35 Irrationality in Game Theory

The history of game theory has been evolutionary rather than revolutionary. Over the years, the fundamental concepts of the theory have been used in an ever-increasing circle of contexts; some of the most important developments consisted of innovative ways of recasting the theory so as to apply to problems that previously appeared not to fit in. A prime example is Harsanyi's historic work (1967–68) on games of incomplete information. Before this work, it had been universally thought that for game theory to be applicable, the payoff matrix of the game must be commonly known. Harsanyi's achievement consisted of formulating games of incomplete information so that they could be seen and analyzed as games of complete information; by this achievement, he opened the door to the development of the enormous fields of informational game theory and economics as we know them today. Similarly, the theory was extended from perfect information (Zermelo) to imperfect information (Borel and von Neumann); from strategic (normal) form to extensive; from two-person to n-person; from zero-sum to non-zero-sum; and from TU (transferable utility) to NTU (non-transferable utility).

The current frontier—the one discussed here—is that of rationality. Briefly, it is suggested that game theory need not assume that the players must be rational (i.e., utility maximizers).

This may sound paradoxical to the reader. After all, rationality is what game theory is all about; game theory without rationality sounds like geology without rocks or biology without life. Yet some of the most challenging problems facing the theory concern the interface between rationality and irrationality; they are about situations that cannot be dealt with on a purely rational basis, much like geological phenomena that depend on living organisms (such as cracking of rocks by plants) or biological phenomena that are the result of the nonliving environment.

1 Approach

The underlying game model from which we take off is that presented in Aumann 1987. It starts by considering the set of all "states of the world," where the specification of a state includes all relevant factors, including the (pure) strategy that each player uses in the given game, and what he knows

This chapter originally appeared in *Economic Analysis of Markets and Games: Essays in Honor of Frank Hahn*, edited by P. Dasgupta, D. Gale, O. Hart, and E. Maskin, pp. 214–227, The MIT Press, Cambridge and London, 1992.

when he decides to do so. It then asks, *suppose* that each player is rational at each state of the world that occurs with positive probability; i.e., that the strategy that that state specifies for each player happens to be one that maximizes his utility given his information. Can we then say anything about the distribution of strategy n-tuples? Does the assumption of rationality imply any specific form for this distribution?

The point of view of this model is not normative; it is not meant to advise the players what to do. The players do whatever they do; their strategies are taken as given. Neither is it meant as a description of what human beings actually do in interactive situations. The most appropriate term is perhaps "analytic"; it asks, what are the *implications* of rationality in interactive situations? Where does it lead? This question may be as important as, or even more important than, more direct "tests" of the relevance of the rationality hypothesis.

The answer given in Aumann 1987 was that it leads to correlated equilibrium—not that the players consciously choose a correlated equilibrium according to which they play, but that, to an outside observer with no information, the distribution of action n-tuples appears as if they had.

What is proposed here is to take this framework, remove the rationality hypothesis, and see where we are led.

Needless to say, I do not simply drop the rationality hypothesis and leave it at that. Rather, I propose to use the epistemic model of Aumann 1987—whose roots extend back to Radner (1968, 1972), to Harsanyi (1967–68), to von Neumann and Morgenstern, to formal epistemologists such as Kripke and Hintikka, and to probabilists such as Kolmogorov—as a very general kind of framework for studying limited rationality; a unified framework for considering rationality in environments that may include irrationality, indeed where irrationality may be of the essence.

2 Formal Description of the Model

Formally, define an *information system* to consist of the following:

(i) a strategic (normal) form n-person game G;

(ii) for each player i, a set whose members, s_i, are called *information states* of i; and

(iii) a function that associates with each information state of each player i

(a) a pure strategy of i in G and

(b) a probability distribution on $(n-1)$-tuples of information states of the other players.

An information system is a simultaneous representation of all the players' uncertainties. Each player makes some definite choice of a pure strategy, based on whatever he knows or believes about customs, history, personalities of the other players, and so on—in brief, based on the state of his information. He knows his own choice; that is what (iiia) says. But he does not know what the others choose, though he has some belief about this. Moreover, he does not know their information states—what information is available to them when they make their choices, what they believe about his choice, and so on.

For simplicity, we confine ourselves to information systems that are finite, i.e., in which each player has finitely many information states.

