JOURNAL OF ECONOMIC THEORY 61, 160–168 (1993)

Notes, Comments, and Letters to the Editor

Finite Automata Play a Repeated Extensive Game

MICHELE PICCIONE

Department of Economics, University of British Columbia, 997-1873 E Mall, Vancouver, British Columbia, Canada V6T 1Z1

AND

ARIEL RUBINSTEIN

Department of Economics, Tel Aviv University, Tel Aviv, Israel and Department of Economics, Princeton University, New Jersey 08544

Received March 9, 1992

This paper analyzes a two-player game in which each player has to choose an automaton (machine) which plays an infinitely repeated *extensive game*. We assume that the preferences of the player depend both on repeated game payoffs and the number of states of their machine. In contrast to repeated normal form games, it is shown that if the stage-game is an extensive game with perfect information, any Nash equilibrium of the machine game will induce a path consisting of a constant play of a Nash equilibrium of the stage-game. *Journal of Economic Literature* Classification number: C72. © 1993 Academic Press, Inc.

1. INTRODUCTION

This paper continues a line of research investigating the notion of bounded rationality within the theory of infinitely repeated games. We will study a model, called the *machine game*, in which two players play an infinitely repeated game by choosing automata (machines) which implement their strategies. We assume that the preferences of each player over machines depend both on the discounted sum of stage-game utilities and on the complexity of the machine which the player chooses.

This line of research was initiated by Rubinstein [9] and Abreu and Rubinstein [2].[1] Abreu and Rubinstein [2] study the structure of Nash equilibria in the machine game. They show that in any equilibrium for the

[1] In these papers the limit of the means case is considered as well.

0022-0531/93 $5.00

machine game there is a one-to-one correspondence between the stage-game actions played by each of the two players. This implies substantial restrictions on the set of equilibrium payoffs for a wide class of games. For example, in the infinitely repeated Prisoner's Dilemma any Nash equilibrium path of outcomes can consist only of outcomes from the set $\{(C, C), (D, D)\}$, or the set $\{(C, D), (D, C)\}$. The proof of Abreu and Rubinstein [2] was improved in Piccione [8]; this improved version is applied in the proof of the main proposition of this paper.[2]

In the above literature, the stage-game is always assumed to be a two-person normal form game. In each period players move simultaneously and, at the end of each round, they obtain full information about their opponent's choice of stage-game strategies. In the current paper, we extend the analysis to the case where the stage-game is a two-person *extensive form* game. We assume that at the end of each play, a player obtains information only about the actual *terminal node* which has been reached. Note that in such a framework, a terminal node does not convey information about the choice of actions at information sets off the path which leads to it. This restriction causes the set of equilibria of the machine game consisting of the infinite repetition of an extensive game Γ to differ substantially from the set of equilibria of the machine game consisting of the infinite repetition of the reduced normal form of Γ.[3]

For example, consider the following two-player normal form game G:

	A	B
A	3, 1	1, 3
B	2, 0	2, 0

For sufficiently large discount factors, the set of Nash (or Subgame Perfect) equilibrium paths includes all those which yield an average payoff vector above (2, 0). It contains, for example, paths that consist of combinations of

[2] Banks and Sundaram [3], Lipman and Srivastava [5], Neme and Quintas [6, 7], also analyze the machine game under the assumption that players not only want to increase their repeated game payoff but also reduce the complexity of their machines.

[3] The sets of perfect equilibria for the infinite repetition Γ and for the infinite repetition of its reduced normal form G are different. The difference emerges from the fact that in the repeated game of Γ the perfection requirement applies to more nodes of decision than in the repeated game of G. Rubinstein and Wolinksy [10] provide several examples of *extensive* games for which the set of subgame perfect equilibrium payoff vectors of the repeated game with discounting is very different from that of the corresponding repeated reduced normal form game even when the discount factor is close to 1. Nevertheless, it is true that with a "dimensionality" condition (such as in Fudenberg and Maskin [4], or, Abreu and Dutta [1]) all feasible and strictly individually rational payoff vectors are subgame perfect equilibrium payoff vectors when the discount rate approaches one.

