and} \quad \mu(q, s) = q_s \quad \text{for } s = C, D.$$

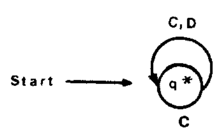

FIGURE 3

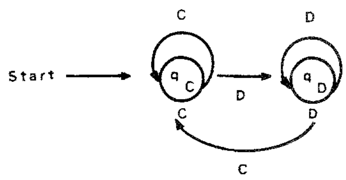

FIGURE 4

(c) The "grim strategy" (play C as long as the other player plays C) is executed by the machine diagrammed in Fig. 5:

$$Q=\{q_C, q_D\}, \quad q^0=q_C, \quad \lambda(q_s)=s, \quad \mu(q_C, s)=q_s, \quad \text{and} \quad \mu(q_D, \cdot)\equiv q_D.$$

(d) The strategy to play C until the other player plays D and then to punish him for three periods before returning to the cooperative behavior needs at least a four-states machine (see Fig. 6):

$$Q=\{q, p_1, p_2, p_3\}, \quad q^0=q, \quad \lambda(q)=C, \quad \lambda(p_h)=D \qquad (h=1, 2, 3),$$
$$\mu(q, C)=q, \quad \mu(q, D)=p_1, \quad \mu(p_h, \cdot)\equiv p_{h+1}, \quad (h=1, 2), \quad \text{and} \quad \mu(p_3, \cdot)\equiv q.$$

2.4. *The Solution Concept*

Let M_1 and M_2 be a pair of machines for the two players. Let $\mathbf{q}^t=(q_1^t, q_2^t)$, $t=1, 2,...$ be the sequence of states of the machines and $\mathbf{s}^t=(s_1^t, s_2^t)$, $t=1, 2,...$ be the sequence of actual plays of the supergame. Since the machines are finite there is a minimal t_2 and $t_2 \geqslant t_1$, such that $\mathbf{q}^{t_1}=\mathbf{q}^{t_2+1}$. Thus at period t_2+1 the pair of states repeats itself for the first time. I refer to the finite sequence $(\mathbf{q}^1,..., \mathbf{q}^{t_1-1})$ as the introductory part of the play by (M_1, M_2) and to the finite sequence $(\mathbf{q}^{t_1},..., \mathbf{q}^{t_2})$ as the cycle of the play of the pairs (M_1, M_2). The length of the cycle is denoted by T, $T=t_2-t_1+1$.

Define the supergame payoff of player i as the average payoff in the cycle, i.e.,

$$\pi_i(M_1, M_2)=\frac{1}{T}\sum_{t=t_1}^{t_2} u_i(\mathbf{s}^t).$$

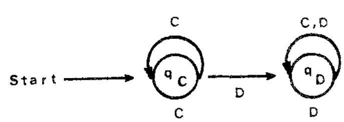

FIGURE 5

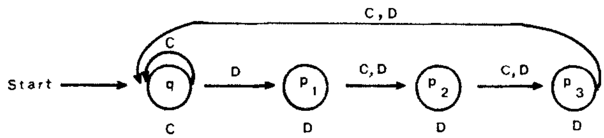

FIGURE 6

Let $M = \langle Q, q^0, \lambda, \mu \rangle$ be a machine. Denote by $|M|$ the number of states in Q; for $q \in Q$ denote by $M(q)$ the machine $\langle Q, q, \lambda, \mu \rangle$, that is the machine M starting at the state q. Let $>_L$ be the lexicographic order on $\mathbb{R}^2$. Define

$$(M_1, M_2) >_i (\bar{M}_1, \bar{M}_2) \quad \text{if } (\pi_i(M_1, M_2), -|M_i|) >_L (\pi_i(\bar{M}_1, \bar{M}_2), -|\bar{M}_i|).$$

Thus the pair of machines (M_1, M_2) is preferred by player i to the pair of machines $(\bar{M}_1, \bar{M}_2)$ if player i achieves in (M_1, M_2) a higher average payoff or if he gets the same average payoff using a machine with less states.