Let us use the term *surmise* for a probability distribution maintained by a player, *theory* for his surmise on the $(n-1)$-tuple of information states of the other players, and *belief* for his surmise on the $(n-1)$-tuple of pure strategies that the other players choose. Thus, (iiib) associates a theory with each information state of each player i, from which we may derive i's belief. From i's theory we may also derive a surmise on the theories of the others, and hence about *their* beliefs. This is called a *second-order* belief; it may be computed explicitly from the information system. Similarly, we may define and compute beliefs of any order. Note that the resulting "belief hierarchy" is derived from the information system, not given exogenously.

We call a player *rational* at a given information state if the pure strategy he chooses at that state maximizes his expected payoff when calculated according to his belief at that state. (See section 7 for a discussion of this definition.)

Let us use the term *state of the world*, or simply *world*, for an n-tuple of information states; it is a complete specification of all relevant parameters. A widely used regularity condition on information systems is that there exists a single probability distribution p on the states w of the world, called a "common prior," such that the theory of each player i at each w is the conditional of p given i's information state at w (see Harsanyi 1967–68, p. 493ff., and Aumann 1987, section 5). Intuitively, p represents the surmise about the state of the world held by an outside observer with no information at all; it is assumed that if the observer had the same information as a player he would make the same surmises. When it exists, a common prior allows a particularly concise description of the information system.

For an illustration, let G be the game with the following matrix:

	L	R
T	1,1	0,0
B	0,0	2,2

(as usual, the first and second coordinates of each entry represent payoffs to the row and column players—say, Alice and Bob—respectively). Consider the information system depicted by

	L	L	R
T	1,1*	1,1*	0,0
T	1,1*	1,1	0,0*
B	0,0	0,0*	2,2*

The first two rows represent two different information states of Alice, both corresponding to her pure strategy T; the last row represents a third information state, corresponding to her pure strategy B. Similarly, the columns represent Bob's information states. There are nine states of the world. Each star represents a probability of 1/6; the other probabilities vanish. Alice's first information state differs from her second in her belief; in the first, she knows that Bob plays his pure strategy L, whereas in the second, she attributes half-half probabilities to Bob's two pure strategies. Similarly for Bob.

Thus, in the world w represented by the top left corner of the 3×3 matrix, Alice plays T, and Bob knows this. But she does not know that he knows it; indeed, she ascribes probability only 1/2 to this event. Similarly for Bob. Finally, each player is rational in w, but neither knows that the other is; indeed, each ascribes probability only 1/2 to the rationality of the other.

(Throughout this paper, the term *know* is used to mean "believe with probability 1." In the current framework one cannot even formulate the stronger meaning of *know*, which allows no exceptions at all, even with probability 0.)

The above definition of an information system is analogous to Harsanyi's (1967–68) definition of a game of incomplete information. There each player may be one of several "types," where a type determines his utility function and a distribution over $(n - 1)$-tuples of types of the others. Here we replace the utility functions by actions; the information states here play the same role as the types there.

Information systems are essentially equivalent to the standard partition models of information in games (as in Aumann 1987). There the primitives are the set of worlds, and n partitions of this set, one for each player. An information state of i is then defined as an atom of i's partition. Formally, information systems are slightly more parsimonious than partition models, and are better suited to our current purposes.

The purpose of this essay is to consider information systems where some players are irrational at some states of the world, and the implications this has for behavior at other states of the world, where all players *are* rational. In particular, we will see how such considerations can resolve certain "paradoxes of rationality" arising from backward induction in "centipede games."

3 High-Order Mutual Knowledge

Recall that an event is called *common* knowledge if all players know it, all know that all know it, and so on *ad infinitum*. Call an event *mutual* knowledge if all players simply know it; *second-order* mutual knowledge if it is mutual knowledge that it is mutual knowledge; and so on for any finite order. Thus, common knowledge is the same as mutual knowledge for all finite orders simultaneously.

In a conversation at Stanford during the summer of 1987, Jay Kadane of Carnegie-Mellon University suggested that it might be worthwhile to investigate the consequences of relaxing the assumption of common knowledge of rationality that underlies Aumann 1987 (see section 1 above). At the time, the framework for carrying out such an investigation was not clear. I suggest that the framework of the previous section is appropriate for this purpose.

