23 Survey of Repeated Games

Work on repeated games may be broadly divided into two categories: repeated games with complete information, and repeated games with incomplete information. While there are important interrelationships between these two categories, each one represents a coherent and more or less separate body of work with its own set of basic ideas and problems. A closely related area, which is currently very active, is that of stochastic games, but this will not be surveyed here.

1. Repeated Games of Complete Information

The theory of repeated games of complete information is concerned with the evolution of fundamental patterns of interaction between people (or for that matter, animals; the problems it attacks are similar to those of social biology). Its aim is to account for phenomena such as cooperation, altruism, revenge, threats (self-destructive or otherwise), etc. – phenomena which may at first seem irrational – in terms of the usual "selfish" utility-maximizing paradigm of game theory and neoclassical economics.

Let G be an n-person game in strategic (i.e. "normal") form. We will denote by G^* the "supergame" of G, i.e. the game each play of which consists of an infinite sequence of plays of G. Unless we indicate otherwise, we will assume that at the end of each "stage" (i.e. a

---

This work was supported by National Science Foundation Grant SOC75-21820-A01 at the Institute for Mathematical Studies in the Social Sciences, Stanford University (IMSSS). It was presented as background material for a one-day workshop on repeated games that took place at the IMSSS summer seminar on mathematical economics on August 10, 1978. Since then there have been several developments, some of them very important, which are not surveyed here.

This chapter originally appeared in *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern*, pp. 11–42, Bibliographisches Institut, Mannheim, 1981.

particular play of G in the sequence), each player is informed of the strategy chosen by all other players at that stage. Thus the information available to a player when choosing his strategy for a particular stage consists of the strategies used by all players at all previous stages. The *payoff* in G^* is some kind of average of the payoffs in the various stages; more precisely, it may be defined as the Cesaro limit of these averages (the limit, as $k \to \infty$, of the average of the payoffs $h_1, \ldots, h_k$ to the first k stages), or as the Abel limit (the limit, as the interest rate r tends to 0, of the normalized current value of the payoff $r \sum_{j=1}^{\infty} \frac{h_j}{(1+r)^j}$). There are technical difficulties with these definitions, because the limits involved need not always exist; however, for the purposes of this brief survey let us ignore these technical difficulties and proceed with the statement of the theorems.

Theorem 1: *The payoff vectors to Nash equilibrium points in the supergame G^* are the feasible individually rational payoffs in the game G.*

Here a payoff vector h is called *feasible* in G, if it is a convex combination of payoff vectors to pure strategy n-tuples in G; i.e. if it is a possible payoff to a *correlated* strategy n-tuple in G. It is called *individually rational* if for each player i,

$$h^i \geqq \min_\tau \max_\sigma H^i(\sigma, \tau),$$

where H is the payoff function in G, σ ranges over the strategies of i in G, and τ over the strategies of the players in $N \setminus \{i\}$. It should be noted that the min max is not necessarily equal to the max min; in general, it is greater. This is because the set of strategies available to $N \setminus \{i\}$ is not necessarily convex. For example, if the players are confined to using mixed strategies, then we would not in general have min max = max min. The min max is the level of payoff below which i cannot be forced by the remaining players, but he cannot necessarily guarantee himself this payoff. If, however, the players in $N \setminus \{i\}$ can correlate their strategies independently of i – e.g., if they can spin a roulette wheel that i cannot see – then the max min and the min max are equal.

Theorem 1 has been generally known in the profession for at least 15 or 20 years, but has not been published; its authorship is obscure. The proof, which rests on the idea of threats of punishment, is in principal quite simple; it is outlined in Appendix 2. We shall call it the "Folk Theorem". More or less, it may be considered the starting point or touchstone for the further developments which will be described below.

The significance of the Folk Theorem is that it relates cooperative behavior in the game G to non-cooperative behavior in its supergame G^*. This is the fundamental message of the theory of repeated games of complete information; that cooperation may be explained by the fact that the "games people play" – i.e., the multiperson decision situations in which they are involved – are not one-time affairs, but are repeated over and over. In game-theoretic terms, an outcome is cooperative if it requires an outside enforcement mechanism to make it "stick". Equilibrium points are self-enforcing; once an equilibrium point is agreed upon, it is not worthwhile for any player to deviate from it. Thus it does not require any outside enforcement mechanism, and so represents non-cooperative behavior. On the other hand, the general feasible outcome does require an enforcement mechanism, and so represents the cooperative approach. In a sense, the repetition itself, with its possibilities for retaliation, becomes the enforcement mechanism.

Branching out from the Folk Theorem there are developments in several directions, which we will take up in roughly historical order. The first arose from an attempt to refine somewhat the notion of "cooperative outcome" on the cooperative side of the Folk Theorem. Specifically, while the set of all feasible individually rational outcomes does represent a solution notion of sorts for a cooperative game, it is relatively vague and uninformative. One would like a characterization, in terms of the supergame G^*, of more specific kinds of cooperative behavior in the single-shot game G; e.g., of the core. This is achieved by replacing the notion of equilibrium by *strong equilibrium,* defined[1] as an n-tuple of strategies for which no *coalition* of players can simultaneously all do better for themselves by moving to different strategies, while the players outside of the coalition maintain their original strategies. We then have

Theorem 2: *The payoff vectors to strong equilibrium points in the supergame G^* are the elements of the β-core of the game G.*

Here the *β-core* of G is the set of feasible payoff vectors x such that each coalition S can be prevented by its complement from achieving for each of its members i a payoff larger than x_i. It may be contrasted with the larger and somewhat more transparent *α-core,* defined as the set of feasible payoff vectors x such that no coalition S can guarantee

[1] In an arbitrary game, not necessarily a supergame.

to each of its members i a payoff larger than x_i. Vis-a-vis the α-core, the β-core plays a role similar to that played by the min-max vis-a-vis the max-min (compare the discussion of Theorem 1 above).

Theorem 2 was proved in 1959 by the writer of these lines (in *Contributions to the Theory of Games* IV, Princeton University Press, 287–324). The original paper is poorly written and difficult to read. We must, however, not be too harsh in our judgements; bitter experience shows that while the basic ideas in this subject may be quite transparent, it is by no means easy to formulate the proofs in a manner that is at once rigorous, elegant, and brief. A rigorous statement of Theorem 2 may be found on page 551 of the *Transactions of the American Mathematical Society* Vol. 98 (1961). An intuitive outline of the proof may be found on page 23 (Theorem 13) of the 1967 Morgenstern Festschrift[2].