The pair (M_1^*, M_2^*) is said to be a *Nash equilibrium* if there is no M_1 or M_2 such that $(M_1, M_2^*) >_1 (M_1^*, M_2^*)$ or $(M_1^*, M_2) >_2 (M_1^*, M_2^*)$. The main definition in this paper is the following:

DEFINITION. A pair of machines (M_1^*, M_2^*) is a *semi-perfect-equilibrium* (SPE) if there is no time t, no M_1, and no M_2 such that

$$(M_1, M_2^*(q_2^t)) >_1 (M_1^*(q_1^t), M_2^*(q_2^t))$$

or

$$(M_1^*(q_1^t), M_2) >_2 (M_1^*(q_1^t), M_2^*(q_2^t)).$$

(Recall that q_i^t is the state of M_i^* at period t where the game is played by the pair of machines (M_1^*, M_2^*).)

Thus in a SPE there is no stage of the play of the game after which one of the players prefers to change his machine. The SPE differs from the Nash equilibrium concept in the requirement that the machine M_i^* be optimal for player i against M_j^* not only at the beginning of the game but also at the start of each repetition. The additional requirements are in the spirit of Selten's subgame perfection and the idea of dynamically consistency. However notice that the option of replacing the machine is not modelled as a part of the game. It appears only in the solution concept and is required only on the equilibrium path.

2.5. *A Discussion of the Solution Concept*

(a) *The Restriction of Strategies to Finite Automata*

The players are restricted to play the repeated game with finite automata. By itself this requirement is quite weak. Consider again the repeated prisoner's dilemma. For every rational convex combination $\alpha_1, \alpha_2, \alpha_3, \alpha_4$, such that

$$(u_1, u_2) = \alpha_1(2, 2) + \alpha_2(0, 0) + \alpha_3(3, -1) + \alpha_4(-1, 3) \geqslant (0, 0)$$

there exists a pair of machines (M_1, M_2) such that

$$(\pi_1(M_1, M_2), \pi_2(M_1, M_2)) = (u_1, u_2)$$

and none of the players can deviate and achieve a higher average payoff even by using a strategy which cannot be executed by a finite machine. The machines follow the well-known Folk theorem's equilibrium: they play (C, C) N_1 times, (D, D) N_2 times, (D, C) N_3 times, and (C, D) N_4 times, where $\alpha_h = N_h/N$, $N = \sum_{h=1}^{4} N_h$. After N periods they start again from the beginning. Where player i deviates, the other player, j, responds by moving to an absorbing state q where $\lambda_j(q) = D$ and $\mu_j(q, \cdot) \equiv q$. Note that this pair of machines is not SPE since given that player i is using M_i, the other player, j, can achieve the average payoff u_j even with a machine which does not include the punishment state.

(b) *A Comparison with Nash Equilibrium*

The "trigger" strategies as well as the "tit for tat" strategies are not even Nash equilibria in the machines game since they include states which are never used. If both players use the machines of Fig. 5, for example, then the state q_D is not used and each player is able to drop this state without affecting the supergame payoff.

The definition of SPE requires further that a solution includes only states which are used infinitely often.

Consider the machine for the repeated prisoner's dilemma which is represented by Fig. 7. If both players use this machine, we get a pair of machines which is a Nash Equilibrium and which is not SPE. Here, the

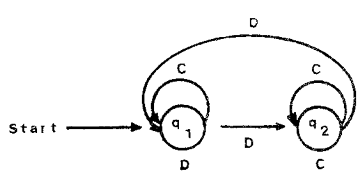

FIGURE 7

punishment phase is used in the introductory part of the play before the players reach the cycle. The instability of this pair of machines is due to the ability of the players to drop the state q_1 after the first period. From the point of view of period 2 the state q_1 is redundant.

(c) *Threats in SPE*

The definition of SPE does not exclude the possibility of players threatening their opponents. However, the punishment should have the feature that it could be executed through states of the machine which will be used anyway in the regular course of the game.

The idea is that if threatening demands resources, and in equilibrium they are not exercised, then it is not optimal to hold the threats unless the threatening machinery has some other functions as well.

