RJ 5270 (54310) 8/27/86
Mathematics

Research Report

REMARKS ON BOUNDED RATIONALITY

Nimrod Megiddo

IBM Research, Almaden Research Center
650 Harry Road, San Jose, CA 95120-6099
and Tel Aviv University, Tel Aviv, Israel

Research Division
Yorktown Heights, New York • San Jose, California • Zurich, Switzerland

RJ 5270 (54310) 8/27/86
Mathematics

REMARKS ON BOUNDED RATIONALITY

Nimrod Megiddo

IBM Research, Almaden Research Center
650 Harry Road, San Jose, CA 95120-6099
and Tel Aviv University, Tel Aviv, Israel

Abstract: The concept of bounded rationality is not well-defined. Several aspects of bounded rationality are discussed. Two different kinds of bounded rationality are distinguished. First, rationality may be bounded in the sense that the player cannot perform all the necessary calculations within the time frame of the game. This applies not only to numerical calculations but also to any kind of information processing, for example, figuring out logical consequences, or self-examination of preferences. Recent studies on the effect of bounded rationality concentrate on this type, studying the effect of limited computational capability on the set of Nash-equilibria of the game. Some objections to the approach taken in these papers in dealing with the question of bounded rationality are presented. In the second type, players may simply not accept that Nash-equilibrium is a necessary consequence of rationality. Such players challenge the Nash-equilibrium simply because it rules out some desirable outcomes. An example is presented followed by a proposed notion of approximate equilibrium (stronger than Radner's) which offers some help in this direction.

REMARKS ON BOUNDED RATIONALITY*

Nimrod Megiddo

IBM Research! Almaden Research Center, San Jose, CA 95120-6099,
Mathematical Sciences Research Institute, Berkeley, CA 94720
and Tel Aviv University, Tel Aviv, Israel

Introduction

The term "bounded rarionality" [Si] has been used informally by economists in various contexts. However. since the notion of full rationality is usually not well-defined. the concept of bounded rationality is not well-defined either. It is more or less clear what we mean by full rationality if only a single decision maker is involved: a fully rational player is one that maximizes his utility. To do that, he must have some beliefs about the world. He is so smart that his beliefs are absolutely complete and consistent and he is aware of them ail the time. Thus, a fully rational player (playing in solitude, or "against nature") has no computational limits. In particular. he can even decide undecidable problems. The case of a single decision maker is simple in this respect just because there are no indeterminacies. That is. for each possible action a, the decision maker has a probability distribution over the states of the world that may result from the action a. In the presence of other deciders, the future state may depend on actions of others. Rational players in the traditional theory of games reason about each other's actions and do not just form subjective probability distributions with respect to all the unknowns in the system. It has usually been assumed that rational players play some Nash-equilibrium combination of strategies, that is, a combination of strategies such that no player can benefit by (unilaterally) deviating from the strategy assigned to him in the combination. It is not clear how rational players reach a Nash-equilibrium, especially in games with no communication, or even in games with communication when they fail to agree on a specific equilibrium. It should be mentioned though that Aumann [A2] has recently shown that if each player is "Bayesian rational" (that is, maximizes his utility against what he ***believes*** to be the resultant of the actions of the other players) then the distribution of outcomes of the game can be achieved as a correlated ***equilibrium*** [A1]. This result however is based on the assumption that all the players have the same prior probability distribution over the states of the world. We discuss this assumption later. The set of correlated equilibria contains the convex hull of the set of Nash-equilibria, and this containment may be strict.

* This paper was presented at the Decentralization Workshop on May 16, 1986 in Berkeley, California.

We discuss here several aspects of bounded rationality. We distinguish two different kinds of this property. First, rationality may be bounded in the sense that the player cannot perform all the necessary calculations within the time frame of the game. This applies not only to numerical calculations but also to any kind of information processing, for example, figuring out logical consequences, or self-examination of preferences. Second, we may consider players who are not limited in their computational ability but who simply do not accept that Nash-equilibrium is a necessary consequence of rationality. These players challenge the Nash-equilibrium simply because it rules out some desirable outcomes. The classical example is of course the finitely repeated prisoners' dilemma game with a large number of rounds, where rational players (that is, ones that always play a Nash-equilibrium) perform very poorly relative to "irrational" ones. This phenomenon occurs in many other games. It seems that both kinds of "bounded rationality" can "resolve" the prisoners' dilemma . However. as we argue later, difficulties similar to the ones present in the finitely repeated prisoners' dilemma may not always be attributed to computational complexity alone.

Let us informally use the term "reasonable player" for a player who does not necessarily care so much about an *exact* equilibrium, but rather looks for practical ways to increase his utility, even at the expense of disequilibrium. Of course, we are immediately faced with a paradox: if "reasonable" but irrational players end up performing better than rational players, isn't there something wrong with our understanding of rationality? We will later suggest a notion of ϵ-equilibrium (stronger than Radner's [Ra1, Ra2]) which we believe offers some help in this direction.

Recent studies on the effect of bounded rationality concentrate on the first type, that is, limited computational ability. This direction, beginning with the papers of Rubinstein [Ru] and Neyman [N] (and further work [MW], [KS], [Z]) studies the effect of limited computational capability on the set of Nash-equilibria of a game. We present below some objections to the approach taken in these papers in dealing with the question of bounded rationality.

Rational play through bounded machines

In this section we consider a model in which strategies are identified with computing machines or computer programs. Our discussion relates directly to the models considered in [N, MW, Z]. The model is roughly as follows. Two players have to choose computing machines (from a well-defined class) to play for them a repeated N-round game. In [N] the machines are automata a_1, a_2, with no more than l_i states in a_i $(i = 1, 2)$, which are chosen after the number N is given. In [MW] the players choose Turing machines, of bounded number of internal states and unbounded memory, for playing any N-round repetition of a given game, where N is given as input to the machines at the start of the game. The latter can be interpreted also as a model in which the players choose computer programs (under some well-defined encoding scheme) of restricted length. A pair of programs (P_1, P_2) of permissible length is said to be in equilibrium if for every program P (of permissible length) the play of (P_1, P_2) is at least as good, from the point of view of player 1, as the play of (P, P_2), and at least as good, from the point of view of player 2, as the play of (P_1, P).