A sort of footnote to Theorem 2 was published in the *Pacific Journal of Mathematics* in 1961 (Vol. 98, 539–552). It asserts that in games of perfect information, the strong equilibrium points of G^* are achievable in *pure* supergame strategies. In spite of the classical connection between pure strategies and perfect information, this is not quite obvious. Suppose x is in the β-core and S is a coalition. The complement of S will certainly have a pure strategy τ in G that prevents S from achieving more than x_i for each of its members i. That is, for each strategy σ of S, there will be a member i of S whose payoff from the pair (σ, τ) will be $\leqq x_i$. But it is conceivable that by jumping around from one σ to another, S could, against τ, get more *on the average* than x_i for each of its members i.

To put it differently, mixing really has two functions: secrecy and convexification. It is fairly clear that the first of these two functions is obviated by perfect information, but not so clear that the second one is.

The proof uses the convexity of $v_\beta(S)$ (the set of S-vectors that the complement of S cannot prevent S from achieving); this in itself is non-trivial – it depends on Kakutani's fixed point theorem (see *Transactions of the American Mathematical Society* 98 (1961), 549). Since x is in the β-core, it is not in the interior of $v_\beta(S)$, and so can be (weakly) separated from $v_\beta(S)$ by a hyperplane with non-negative coefficients. These coefficients are used to combine the payoffs of S's members into a single number, which is taken as the payoff to a 2-person zero-sum game of perfect information between S and its

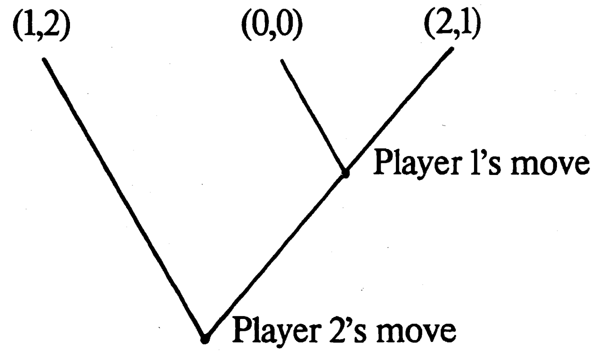
[2] *Essays in Mathematical Economics in Honor of Oskar Morgenstern,* edited by Martin Shubik, Princeton University Press, Princeton, N.J., 1967.

complement. In case S defects, the complement uses a pure optimal strategy in this 2-person zero-sum game.

Soon after John Harsanyi's landmark work in the mid-sixties on the general theory of games of incomplete information[3], attention in the theory of repeated games shifted to the incomplete information case. This extensive work will be surveyed separately in Section 2 below.

In the mid-seventies interest in the complete information case started reviving. We mention first the sociologically and biologically oriented applications, such as the papers of Hammond[4] and Kurz[5] on Altruism, and J. Maynard Smith's[6] work on evolution. However, we prefer to postpone discussing these provocative works until after we have described the more recent theoretical developments.

The first of these centers on Reinhard Selten's concept of *perfect equilibrium point*[7]. Experience with Nash equilibrium points in games played over time leads to the conclusion that they can embody elements not usually associated with a purely non-cooperative theory. Specifically, a Nash equilibrium point may dictate a choice that is irrational in a given situation, which makes sense only in terms of deterring the other players from making moves leading to that situation. In other words, it permits empty threats, threats that could never rationally be carried out when the chips are down; and this even when the other players know it to be so. A simple but revealing example is the extensive game with the tree



[3] As often happens in our business, the work became widely known several years before it was published in 1967–68 (*Management Science* 14, 159–182, 320–334, 486–502).

[4] "Charity: Altruism or cooperative egotism?" in E. Phelps, ed. (1975), *Altruism, Morality, and Economic Theory*, Russell Sage Foundation, New York.

[5] "Altruistic Equilibrium", in *Economic Progress, Private Values, and Public Policy* (The Fellner Festschrift), North Holland, 1977; see also Footnote 11.

[6] "Evolution and the Theory of Games", *American Scientist* 64 (1976), 41–45.

[7] "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games", *International Journal of Game Theory* 4 (1975), 25–55.

in which (L, L) is a Nash equilibrium point; Player 1 goes left because of 2's rather unconvincing threat spitefully to go left if 1 goes right.

To deal with this problem, Selten introduced his concept of perfect equilibrium point. It is based on what might be called the "trembling hand principle". Assume that whenever a player makes a choice, the outcome will be as he intended with a probability that is almost, but not quite, one; each of the other possibilities – those that he did not intend – will occur with an infinitesimal probability. This infinitesimal probability, while it is not large enough to affect the payoff significantly, does make every situation (i.e., every vertex in the game tree) a possibility that must be reckoned with; it thus eliminates empty threats and thereby drastically reduces the set of equilibrium points.

Two quick applications:

(i) A perfect equilibrium point cannot assign positive probability to any "weakly dominated" strategy[8].

(ii) In a game of perfect information, the perfect equilibrium points are precisely those obtained by the usual "dynamic programming" procedure, i.e., the one in Zermelo's classic paper on chess. (In non-zero sum games there *are* others – many others!)

In a sense, much of the theory of repeated games is an attempt to cut down, in one way or another, the bewildering wealth of equilibrium payoffs provided by Theorem 1. Thus Theorem 2 obtained a drastic reduction by going to strong equilibria – a reduction that may be too drastic, as the core is frequently empty. The proof of Theorem 1 (see Appendix 2) indicates that a reduction can also be expected by going to perfect equilibria, since the proof depends heavily on "punishments" that may well be harmful to the punisher; they thus constitute precisely the kind of irrational threat that is excluded by the perfectness notion. Unfortunately, we have

Theorem 3: *The payoff vectors to perfect equilibrium points in the supergame G^* are the same as the payoff vectors to the Nash equilibrium points.*

In other words, no reduction is achieved by going to perfect equilibria. But while the result itself is certainly disappointing, the proof is actually quite revealing. The equilibrium is held together by an infinite regress of threatened punishments, in which a player who does not

---

[8] σ_1 *weakly dominates* σ_2 if σ_1 yields at least as much as σ_2 to the player using it no matter what the other players do, and more for at least one $(n-1)$-tuple of strategies of the other players.

punish a defector as he should is in turn punished by the defector for not punishing him. Thus a motorist stopped by the highway patrol may refrain from offering the patrolman a bribe for fear of being turned in by him; and the patrolman would probably indeed turn him in, for fear of being himself turned in by the motorist otherwise. Much of whatever stability society may possess is perhaps traceable to this kind of perfect equilibrium.