The above considerations have some similarity to phenomena frequently observed in real life: social institutions, various types of organizations, and human abilities degenerate or are readily discarded if they are not used regularly.

(d) *The Trade-off Between the Supergame Payoff and the Procedural Costs*

A more natural model might allow a real trade-off between the supergame payoff and the procedural costs. Here a most extreme case is studied. The players care lexicographically primarily about the average payoff and only secondarily about the measure of complexity of the machine. This is done mainly for the sake of simplicity. It is an approximation of a situation in which the magnitude of the procedural costs is small relative to the supergame payoff. The strength of this assumption is that it allows a deviating player to achieve a higher payoff even using a very complicated machine.

(e) *The Measure of Complexity*

The measure of complexity of a machine which is used in this paper is the number of states held in the machine. This measure remain the measure for complexity throughout the course of the game. Thus, for example, it is not cheaper to maintain states if they were used in the past. The definition of SPE is motivated by the scenario described in the introduction. A player "pays" a "very small" fee per state per period that the state is held in its machine. Operating the machine with less states is desirable since it saves future costs associated with maintaining the extra states.

As mentioned earlier the measure of complexity used here is very naive. Other possible complexity measures may take into account the complexity of the transition function.

3. Properties of SPE

In this section two properties of SPE are proved. First, we shall show that during a cycle no player will repeat the same states twice. Note that the "memory" of a player is embedded in the name of the state. This conclusion means, therefore, that at a solution each player keeps track of his exact position in the cycle.

Second, we shall show that in a solution there is full coordination in timing the switching of the one-shot game strategies. Whenever one player changes his strategy, the other player changes his strategy too. This fact has a clear implication to repeated 2×2 matrix games, in which case at a solution the G-outcomes are all on one of the matrix's two diagonals.

In the following, let $\langle M_1^*, M_2^* \rangle$ be a pair of machines with a cycle $(\mathbf{q}^{t_1}, \ldots, \mathbf{q}^{t_2})$. Denote $\pi_i(M_1^*, M_2^*)$ by π_i^*. Given two number k and l between t_1 and t_2, define the set $T(k, l) \subseteq \{t_1, \ldots, t_2\}$ as the set of periods from k to l in the oriented cycle from t_1 to t_2; that is,

$$T(k, l) = \begin{cases} \{k, k+1, \ldots, l\} & k \leqslant l, \\ \{k, k+1, \ldots, t_2, t_1, t_1+1, \ldots, l\} & k > l. \end{cases}$$

Let $A_i(k, l)$ be i's average payoff in $T(k, l)$.

Proposition 1. *If $\langle M_1^*, M_1^* \rangle$ is a solution then:*

(a) *For all i, the states $q_i^{t_1}, \ldots, q_i^{t_2}$ are distinct.*

(b) *For all $t \geqslant t_1$, $\lambda_1^*(q_1^t) = \lambda_1^*(q_1^{t+1})$ if and only if $\lambda_2^*(q_2^t) = \lambda_2^*(q_2^{t+1})$.*

The proof of the proposition is divided into three simple Lemmas:

Lemma 1. *If $q_1^{k_1} = q_1^{k_2}$ for some $t_1 \leqslant k_1 < k_2 \leqslant t_2$ then $A_2(k_1, k_2 - 1) = \pi_2^*$.*

Proof. Assume that $A_2(k_1, k_2 - 1) > \pi_2^*$. (A similar proof is needed for the case where $A_2(k_1, k_2 - 1) < \pi_2^*$). Let us define a machine, M_2, for player 2 such that $\pi_2(M_1^*, M_2) = A_2(k_1, k_2 - 1)$. The machine M_2 includes $k_2 - 1$ states $p^1, \ldots, p^{k_2 - 1}$. Its initial state is p^1. The function λ_2 of the machine M_2 mimics M_2^*, that is, $\lambda_2(p^k) = \lambda_2(q_2^k)$ for all $1 \leqslant k \leqslant k_2 - 1$. The transition function, μ_2, satisfies $\mu_2(p^k, \lambda_1^*(q_1^k)) = p^{k+1}$ for $k < k_2 - 1$ and $\mu_2(p^{k_2 - 1}, \lambda_1^*(q_1^{k_2 - 1})) = p^{k_1}$. Clearly the cycle of (M_1^*, M_2) is $((q_1^{k_1}, p^{k_1}), \ldots, (q_1^{k_2 - 1}, p^{k_2 - 1}))$ and $\pi_2(M_1^*, M_2) = A_2(k_1, k_2 - 1) > \pi_2^*$.