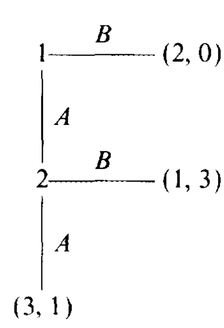
(A, A) and (A, B) which assign a sufficiently large weight to (A, A). This conclusion no longer holds for Nash equilibria if we assume that players play the machine game defined in Abreu and Rubinstein [2]. Any equilibrium play of the machine game results in an introductory phase and a cyclical phase which are composed only of combinations of (A, A) and (B, B). If the discount factor is sufficiently close to one and complexity costs are sufficiently small, any combination of (A, A) and (B, B) in the cyclical phase can be generated on an equilibrium path.

We can construct an equilibrium of the machine game in which the cyclical phase consists of a constant play of (A, A) by assuming that players choose two identical machines such as the following (q_1 is the initial state):

		Transition after	
State	Output	A	B
q_1	B	q_1	q_2
q_2	B	q_1	q_3
q_3	B	q_1	q_4
q_4	A	q_4	q_1

In this equilibrium, each machine starts with a show of power in which it plays B for three periods, after which it switches to a constant play of A. This is supported as an equilibrium by player 1's threat to punish any deviation made by player 2 with a three period play of B and by an analogous threat of player 2. Note that it is essential that player 1 monitors player 2 during the first three periods of the game despite the fact that player 2 does not gain by deviating. If in any of these periods player 1 does not deter a deviation by player 2, the latter can use the same state as in period 4 during the initial phase and thus reduce the complexity of his machine.

The game G is the reduced normal form game of the extensive game Γ:



Consider the machine game in which Γ, rather than G, is repeated over time and suppose that at the end of each period players are informed only of the terminal node which they have reached. It can be shown that the only equilibrium outcome path consists of a constant play of the stage-game Nash equilibrium (B, B). Part of the intuition behind this result lies in the reason why the above equilibrium for the machine game with G is not an equilibrium for the machine game with Γ. If player 1 chooses B, he cannot verify whether player 2 would play B in the event that player 2's decision node were reached. Thus, player 1 cannot monitor player 2 through the "introductory phase," and since player 2 does not have the ability to punish, the equilibrium collapses.

This paper provides a generalization of this result. It is shown that if Γ is an extensive game with perfect information, any equilibrium outcome path for the machine game consists of an infinite repetition of a Nash equilibrium of Γ irrespective of the discount factors.

We believe that this result is of interest in and of itself. However, the model also serves two additional purposes. First, it examines the notion of "strategic complexity" in extensive games. Second, it demonstrates that when bounded rationality considerations are introduced, a significant difference between an extensive game and its reduced normal form may arise.

2. The Model

We now give a formal presentation of the model. Let Γ be a two-player extensive-form game with perfect recall. Let S_i denote the set of pure strategies for player i and let S denote $S_1 \times S_2$. The set of terminal nodes of the game tree is denoted by E. Let $h_i(e)$ be the payoff obtained by player i upon reaching the end-node e and let $E(s)$ be the end-node which is reached when the strategy profile s is played. U_i denotes the set of information sets of player i and $A(u_i)$ is the set of actions available to player i at $u_i \in U_i$. Given $s_i \in S_i$, let $s_i(u_i)$ denote the action induced by s_i at u_i.

The machine game is constructed as follows. We assume that players play the infinite repetition of Γ by choosing machines and that player i evaluates his stream of payoffs by using a discount factor $0 < \delta_i < 1$. Several options are available to extend the definition of a machine used in a repeated normal form game to the model under consideration. Recall that if the stage game is a normal game G, a machine is defined as a four-tuple $M_i = \langle Q_i, q_i^1, \lambda_i, \mu_i \rangle$, where Q_i is a finite set of states ; q_i^1 is the initial state; $\lambda_i : Q_i \to A_i$ is the output function, where A_i is the set of actions of player i in G; and $\mu_i : Q_i \times E \to Q_i$ is the transition function, where $E = A_1 \times A_2$. Note that each state is assigned one action and state transitions can occur at the end of a period after the outcome of that period has been realized.