This theorem was demonstrated in 1976 by Rubinstein[9] and by Aumann and Shapley[10] independently, with relatively minor differences in formulation. It has not yet been published.

We come next to the subject of discounting. Research in undiscounted supergames has a slightly disconcerting never-never quality. What a person actually does at any particular stage doesn't directly affect his payoff *at all.* In certain evolutionary applications, where all that seems to matter is long-term survival, such a model is perhaps just what is called for. In the more usual applications, though, it seems rather extreme. There literally is all the time in the world to do anything you might want to do (in the nature of signalling, punishing, repenting and all the other activities that the denizens of this stylized world engage in). Time, far from being money, is for all intents and purposes free.

One good example is the treatment of randomized punishments. Punishing a defector may involve holding him to his min max or close to it, and this in turn will in general involve the use of randomized strategies. This leads to no particular problems when done in connection with Theorem 1. In Theorem 3, though, a punisher who fails to punish must himself be punished, if necessary by the defector. But if the punisher is using randomized strategies and the defector can only observe the pure realizations, how can he tell whether the punisher is doing what he should? In real time this certainly leads to statistical decision problems that are far from trivial. Here, though, it doesn't. Any deviation can be detected with an arbitrarily high probability if one can make as many observations as one likes; and in this case we indeed can, and still have all the time in the world for any subsequent (second-order) punishments that may be called for.

One way to make time "bite" is to substitute a large but finite num-

---

[9] Subsequently published as "Equilibrium in Supergames," in N. Megiddo (ed.), *Essays in Game Theory in Honor of Michael Maschler*, Springer-Verlag, 1994, 17–27.

[10] Subsequently published as "Long-Term Competition—A Game-Theoretic Analysis," ibid., 1–15 [Chapter 22].

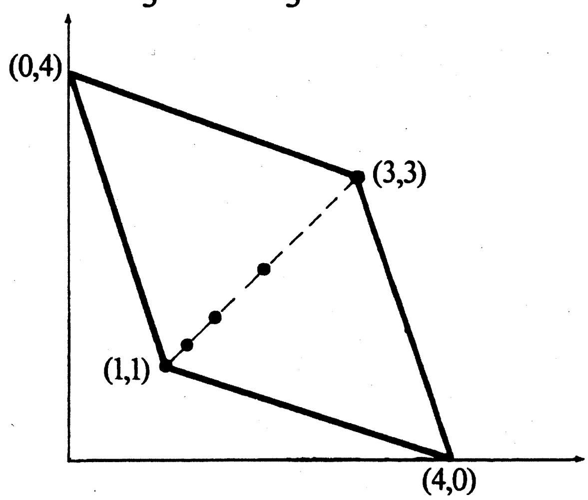
ber of stages for the infinite number we have been using up to now. But the presence of a last stage which is recognized as such by all players, aside from being unrealistic, creates unnatural terminal effects which propagate themselves backwards and grossly distort the entire analysis. Besides, one cannot have a treatment that is in any sense stationary when the number of stages is fixed and finite.

A more natural way to take account of the value of time is by discounting. A discount factor may be viewed either as a stop probability or a measure of impatience or a combination of the two. It turns out that for small discount factors the set of equilibrium payoffs is close (in the Hausdorff topology) to that in the zero discount case. As the discount grows, though, the situation changes, and we get some noteworthy effects.

Consider, for example, the prisoner's dilemma

3,3	0,4
4,0	1,1

In any finite repetition the only equilibrium payoff is $(1,1)$. This is also an equilibrium payoff in *every* discounted version. As the discount rate $\delta=\dfrac{r}{1+r}$ grows from 0 to 1, the set of equilibrium payoffs gradually changes, until only $(1,1)$ remains. However, it does not change continuously. Certain payoff vectors which are close to $(1,1)$ for at least one of the players are the first to disappear. As δ grows the set changes in a rather complicated manner, and then, when $\delta=\frac{2}{3}$, we get a set that includes the denumerable set $\{(3,3), (1\frac{2}{3}, 1\frac{2}{3}), \ldots, (1+\frac{2}{3^{n-1}}, 1+\frac{2}{3^{n-1}}), \ldots\} \cup \{(1,1)\}$ pictured here.



Immediately after that, everything disappears except (1,1) (compare Kurz's second Altruism article[11], in which there is a similar effect).

Also the set of perfect equilibrium payoffs is continuous near $\delta=0$, but for $\delta>0$ it will in general be unequal to the set of equilibrium payoffs. Lloyd Shapley and I have worked out an example; see Appendix 3. As might be expected, the equilibrium nature of a payoff depends on the available punishments. The perfect equilibrium nature of a payoff, however, is a more subtle matter; it depends not only on the available punishments, but also on their *costs* to the punisher. Thus there emerges something conceptually akin to the Nash variable threats bargaining model.

The next item on our shopping list is a critical examination of Theorem 2, which "predicts" an outcome in the core. To obtain such a result we had to go from Nash equilibrium points to strong equilibrium points. Now in a sense this is begging the question; strong equilibrium involves group action in its definition, and this already assumes a measure of cooperation. Somehow, one feels that in a sufficiently detailed model, the joint action necessary to achieve a strong equilibrium should *follow* from considerations of individual utility maximization, rather than having to be assumed.

The problem already presents itself in the search for an individualistic justification for Pareto optimality (or efficiency) in terms of repeated games. For a specific example, let us return to the prisoner's dilemma discussed above. In the models we have discussed up to now, whether they are limiting average or discounted with a moderate discount rate, efficient outcomes have always been *possible* as equilibrium outcomes. But invariably, other outcomes – including the highly inefficient (1,1) – have also been possible. Many of us feel strongly that (1,1) is not a rational outcome of a repeated prisoner's dilemma. Can't we somehow narrow down the definition of equilibrium so that it yields efficient outcomes only, while at the same time involving only individualistic concepts?