The model can be interpreted in at least two different ways:
(i) **Rational play.** Under this interpretation. fully rational players **voluntarily** restrict themselves to play the game through computer programs. However, each can optimize the particular choice of hi3 program subject to the constraints that were agreed upon. This scenario is somewhat like a limited binding agreement. The original game does not allow for such agreements but the players would do better (in the sense of equilibrium payoffs) if they could commit themselves to play through programs of restricted size. Each player must be able to verify that his opponent does not cheat.
(ii) **Bounded rationality.** Under this interpretation. players' rationality is bounded in the sense that they cannot consider strategies other than those that can be played by programs not longer than l bits (under some fixed encoding scheme) and the number l is common knowledge. For brevity we refer to these as I-programs or I-strategies. Here, the limited capabilities are given *exogenously*.

Under both interpretations the intricacies of the game are migrated into a "preprocessing game" where the players have to consider the space of all possible machines. The sense of the equilibrium is that for none of the players there exists (in the space of possible machines) a better response to the machine chosen by the opponent. Thus, if the players find (together) an equilibrium pair of machines, they can use it. This makes sense under the first interpretation; there it is implicitly assumed that the players communicate in order to reach an agreement, and are supposed to be more "rational". With regard to the second interpretation one may ask:

> If the players cannot consider strategies that are not playable by 1-programs, are they rational enough to ***find*** an equilibrium pair of such strategies or even just ***verify*** that a given pair is in equilibrium?

This question is sharpened considerably if the players are supposed to find an equilibrium pair of mixed strategies. that is, an equilibrium pair of **probability *distributions*** over the space of s-state automata or even Turing machines. It seems that (under the bounded rationality interpretation) the definition of equilibrium is justified only in case the players can consider only l-strategies, and at the same time can reason about equilibrium with respect to the set of all l-strategies. It is hard to accept such a dividing line between what a player can and cannot do.

In the context of Turing machines there is a fundamental problem with time. The game is not well-defined if the players may submit machines that, under certain circumstances, do not halt. Even if the players are required to submit halting machines, there is still the problem of how to verify that these machines halt. Thus, one is led to consider a model where players have to submit machines that halt within a certain given amount of time T whenever they are called upon to play. By imposing a strict time limit we of course cut the space of strategies in a rather arbitrary way. Moreover, it is precisely this strict constraint that enables cooperation in equilibrium. The underlying idea is that, no matter what the scarce resource is, players design machines that waste the entire amount of available resource so that they cannot perform even simple tasks like counting the number of stages. A possible way out of this problem is probably by introducing costs of computing time. This will of course eliminate non-halting machines but will also distort the whole picture since efficient computation then becomes a goal in itself.

Can machines be considered as players?

We find it difficult to interpret the programs of the preceding section as *players* since, once chosen, they play a *single* strategy. We prefer to regard a program as a *player* relative to a certain game if it *can* play different strategies of the game but *selects* one of them after some processing. Of course, any deterministic program can in conclusion be identified with a single strategy relative to any single game it can play. Moreover, such a program is playing a single strategy in the "universal" game, that is, the game where a player is given an instance of a game (from a certain class, using a certain encoding scheme) and has to select a strategy for it. In other words, the process of selecting a strategy for an input game is itself a single strategy in the universal game. An interesting question is whether or not a program can judge the single strategy that determines its own operation. Of course, the program does not have "free will", so it tests itself only if it is "programmed" to do so. We might say that a program (for playing games) is "rational" if it always searches for profitable deviations from a strategy it has designed for itself (for an input game) before it actually plays it. However, the search procedure is subject to the constraints of the game and different search procedures show different degrees of sophistication. Interestingly, sometimes it is preferred that certain profitable deviations will not be discovered (for example, in the context of the prisoners' dilemma). We believe that the concept of bounded rationality should be developed with respect to *classes* of games rather than singles ones. The prisoners' dilemma is not resolved by proving the existence of (or even exhibiting) a pair a programs that are in equilibrium relative to a class of strategies of bounded complexity. Such programs may not be able to play other games, and if they do they are likely to play it without much sense. A more desirable resolution may be in the form of programs that play games from a large class; they have to do "the best they can" in any game. For example, it may be reasonable to require that they play optimal strategies in small two-person zero-sum games.

In light of the foregoing discussion, let us consider a player as a "program" that receives a description of a game from a certain class as input, does some preprocessing, and then plays the game.

Let us consider, as an example, the class of all two-person finitely repeated (2×2)-games. Consider a "universal* game as follows. Nature chooses a (2×2)-game G and a number N. The players then repeatedly play G for N rounds, where both G and N are common knowledge. In traditional game theory the process by which nature chooses G and N is irrelevant. The players have unlimited computational resources for analyzing and playing any game so they can postpone the analysis of the game until after they have actually been informed of G and N, and then play their "optimal" strategies. Of course, if the game is not zero-sum, optimality of strategies has yet to be defined. The situation is different when the players have to write *finite* programs for playing any pair $(G,$ A'). Obviously, they cannot submit an infinite list of strategies for the instances (G,A'). That in itself does not imply that they cannot play optimally. For example, it is easy to write a short computer program that plays optimally any pair (G,N) where G is any (2×2)-*zero-sum* game. Here, the process by which nature selects G and N is irrelevant if the programs may be sufficiently long. In general, when the computational resources are restricted, the choice of a program must depend on what is believed to be the probability distribution of the pairs (G,N). Intuitively, if the restrictions imply that the program cannot play the "best" in all games, then the choice of instances in which it will play suboptimally must depend on the probability distribution over games.

In the traditional theory of games, a (complete) strategy for a player in a game in extensive form is a mapping that assigns a valid action to each information set of the player, that is, it is a complete plan of what exactly to do in each decision situation that may arise in the game. It is not necessary (and usually even impossible) for a player to design a complete strategy for the game in advance. Obviously, there is no need to design a complete strategy since during a single play only responses to the opponent's actions are required. A computer program that can play a game G (among some other games) does not have to contain (even implicitly) a complete strategy for G, since the actions it would "like" to take in hypothetical situations may change as a result of information or experience gained during actual play of G or other games. This motivates the following definition.