We interpret an extensive game Γ as a representation of a strategic situation in which a player does not have to calculate himself the information set which he is at. This information is given to a player by an external source (the "master of the game"). We therefore require that each state be assigned an action for each information set of Γ and we define the output function to be $\lambda_i: Q_i \times U_i \rightarrow A(U_i)$, where $\lambda_i(q_i, u_i) \in A(u_i)$.

With regard to state transitions, a natural extension is to allow a change in the state each time new information arrives, i.e., when a player is told either that a particular information set has been reached or that he is at the end of a period and is informed about the terminal node that the play of game Γ has led to. We can then define the transition function $\mu_i: Q_i \times (U_i \cup E) \rightarrow Q_i$ with the convention that the transition of a state occurs at the time an information set is reached and *before an action is taken*. Thus, in contrast to the case of the infinite repetition of G, a change in the state can occur during the play of the one-shot game as well as at the end of a play. This contrast is, however, redundant in this framework. Without loss of generality, we can restrict our attention to *simple* machines where the transition function is such that a change in the state occurs only upon reaching an end-node of Γ. This is due to the following. Consider a state $q_i \in Q_i$ and recall that we limit our discussion to games with perfect recall. If the machine M_i is at q_i at the beginning of a play of Γ then, for every information set u_i, the machine M_i associates with q_i a unique arrival state $q(q_i, u_i)$ and an action $\lambda_i(q(q_i, u_i), u_i)$. Also, for every end-node e, there is a unique state $q(q_i, e)$ which the machine associates with q_i. The state $q(q_i, e)$ will initiate the play of player i for the following period given that the terminal node obtained by the previous play of Γ is e. Thus, defining output at state q_i to be the Γ-strategy $(\lambda_i(q(q_i, u_i), u_i)\}_{u_i \in U_i}$ and defining the transition after q_i and the terminal node e to be $q(q_i, e)$, we obtain a simple machine which does not switch states along the play of Γ and is equivalent to M_i in the sense that after every history it plays the same actions as M_i.

In what follows, we then define a *machine* for player i as a four-tuple $M_i = \langle Q_i, q_i^1, \lambda_i, \mu_i \rangle$, where Q_i is a finite set of states, q_i^1 is the initial state $\lambda_i: Q_i \rightarrow S_i$ is the output function, and $\mu_i: Q_i \times E \rightarrow Q_i$ is the transition function. A transition in a state occurs at the end of each period after the terminal node is announced.

A pair of automata (M_1, M_2) induces a sequence of stage-game strategy pairs (s^t). Let $\pi_i(M_1, M_2) = \sum_{t=1}^{\infty} \delta_i^{t-1} h_i(E(s^t))$ be i's repeated game payoff resulting from the pair (M_1, M_2). The complexity of M_i is assumed to be the number of states in M_i and is denoted by comp(M_i). We assume that the preferences of the players depend only on the repeated game payoff and the complexity of their own machine and that player i strictly prefers (M_1, M_2) to (L_1, L_2), which is denoted by $(M_1, M_2) >_i (L_1, L_2)$, whenever

(i) $\pi_i(M_1, M_2) > \pi_i(L_1, L_2)$ and $\text{comp}(M_i) = \text{comp}(L_i)$ or

(ii) $\pi_i(M_1, M_2) = \pi_i(L_1, L_2)$ and $\text{comp}(M_i) < \text{comp}(L_i)$.

No additional assumptions concerning the tradeoff between "utility payoff" and "complexity" are required to derive the results in this paper.

A pair of automata (M_1, M_2) is a *Nash equilibrium* of the machine game if there is no player i and machine M_i' such that $(M_i', M_j) >_i (M_i, M_j)$.

Remark. An alternative view of an extensive game is that of a situation in which players do not know their position unless they "calculate" it. We felt, however, that the cost of obtaining information about past moves in the one-shot game is significantly smaller than the cost of holding information about previous plays of the game. For example, if Γ is a "take it or leave it" game where player 1 is the offerer, it is reasonable to assume that player 2 knows what the offer is that he has to respond to at the time he has to say "Y" or "N", even though he may "immetiately" forget this information.

3. The Nash Equilibrium of the Machine Game

Our first result is identical to one obtained in Abreu and Rubinstein [2] for the case of repeated normal form games:

LEMMA. *If (M_1, M_2) is an equilibrium for the machine game then* $\text{comp}(M_1) = \text{comp}(M_2)$.