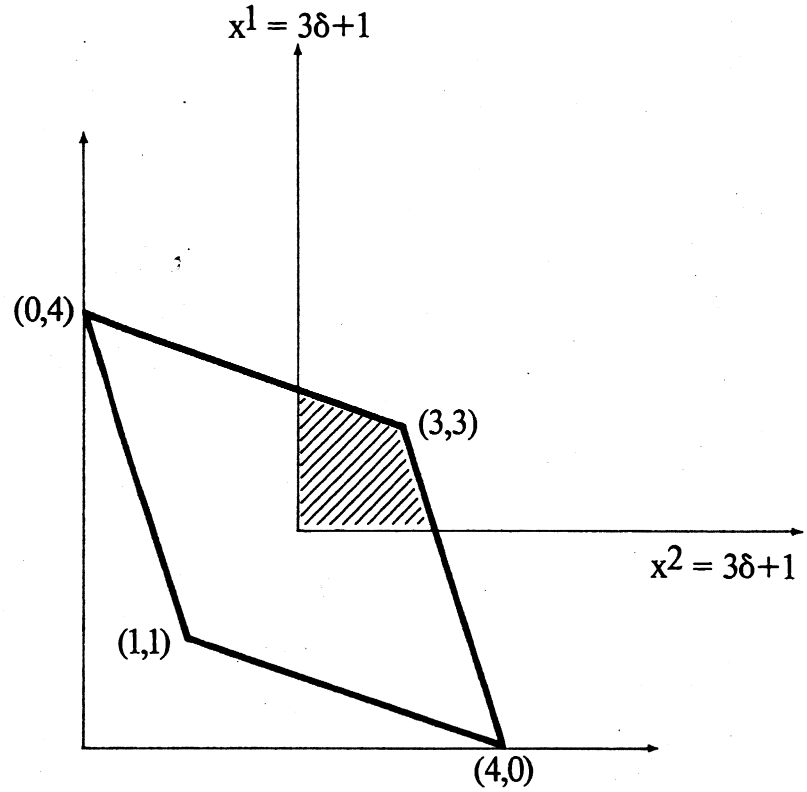
One approach is that of Kurz's second altruism paper, which considers *stationary* equilibria[12] in repeated games. In the prisoner's dilemma presented above there are actually only two stationary equilibrium outcomes (when $0 \leqq \delta \leqq \frac{2}{3}$), namely (1,1) and (4,4). However,

---

[11] "Altruism as an Outcome of Social Interaction", *American Economic Review* 68 (1978), 216–222.

[12] Equilibria yielding a flow of payoffs not depending on the serial number of the period.

if we consider the variant in which the choice of mixed strategies at each stage is observable, then the set of stationary equilibrium outcomes is considerably larger. In fact, we get the shaded corner in the following figure, *plus* the point (1,1). As δ grows the corner shrinks; finally (at $\delta = \frac{2}{3}$) we get only (3,3), plus, unfortunately, (1,1). Thus although discounting appears to eliminate more and more inefficiency, it fails to eliminate the ultimate, "mulish" inefficiency in which the players just keep playing "double-cross" no matter what happens.



To try to get around this, let us examine more carefully the concept of stationarity. This is not captured adequately by simply demanding a constant pay-off flow. What we want is stationary *strategies*, not just stationary payoffs. Now the simplistic definition of stationary strategy in which the players choose the same stage-strategy (or *action*) at each stage is obviously unsatisfactory, since it does not allow for any response to what was done at previous stages. The proper definition would seem to be that a player's action at a given stage depends only on the history, not on the serial number of the stage; but this is also unsatisfactory, since the serial number of the stage can be read off from the history (it is simply its length). To make it work, we would need time stretching infinitely backwards, a concept that has been considered in a game context[13], but is associated with great difficulties.

One way to handle this is to assume that each player has a finite memory. More precisely, his mind has a finite number of "states". Before each stage, the player may change the state of his mind in a way that depends only on the previous state and the previous action of the other player; he must then play an action that depends only on the (new) state.

It should be noted that this formulation does *not* mean that memory extends only a bounded time backwards. For example, the "grim" strategies under which a player gets doublecrossed forever if he even once doublecrossed his friend are certainly admissible. It does, however, put a bound to the complexity a strategy can have and enables an analysis in the framework of finite games.

Mordecai Kurz, Jonathan Cave and I have analyzed one very special case of this, which one may call "memory zero"; see Appendix 5. In this case, the new state may not depend on the old one but only on the previous action of the other player. Even then the analysis is not trivial; we get eight pure strategies on each side (four possibilities for the steady state multiplied by two for the initial move). It turns out that in the undiscounted case, there are several equilibrium outcomes; but successive[14] weak domination eliminates all but one, namely (3, 3). Unfortunately, in the discounted case this no longer holds true. And though there is no clear connection between perfect equilibrium and successive weak domination (such as the connection with ordinary weak domination mentioned in (i) above), this result does raise the possibility that in the general finite memory undiscounted case, at least the worst "mulish" behavior will not appear as a perfect equilibrium.

Finite memory has some conceptual ties to Radner's bounded rationality[15]. In addition we cite the very provocative recent paper of Robert Rosenthal[16], which also uses a finite state space hypothesis.

---

[13] Schwartz, G. (1974), "Randomizing when Time is Not Well-Ordered", *Israel Journal of Mathematics* 19, 241–245.

[14] If one eliminates all weakly dominated strategies from a game, one obtains a new game to which one can apply the process of weak domination; this can be iterated arbitrarily often. For a concrete example see Appendix 5.

[15] R. Radner (1986), "*Can Bounded Rationality Resolve the Prisoners' Dilemma*", in Mas-Colell, A. and Hildenbrand, W. (editors) (1986), *Essays in Honor of Gerard Debreu*, Elsevier Science Publishers B.V.: Amsterdam, pp. 387–400.

[16] "Sequences of Games with Varying Opponents", *Econometrica* 47 (1979), 1353–1366.

Rather oddly, Rosenthal's results are quite different from ours; he always gets the double-double cross as an equilibrium, and the "friendly" solution only in isolated instances.

We come now to the applicative side, and in particular to the contributions of Hammond, Kurz, and Maynard Smith mentioned above. Hammond and Kurz use the theory of repeated games to account for altruism, defined as the act of giving some good to another player without necessarily receiving anything in return. Kurz's "Altruistic Equilibrium"[17] is essentially an equilibrium in a repeated game in which the players keep giving each other goods, with the punishment strategy ready in the background if they should ever refuse[18].