Indeed, in such a framework it is quite easy formally to describe a situation in which there is iterated mutual knowledge of rationality up to any given order, but not common knowledge. Thus, consider a variant of the example in the previous section, with the same game G as there, but with the information system

	L	L	L	L	R
T	1,1*	1,1*	1,1	1,1	0,0
T	1,1*	1,1	1,1*	1,1	0,0
T	1,1	1,1*	1,1	1,1*	0,0
T	1,1	1,1	1,1*	1,1	0,0*
B	0,0	0,0	0,0	0,0*	2,2*

where each star represents a probability of 1/10, and the other probabilities vanish. Then in the state of the world represented by the top left corner there is second-order—but not third-order—mutual knowledge of rationality.

Common knowledge of rationality seems a very strong assumption. Even mutual knowledge of rationality is quite strong. A player may be rational himself, but how can he know for sure that another person is? Even if he does, it seems highly unlikely that he will know for sure, without the slightest doubt, that the other player knows that he is rational. Yet this would only be second-order mutual knowledge of rationality; common knowledge of rationality is far stronger.

The question therefore arises as to what happens when we relax this requirement just a little, e.g. by allowing a small measure of doubt to enter at a high level of mutual knowledge. Suppose, for example, that both players are rational, both know for sure that both are rational, and both impute very high probability to this being so (i.e., to both knowing for sure that both are rational); but they do not know it for *sure*. Can this make a big difference in the outcome of a game?

We will see that it can make a very big difference indeed.

4 Paradoxes of Backward Induction

Among the most disturbing counterintuitive examples of rational interactive decision theory are the paradoxes of "backward induction." The best known example is the finitely repeated prisoner's dilemma, in which all

Nash equilibria (and also all correlated equilibria) dictate playing "greedy" ("defect") at each stage, no matter how often play is repeated. Rosenthal (1982) constructed a particularly simple class of backward-induction paradoxes—the "centipede games"—a striking instance of which, due to Megiddo (1986), is as follows: \$10.50 is lying on a table. Alice has the option of taking \$10, leaving 50 cents for Bob. If she does, the game is over. If not, the amount on the table is increased tenfold, to \$105, and it is Bob's turn to play: he can take \$100, leaving \$5 to Alice. If he does, the game is over. If not, the amount on the table is again increased tenfold. This continues for a total of at most three rounds, and then ends. If play has not terminated before, then at the last stage Bob has the opportunity of taking \$1,000,000, leaving \$50,000 for Alice; after that, play is over, whether or not Bob takes the opportunity.

At all Nash equilibria of this game, Alice "goes out" immediately —i.e., she takes the \$10 at her first move, leaving 50 cents for Bob. This is also so for all correlated equilibria.

Some readers may feel that, upon consideration, Megiddo's example is not so terribly counterintuitive; that although Alice may feel quite frustrated, a considered analysis will nevertheless lead her in the end to pick up the money at the first opportunity. After all, there is a difference between frustration and paradoxicality; the players of a one-shot prisoner's dilemma will certainly feel frustrated, but the logic of that situation *does* inexorably point to playing "greedy," and the time has long passed since this was considered paradoxical. One might even say that properly digesting Megiddo's example makes it easier to accept the equilibrium outcome of, say, the 100-times-repeated prisoner's dilemma.

But others will feel that if this is rationality they want none of it—or, more to the point, that it represents an approach that is of little practical interest, at least in this example.

5 A Resolution of the Backward-Induction Paradoxes

What is proposed here is to apply the new kind of bounded rationality suggested in section 3, in which all agents are in fact perfectly rational but there is some breakdown in the commonality of the knowledge that they are. Typically, we will be interested in situations where there is iterated knowledge ("Alice knows that Bob knows that Alice knows...") of rationality up to a specified level, but no further. We will find that an extremely small breakdown in the commonality of the knowledge of rationality is enough to justify the kind of behavior that most of us would consider intuitively "reasonable."

The Megiddo example appears in matrix form as follows:

Bob

Alice		**a**	**b**	**c**	**d**
	A	10 / .50	10 / .50	10 / .50	10 / .50
	B	5 v / 100	1,000 w / 50	1,000 / 50	1,000 / 50
Alice	**C**	5 / 100	500 x / 10,000	100,000 y / 5,000	100,000 / 5,000
	D	5 / 100	500 / 10,000	50,000 z / 1,000,000	0 / 0

The rows and columns represent pure strategies of Alice and Bob, respectively. They are arranged in the order in which they prescribe "going out;" thus, Alice's third row represents the strategy in which she takes the money—"goes out"—when (and if) it is offered to her for the third time, but not before. In each entry, the upper number represents Alice's payoff, the lower one Bob's.