Proof. Consider the equilibrium machine M_j for player j and a policy function $b_i: Q_j \to S_i$ which maximizes the discounted flow of player i's stage-game utilities given M_j. Consider the machine $M_i' = \{Q_j, q_j^1, \lambda_i', \mu_i'\}$ for player i defined by $\lambda_i'(q_j) = b_i(q_j)$ and $\mu_i'(q_j, \cdot) = \mu_j(q_j, E(b_i(q_j), \lambda(q_j)))$. The machine M_i' implements the optimal policy function b_i and is such that $\text{comp}(M_i') = \text{comp}(M_j)$. It follows that the equilibrium machine M_i has to satisfy $\text{comp}(M_i) \leqslant \text{comp}(M_j)$. Equality follows by a symmetric argument. ∎

PROPOSITION 1. *Suppose (s^t) is a sequence of profiles of Γ-strategies induced by an equilibrium of the machine game. If there exist $s_1 \in S_1$ and $r \neq k$ such that*

$$E(s_1, s_2^r) = E(s_1^r, s_2^r) \quad \text{and} \quad E(s_1, s_2^k) = E(s_1^k, s_2^k), \tag{$*$}$$

then $E(s_1^r, s_2^r) = E(s_1^k, s_2^k)$.

Discussion. Proposition 1 claims that if there is a Γ-strategy s_1 such that, when played with s_2^r and s_2^k, it induces the terminal nodes which are realized in periods r and k, then those terminal nodes must be identical. Intuitively, if they were not identical, player 1 could save at least one state by choosing a machine which replaces two distinct states with a single one, the output of which is s_1, and which uses the different terminal nodes to perform the "correct" transition to the subsequent states. Of course, an analogous claim holds for player 2.

Proof. Suppose that (M_1, M_2) is an equilibrium for the machine game and that there are two periods r and k and a Γ-strategy s_1 such that $E(s_1, s_2^r) = E(s_1^r, s_2^r)$ and $E(s_1, s_2^k) = E(s_1^k, s_2^k)$. It is easy to show that, given player 2's machine M_2, there exist an optimal policy for player 1, $b_1: Q_2 \to S_1$ such that $b_1(q_2^r) = b_1(q_2^k) = s_1$. We now show that $E(s_1^r, s_2^r) \neq E(s_1^k, s_2^k)$ leads to a contradiction. We construct a machine $M_1' = \{Q_1', q_1', \lambda_1', \mu_1'\}$ in which the state set is $Q_1' = \{Q_2 - \{q_2^r, q_2^k\}\} \cup \{\hat{q}\}$, the initial state is q_2^1 if $q_2^1 \notin \{q_2^r, q_2^k\}$ and $\hat{q}$ otherwise, the output function is such that $\lambda_1'(q_2) = b_1(q_2)$ for $q_2 \neq \hat{q}$ and $\lambda_1'(\hat{q}) = s_1$, and the transition function is such that

for $q_2 \notin \{q_2^r, q_2^k\}$

$$\mu_1'(q_2, \cdot) = \begin{cases} \mu_2(q_2, E(b_1(q_2), \lambda_2(q_2))) & \text{if } \mu_2(q_2, E(b_1(q_2), \lambda_2(q_2))) \notin \{q_2^r, q_2^k\} \\ \hat{q} & \text{if } \mu_2(q_2, E(b_1(q_2), \lambda_2(q_2))) \in \{q_2^r, q_2^k\} \end{cases}$$

for $q_2 \in \{q_2^r, q_2^k\}$

$$\mu_1'(\hat{q}, E(s_1, \lambda_2(q_2))) = \begin{cases} \mu_2(q_2, E(s_1, \lambda_2(q_2))) & \text{if } \mu_2(q_2, E(s_1, \lambda_2(q_2))) \notin \{q_2^r, q_2^k\} \\ \hat{q} & \text{if } \mu_2(q_2, E(s_1, \lambda_2(q_2))) \in \{q_2^r, q_2^k\} \end{cases}$$