Definition. Consider a game G in extensive form. A *partial* strategy for player i is a mapping from a *subset* of the set of information sets of player i into respectively legal actions. In simple words, a partial strategy is a plan of what to do in *some* cases.

We note that the partial strategy of a player may vary during the play of a game and, in particular, be affected by other players' actions.

Consider, for simplicity, a 2-person game in extensive form. Some pairs of partial strategies are **decisive** in the sense that they are sufficient for completing a full play of the game. Obviously, the definition of an equilibrium can be extended to sets of partial strategies.

Definition. Let A_i denote a set of partial strategies for player i $(i = 1, 2)$. A decisive pair of partial strategies (σ^1, σ^2) $(\sigma^i \in A_i)$ is in equilibrium relative to (A_1, A_2) if there exists no $\tau^1 \in A_1$ such that (τ^1, σ^2) is decisive and preferred by player 1 over (σ^1, σ^2), and there exists no $\tau^2 \in A_2$ satisfying similar conditions with respect to player 2.

The idea of this concept is that a "de facto" state of equilibrium may prevail just because players are *not aware* of better partial strategies against each other. Of course. the play is not well-defined when one of the programs does cannot take a legal action.

Irrational subjective probabilities

Bayesian decision theory assumes agents have prior probability distributions over the states of the world. This fundamental assumption typically invokes a debate about how agents pick their priors and whether priors can be different for different agents. Subjective probability is, on the one hand, an expression of the agent's state of information about the world. Thus, if a coin has been tossed and the agent has not received any clue about the outcome, then his prior ought to be **50%** for heads and 50% for tails. On the other hand. subjective probability is also an expression of the degree of confidence in a certain outcome, to the extent required for taking some action. In other words, subjective probability can be viewed as an expression of preferences. For example, an agent has a subjective probability of at least p that it will rain the following day if he agent prefers betting on this event to betting on another event which has an objective probability of p.

The interpretation of subjective probability as an expression of preferences is often used to convince people that they do have subject probabilities with respect to almost everything. It is easier to convince people that they have preferences (including indifferences). Using a "binary search" type of interrogation, one can measure subjective probabilities of subjects with respect to certain events. However, these probabilities may not be additive just like preferences may not be transitive. In Bayesian theory, agents are assumed to have complete subjective distributions over the states of a certain world. However, with both interpretations of subjective probability it is hard to justify this assumption when dealing with real people. An agent may be well aware that his beliefs about the outcome of a coin toss are **50-50,** but it certainly takes some thought (and time) to realize that a consequence of this belief is that the probability that in 10,000 independent coin tosses less than **5150** will come out heads is more than 98%. Thus, if an agent is called upon to make a decision related to the latter event, and he does not have enough time to analyze the consequences of his basic lack of information, then he may act according to beliefs which do not follow from his state of information.

Consider the following example. A subject is offered to choose within a minute between winning an expensive prize with probability p or winning it provided the number of heads in a random sample of 10,000 independent coin tosses is less than 5150. For a "reasonable" subject we may assume the existence of a certain threshold p^o such that he would prefer the lottery if $p > p^o$ and the coin experiment if $p < p^o$. Thus, a reasonable subject may have an implicit subjective probability p^o for the event of less than 5150 heads. but this prohability may depend on the amount of time he has for thinking about the offer. However, the subject does not necessarily know what exactly his subjective probability is. For certain values of p the decision may be easy. For such values the subject can make the "right" decision without dealing with the question of what his subjective probability is. On the other hand, in extreme cases. the subject may realize what his subjective probability is by coping with the decision problem. However. the implicit subjective probability depends on the amount of time and the characteristics of the individual. If the agent is a computing machine then this implicit subjective probability is actually determined by the program.

In the latter example the subjective probabilities depended on time due to the need to compute. However, subjective probabilities may depend on time simply because the subject needs to interrogate himself and maybe analyze the event under consideration more carefully. It is also interesting to observe that the implicit subjective probability may depend not only on the event and the time but also on the prize. Both the calculations and general considerations may be different for different values of prizes.

What do machines know?

In a discussion of behavior of players in a game or processors in a distributed system, we often use phrases like "player i knows that...". or "player i can find out what is...", etc. We tend to attribute to processors certain "human" characteristics. Numerous theories of knowledge have been proposed (see [HI).A fundamental difficulty is the problem of closure under logical deductions. Agents are traditionally assumed to know all the logical consequences of the things they know, regardless of how hard it is make the deductions. However, recently a new theory has been proposed in [GMR], where agents are assumed to know only results of polynomial-time computations.

Knowledge of individuals can be discussed at different levels of detail. However, if we model players as machines, there is an implicit notion of knowledge which is, on the one hand, easy to define but, on the other hand, hard to work with.

In our opinion knowledge should always be related to time. In any test of an agent's knowledge, we must specify the amount of time given to the agent to come up with an answer to a query or with any other form of evidence that he knows the answer. Knowledge which cannot become evident is as interesting as as book that can never be read. Consider, for example, knowledge in the contest of arithmetic calculations. Almost everybody knows that $1+1=2$ but what do we mean when we say that i knows that $76982456+73967858=150950314$? It certainly takes some time to find the answer. Thus, it would be more accurate to say something like "player i would be able to tell within 30 seconds that $76982456+73967858=150950314$". In view of this example, it is clear that one should, at least implicitly, associate response times with claims about knowledge. For instance, "i knows how to set the Rubic cube in less than three minutes" or "it will take me no more than ten seconds to recall a certain phone number". On the other hand. one may never be able to come up with the right answer if the latter requires knowledge of some concealed information.

We might say that agent i knows$[t]$ whether proposition ϕ is true or false if, assuming nothing in the environment will change, the agent will tell whether ϕ is true or false within t time units from the time the query is presented. Of course, there is still the issue of the language in which the question is presented. At the lowest level, computing machines interact with the outside world by sending and receiving bits. It also seems that living creatures work with bits. It is easy to define knowledge in terms of strings of, say, 0's and 1's, but the definition is not particularly enlightening. One can think, for example, of a Turing machine where the work of the machine when given a certain query is determined by the description of its control, the current position, and the query which is fed into it.