Consider now the information system depicted in the above matrix. Each player has just one information state for each of his pure strategies. There is a common prior, with non-negative probabilities v, w, x, y, and z; all other probabilities vanish. Note that with probability 1 Alice "stays in" on her first move.

(When one is discussing an example, as here, common priors are a desideratum, not a restriction. In theorems one strives for generality; in examples, for "regularity," for an absence of pathology. The point is that the backward-induction paradoxes can be resolved even if we restrict ourselves to common priors; *a fortiori* if we don't.)

For simplicity, consider first the case where $x = 0.99447$, $y = 0.0055$, $z = 0.00003$, and $u = v = w = 0$. Note that Bob is rational at all three states of the world that occur with positive probability; as for Alice, she is rational at x and at y, but not at z. Thus, Bob is rational with probability 1, and Alice with a probability very close to 1; the probability that she is irrational is only 3 in 100,000.

Suppose now that the true state of the world happens to be the one that has the overwhelming probability, namely x. Then both Alice and Bob are rational; moreover, each of them knows that both are rational. Indeed, it is common knowledge that Bob is rational. And he still gets $10,000, which is surely the most he could reasonably hope for.

One can ring the changes on this in many ways. The underlying point is that for Bob to be rational everywhere, and Alice everywhere except at z, it suffices that, of the ratios v/w, w/x, x/y, and y/z that are not 0/0, none exceeds 198. The probability is then 1 that Bob is rational and gets at least $50; and at the state of the world corresponding to v, both players are rational, both know this (with certainty), both know that they know this, and both know that they know that they know this, and yet Bob collects $100. And though Alice is irrational with positive probability, that probability may be tiny: when all the above ratios are 198, the probability z that Alice is irrational (as estimated by an outside observer) is 0.000000000648—less than one in a billion. Multiplying this probability by the size of Alice's "loss" (or rather forgone profit) at z, namely $50,000, we get an "expected irrationality" of less than 1/300 of a cent, an utterly insignificant sum.

Similar considerations hold for the finitely repeated prisoner's dilemma. With extremely small—practically speaking, negligible—overall probabilities of irrationality, one can, in the 100-fold repetition, justify "cooperating" until at least the 85th or the 90th stage, and even beyond. This kind of result jibes well with observed experimental behavior, in which subjects start "defecting" more and more as the game approaches its end.

This attractive resolution for these paradoxes gives a rigorous justification to the elusive idea that, whereas one should certainly play rationally at the end, it seems somehow foolish to act from the very beginning in the most pathologically pessimistic, "play it safe" way. The above analysis shows that, contrary to what had been thought, one may act in a way that is rational—even with a high degree of mutual knowledge of rationality—and yet is quite profitable. And it reflects the fact that in real interactive situations there is a great deal of uncertainty about what others will do, to what extent they are rational, what they think about what you think and about your rationality, and so on.

6 Measuring Irrationality

In this section we confine our attention to information systems with a common prior, to which the term *probability* will refer.

The methodology of section 5 lends itself well to the task of measuring irrationality. We have already used one such measure: "expected irrationality." Formally, this is defined for each player i separately, as follows: At each of i's information states, one multiplies the probability of that state by the difference between i's expected payoff there and the maximum expected payoff that he could have gotten by changing his strategy; then one sums over all of i's information states. This measure of irrationality should be useful in comparing various approaches to approximate rationality, including the ε-rationality of Radner (1980, 1986), the "perturbations" of Kreps, Milgrom, Roberts, and Wilson (1982), and the approach of this paper. What is being suggested is that, in evaluating a specified solution in a specified context, one should take into account both the amount of irrationality (as does Radner) and its probability (as do Kreps et al.). There is an obvious tradeoff between the two that is not captured by either one separately: a player may be willing to tolerate a higher probability of an irrationality that results in the loss of \$1 than one costing \$1 million.

The information systems of section 5 involve "correlated" strategies: each player's choice of a pure strategy depends on his belief, with different beliefs corresponding to different choices. Effects that are similar insofar as expected irrationality is concerned can be obtained in the independent case—with "mixed" strategies, so to speak. Thus, suppose that at each of his information states Bob has the same belief about Alice's choice, represented by the probabilities $A, B, \ldots$ (see the matrix in section 5). Similarly, let $a, b, \ldots$ represent Alice's belief at each of her information states. Let $B/C = C/D = a/b = b/c = 198$, and let $A = d = 0$. Then, as in the second example in section 5, Bob's expected irrationality is 0, and Alice's is less than 1/300 of a cent. (Alice is irrational only when choosing row D.) The *probability* of irrationality is greater than in section 5; it is $D = 0.0000253$, rather than $z = 0.000000000648$. But the *size* of the irrationality (the loss in expected payoff from making that choice rather than maximizing) is smaller; not \$50,000, but \$50,000 $\times$ c. Hence, the expected irrationality is \$50,000 $\times$ cD. Since $c = D$, this works out to \$0.0000320, as compared with \$0.0000324 in the second example of section 5.

