GAMES AND ECONOMIC BEHAVIOR 1, 213–221 (1989)

Bounded versus Unbounded Rationality: The Tyranny of the Weak*

ITZHAK GILBOA†

Department of Managerial Economics and Decision Sciences, J. L. Kellogg Graduate School of Management, Northwestern University

AND

DOV SAMET‡

Department of Managerial Economics and Decision Sciences, J. L. Kellogg Graduate School of Management, Northwestern University, Evanston, Illinois 60208, and Recanati School of Business, Tel-Aviv University, Tel-Aviv 69978, Israel

We examine the case of a two-person repeated game played by a boundedly rational player versus an unboundedly rational opponent. The former is restricted to strategies which are implementable by connected finite automata. It is shown that the "rational" player has a dominant strategy, and that in some cases the "weaker" (boundedly rational) player may exploit this fact to "blackmail" him. It is also shown that for a repeated zero-sum game, the rational player has a strategy which drives the automaton player's limit payoff down to his security (maxmin) level, even if he may choose any finite automaton. © 1989 Academic Press, Inc.

1. INTRODUCTION

The concept of bounded rationality (or "limited rationality"), originally introduced by Simon (1972, 1978), and also discussed in Radner (1986), was recently formalized for the case of repeated games using the model of a finite automaton. (See Aumann (1981), Rubinstein (1986), Neyman (1985), Kalai and Stanford (1985, 1988), and others.)

* We thank Ehud Kalai, Ehud Lehrer, and an anonymous referee for comments and references.

† Partial support from the National Science Foundation (Grant IRI-8814672) is gratefully acknowledged.

‡ Partial support from the National Science Foundation (Grant SES-8720342) is gratefully acknowledged.

0899-8256/89 $3.00

The very notion of boundedness invokes quantitative and comparative questions. For instance, if one has a certain measure of rationality, it makes sense to ask which of the players in the game is "more rational." We also have a particular interest in situations of asymmetric rationality, which were also discussed in Neyman (1985) and Ben-Porath (1985). In both papers, each player is restricted to choosing finite automata not exceeding a certain size where their bounds differ from each other (Ben-Porath discussed only zero-sum games).

In this paper we examine a qualitative rather than quantitative difference in players' rationality, namely, a case in which one player is restricted to the choice of a finite automaton (of arbitrary size), while the other is not.

We further restrict the set of strategies of the boundedly rational player: we shall assume that the automata among which he has to choose are only those which are "connected" in the following sense: for any pair of states of the automaton (which are to be thought of as "states of mind") there exists a sequence of moves of the opponent, such that, beginning at the first state, the automaton will end up at the second. The intuitive meaning of connectedness is the absence of dramatic, irrevocable moves; regardless of the past, there is always hope for any possible future. In particular, this assumption excludes "vengeful" strategies; however "angry" the automaton may be, it can always be appeased.

We note that the set of connected automaton strategies is (strictly) larger than that of bounded-recall strategies studied by Aumann and Sorin (1989) and Lehrer (1988a,b). It turns out that, in general, the automaton player can exploit his weakness and use a "blackmail" strategy, thus gaining all the surplus of cooperation: since it is common knowledge that he is only boundedly rational while his opponent is "stronger" (in the sense of "unbounded rationality"), he can use a reliable threat, while it is advisable for the rational player—who is known to be "wise" enough to understand the threat and to be able to choose the best response to it—to comply with the scheme the automaton dictates. This weakness is *not*, and cannot be interpreted as, a matter of strategic choice. A player commonly known to be rational cannot choose to forego a subset of his strategies and simulate a weaker player.