Any query is presented at the lowest level as a string of bits. This is true not only for computing machines but probably also for living creatures. The query invokes some processing by the machine which then may or may not respond to the query. Let's assume the existence of a well-defined correspondence from query strings to admissible strings of answers. We might say that what the machine currently knows$[t]$ is the set of all the answers to queries that it would answer correctly (if the query is presented) within t time units from the time the query is presented. It is interesting to observe that by presenting a query we might change the knowledge state of the machine altogether. This may happen either because the query itself may reveal information or because the machine may reorganize itself during the processing of the query. Response time to future queries may depend very heavily on such reorganization.

Our interest here is limited to finite machines, that is, a finite number of bits always suffices for describing the machine and its instantaneous situation. We believe living creatures also fit into this framework. Furthermore, we are interested in knowledge about facts with finite encoding. The model is of course too detailed. In general, unless we make some assumptions about our computing machines, what the machine knows or does not know seems very arbitrary. Thus, we will not be able to proceed without further assumptions on the behavior of our machines. In principle, every machine works according to a program, reacting to input received from the environment. Input from the environment may also include programs. This is one way for the main program to learn to make better decisions. To develop a theory of decision making and games along these lines, we must have some idea on how our brain works. For example, when a mathematician attempts to prove a certain theorem, his brain is probably guided by some "program". Moreover, different people may have completely different programs for doing that. How can we develop an axiomatic theory and reason about these different processes without some basic understanding of how the brain works?

The common prior assumption

In the literature on games with incomplete information, starting with Harsanyi [Ha], it is usually assumed that the prior probability distributions of players, with respect to the state of the world. are identical. The assumption is justified by the argument that if players have different priors then these are not priors but rather posteriors, resulting from different information that has already been given to the players in some form. The discussion in [A2] leads to the following framework for a game. The players may be considered as identical "empty shells" before the game starts. Each shell has the same subjective prior probability distribution over the states of the world. This distribution is therefore "objective". The shells are then filled, by some random process, with the individual characteristics of the actual players. Before this process takes place there is no association between any shell and the contents to be poured into it.

It is interesting to examine this question in a model where the players are identified with computing machines. Here, the "empty shell" concept is quite natural. We can imagine a world where, in the beginning, each player is identified with a certain computer and all the computers are identical. Then, by some random process. the computers are loaded with some software and they start playing. Before any software is loaded. the computer is just an "empty shell". It cannot communicate with the outside world. It may be justified to say that, from the point of view of this empty shell, the probability of heads or tails in a coin toss is 50%. It is not clear that we may assume this empty shell has as subjective probabilities all the probability-theoretic consequences of this assumption, for example, for many coin tosses. Although it is not clear what the right model of priors is. it seems right to assume that the priors are identical. The empty shell is not called upon to make any decision so it does not really matter what its priors are. It is only later, after the software has been loaded, that the computer may need to consider the priors that prevailed earlier, in order to compute its posteriors. It takes a certain amount of computation to construct the prior from the basic assumption and the way this computation is done depends of course on the software. We arrive at the conclusion that, since different computers are loaded with different softwares, different players may eventually use *different* priors in their calculations of the posteriors.

We may view the process of loading the software as an analogue of the scenario in games with incomplete information, where in the beginning nature tells each player what its "type" is ([Ha, MZ]). Following [MZ], we may think of states of the world as $(n+1)$-tuples that specify the state of nature and the "types" of the n players. **A** "type" of a player is characterized by its beliefs about nature and the types of the other players. In other words, a state of the world can be "implemented" by fixing the state of nature and letting each player know its own type. The Bayesian approach says that each player's belief about the state of the world constitutes a posterior probability distribution (based on his prior distribution), given the information he has received. Suppose each player has a **prior** probability distribution with respect to the state of the world. **so** that once the player is informed of his own type. this prior is updated into a posterior distribution. It is easy to ***see*** that the prior of a player is in fact determined by his prior distribution with respect to his own type. since once his type is fixed. the rest of his beliefs are fixed too. Indeed, any distribution with respect to the type of player i can be used to define a prior for i that is consistent with the posteriors associated with the different types. As pointed out in [MZ], it is not always true that players' posteriors are derived from the same priors. Some consistency about the types revealed to the players has to be assumed. In our case, this means that the different pieces of software loaded into different machines have to be consistent with some common prior. The question of where the software is coming from is inevitable. Notice that this is not the the usual software the people develop but rather the very initial that is loaded into the empty shells.

'Bounded rationality" without complexity

Games in extensive form and. in particular, repeated games. are inherently complex for computation because of the large number of strategies. When a **2 x 2** two-person game is repeated N times, the number of strategies for each player in the N-round game is 2^{2^N-1}. This means that explicit enumeration of all strategies is not feasible in most practical situations. However, implicit enumeration may sometimes be feasible. This high computational complexity has motivated the study of play where not all the strategies are playable. Neyman [N] shows that if the play is performed through finite automata (with a suitable number of states) then the prisoners' dilemma is "resolved".

We believe the theoretical difficulties with games like the finitely repeated prisoners' dilemma are due only in part to the high computational complexity. Similar difficulties arise in games which are not at all computationally complex. Consider, for example, the following game which may be called "Share-or-Quit". This is a finitely staged game which is not a repeated one. It proceeds as follows. First, player

1 has to choose between (i) taking 1 cent for himself and giving 1 cent to player 2 ("Share"), and (ii) taking 2 cents for himself ("Quit"), in which case the game ends immediately. If player 1 chooses to share then the game continues. Player 2 is then asked to choose between (i) sharing 20 cents with player 1 and (ii) taking the 20 cents for himself, thereby ending the game. If he chooses to share then player 1 is asked to choose between (i) sharing 2 dollars with player 2 and (ii) ending the game by taking the 2 dollars for himself. The game continues in this way, with amounts increasing by a factor of 10 from round to round (that is. playing for 20, 200, 2000, etc.) until either a player chooses to quit or the amount has reached 2 million dollars.