Finally, we have Maynard Smith's "Evolution and the Theory of Games". It is of course unfortunate that game theorists have communicated so poorly with the world at large that more than a score of years after the focus of research shifted away from zero-sum games, an eminent scientist can still write that anything that is not minimax is not game theory. But this is a minor point; the paper is highly interesting and well worth studying (again, I am indebted to Arrow for sending me to this). Presumably Maynard Smith's simple animal contests could be put in the framework of small-memory models in the sense we defined above; one could then get a more systematic search for evolutionarily stable strategies than appear in his paper. It is clear that there is quite a bit of room for communication and interaction between biologists and game theorists[19].

---

[17] Balassa, B. and Nelson, R. (editors) (1977), *Economic Progress, Private Values, and Public Policy,* North Holland, 177–200.

[18] Do unto me kindness and *truth,* please do not bury me in Egypt" (Genesis 47, 29); the Patriarch Jacob is speaking to his son Joseph. The eleventh century commentator Rashi explains that "kindness to the dead is *true* altruism (Khessed shel emmet), as the doer does not expect any quid pro quo". Since then, the phrase "true altruism" has been traditionally used for services connected with burial. Perhaps this too can be accounted for "rationally", by an intergenerational equilibrium model; but it is certainly a more complex matter. A similar problem, though in an entirely different context, was brought up by Ken Arrow in a recent conversation. When Al Capone was released from jail after serving close to ten years, he was physically and mentally a broken man. Nevertheless he was well and respectfully cared for by the mob. More generally, of course, the whole matter of care for the aged and the infirm comes under this heading.

[19] Recently Reinhard Selten has taken up the evolutionary theme. See "A Note on Evolutionarily Stable Strategies in Asymmetric Animal Conflicts", *Journal of Theoretical Biology* 84 (1980), 93–101.

2. Repeated Games of Incomplete Information

The stress here is on the strategic use of information – when and how to reveal and when and how to conceal, when to believe revealed information and when not, etc.

The subject has broken itself up naturally into the zero-sum and the non-zero-sum case. That is because problems of revelation and concealment occur already in the zero-sum case, and can be studied there "under laboratory conditions", without the distractions and complications of cooperation, punishments, incentives, etc. A mathematical theory of considerable elegance, depth, and scope has grown up in the last fifteen years, almost all of which is concerned with the zero-sum case (Appendix 6 is a partial bibliography, to which we will refer throughout this part). This theory still has some fascinating open problems, to which I will allude in the sequel. For the purpose of economic applications, though, the non-zero-sum theory is more interesting and challenging; it is also far less complete, so that there is more room for spadework.

One cannot understand the non-zero-sum theory without first reviewing the zero-sum theory, which we now proceed to do. One of two 2-person zero-sum games, A and B, is being played repeatedly; Player 1 knows which of the two it is, but Player 2 only knows the probability p that it is A (this probability is common knowledge). This can be viewed as a complete information game G^* in which nature chooses A or B with probabilities p and 1-p respectively, and informs Player 1 but not Player 2 of the choice; the game is then played repeatedly.

The central fact of the theory is that the situation cannot be understood unless p is allowed to vary. Denote by $u(p)$ the value of $pA+(1-p)B$, i.e. the one-shot (or many-shot) game when Player 1 ignores his information. Denote by $v(p)$ the value of G^*. Let Cav u be the least concave function that is $\geqq u$ for each p.

Theorem 4: $v(p) = \text{Cav}\, u(p)$.

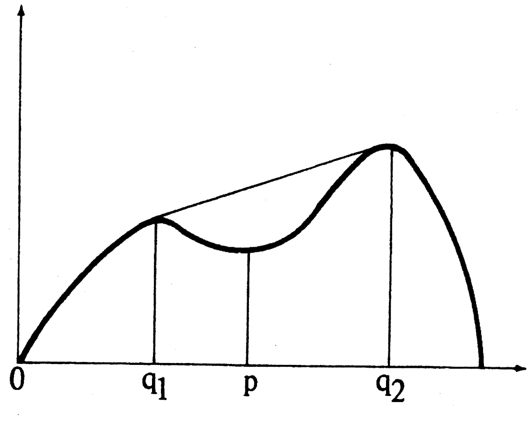
To prove this, one starts out with a very simple argument, based on the fact that in a zero-sum situation information can't hurt you, which shows that $v(p)$ must be concave. Next, note that since Player 1 can choose to ignore his information, we must have $v(p) \geqq u(p)$. Since $v(p)$ is concave, it follows that $v(p) \geqq \text{Cav}\, u(p)$.

The opposite inequality is slightly trickier. By the minimax theorem there is a mixed strategy that *guarantees* $v(p)$ to Player 1; therefore he may as well announce it (and use it). Given this strategy, Player 2 can deduce a sequence of posterior probabilities $(p_1, p_2, p_3, \ldots)$ for the contingency that nature originally chose A. The p_k are random variables because they depend on which pure actions were chosen by Player 1 in the various stages; we have $Ep_k = p$ for each k. Now it can be shown that any infinite sequence of probabilities that are conditioned on more and more information converges with probability 1; that is, there is a random variable q such that $p_k \to q$ with probability 1 as $k \to \infty$ (q need not be 0 or 1, and need not be constant). What this means is that after a while, Player 1 will have revealed just about all the information he is ever going to reveal (which is perhaps obvious anyway). From then on he must play the same (or almost the same) whether A or B were chosen – otherwise he would be revealing more. At this point Player 2 thinks that nature has chosen A with probability close to q and Player 1 is essentially ignoring his information. Therefore, Player 1 cannot do better than $u(q)$. On average, therefore, he cannot do better than $Eu(q)$. But since $Ep_i = p$ and $p_i \to q$ almost surely, we have $Eq = p$. Hence by Jensen's inequality,

$$Eu(q) \leqq E\,\mathrm{Cav}\,u(q) \leqq \mathrm{Cav}\,u(Eq) = \mathrm{Cav}\,u(p),$$

and the "proof" is complete.

The odd thing about the whole procedure is that it is all imaginary; Player 2 cannot *really* assume that Player 1 is using any particular optimal strategy or for that matter an optimal strategy at all. He will *not* be computing any posteriors and must guard his flanks by a completely different procedure, called the *Blackwell strategy*. We do, however, get a guide as to how far Player 1 can go in revealing the true game being played (i.e., in playing differently contingent on A and on B). In fact, what he must do is find two probabilities q_1 and q_2 such that there is an α in $[0,1]$ with $p = \alpha q_1 + (1-\alpha) q_2$ and $u(p) = \alpha u(q_1) + (1-\alpha) u(q_2)$ (see Figure).