At this point one should note that although instances of the "tyranny of the weak" phenomenon are well-known, our results characterize a general type of weakness which turns out to be beneficial in all repeated games. To illustrate this point, let us consider the "battle of the sexes" game given by the following matrix:

		Player II: L	Player II: R
Player I	T	(2,1)	(0,0)
Player I	B	(0,0)	(1,2)

Note that the game is symmetric, so that there is no way to explain the choice of either of the two pure strategy equilibria (T,L) and (B,R) (neither in the framework of one-shot games nor in that of repeated ones). However, if we introduce rationality asymmetry, and assume that (only) one of the players, say, Player I, is restricted to connected finite automata strategies, there is only one "reasonable" equilibrium: the automaton player may safely choose an automaton which always plays T. Player II is clever enough to understand player I's threat ("I'll play T even if you play R, so you'd better play L"), and to realize he has no choice (but playing L). Player II may, of course, threaten to choose R, no matter what his opponent does. But he has no way of making this threat reliable; it does not matter how loudly he promises to play R, nor how solemn are the oaths he takes: it will always be common knowledge that at the very last moment he can switch back to a more flexible strategy complying with Player I's stubborn behavior.

Indeed, "conventional wisdom" predicts the outcome (T,L) if Player I announced first that he would play T, provided he has some reputation for being "stubborn." But this argument, which resembles ours on the intuitive level, remains out of the model of the game. (See, for example, Luce and Raiffa (1957).) In our model we formalize the notion of "stubbornness" and we do not resort to preplay communication: the automaton player does not have to announce what his choice is going to be: he simply chooses it, and lets the rational player learn what it is.

It will turn out to be a dominant strategy for the rational player to find out what automaton he is playing against since this automaton is known to be forgiving (i.e., connected). Otherwise, the experimentation itself may be a fatal mistake, an offense never to be condoned by the automaton. Therefore, in the absence of connectedness, the results described above do not have to hold.

The remainder of this paper is organized as follows. In Section 2 we introduce the formal model and state the results. The proofs of these are to be found in Section 3. Finally, Section 4 contains some brief remarks regarding extensions to mixed strategies spaces.

2. Framework and Results

We consider two-person games with players denoted as Player a (for automaton) and Player r (for rational). A *stage game* consists of two finite and nonempty *action sets* S_a and S_r and two *payoff functions* u_a, u_r: $S \to \mathbb{R}$ where $S = S_a \times S_r$. The *infinitely repeated game* G^∞ consists of two strategy sets Σ_a^G and Σ_r^G and two payoff functions, π_a and π_r, which we now proceed to define.

A *history of player i of length k* for $i \in \{a, r\}$ and $k \geq 1$ is an element of the Cartesian product S_i^k. We define a unique history (of either player) of length zero denoted by Λ. A *strategy* for player i ($i \in \{a, r\}$) is a function σ

from the set of all histories of the other player into S_i. Σ_r^G will be the set of all strategies of Player r. However, Player a will be restricted to "automaton strategies." In order to define these, we first define an automaton:

An *automaton A of Player a* is a quadruple, $A = (Q, q_0, \delta, \lambda)$, where:

(1) Q is a finite and nonempty set of *states*;

(2) $q_0 \in Q$ is the *initial state*;

(3) $\delta: Q \times S_r \to Q$ is the *transition function*;

(4) $\lambda: Q \to S_a$ is the *behavior function*.

The transition function δ can be extended (in a natural way) to the set of all pairs (q, h_r)—where $q \in Q$ and h_r is a history of Player r—as follows: $\delta(q, \Lambda) = q$ for all $q \in Q$, and for $k \geq 1$ and $h_r = (s_r^1, s_r^2, \ldots, s_r^k)$ let $\delta(q, (s_r^1, \ldots, s_r^k)) = \delta(\delta(q, (s_r^1, \ldots, s_r^{k-1})), s_r^k)$. It follows from this definition that for each $j < k$, $\delta(q, (s_r^1, \ldots, s_r^k)) = \delta(\delta(q, (s_r^1, \ldots, s_r^j)), (s_r^{j+1}, \ldots, s_r^k))$.

An *automaton strategy induced by the automaton* $A = (Q, q_0, \delta, \lambda)$ is the strategy α_A (for Player a) defined by $\alpha_A(h_r) = \lambda(\delta(q_0, h_r))$ for each history h_r of Player r. (We also say that α_A is *computed* or *defined* by A.)