It is easy to prove by backwards induction that. in any Nash-equilibrium of this game. player 1 must quit in the first round. This equilibrium is also obtained by iterated elimination of (weakly) dominated strategies. Moreover. the restrictions of these eliminated strategies to residual subganies are strictly dominated. We believe that if this game were put into an experiment, subjects would not play the Nash-equilibrium. It is interesting to note that the number of strategies in "Share-or-Quit" grows only **linearly** with the number of stages. Thus, with the particular numbers given above it is even easy to write the normal form of the game explicitly. In fact, a strategy in this game amounts to the choice of the stage in which the player plans to quit if given the chance to. It is thus trivial to implement here every strategy even on a very primitive machine. Note that the only way for a player in this game to convey information to the opponent is by continuing to play. So, in a certain sense, any strategy can be played without further computation.

The Bayesian analysis of "Share-or-Quit" is rather trivial. Each player starts with a prior probability distribution with respect to the stage in which the opponent will quit if given the chance to. If the player has already chosen a strategy (possibly a mixed one) then he also has such a prior with respect to the stage in which he himself will quit if given the chance to. These subjective probabilities are then updated in the obvious way as the game proceeds. A rather minimal assumption of rationality is that if a player gets to play the last stage then he "quits". Furthermore, this property is common knowledge. However, if this line of thought is pursued then one reaches a conclusion that, among "rational" players, it should be common knowledge that whenever in this game a player is called upon to play, the player must quit. Of course, if a player believes with a certain probability that his opponent will not quit then it pays for him to share.

On the equilibrium hypothesis

The tradition in game theory has been that in order for a theory (of how to play noncooperative games) to be acceptable by all the players, it is necessary that it should lead to a Nash-equilibrium. Let's call this the ***equilibrium hypothesis.*** The question of what is a good theory arises not only in economics but also in computer science. where principles of distributed computer systems have to be developed. Individual processors have to be endowed with some rationality. The equilibrium hypothesis is justified as follows. If a theory leads to a disequilibrium then at least one player has an incentive to deviate from what the theory prescribes. Moreover, being aware of this possibility, other players may not be willing to follow the theory any more. It is very important to notice that implicit in this hypothesis is an assumption that at least one of the players is ***capable*** of recognizing a disequilibrium. If a theory recommends a disequilibrium, but it is common belief among the players that it is an equilibrium then the theory may hold. Interestingly, the players do not have to believe that the recommended plan is an equilibrium. It is enough to have common belief that no player will be able to figure out a profitahle deviation. We will discuss this aspect later. Of course. it may he argued that the actual play is an equilibrium relative to the conditions of unawareness of a possibility to deviate profitably, hut we still do not have a theory to deal with what players choose to know (either by collecting information or by processing available information). Obviously. we cannot refute a claim that whatever players do is an equilibrium relative to an appropriate framework. In fact, one may learn about the framework by looking at the "equilibrium". For example, we may infer about the "cost" of calculations for a given individual by measuring the inefficiency of his buying habits. As another example, consider the situation of users of a computer system that rely on the security of passwords. Using passwords is not a Nash-equilibrium in the usual sense since, in principle, any user can find out any other user's password. However, we can learn about the relative importance of privacy to users of a certain system by looking at the security arrangements that seem to work there.

Complexity-based equilibrium

In this section we examine the difficulties in defining an equilibrium based on time-complexity. We adopt the finitely repeated prisoners' dilemma game as a generic example. Let's fix the following matrix as a stage of the game:

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} 2,2 & -1,3 \\ 3,-1 & 0,0 \end{pmatrix} \end{array} .$$

We start by examining the proof that every equilibrium pair in the finitely repeated prisoners' dilemma game results in the steady defection. that is, a play of (D, D). Consider any pair of strategies (Σ^1, Σ^2) for the players in an N-round prisoners' dilemma game, and suppose this pair does not lead to the steady play of (D, D). Assume round i is the last time player 2 is supposed to play C. If player 1 is supposed to play C in round i, or any round thereafter, then he can make a profit (not only locally but also in the N-round game) by switching to playing D from round i through the end. If player 1 is supposed to play

D from round i to the end then player 2 can make a profit by playing D in round i. However, in order for any of the players to make such a profit, he has to know *when* to deviate profitably, that is, he must recognize the ***critical* round** when it comes.

Some of the strategies for the N-round game have short descriptions, that is, a player may be able to write a short computer program that will play for him even when the number of rounds is huge. The actual play is of course determined when *both* players submit their computer programs. Thus, consider the following scheme for playing the N-round game. The players would like to agree on some sequence consisting of (C,C)'s an (D,D)'s. For example, they can rely on some pseudo-random number generator. Their idea is to play according to the chosen sequence as long as both have done so. and otherwise to switch to playing (D, D) through the end. Obviously. only the sequence of $(D.D)$ is in equilibrium. However, depending on the way the sequence is generated and described. there may be room for improvement by playing relatively complex sequences. Whatever sequence the players choose to play, they should be able to compute the "next" element of the sequence before the next round of the game takes place, or else the sequence' would not be operational at all. On the other hand, the players should not be able to recognize the last round in which they are supposed to play (C, C), before that round actually takes place. Assuming sequences with uniform cost of computing the next element, we see that such a scenario may be possible for N sufficiently large. The idea presented here is appropriate only for situations in which the rounds of the game take place at predetermined times. Suppose the rounds are played at times $1, 2, 3, \ldots$ and the amount of time it takes to compute the next element of the sequence is r $(r < 1)$. This means that each player must submit his action for round i by time i and he has enough time to compute what this action should be. It follows that the sequence is operational for any number of rounds. Assuming none of the players can speed up the computation. it takes rN time to compute the entire sequence. So, if the critical round occurs before time rN then it cannot be recognized by any player. The feasibility of this idea depends on the existence of a sequence $a = \{a_1, a_2, \ldots\}$ $(a_k \in \{C, D\})$ where it must take a known uniform amount of time $t(n) = nr$ to compute a_{k+n} from a_k. Then, if r is too small. we can select subsequences of the form $b_k = a_{mk}$ so the nest element of the subsequence is computed in $r' = mr$ time units. Furthermore, a similar idea can be used to increase the frequency of playing (C,C).