Lemma 2. *For all i and for all $t_1 \leqslant k_1 < k_2 \leqslant t_2$, $q_i^{k_1} \neq q_i^{k_2}$.*

Proof. Let (k, h) be the lexicographically minimal pair of integers $k > k - h = l \geqslant t_1$, such that there exists a player i satisfying $q_i^l = q_i^k$. Assume that $i = 1$; $q_2^l \neq q_2^k$, otherwise the cycle is shorter. The number of the states of M_2^* is at least $h+1$ unless $k=t_2+1$. Next define an h-state machine, M_2, ensuring that $\pi_2(M_1^*, M_2)=A_2(l, k-1)$ (and by Lemma 1 also equal to π_2^*). The machine M_2 includes the states $p^l, \ldots, p^{k-1}$. The function λ_2 satisfies $\lambda_2(p^j)=\lambda_2^*(q_2^j)$ and the transition function μ_2 satisfies $\mu_2(p^j, \lambda_1^*(q_1^j))=p^{j+1}$ for $j<k-1$ and $\mu_2(p^{k-1}, \lambda_1^*(q^{k-1}))=p^l$. The initial state of M_2 is p^l. The cycle of $(M_1^*(q_1^l), M_2)$ is $((q_1^l, q_2^l), \ldots, (q_1^{l+h-1}, q_2^{l+h-1}))$ and thus $\pi_2^*(M_1^*(q_1^l), M_2)=\pi_2^*$ as long as $k \neq t_2+1$, $|M_2|<|M_2^*|$, therefore, $k=t_2+1$ and $q_i^{k_1} \neq q_i^{k_2}$ for all $t_1 \leqslant k_1 < k_2 \leqslant t_2$ and for all i.

LEMMA 3. *There is no $j \geqslant t_1$ such that $\lambda_1^*(q_1^j) \neq \lambda_1^*(q_1^{j+1})$ and $\lambda_2^*(q_2^j)=\lambda_2^*(q_2^{j+1})$.*

Proof. If otherwise, player 2 would deviate at t_1 by a machine M_2 which satisfies $\pi_2(M_1^*, M_2)=\pi_2^*$ and $|M_2|<|M_2^*|$. The machine includes all the states of M_2^* with the exception of q_2^{j+1}. The initial state of M_2 is $q_2^{t_1}$ if $j \neq t_2$, and is $q_2^{t_2}$ if $j=t_2$. The function λ_2 is as λ_2^*. The transition function of M_2 is modified from μ_2^* such that $\mu_2(q_2^j, \lambda_1(q_1^j))=q_2^j$ and $\mu_2(q_2^j, \lambda_1(q_1^{j+1}))=q_2^{j+2}$. Note the simple idea behind the proof. In M_2^* player 2 uses q_2^{j+1} only for counting the periods. Since player 1 is behaving differently in periods j and $j+1$, player 2 could avoid the need to use q_2^{j+1} and instead rely on the "free" service that player 1 provides him by switching from $\lambda_1^*(q_1^j)$ to $\lambda_1^*(q_1^{j+1})$.

4. THE REPEATED PRISONER'S DILEMA

In this section, Proposition 1 is used for characterizing the SPE payoffs in the repeated prisoner's dilemma.

PROPOSITION 2. *The pair (π_1^*, π_2^*) is a solution's payoff if and only if $(\pi_1^*, \pi_2^*)=(0,0)$ or where there is a rational number α such that $(\pi_1^*, \pi_2^*)=\alpha(3,-1)+(1-\alpha)(-1,3)>(0,0)$.*

Proof. Let us start with the "necessary" part. Assume that (M_1^*, M_2^*) is a solution such that $\pi_i(M_1^*, M_2^*)=\pi_i^*$. If $\pi_i^* \leqslant 0$ then M_i^* must be the single state machine which plays D constantly. In such a case M_j^* must be the same as M_i^* and $(\pi_1^*, \pi_2^*)=(0,0)$.