We now define connected automata. Given an automaton $A = (Q, q_0, \delta, \lambda)$, and given two states $q, \bar{q} \in Q$, we say that $\bar{q}$ is *accessible* from q (and write $q \to \bar{q}$) if there exists a history h_r such that $\delta(q, h_r) = \bar{q}$. (Note that accessibility is a reflexive and transitive binary relation on Q.) Two states, q and $\bar{q}$, are *mutually accessible* (denote by $q \leftrightarrow \bar{q}$) if both $q \to \bar{q}$ and $\bar{q} \to q$. (Note that mutual accessibility ($\leftrightarrow$) is an equivalence relation on Q.) The automaton A is said to be *connected* if all the states in Q are mutually accessible. (That is, Q is an equivalence class of $\leftrightarrow$.) In this case, the strategy α_A will also be said to be "connected."

The set of strategies of Player a, Σ_a^G is the set of connected strategies, i.e., those defined by connected automata.

We now define the payoff functions π_a and π_r. We begin with the following definition: a *path p* in G^∞ is a sequence $(s_k)_{k\geq 1}$ where $s_k \in S$ for each k. Any two strategies, σ_a and σ_r of Players a and r, respectively, define a path in G^∞, denoted by $P(\sigma_a, \sigma_r) = (P_k(\sigma_a, \sigma_r))_{k\geq 1}$, as follows: $P_1(\alpha_a, \sigma_r) = (\sigma_a(\Lambda), \sigma_r(\Lambda))$ and for $k > 1$ let $P_k(\sigma_a, \sigma_r) = (\sigma_a((s_r^j)_{j=1}^{k-1}), \sigma_r((s_a^j)_{j=1}^{k-1}))$ where $(s_a^j, s_r^j) = P_j(\sigma_a, \sigma_r)$ for $1 \leq j < k$. We now define π_a, π_r: $\Sigma_a^G \times \Sigma_r^G \to \mathbb{R}$ by

$$\pi_i(\alpha, \sigma) = \liminf_{k\to\infty} (1/k) \sum_{j=1}^{k} u_i(P_j(\alpha, \sigma))$$

for $i \in \{a, r\}$. This completes the definition of G^∞.

A strategy $\sigma \in \Sigma_r^G$ is *dominant* if $\pi_r(\alpha, \sigma) \geq \pi_r(\alpha, \sigma')$ for each $a \in \Sigma_a^G$ and $\sigma' \in \Sigma_r^G$.

We are now equipped to phrase our first result:

THEOREM A. *For each game G there is a dominant strategy of the rational player,* $\sigma_d \in \Sigma_r^G$, *such that for each* $\alpha \in \Sigma_a^G$, $\lim_{k\to\infty} (1/k) \sum_{j=1}^{k} u_i(P_j(\alpha, \sigma_d))$ *exists for* $i \in \{a, r\}$ *(hence equaling* $\pi_i(\alpha, \sigma_d)$*). Moreover, there exists an algorithm computing* σ_d *as a function of G.*

We next turn to study the equilibria resulting when the rational player uses one of his dominant strategies defined above.

A pair of strategies, $(\alpha, \sigma) \in \Sigma_a^G \times \Sigma_r^G$, is an *ε-equilibrium* for $\varepsilon \geq 0$ if for each $\alpha' \in \Sigma_a^G$ it is true that $\pi_a(\alpha, \sigma) \geq \pi_a(\alpha', \sigma) - \varepsilon$, and for each $\sigma' \in \Sigma_r^G$ we have $\pi_r(\alpha, \sigma) \geq \pi_r(\alpha, \sigma') - \varepsilon$. A 0-equilibrium will also be called an *equilibrium*.

Now denote by v_r the individually rational payoff Player r in pure strategies, that is:

$$v_r = \min_{s_a \in S_a} \max_{s_r \in S_r} u_r(s_a, s_r).$$

Let C be the convex hull of the set of cooperatively feasible payoffs $\{(u_a(s), u_r(s)) | s \in S\}$.