The preceding discussion leads to the following abstract definition of equilibrium based on time-complexity. However, the traditional model of games in extensive form has to be refined. First, we have to assume that each player has a certain computational power. We prefer to leave this concept in a rather abstract form since it depends not only on availability of hardware and software but **also** on skills, experience and many other factors. Second. we need to introduce **lime** into the model. Specifically, together with every decision position (information set) P_i for player i we associate a "time limit" $r(P_i)$. The interpretation is that player i has to choose an action at P_i within $r(P_i)$ time units from the time he is informed the play reached the position P_i. A strategy for player i is a procedure (feasible relative to i's computational power and the time limits) that selects an action for each decision position of i. In certain games it may be too restrictive to require that a strategy obeys the time limit in all possible decision positions. Alternately, we can talk about a mutually feasible n-tuple of strategies in an n-person game. In this case it is required only that the resulting play be feasible, that is, only the decision positions that actually occur during the play are taken into account. The following definition is only intuitive since it uses terms that have not been defined precisely.

Definition. An n-tuple of mutually feasible strategies is in equilibrium if no player can profitably deviate from it, that is. no player can change his strategy and increase his own payoff without destroying the property of mutual feasibility.

It seems that playing according to a suitable sequence (as esplained above) in the N-round prisoners' dilemma game constitutes an equilibrium in the sense of our definition. More specifically, it is not feasible for either player to compute the critical round. Of course, this is not an equilibrium in the usual sense. Although the sequence is determined by some computational procedure (known to the players), some decisions based on the sequence are not feasible within the available computational power. For example, there is not enough time to compute the critical round. Assume each player has a subjective probability distribution with respect to the critical round. This distribution may change during the play. Players' beliefs give rise to probabilistically critical rounds. We can say that round j is the critical from the point of view of player i if it is the last round to which i assigns a positive subject probability to be the critical round. If player i could modify his strategy to follows the sequence up to round j and then play D through the end, then he would make a positive expected profit with respect to his subjective distribution, provided the other player plays according to the sequence.

However, the complexity-based equilibrium is not well-defined and we can prove (even without subjective probabilities) that every equilibrium must result in the steady play of (D, D). It is obvious that in equilibrium the players must play D in the last round regardless of what the sequence prescribes. If they do not then a profitable deviation is readily available. Assume, by induction, that in equilibrium the players must play D in the last k rounds regardless of what the sequence prescribes. Obviously, it is impossible in equilibrium that they play C in round $k+1$ from the end since, again, by the induction hypothesis a profitable deviation is now readily available. The "paradox" follows from the fact that we did not define what it means that the player cannot find the critical round that is determined by the sequence. In fact, what we saw is that there cannot be a critical round if the sequence induces an equilibrium. In a sense the traditional definition of equilibrium is too restrictive. The *existence* of a profitable deviation suffices for destroying an equilibrium. It does not matter how hard it is to find such a deviation. The player may perform the deviation even without prior computation. It seems that the right way to approach these issues is to assign probabilities to the event that the player does some computation that leads to the deviation and to the event that he discovers the deviation by luck.

The cost of computation is a very important factor. Obviously, if the cost of analyzing the pair of programs (for example, by running them against each other) is higher than the anticipated profit from a profitable deviation expected to be discovered, then an equilibrium has been reached. Moreover, when the players actually play the prisoners' dilemma game, it may become apparent during some round i that from round i_0 and on one of the players will steadily play D. However, such a discovery cannot affect what has been accomplished before round i.

Consider. for example. the finitely repeated prisoners' dilemma game. where payoffs are denominated in cents as stated above. Thus, each player receives 2 cents if both players cooperate: each receives 0 cents if both defect, and if they play differently then the defector receives 3 cents while the cooperator has to pay 1 cent. Obviously each player can be secure (by defecting all the time) not to lose even 1. Suppose the number of rounds is very large, say. 10^8. Then, there is a potential for winning hundreds of thousands of dollars in this game. If both players follow a well-mixed sequence of C's and D's then each of them makes a profit of hundreds of thousands of dollars. However, the pair is not in equilibrium and the players may well know this fact. Each player can make a profit of one cent (assuming the other player does not do the same) if he computes the largest i ($i \leq 10^8$) such that $a_i = C$ and deviates accordingly. The cost of the computation may not worth it. This suggests an alternate definition of equilibrium.

Definition. An n-tuple of strategies in an n-person game is in equilibrium if for each player, the cost of computing a profitable deviation is larger than the anticipated profit.

It seems that the players are happier when they cannot compute the critical round in advance since this lets them make huge profits! Again, the classical definition of equilibrium does not deal with the problem of computing the profitable deviation whose very existence destroys an equilibrium.

Approximate equilibrium

It is well-known that Nash-equilibrium points may sometimes be very inefficient in the sense of Pareto. Moreover, all the equilibria may lie far from the Pareto-optimal set. It is interesting to note that sometimes there exist "approximate equilibria" lying close to the Pareto-optimal set whereas, all the exact equilibria lie far from that set.

We consider again the finitely repeated prisoners' dilemma game, where payoffs are denominated in cents as above. Suppose the players contemplate playing some sequence of (C,C)'s and (D,D)'s (possibly all (C,C)'s) with the provision that if one deviates then they immediately switch to playing all (D,D). Unless the sequence consists only of (D,D)'s, there is a critical round in which a player may deviate profitably. However, it is common knowledge that if one of the players plays the contemplated strategy and the other one deviates (in the critical round) then the "faithful" player is worse-off only by 3 cents. Thus, this is an "+equilibrium" in the sense of Radner [Ra1, Ra2], with which reasonable players may feel comfortable since (it the frequency of C is sufficiently large) it yields profits of hundreds of thousands of dollars (as opposed to zero in the unique exact equilibrium), and the only profitable deviation yields 1 cent *to* the deviator. Radner defines an *ϵ-equilibrium* as a combination of strategies such that no player can gain more than an ϵ by (unilaterally) deviating from the combination. It is interesting to note another property which is not stated as a condition in Radner's c-equilibrium, that is, the loss to the faithful player is also small (only 3 cents). Also, if both players perform the "profitable" deviation then each loses **2** cents.