Thus, there is no appreciable difference in expected irrationality between the two systems; it is negligibly small in both (in fact, slightly smaller in the independent case). What makes the example of section 5 more attractive is that at v, say, it displays a much higher level of mutual knowledge of rationality. By contrast, in the independent case there is no mutual knowledge of rationality at all; at v (indeed, at all states), Bob ascribes positive probability to Alice's being irrational.

We may define the *degree* of irrationality of a player i at state w of the world as minus the maximum level of mutual knowledge of i's rationality at w. This is to be distinguished from expected irrationality as defined above, which measures the "amount" of irrationality in the system. Thus, whereas the example of section 5 has a slightly higher "amount" of irrationality than the "independent" example, it has a far lower degree.

Expected irrationality is a global measure; it applies to the information system as a whole. Degree of irrationality is a local measure; it applies to a specific state of the world. It would be desirable also to develop a local measure of expected irrationality, conditional on a specific state of the world and on the information of the players at that state. This would enable us to refer to the "amount of irrationality" that is needed in the system to justify the choice of a certain strategy pair by the players (like x in Megiddo's game), even though under the circumstances both players' choices are completely rational.

7 A Difficulty with the Definition of Irrationality

The framework of section 2 requires each player, at each state of the world, to have well-defined probabilities for the other players' information states, pure strategies, and so on. Yet it permits him sometimes to be irrational, i.e., to fail to maximize expected utility. This raises a difficulty, since personal probabilities are usually *defined* via utility maximization (Savage 1954); nonmaximizers of utility don't have probabilities.

Consider first the case when there is a common prior, as in sections 5 and 6. The underlying idea of common priors is that reasonable probability assessments are or should be based on information, so that people with exactly the same information "should" have the same probabilities. Thus, one can think of a common prior as representing the assessment of a rational outside observer who has no personal interest in the outcome of the game and no private information. If he then gets the same private information as a player who is rational, he will entertain the same probabilities as that player; and if the player is not rational, it is natural to think of the observer's probabilities as the "right" ones.

In the general case, when there need not be a common prior, the difficulty may be resolved by revising the definition of an information system (section 2). Item iiib of the definition specifies a "theory" (a complete system of probability assessments) for each player at each of his information states. The revised definition specifies whether or not that player is rational at that information state, and specifies a theory for him only when he is rational. This enables us to formulate and prove general theorems, and to define the *degree* of irrationality (see section 6); its disadvantage is that it does not enable us to measure the expected *amount* of irrationality, as in sections 5 and 6.

8 Relations with "Crazy Perturbations"

In the "crazy perturbation" literature (e.g., Kreps et al. 1982; Kreps and Wilson 1982; Milgrom and Roberts 1982; Fudenberg and Maskin 1986; Aumann and Sorin 1989; Fudenberg and Levine 1989), one looks for Nash equilibria of a repeated game in which there is a small exogenous probability that a player plays irrationally but in such a way as to motivate the other player to play in some specific way—e.g., a mutually beneficial way. Most of the results say that, in one sense or another, the rational types tend to mimic those irrational types that are in some sense "best" for the player who does the mimicking. Intuitively, one might say that the rational types "disguise" themselves as irrational; they make believe they are crazy, thus "forcing" the other player to play accordingly (i.e., to maximize against the selected irrational type).

In general, this kind of theorem gives rather sharp results; it says that *all* Nash equilibria of the perturbed game have payoffs in a sharply defined class. They thus go in the opposite direction from many of the theorems on repeated games, in which results akin to the folk theorem lead to very diffuse sets of equilibrium payoffs.

Consider now a situation in which there is mutual knowledge of some given order that all players are rational, but not common knowledge; and that if the players are not rational, then they play some specified strategy, or a strategy in some specified set (such as bounded recall strategies, as in Aumann and Sorin 1989). Can theorems of the above type still be proved in this more general context? The model outlined in section 2 provides an appropriate general framework in which to investigate this question.