Let w_a denote the maximal payoff for Player a (in mixed strategies), which is feasible, subject to the restriction that his opponent is kept above his individually rational payoff, i.e.,

$$w_a = \max\{x | \exists\ y \text{ s.t. } (x, y) \in C \text{ and } y \geq v_r\}.$$

We now have:

THEOREM B. *Assume that for some* $s \in S$, $u_r(s) > v_r$. *Then for any* $\varepsilon > 0$ *and any dominant strategy of Player r,* σ_d, *there exists an ε-equilibrium* (α, σ_d) *in* G^∞ *and for each such ε-equilibrium*

$$w_a - \varepsilon \leq \pi_a(\alpha, \sigma_d) \leq w_a.$$

Moreover, if for some $s \in S$, $u_a(s) = w_a$ *and* $u_r(s) > v_r$, *then for each dominant strategy of Player r,* σ_d, *there exists an equilibrium* (α, σ_d) *in* G^∞, *and for each such equilibrium* $\pi_a(\alpha, \sigma_d) = w_a$.

We now turn to the case in which the automaton player is not restricted to connected automata. Let $\overline{\Sigma}_a^G$ be the set of all strategies induced by the finite automata, and denote by $\overline{G}^\infty$ the repeated game in which Player a's strategy set Σ_a^G is replaced by $\overline{\Sigma}_a^G$. As we noted in the Introduction, there are no dominant strategies for the rational player in this case. Yet Player r has a "strong" strategy which guarantees him his individually rational payoff for *any* strategy of his opponent. Moreover, that strategy is recursive and the algorithm that computes it does not depend on the stage game.

THEOREM C. *There exists an algorithm which computes, for each game G given as datum, a strategy* $\sigma_r \in \Sigma_r^G$ such that $\pi_r(\alpha, \sigma_r) \geq v_r$ *for each* $\alpha \in \overline{\Sigma}_a^G$. *Moreover, the sequences of average payoffs (for both players) generated by* α *and* σ_r *converge for any* α.

We shall use this result in the analysis of the cases in which there are no mutual benefits from cooperation, i.e., the case of zero-sum games. Note that for this class of games, the individually rational payoff of Player r (v_r) is exactly the security level of Player a in pure actions. In other words, this is the minimal loss Player a can assure himself by choosing a pure action. We therefore obtain the following result:

THEOREM D. *If G is a zero sum game, then the values of the games* G^∞ *and* $\overline{G}^\infty$ *exist, and they are both equal to the security level of Player a in pure actions* (v_r).

Clearly, the lack of cooperation opportunities is detrimental to the automaton player, since his weakness cannot be exploited in a purely competitive game.

Results similar to Theorem D were obtained by Ben-Porath (1985) for two-automaton players (in a zero-sum game), where each player has a bound on the number of states his automaton uses, and the bound of one of them is relatively large enough.

We note that our results do not extend to n-person games. Of course, if more than one player is unboundedly rational, there is no reason for us to expect that any of them will have a dominant strategy. However, even if we assume only one of the players to be a rational player (while all the others are restricted to connected automata), our results no longer hold. Indeed, $(n - 1)$ finite automata may be simulated by a single one, thus reducing the n-person game to a two-person game. But the simulating automaton need not be connected. In fact, one may easily find examples of connected automata which, playing together against another player, implement a "vengeful" strategy.

3. PROOFS OF THE THEOREMS

For brevity, we will only sketch the proofs. More complete versions are to be found in Gilboa and Samet (1987).

3.1. *Proof of Theorem* A. We begin with two simple observations which will be given without proof.

Given two automata, $A = (Q, q_0, \delta, \lambda)$ and $A' = (Q', q_0', \delta', \lambda')$ (for player a in G^∞), and a history h_r of Player r, A and A' are *equivalent with respect to* (*w.r.t.*) h_r if for any other history $\bar{h}_r$ we have $\lambda(\delta(q_0, h_r \circ \bar{h}_r)) = \lambda'(\delta'(q_0', h_r \circ \bar{h}_r))$ where $\circ$ denotes string concatenation.