Radner's ϵ-equilibrium may be convincing in the case of the prisoners' dilemma game. However, in order to justify play according to it in general (even in 2-person games), the players must rely on some complicated reasoning as follows. Each player has to be convinced that the opponent will not try to gain an additional ϵ. Also, each player has to be convinced that the opponent is convinced that he himself will not try to gain an additional ϵ, and to be convinced that the opponent is convinced that he is convinced that the opponent would not try..., and so on. Of course, the question remains why ***should*** they be convinced?

We propose here a stronger definition of an ϵ-equilibrium for 2-person games. There are various ways to extend the definition to n-person games but we prefer not to this in the present paper. Our definition is inspired by the difficulties with the finitely repeated prisoners' dilemma . We believe that, in general. is no reason for players to be convinced that others would not try to gain an additional ϵ. However, in the prisoners' dilemma game there is no need to be convinced at all. The fact is that the each player does not lose much if his opponent is greedy and tries to gain an additional ϵ. Thus, why not incorporate this property as an additional condition for an ϵ-equilibrium? Denote by $H_i(\sigma^1,\sigma^2)$ $(i=1,2)$ the payoff to player **i** when a pair of strategies (σ^1,σ^2) is played. We give below a revised definition of ϵ-equilibrium. We would have called it *strong* ϵ-equilibrium to distinguish it from Radner's ϵ-equilibrium but, unfortunately, the concept of strong equilibrium already exists elsewhere in game theory.

Definition. A pair (σ^1,σ^2) of strategies in a 2-person game constitutes an *ϵ-equilibrium* if the following conditions, as well as ones obtained by interchanging the roles of 1 and 2, hold:
(i) for any strategy τ^1 of player 1,

$$H_1(\tau^1,\sigma^2) \le H_1(\sigma^1,\sigma^2)+\epsilon$$

(ii) if τ^2 is a strategy of player **2** such that

$$H_2(\sigma^1, \tau^2) > H_2(\sigma^1, \sigma^2) ,$$

then

$$H_1(\sigma^1, \tau^2) \geq H_1(\sigma^1, \sigma^2) - \epsilon .$$

We could strengthen the definition with a third condition, that for any profitable deviation of one player, the other player best response of the other player cannot get the latter more than c profit. This would still be valid in repeated games. We could add more complicated conditions of this kind up to any order. For simplicity, let's not do that though.

Intuitively. our t-equilibrium prevails if unilateral profitable deviations yield only small profits to the deviator and cause only small losses to the other player. Obviously, any Nash-equilibrium is an c-equilibrium in our sense, for any $\epsilon \geq 0$. since it allows for no unilateral profitable deviations. Notice that we do require the profit to the deviator to be small; otherwise, one player can be almost sure that the other player will deviate and therefore he may reconsider his own plans. On the other hand, in an c-equilibrium each player is satisfied that his opponent does not have a great incentive to deviate, and even if he does the loss is small. The idea is that if the players have "agreed" to play a certain c-equilibrium, then they only have to be convinced that none will perform a *losing* deviation from it. The players are not "insured" against such deviations by each other. This is of course also the case with exact equilibrium. It is interesting to note that even with exact equilibria it is sometime difficult to argue that players will not perform losing deviations. Consider. for example, the following game:

$$\begin{array}{cc} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} T \\ B \end{array} & \begin{pmatrix} 1000{,}1000 & a, 100 \\ 100, a & 100{,}100 \end{pmatrix} \end{array} .$$

Consider the point $(1000, 1000)$. It is Pareto optimal and also dominates every other equilibrium in the sense of Pareto. It is very natural to expect that if the players can communicate then they would agree to play (T, L). However. if a is small then the risk to each player is large. After all, even if the players agreed to play (T, L), each may still have some small positive subjective probability that his opponent will not fulfill the agreement. If a is sufficiently small, a suspicious player may prefer not to fulfill the agreement. and perform a losing deviation. It is interesting that suspicions may be "indirect". A player may. for example, suspect that his opponent suspects that he is going to defect. or he may suspect that his opponent suspects that he suspects that the opponent is going to defect, an so on.

In view of the example, the ϵ-equilibrium concept cannot be criticized more than the exact equilibrium on the grounds that it assumes the players do not perform losing deviations. Assuming the opponent will not perform a losing deviation, the player knows he is playing within ϵ of his best response and will not lose more than ϵ if the opponent deviates. One is of course tempted to continue the analysis and ask why should a player not play his best response. One possibility is as follows. Since ϵ is small, the player does not know for sure what the opponent is going to do. The player may, at best, have beliefs with respect to the opponent's action. As suggested earlier, it takes time to form the beliefs, and it is likely that there is some "cost" (in the form discomfort or fatigue) associated with the thinking about what the opponent is going to do, especially when his deviations are not going to change much anyway. If thinking about the problem cannot lead to a change in payoffs by more than ϵ then the player may be reluctant even to start thinking about it. The 6-equilibrium idea may be criticized with the argument that theorists should perhaps develop better models where players do optimize subject to all the constraints they have. Thus. if it costs to compute then it the cost should he built into the model. The answer is that the ϵ-equilibrium is indeed an exact equilibrium in some a modified game which is much harder to describe precisely. Suppose both players have been advised to play a certain ϵ-equilibrium and the advice is common knowledge. If ϵ is sufficiently small then perhaps the best thing to do is follow the advice and not think about the problem any more, since the cost of thinking is not worth it. There is of course no reason not to go after the additional ϵ but it may not be worth it to think about what the opponent is going to do and what he thinks we are going to do, and so on. The problem is however more complicated if the player is a mathematician, a philosopher or a theoretical economist, who actually *enjoys* thinking about the problem.