After the introductory part the outcomes must, according to Proposition 1, include (C, C) and (D, D) only or (C, D) and (D, C) only. Assume that in the cycle only (C, C) and (D, D) are played and that (C, C) is played in the cycle at least once. Let k_1 satisfy the condition that $(\lambda_1^*(q_1^{k_1}), \lambda_2^*(q_2^{k_1}))=(C, C)$. Let $\hat{k}_2$ satisfy that $\mu_2(q_2^{k_1}, D)=q_2^{\hat{k}_2+1}$. Then $A_1(k_1+1, \hat{k}_2)>\pi_1^*$, since otherwise $A_1(\hat{k}_2+1, k_1) \geqslant \pi_1^*$ and player 1 can deviate profitably by making the change in M_1^* such that $\lambda_1(q_1^{k_1})=D$ and

$\mu_1(q_1^{k_1}, C) = q_1^{\hat{k}_2+1}$. Since $A_1(k_1+1, \hat{k}_2) > \pi_1^* > 0$, we can choose $k_2, k_1+1 < k_2 \leqslant \hat{k}_2$ to satisfy that $A_1(k_1+1, k_2) > \pi_1^*$ and $(\lambda_1^*(q_1^{k_2}), \lambda_2^*(q_2^{k_2})) = (C, C)$.

In the same way we can continue to choose a sequence $k_1, k_2, k_3, k_4, \ldots$. Because of the finite number of states in the machines we eventually reach $l \neq m$ such that $k_l = k_m$. Then the contradiction to $\pi_1(M_1^*, M_2^*) = \pi_1^*$ is straightforward.

Finally, let us turn to the sufficiency part of the proposition. Let N_1 and N_2 be two natural numbers such that $(\pi_1^*, \pi_2^*) = (1/N)[N_1(3, -1) + N_2(-1, 3)] > (0, 0)$, where $N = N_1 + N_2$. Let us construct a solution (see Fig. 8) whereby in the course of its cycle the players will play (D, C) N_1 times and (C, D) N_2 times. Define $M_1^* = \langle Q_1^*, q_1^*, \lambda_1^*, \mu_1^* \rangle$, $Q_1^* = \{q_1^1, \ldots, q_1^N\}$, $q_1^* = q_1^1$,

$$\lambda_1^*(q_1^k) = \begin{cases} D, & k \leqslant N_1 \\ C, & N_1 < k \leqslant N' \end{cases}$$

$$\mu_1^*(q_1^k, s_2) = \begin{cases} q_1^1, & s_2 = D; \quad k \leqslant N_1, \\ q_1^{k+1}, & s_2 = C; \quad k \leqslant N_1, \\ q_1^{k+1 (\mathrm{mod}\, N)}, & s_2 = C, D; \quad N_1 < k \leqslant N. \end{cases}$$

Similarly define M_2^*. Clearly (M_1^*, M_2^*) has the desired N periods cycle and neither of the players can deviate and increase his average payoff. Let us verify that for a player to achieve π_i^* he needs a machine with at least N states.

Assume that M_2 is a machine for player 2 such that $\pi_2(M_1^*, M_2) = \pi_2^*$. It follows that the length of the cycle of the play of the game by M_1^* and M_2 is at least N. Player 1 must play C in the cycle at least once. Therefore, M_1^* reaches one of the states $q_1^{N_1+1}, \ldots, q_1^N$ at least once during the cycle of (M_1^*, M_2). Due to the structure of M_1^* it must pass sequentially through the block $[q_1^{N_1+1}, \ldots, q_1^N]$. In order to return to $q_1^{N_1+1}$ it has to go through all the states $[q_1^1, \ldots, q_1^{N_1}]$. Therefore the sequence of player 1's states in the cycle of (M_1^*, M_2) must each be composed of blocks of the type $[q_1^1, \ldots, q_1^{k_1}]$