For the finitely repeated prisoner's dilemma, this question was treated by Kreps et al. (1982). Suppose that it is first-order mutual knowledge that the players are rational: each player is rational and knows that the other is, but does not know that the other knows that he is. In particular, he thinks that the other may think that, with some fixed small probability, he is a tit-for-tat (TFT) automaton. In this situation, it would appear that each player will still be motivated to pretend that he is TFT, since he doesn't know that the other knows that he isn't. Once he plays TFT, the other is motivated to go along, to the mutual benefit of both.

Now what happens when mutual knowledge of rationality is taken one step further—i.e., when not only does each player know that the other is rational, but each also knows that the other knows that he is? Then the reasoning appears to break down; each player might feel that there is no sense in pretending to be crazy, since he knows that the other knows that he is rational.

But further analysis indicates that the reasoning survives this step as well. A player playing TFT indeed knows that the other knows that he is pretending; but she (the other) needn't know this. He therefore figures that she may well go along, in order to make him think that she believes that he is indeed crazy, and so encourage him to continue with the mutually beneficial TFT. So he tries, and she indeed goes along, for precisely that reason.

Indeed, it appears that one can carry mutual knowledge of rationality to any finite level short of common knowledge, and still get the same effect: that the players will be motivated to play mutually beneficial but seemingly irrational strategies.

Most of us have experienced situations where some harmful fact is perfectly well known but is studiously overlooked by everybody. In this case, the harmful fact is the players' rationality (!). More precisely, the fact itself need not be harmful, but common knowledge of it would be. The above approach enables us to understand this phenomenon within the context of the theory.

This reasoning presupposes the kind of setting where the perturbation is mutually beneficial. In other settings where the "crazy perturbation" methodology has been applied, such as the chain-store paradox, we do not know whether the results extend.

Dedication

Lovingly dedicated to Frank Hahn, who has always had a penchant for the irrational.

Acknowledgments

Research partially supported by NSF grant IRI-8814953 at the Institute for Mathematical Studies in the Social Sciences, Stanford University. I am grateful to Adam Brandenburger for some important input, and, in particular, for pointing out a serious error in a previous version.

References

Aumann, R. J. 1987. "Correlated equilibrium as an expression of Bayesian rationality." *Econometrica* 55: 1–18 [Chapter 33].

Aumann, R. J., and S. Sorin. 1989. "Cooperation and bounded recall." *Games and Economic Behavior* 1: 5–39 [Chapter 24].

Fudenberg, D., and D. Levine. 1989. "Reputation and equilibrium selection in games with a single patient player." *Econometrica* 57: 759–778.

Fudenberg, D., and E. Maskin. 1986. "The Folk Theorem in repeated games with discounting and with incomplete information." *Econometrica* 54: 533–554.

Harsanyi, J. C. 1967–68. "Games of incomplete information played by Bayesian players" (parts I, II, III). *Management Science* 14: 159–182, 320–334, 486–502.

Kreps, D., P. Milgrom, J. Roberts, and R. Wilson. 1982. "Rational cooperation in the finitely repeated prisoner's dilemma." *Journal of Economic Theory* 27: 245–252.

Kreps, D., and R. Wilson. 1982. "Reputation and imperfect information." *Journal of Economic Theory* 27: 253–279.

Megiddo, N. 1986. Remarks on Bounded Rationality. IBM Research Report RJ 54310.

Milgrom, P., and J. Roberts. 1982. "Predation, reputation and entry deterrence." *Journal of Economic Theory* 27: 280–312.

Radner, R. 1968. "Competitive equilibrium under uncertainty." *Econometrica* 36: 31–58.

Radner. R. 1972. "Existence of equilibrium plans, prices, and price expectations in a sequence of markets." *Econometrica* 40: 289–303.

Radner, R. 1980. "Collusive behavior in non-cooperative epsilon-equilibria in oligopolies with long but finite lives." *Journal of Economic Theory* 22, 136–154.

Radner, R. 1986. "Can bounded rationality resolve the prisoners' dilemma." In *Essays in Honor of Gerard Debreu*, ed. A. Mas-Colell and W. Hildenbrand. North-Holland, pp. 387–400.

Rosenthal, R. 1982. "Games of perfect information, predatory pricing, and the chain store paradox." *Journal of Economic Theory* 25: 92–100.

Savage, L. J. 1954. *The Foundations of Statistics*. Wiley.