The ϵ-equilibrium can also be looked at as a form of an illusion that players prefer to be in. There is a phenomenon in decision making, that people refuse to consider certain possibilities simply because they don't like to think about them. One example is the way people deal with insurance. People sometimes enjoy the belief that they are fully covered under the policy they are holding even though they are actually not. Although buying additional insurance would make sense, people sometimes prefer not to study the list of exclusions from coverage too carefully because they do not want to be persuaded to buy more insurance.

An obvious question with respect to the c-equilibrium concept is what is the actual value of ϵ. Obviously, any outcome is an ϵ-equilibrium for ϵ sufficiently large. However, we can make asymptotic statements without an exact value for E. For example. in a finitely repeated game with N rounds, for every ϵ there is an N for which "good" ϵ-equilibria exist. In this sense ϵ may be treated as an infinitesimal. Asymptotically, with large N and small ϵ the set of 6-equilibria approximates the set of individually rational payoffs. The proof of the "folk theorem" is essentially the same as the one given by Radner **[Ra2]**.

We should also note that ϵ has to be given the meaning of a *relative* loss or gain in utility. Equivalently, we may apply the concept only after the utility scales have been normalized.

Conclusion

We have attempted to show that bounded rationality is a nontrivial concept. To reason about the behavior of bounded rational agents, we have *to* know something about the "operating systems" that drive them. Computing time is an important factor but there are other constraints. For example, it is not clear that every agent, with sufficient amount of time, will be able to set the Rubic cube. The problem of enumeration in a combinatorial system is in itself very challenging. If a player has to solve a complex problem. he might just get too tired in the process even though there is ample time.

We do not know whether humans are no more than sophisticated computing machines. We have estimates on the number of brain cells, and we can of course estimate the number of elementary particles. but we do not know for sure that the thinking process can be fully described at the elementary particles level. In any case, even if humans are indeed sophisticated finite machines, they operate with rules that are still not very well understood. There are emotional and ethical aspects involved in the decision making process which, at the moment, we do not know how to quantify. For example, a person may not bother to pick up a penny from the floor but yet get upset if the gas gas station attendant neglects to give him back one penny when he pays 10for9.99 worth of gas.

Computer programs that people write usually do not modify themselves. It is conceivable that the brain programs do modify themselves. Moreover, living creatures reproduce themselves and are subject to evolutionary processes. Information is passed from generation to generation not only genetically but also simply by co-existence. This includes not only data but also skills, opinions and tastes. Thus, although individuals are mortal, there is continuity and evolution that sometimes justifies treating a sequence of generations as one player.

Experiments with human subjects often reveal "irrational" behavior. Obviously, experiments and questionnaires cannot represent the behavior in real-life situations. The main criticism is that subject do not have real incentives to make good decisions. This is true if the rewards are too small. It is argued that the subjects would behave differently if their decisions would have serious consequences. On the other hand. unlike machines, humans sometimes get too nervous when they have to make very serious decisions. and this of course affects their performance.

Theorists attempt to identify simple principles ("laws of nature") that explain, at least conceptually, how systems work. It is not at all clear that human decision making processes can be explained by a short list of principles. There exist computer programs that cannot be compressed. It may well be that the description of what an individual would do in different situations must be as long as a full description of **his** brain.

References

[A1] R. J. Aumann, "Subjectivity and correlation in randomized strategies", ***Journal of Mathematical Economics*** 1 (1973) 67-96.

[A2] R. J. Aumann, "Generalized equilibrium as an expression of Bayesian rationality", manuscript. June 1985.

[GMR] S. Goldwasser, S. Micali and C. Rackoff, "The knowledge complexity of interactive proof systems", Proceedings of the 17th Annual ACM Symposium on Theory of Computing (1985). ACM, New York, 1985, pp. 291-304.

[H] J. Y. Halpern, Ed., ***Theoretical aspects of reasoning about knowledge,*** Proceedings of the 1986 conference, Morgan Kaufmann Publishers, Inc., Los Altos, California, 1986.

[HU] J. E. Hopcroft and J. D. Ullman, ***Introduction to automata theory, languages, and computation,*** Addison-Wesley, Reading, Mass., 1979.

[KS] E. Kalai and W. Stanford, "Finite rationality and interpersonal complexity in repeated games", report, Kellogg Graduate School of Management, Northwestern University, 1985.

[KMRW] D. hi. Kreps, P. Milgrom, J. Roberts and R. Wilson. "Rational cooperation in the finitely repeated prisoners' dilemma". ***Journal of Economic Theory*** 27 (1982) pp. 245-252.

[MW] N. Megiddo and A. Wigderson, "On play by means of computing machines", in: ***Theoretical aspects*** of ***reasoning about knowledge,*** Proceedings of the 1986 conference, J. Y. Halpern, ed., Morgan Kaufmann Publishers, Inc., Los Altos, California, 1986, pp. 259-274.

[MZ] J.-F. hlertens and S. Zamir. "Formulation of Bayesian analysis for games with incomplete information", ***International J. Game Theory*** 14 (1985) 1-29.

"] A. Neyman, "Bounded complexity justifies cooperation in the finitely repeated prisoners' dilemma", manuscript, February 1985.

[R1] R. Radner, "Can bounded rationality resolve the prisoners' dilemma?", Bell Laboratories Economics Discussion Paper No. 143, August 1979.

[R2] R. Radner. "Collusive behavior in noncooperative epsilon-equilibria of oligopolies with long but finite lives, **Journal** of **Economic Theory, 22** (1980) 136-154.

[Rap] A. Rappaport, Ed., *Game* **theory** ***as a theory*** **of conflict rerolution.** Dordrecht-Holland/Boston: Reidel, 1974.

[Ru] A. Ruhinstein. "Finite automata play - The repeated prisoner's dilemma", Research Report No. 149, Department of Economics, The Hebrew University, Jerusalem, Israel, 1985.

[Si] H. A. Simon. "Theories of bounded rationality", in: C. B. McGuire and R. Radner, eds., **Decision** ***and organization***, North-Holland, Amsterdam, 1972.

[Z] E. Zemel, "On communication, bounded complexity and cooperation", manuscript, Kellogg Graduate School of Management, Northwestern University, November 1985.