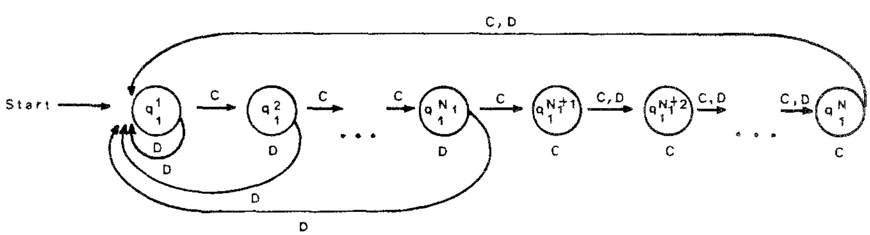

FIGURE 8

$[q_1^1,\ldots, q_1^{k_2}],\ldots, [q_1^1,\ldots, q_1^{k_l}]\ [q_1^{N_1+1},\ldots, q_1^N]$, where $1 \leqslant k_1,\ldots, k_l \leqslant N_1$ and $k_l = N_1$. In this way player 2 can only meet the average π_2^* by ensuring that player 1 will use his states in the cycle in blocks of

$$[q_1^1,\ldots, q_1^{N_1}]\ [q_1^{N_1+1},\ldots, q_1^N].$$

For this a necessary requirement for M_2 is that player 2 plays C whenever 1's state is one from among $q_1^1,\ldots, q_1^{N_1}$ and for achieving the average π_2^* player 1 must play D whenever 1's state is taken from among $q_1^{N_1+1},\ldots, q_1^N$.

Let $[p^1,\ldots, p^N]$ be the states of M_2 which are used by player 2 parallel to one of the appearances of a series of states $[q_1^1,\ldots, q_1^{N_1}]\ [q_1^{N_1+1},\ldots, q_1^N]$ in the plays of the game. Clearly $p^{N_1} \neq p^{N_1+1}$. Therefore, $p^{N_1-1} \neq p^{N_1}$ since they compel different μ-responses to the choice of C by player 1, $\mu_2(p^{N_1-1}, C) = p^{N_1}$ while $\mu_2(p^{N_1}, C) = p^{N_1+1}$. Since λ_2 differs on p^{N_1-1} and p^{N_1+1} it also follows that $p^{N_1-1} \neq p^{N_1+1}$. Repeating this argument it is easy to show that all $(p^h)_{h=1}^N$ are different.

References

1. R. J. Aumann, Survey of repeated games, *in* "Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern," pp. 11–42, Bibliograph. Inst., Zurich, 1981.
2. C. Futia, The complexity of economic rules, *J. Math. Econ.* **4** (1977), 289–299.
3. H. W. Gottinger, "Coping with Complexity," Reidel, Dordrecht, 1983.
4. E. Green, Internal costs and equilibrium: The case of repeated prisoner's dilemma, mimeo, November 1982.
5. J. E. Hopcroft and J. D. Ullman, "Introduction to Automata Theory, Language and Computation," Addison–Wesley, Reading, Mass., 1979.
6. R. D. Luce and H. Raiffa, "Games and Decisions," Wiley, New York, 1957.
7. T. A. Marschak and C. B. McGuire, Economic models for organization design, unpublished lecture notes, 1971.
8. R. Radner, Can bounded rationality resolve the prisoner's dilemma?, mimeo, 1978.
9. A. Rubinstein, Equilibrium in supergames with the overtaking criteria, *J. Econ. Theory* **21** (1979), 1–9.
10. H. A. Simon, Theories of bounded rationality, *in* "Decision and Organization" (C. B. McGuire and R. Radner, Eds.), North-Holland, Amsterdam, 1972.
11. H. A. Simon, On how to decide what to do, *Bell J. Econ.* **9** (1978), 494–507.
12. S. Smale, The prisoner's dilemma and dynamical systems associated to non-cooperative games, *Econometrica* **48** (1980), 1617–1634.
13. H. R. Varian, Complexity of social decisions, mimeo, 1975.