Now let us be given two coins with different parameters r^A *and* r^B. Suppose Player 1 tosses the coin with parameter r^A if A was chosen by nature, and the coin with parameter r^B if B was chosen. If Player 2 could observe the outcome of the toss but not which coin was tossed, he could deduce posterior probabilities for A. It can be show that r^A and r^B can be chosen so that the probability that the coin comes out heads is α, the posterior probability of A given heads is q_1 and the posterior probability of A given tails is q_2. An optimal strategy for Player 1 is then to toss such a coin *once,* and from then on play optimally in $q_1A+(1-q_1)B$ or in $q_2A+(1-q_2)B$ according as the coin came out heads or tails. Though he is not actually announcing the outcome of the toss, the effect is the same.

(It should be noted that this gives an alternative proof for $v(p) \geqq \operatorname{Cav} u(p)$.)

The zero-sum theory has been developed in much more general situations. Before we describe some of these developments, though, it is necessary to clarify the definition of G^*. Two approaches suggest themselves: *limit of value* means that we consider the value of an n-stage game, divide by n, and then let $n \to \infty$; *value of limit* means that we define the payoff by some kind of limiting average of the payoffs and consider the value of the infinite-stage supergame.

Under either approach one could also use the present value of payoff and let discount tend to zero, but in the context of repeated games this has not been studied much[20].

Intuitively, limit of value is appropriate when we have a large number of stages, and we know how many. Optimal strategies developed under this approach often depend critically on the serial number of the stage. Value of limit is appropriate when there is a large number of stages and we have no clear idea of how many there are. Optimal strategies here are in some sense more stationary than in the limit of value case.

For Theorem 4, all approaches are equivalent. The theorem (and the equivalence between all approaches) still holds when the two games are replaced by any fixed finite number k, but one player still is told exactly which game is being played.

A more significant generalization is to the case when the players

[20] Though in the context of stochastic games it has been; see Bewley and Kohlberg in *Mathematics of Operations Research* 1 (1976), 197–208 and 321–336. Also repeated games with *fixed* non-zero discount have been studied; more on this below. Some of the limiting average results go through without too much difficulty for the limiting discount case; see, for example, Mertens and Zamir (1971–72).

are not told, at the end of each stage, exactly what one-stage strategy was used by the opponent (cf. Appendix 1, Problem 2). A typical example of this is when the individual stages consist of extensive games, in which case one would know at most the moves made by the opponent, and perhaps not even that; but certainly not the (one-stage) *strategy*. The information the players receive at the end of a stage may depend only on the game being played and on the strategies chosen by nature; i.e. in addition to the payoff matrix there is an "information matrix" associated with each game. The central concept here is that of a "non-revealing strategy", defined as a one-stage mixed strategy which, when used by the informed player, leads to the same distribution of information for the uninformed player no matter which game is actually being played (though the informed player may make use of his information). It can be shown (Aumann and Maschler (1968)) that G^* still has a value in this case; in fact, if $u(p)$ is taken to be the value of the one-stage game in which Player 1 is restricted to using non-revealing strategies, then Theorem 4 applies. The equivalence between all approaches still holds here.

When we abandon the assumption of full information for one player, things get more complicated. First let us consider the value of limit approach. The first, archetypical case (independent types and full revelation of stage strategies) was considered by Stearns (1967), who derived a formula for the sup inf and inf sup payoffs of these games; from this formula it follows that the value will in general decisively *fail* to exist (i.e. inf sup$>$sup inf)! Later Mertens and Zamir (1980) greatly generalized Stearns's result by removing the assumption of independent types and, with some restrictions, that of stage-strategy revelation.

The limit of value approach to these games was developed by Mertens and Zamir (1971–72) and Mertens (1971–72), (1973), who showed that in general the limit of the value *does* exist, and gave ways by which it may be calculated. Kohlberg, Mertens and Zamir have relaxed the informational conditions and explored the subject thoroughly in various directions (see Appendix 6) that I will not specify further here.

This recital of results may sound somewhat dry; perhaps the reader is deluded into thinking that what we are discussing here are fairly straightforward extensions and variations of two or three basic ideas. Nothing could be further from the truth. Many of the works in this area coming after the ACDA (Arms Control and Disarmament Agency) period of 1966–68, far from simply using and embellishing the earlier ideas, were in fact tremendous tours de force of ingenuity, beauty, and depth. The problems attacked were often open for many years in spite of concerted attempts to solve them; the mathematical ideas employed are among the most surprising and original in the theory of games.

Having gotten this off our chest, let us continue with the survey. In our "proof" of Theorem 4 we described an optimal strategy for the informed player but indicated that the construction of an optimal strategy for the uninformed player was a more complex matter. The construction of such a strategy (called a *Blackwell* strategy because of its reliance on a paper of Blackwell[21] that is fundamental in much of repeated game theory) in a very rudimentary case was done in Aumann and Maschler (1966), and was subsequently extended to much more general situations in Stearns (1967) and in Kohlberg (1975a).

The case of a fixed non-zero discount was considered by Mayberry (1967); his results are rather startling, in that he found that even in the simplest cases, $v(p)$ will be a concave function of p that is not differentiable at any rational value of p (though it must be differentiable a.e. because of its concavity).

The results of Ponssard (1975a), (1975b), (1976), Ponssard and Zamir (1973), Ponssard and Sorin (1980) have mainly concerned finite games or games repeated a fixed finite number of stages; some of these results can be considered as complementing the theory surveyed here, which concerns mainly the case of "many" stages.

I would like to end this survey of the zero-sum situation by proposing two problems. The first has to do with the speed of convergence (in the "limit of value" approach). Returning to the situation of Theorem 4, it can be proved that the speed of convergence is always $O(1/\sqrt{n})$; it is often much smaller. Zamir (1971–72) found an example in which the order of magnitude $1/\sqrt{n}$ is actually attained. This order of magnitude is also suggested by the following intuitive considerations: Suppose we have $u(p) = \operatorname{Cav} u(p)$, so that aside from "fine tuning", the informed player must ignore his information. This means that he will play the same mixed stage-strategy at each stage. Now there will always be a natural random variation in the average payoff. The central limit theorem indicates that this variation will be of the order of $1/\sqrt{n}$. One can think of Zamir's example as one in which the informed player can take advantage of the natural varia-

[21] "An Analog of the Minimax Theorem for Vector Payoffs", *Pacific Journal of Mathematics* 6 (1956), 1–8.

tion, i.e. play a strategy that masquerades as a constant mixed stage-strategy, but actually varies slightly from it. The purposeful variation is of exactly the order that the uninformed player might expect as random, but is actually used to the hilt to the advantage of the informed player.