M_1' keeps track of the states of M_2 except for states q_2^r and q_2^k in which case it uses the same state $\hat{q}$. Since $E(s_1, s_2^r) \neq (E(s_1, s_2^k)$, M_1' can switch from state $\hat{q}$ to states q_2^{r+1} and q_2^{k+1}. Thus, M_1 implements b_1. By Lemma 1, $\text{comp}(M_1') = \text{comp}(M_1) - 1$ which is a contradiction. ∎

We are now prepared to present the main proposition:

PROPOSITION 2. *If Γ is a game of perfect information then any equilibrium of the machine game consists of an infinite repetition of a Nash equilibrium of Γ.*

Proof. Let (s'_1, s'_2) and (s''_1, s''_2) be two pairs of Γ-strategies and suppose that both are observed along the path induced by an equilibrium of the machine game. We first show that $E(s'_1, s'_2) = E(s''_1, s''_2)$. Suppose not and let $u^* \in U_i$ be the first information set in which the two pairs of strategies induce different actions. By the perfect information assumption, we can decompose player j's set of information sets, $j \neq i$, into mutually exclusive sets A_k, $k = 1, ..., 4$ such that

A_1 contains all of j's information sets which precede u^*

A_2 contains all of j's information sets which succeed $s'_i(u^*)$

A_3 contains all of j's information sets which succeed $s''_i(u^*)$

A_4 contains all other information sets of j.

Let s_j be a strategy for player j which agrees with s'_j and s''_j on A_1, with s'_j on A_2 and with s''_j on A_3 and which is defined arbitrarily on A_4. Obviously, $E(s_j, s'_i) = E(s'_j, s'_i)$ and $E(s_j, s''_i) = E(s''_j, s''_i)$. Then, by Proposition 1, $E(s'_1, s'_2) = E(s''_1, s''_2)$. Since any equilibrium induces a constant play of a Γ-strategy pair, any equilibrium machine's set of states is a singleton set. It follows that the pair of Γ-strategies played by the machines must be a Nash equilibrium of Γ since, otherwise, one of the players could deviate profitably with a one-state machine. ∎

Discussion. The main difference between the repetition of an extensive game with perfect information and the repetition of its reduced normal form is that while in the latter a player observes his opponent's entire strategy, in the former a player observes only that part of the strategy which is realized. Once complexity considerations are included in the model, it is the lack of monitoring of behavior off the equilibrium path which causes the collapse of non-degenerate equilibria. The intuition behind the result is the following: if there is an equilibrium path with two different Γ-outcomes, there is a player who plays different Γ-strategies to implement the two outcomes and a player who can fulfill his role in the implementation of the outcomes with one Γ-strategy. Therefore, rather than holding extra states, the latter player can rely on the former player's actions, as revealed in the play of Γ, to obtain the information which is needed to conform to the anticipated routine. However, when using a smaller machine, a player loses the potential for "controlling" the opponent and the equilibrium collapses.

References

1. D. Abreu and P. Dutta, The folk theorem for discounted repeated games: A new condition, mimeo (1991).
2. D. Abreu and A. Rubinstein, The structure of Nash equilibrium in repeated games with finite automata, *Econometrica* **56** (1988), 1259–1282.
3. J. Banks and R. Sundaram, Repeated games, finite automata and complexity, *Games Econ. Behav.* **2** (1990), 97–117.
4. D. Fudenberg and E. Maskin, A non-cooperative equilibrium in repeated games with discounting and with incomplete information, *Econometrica*, **54** (1986), 533–554.
5. B. Lipman and S. Srivastava, Informational requirements and strategic complexity in repeated games, *Games Econ. Behav.* **2** (1990), 273–290.
6. A. Neme and L. Quintas, Equilibrium of repeated games with cost of implementation, *J. Econ. Theory* **58** (1992), 105–109.
7. A. Neme and L. Quintas, "Subgame Perfect Equilibrium of Repeated Games with Cost of Implementation," mimeo (1990).
8. M. Piccione, Finite automata equilibrium with discounting, *J. Econ. Theory* **56** (1992), 180–193.
9. A. Rubinstein, Finite automata play the repeated prisoner's dilemma, *J. Econ. Theory* **39** (1986), 83–96.
10. A. Rubinstein and A. Wolinksy, A remark on infinitely repeated extensive games mimeo, Tel Aviv University and Northwestern University (1991).