LEMMA 1. *Let $A = (Q, q_0, \delta, \lambda)$ and $A' = (Q', q_0', \delta', \lambda')$ be two automata and let h_r be a history of Player r. If A and A' are not equivalent w.r.t. h_r, then there exists a history of Player r, $\bar{h}_r$, of length* $\text{Max}(|Q|, |Q'|)$ *or less, such that* $\lambda(\delta(q_0, h_r \circ \bar{h}_r)) \neq \lambda'(\delta'(q_0', h_r \circ \bar{h}_r))$.

The second lemma states that there exists an algorithm which, given an automaton (as input datum), computes a best strategy against it in the repeated game. To be both more specific and more precise, let us define for an automaton $A = (Q, q_0, \delta, \lambda)$ in a game G^∞ a *cycle of length* k to be the ordered pair $((s_r^1, \ldots, s_r^k), (q^1, \ldots, q^k))$ where $(s_r^1, \ldots, s_r^k)$ is a history of Player r and $(q^1, \ldots, q^k)$ is a sequence of disjoint states in Q, such that $\delta(q^j, s_r^j) = q^{j+1}$ for $1 \leq j \leq k$ and $\delta(q^k, s_r^k) = q^1$. Clearly, a length of a cycle cannot exceed $|Q|$. We may now state

LEMMA 2. *There exists an algorithm which, given a stage name G and an automaton $A = (Q, q_0, \delta, \lambda)$ as data, computes a cycle $((s_r^1, \ldots, s_r^k), (q^1, \ldots, q^k))$ such that* $\max_{\sigma \in \Sigma_r^G} \pi_r(\alpha_A, \sigma) = (1/k) \sum_{j=1}^k u_r(\lambda(q^j), s_r^j)$.

We now turn to describe the algorithm of Theorem A. Let there be given a stage name G, and let $\{A_n\}_{n\geq 1}$ be an effective enumeration of Σ_a^G.

Our algorithm operates as follows. There are infinitely many phases, 1, 2, ... At phase n the algorithm has a putative opponent $\bar{A}_n (\bar{A}_1 = A_1)$; it plays against it optimally (according to Lemma 2) for k_n stages of the repeated game (where k_n will be specified in the sequel) and considers the next automaton, A_{n+1}. It compares $\bar{A}_n$ and A_{n+1}. If A_{n+1} is inconsistent with the history played so far, denoted h_n, or is equivalent w.r.t. h_n to $\bar{A}_n$, it sets $\bar{A}_{n+1} = \bar{A}_n$. Otherwise it plays a sequence of moves that would distinguish between the two (according to Lemma 1). If $\bar{A}_n$ is inconsistent with Player a's moves it sets $\bar{A}_{n+1} = A_{n+1}$, otherwise $\bar{A}_{n+1} = \bar{A}_n$.

It is only left to set k_n to be $n \cdot H_n$, where H_n is the number of stages played until now (i.e., the length of h_n), plus the number of stages that will be required to distinguish between the current guess, $\bar{A}_n$, and the next candidate, A_{n+1}.

It is obvious that this algorithm obtains the maximal payoff (and that the sequence of payoffs converges) for every actual opponent. ∎

3.2. *Proof of Theorem* B. Given $\varepsilon > 0$ construct an ε-equilibrium as follows: choose a cycle of moves with an average payoff (x, y) where $x \geq w_a - \varepsilon$ and $y > v_r$. Consider an automaton that plays this cycle if Player r does, but otherwise retaliates by long (but finite) punishment moves, yielding the payoff v_r to Player r. Thus, an automaton may be connected even though it threatens to punish Player r for long enough periods for

Player r to comply with the dictated cycle. Since Player a may choose strategies that yield him a payoff arbitrarily close to w_a (for any dominant strategy of his opponent), it follows that any ε-equilibrium (α, σ_d) must satisfy $w_a - \varepsilon \leq \pi_a(\alpha, \sigma_d) \leq w_a$.