The interesting thing is that the error term is not always of the order $1/\sqrt{n}$, even when $u(p) = \text{Cav}\, u(p)$. So this kind of "masquerading" is not always possible.

In a startling piece, Mertens and Zamir (1976) have shown that *in the particular original example* of Zamir, there is even a much closer connection to the Central Limit Theorem than that outlined above. The normal distribution actually makes an explicit appearance. Obviously there is something behind this that is much more general than this particular example. But what is it?

It's worth stating that Mertens and Zamir's proof has absolutely *nothing* to do with the above intuitive reasoning. Somehow, inexplicably, the normal distribution springs forth from some differential equation. It is all very mysterious and begs for an explanation (in terms of a general theorem).

The second problem is also based on a recent paper of Mertens and Zamir. Almost all of the work on this subject uses reasoning that is in some way related to that used in our above "proof" of Theorem 4; i.e. there hovers in the background some imaginary posterior probability of the uninformed player which takes the place, at each stage, of the original exogenously given p. But there is a class of games, called *games without a recursive structure,* in which this reasoning somehow doesn't get off the ground. They are very difficult, but Mertens and Zamir (1976) managed to solve *one* such game. The door to this class of problems has opened a crack, but it doesn't seem to budge any more. Opening it wide would involve the development of a radically new technique completely unrelated to what we know now.

We come finally to the non-zero-sum incomplete information case. In the first instance one simply seeks a characterization of equilibrium payoffs, in the spirit of the Folk theorem for the complete information case (Theorem 1). Obviously the incentive-compatibility of revelations is a central issue here, and one could perhaps have hoped for a solution related to the IIC set of Roger Myerson[22]. Unfortunately the issue seems much more complex.

Let us confine ourselves to the 2-person case in which there is com-

[22] "Incentive Compatibility and the Bargaining Problem", *Econometrica* 47 (1979), 61–73.

plete information on one side and the stage strategies become known after each stage. This was attacked in Aumann, Maschler and Stearns (1968), but even in this very simple case only with partial success. Specifically, a sufficient characterization was found, but it was shown not to be necessary. There were even cases in which there are no equilibrium points satisfying the sufficient condition, though there *were some* equilibrium points. To this day it is not known whether even the simplest such games ($k=2$, i.e. there are only two possibilities for the true game) necessarily possess any equilibrium points.

The basic, inescapable complication in this business is that revelation and signalling get all mixed up and interfere with each other. By *revelation* we mean what happens when the informed player makes use of his information, enabling the uninformed player, at least in principle, to make deductions about the true state of nature. By *signalling* we mean what the informed player actually wants to tell the uninformed player. The desire to signal does not imply an ability to reveal; the signal would have to be incentive compatible, and it will not always be. Though this is well known, a simple example is

4,4	0,0	Prob = 0.4

5,0	4,4	Prob = 0.6 .

Player 1 has no strategic choice, but he knows which game is being played; he can try to signal it to Player 2, who must choose a column. A priori Player 1 would like to signal the truth, and Player 2 would like to believe him; but obviously he can't, even if in fact Player 1 is telling the truth.

Appendix 1

Problems Regarding the Complete Information Case

1. Discounting, Perfect Equilibrium, and the Cost of a Threat

I would give first priority to the complete information case with a (fixed) positive discount. For reasons explained in the body of this paper, the positive discount case is in many ways conceptually more attractive than the zero discount case. It is also mathematically more tractable; we have an unambiguously defined payoff function, and the strategy pairs form a compact space on which the payoff function is continuous. The preliminary results (see Appendix 3) indicate that in this case the perfect equilibrium points involve a close relationship between the effectiveness of a punishment (or threat) and its costs to the threatener. This relationship is certainly worth exploring to the hilt.

2. Imperfect Information at the End of a Stage

I don't mean incomplete information in the sense that we don't know the payoff, but imperfect information at the end of each stage, in the sense that each player gets only partial information about the strategies used by the other players. Consider, for example, an altruistic equilibrium in Kurz's sense[23] involving three or more people. Suppose that the donor never becomes known; the recipient only gets to know that he received a gift, not from whom. (Maimonides lists eight levels of charity, of which the anonymous donor who lends the recipient money in order to enable him to make his own livelihood is the highest. Aware of human frailties, he admits the lesser levels as well.) It is not difficult to see that the Altruistic Equilibrium survives this change. It's somewhat less clear whether the Altruistic Perfect Equilibrium survives it, and it seems perfectly clear[24] that the Altruistic Strong Equilibrium will *not* survive it. In more general games, even the ordinary equilibrium will not survive: if you do not know who is defecting, whom will you punish? Can one get general characterizations in this case?

3. Intergenerational Equilibrium

The need for intergenerational models was described in Footnote 20 (Jacob & Joseph, etc.). Offhand, I see no reason why it should not be possible to carry this out, but it would certainly be worthwhile to do it explicitly.

4. Extreme Altruism

A more difficult problem is that posed by Kurz's example of the child in the fire ("Altruistic Equilibrium", p. 182). Kurz assumes that the probability is 90% that both the rescuer and the child will die, but this seems extreme; it is unlikely that a model based on rationality at any level could account for a rescue in such a situation. But even if the probability of death is 40% (for rescuer and child), it does not seem easy to account for a rescue on a purely individualistic basis, as-

[23] See footnote 17.

[24] Famous last words.

suming that the rescuer does not think it likely that *he* will ever have to be similarly rescued.

We may be too hung up on the importance of the individual human being. In fact, groups may be more important. Biologists stress the survival of individual genes; even if we do not wish to assume a biological gene for altruism, there might be a sort of socio-educational "gene" for it – an idea that is transmitted from generation to generation and survives because the groups who practice and propagate it survive.

5. Finite Memory Models

These were described in the body of the text and certainly bear further investigation. In addition to its intrinsic interest, a finite memory model might make many of the other problems easier to handle.

A particular case of a finite memory model is one in which only actions that are at most a fixed finite number of steps back can be considered ("let bygones be bygones"). This would rule out "grim" strategies. The zero-memory model is of course of this kind. If the general finite-memory models turn out intractable, perhaps these would be less so.