As for the "moreover" part, assume that there is a $s \in S$ such that $u_a(s) = w_a$ and $u_r(s) > v_r$. Choose a cycle consisting of s alone. The automaton defined by this one-element sequence will be in equilibrium with any dominant strategy σ_d, and will yield Player a the average payoff, w_a. ■

The proof of Theorem C is very similar to that of Theorem A, and Theorem D follows immediately.

4. Possible Extensions to Mixed Strategies

In the previous analysis we assumed that the boundedly rational player is restricted to choose a pure strategy implementable by a connected finite automaton. One may wonder to what extent do the results depend on the pure strategy assumption.

First, we note that the results hold even if the automaton player is allowed to randomize over connected automata: a dominant strategy σ_d of the rational player will still be dominant, since whatever is the automaton chosen by the opponent's mixed strategy, σ_d will end up playing optimally against it.

The analysis becomes more sophisticated if we allow the automata to be random themselves. This can be done by introducing random behavior functions, random transition functions, or both. We will assume that random behavior and/or transition functions are defined by rational probabilities, so that the set of automata remains countable.

If the transition function is deterministic, the results may be easily adapted: when a dominant strategy compares two possible automata it may not necessarily be able to exclude one of them, but it can do so with arbitrarily small probability of being wrong. Thus ε-dominant strategies surely exist. In fact, by reconsidering "excluded" automata with a decreasing sequence of "arbitrarily small" probabilities one can also achieve optimality.

The case of a random transition function is more complicated. Without loss of generality one may assume that the behavior function is deterministic, since a random-behavior automaton may be simulated by a deterministic-behavior one (with more states). However, the comparison of such automata is more complicated because the comparing algorithm does not know at which state each automaton is (assuming it is the actual opponent); rather, it has a distribution over the automaton's states. Also, the computation of an optimal strategy (against a specific opponent) is no longer straightforward. We conjecture that our results have natural equivalents in this case as well, but this analysis is beyond the scope of this paper.

REFERENCES

AUMANN, R. J. (1981). "Survey of Repeated Games," in *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern*, pp. 11–42. Manheim/Weir/Zurich: Bibliographisches Institut.

AUMANN, R. J., AND SORIN, S. (1989). "Cooperation and Bounded Recall," *Games Econ. Behav.* **1,** 5–39.

BEN-PORATH, E. (1985). "Repeated Games with Bounded Complexity," manuscript.

GILBOA, I., AND SAMET, D. (1987). "Bounded versus Unbounded Rationality: The Strength of Weakness," Foerder Institute of Economic Research working paper.

KALAI, E., AND STANFORD, W. (1985). "Equally Sophisticated Players: On the Complexity, Memory and Automation of Repeated Game Strategies," manuscript.

KALAI, E., AND STANFORD, W. (1988). "Finite Rationality and Interpersonal Complexity in Repeated Games," *Econometrica* **56,** 37–410.

LEHRER, E. (1988a). "Repeated Games with Stationary Bounded Recall Strategies," *J. Econ. Theory* **46**(1), 130–144.

LEHRER, E. (1988b). "*n* Players with Bounded Recall in Infinitely Repeated Games," manuscript.

LUCE, R. D., AND RAIFFA, H. (1957). *Games and Decisions*. New York: Wiley.

NEYMAN, A. (1985). "Bounded Rationality Justifies Cooperation in the Finitely Repeated Prisoner's Dilemma," *Econ. Lett.* **19,** 227–229.

RADNER, R. (1986). "Can Bounded Rationality Resolve the Prisoner's Dilemma?" in *Essays in Honor of Gerard Debreu* (A. Mas-Colell and W. Hildenbrand, Eds.). Amsterdam: North-Holland.

RUBINSTEIN, A. (1986). "Finite Automata Play the Prisoner's Dilemma," *J. Econ. Theory* **39**(1), 83–96.

SIMON, H. A. (1972). "Theories of Bounded Rationality," in *Decision and Organization* (C. B. McGuire and R. Radner, Eds.). Amsterdam: North-Holland.

SIMON, H. A. (1978). "On How to Decide What to Do," *Bell J. Econ.* **9,** 494–507.