6. Learning and Efficiency[25]

Kurz's speeder ("Altruistic Equilibrium", p. 184) is signalling other motorists to speed as well; he is teaching them that the true equilibrium in which the world finds itself involves more speeding than they thought. Conversely, his voter is teaching people in general to vote (presumably Pareto superior to not voting). There is a theorem here which Kurz and I have been pursuing for years without success. The "theorem" says that if an equilibrium point (in the undiscounted game) has several possible payoffs with positive probability each, then no one can Pareto dominate another; each equilibrium is efficient among those that can actually occur. The idea of the "proof" is to look at an equilibrium point not as a conscious choice of strategies by n people, but as a sort of probability estimate of what the world really looks like (see Appendix 4). In a repeated game, the players will revise this estimate as time progresses. If two outcomes A and B with A Pareto dominating B were possible, then in a state of the world leading to B, each player would try to "teach" the others that the true

[25] In this connection see also Kurz's "Altruism as an Outcome of Social Interaction", and Tomioka's "On the Bayesian Selection of Nash Equilibrium", (Chapter V of his thesis).

state leads to A (even though they may think that this is highly unlikely). Such teaching would be learned with alacrity by the other players.

To make this work one would have to make some kind of assumption under which each of the two equilibrium outcomes in question always retains a positive probability no matter what the players do at each stage.

Can this "theorem" be made into a theorem?

Appendix 2

Outline of Proof of the Folk Theorem[26]

It is easy to prove that the payoff vectors to Nash e.p.'s in the supergame G^* are feasible and individually rational in G; the more significant part is the converse. Assume for simplicity that $n=2$ (there are just two players). Suppose h is a feasible individually rational payoff vector. Here we may write $h=\sum_{i=1}^{k} \alpha_i h_i$, where the α_i are non-negative weights that sum to 1, and all the h_i are payoff vectors corresponding to pure strategy pairs in G. Suppose first that the α_i are rational, and express them in the form $\alpha_i=\frac{p_i}{q}$, where p_i are positive integers and q is their sum. The payoff vector h can then be achieved as a limiting average in G^* by having the players play, for p_1 consecutive periods, an n-tuple that achieves h_1, then for p_2 consecutive periods an n-tuple that achieves h_2, and so on; after q periods, we start again from the beginning.

If the α_i are irrational, the same effect can be attained approximating to them by rational numbers and playing once through each approximation in turn, to yield the desired limiting average.

This procedure, however, does not yet describe a Nash equilibrium point in G^*, and in fact does not even describe a pair of supergame strategies. A supergame strategy must describe each player's responses to the other player's actions not only when he "plays along", but also when he "defects" from a prescribed course of play, such as the one described above. This is where the requirement that h be individually rational comes in.

---

[26] Adapted from a manuscript by R. Aumann and L. Shapley [See Chapter 22].

Since h is individually rational, the payoff h^i of each player i is at least equal to his min max value (see the discussion following the statement of Theorem 1 in the text). Let τ be a mixed strategy for the other player j that holds i down to his min max value; a fortiori, j can, by playing τ, guarantee that i will not receive more than h^i.

We may now describe an e.p. in G^* as follows: the players start by playing to obtain an average payoff of h, as outlined above. If at any stage a player i "defects" – i.e., does not play the prescribed choice in G for that round – then starting from the next round, the other player j plays the mixed strategy τ. This will hold i down to at most h^i, so that he will have gained nothing by his defection. Thus, h is indeed an e.p.

In this demonstration one can see clearly the role that "threats" play in enforcing a prescribed "cooperative" outcome. Repetition of G enables "endogenous enforcement" of agreements, and thus the basically noncooperative notion of Nash equilibrium points becomes cooperative in the supergame.

Appendix 3

Discounted Payoffs in Repeated Games: Discussion of an Example

by R. J. Aumann and L. S. Shapley

Please see Section 5 of Chapter 22 (pp. 403–408) of this volume.

Appendix 4

A Bayesian View of Equilibrium

Let n be a positive integer. Define an n-person *information structure* to consist of

(i) A measurable space (Ω, B) *(the states of the world)*, and

(ii) An n-tuple $I=(I_1, \ldots, I_n)$ of sub-σ-fields of B (I_i is i's *information σ-field*).

Let G be a game with n players, a finite set S_i of pure strategies for each player i, and a payoff function H: $S_1 \times \ldots \times S_n \rightarrow R^n$. Given an information structure (Ω, B, I), a *strategy* of player i is an I_i-measurable function from Ω to S_i. An *equilibrium point* in G is defined to consist of an information structure (Ω, B, I), a probability measure p on (Ω, B), and an n-tuple $s = (s_1, \ldots, s_n)$ of strategies, such that for each j and each strategy t_j of j, we have

$$(*) \qquad E H_j(s_{-j}, t_j) \leqq E H_j(s)\,,$$

where E is the expectation operator w.r.t. the measure p, and (s_{-j}, t_j) stands for $(s_1, \ldots, s_{j-1}, t_j, s_{j+1}, \ldots, s_n)$.

In this definition it is possible to replace the single "objective" probability measure p by n different "subjective" probability measures p^i. In that case j's expectation operator E would have to be replaced by E_j, but otherwise the definition would not be affected. We refrain from using the more general set-up only for notational simplicity.

There are several ways of viewing an equilibrium point. One is as a self-enforcing agreement. The players agree to use a randomizing device described by (Ω, B, p), to be informed of the outcome in accordance with the I_i, and to act in accordance with the s_i; once having agreed to this, it will be to the advantage of no players to renege. This is the point of view taken in [1] (see especially Section 9b).

Another point of view is that of "rational expectations". Each point ω in Ω is thought of as constituting a complete objective description of a state of the real world. Such a description would of course include the choice $s_j(\omega)$ of each player j in the game G, as well as j's state of information (i.e. the set of all states of the world that j cannot distinguish from ω). Thus both the information structure and the equilibrium strategies s_i are in a sense tautologically defined. What is not tautological is the probability measure p, which describes how each player views the world, probability-wise. In assuming that this is the same for all players, we are putting ourselves in Harsanyi's "consistent case" [2]; however, as noted above, this is not an important assumption for us.

This formulation does away with the dichotomy usually perceived between the "Bayesian" and the "game-theoretic" view of the world. From the Bayesian viewpoint, subjective probabilities should be assignable to everything, including the prospect of a player choosing a certain strategy in a certain game. The so called "game-theoretic" viewpoint holds that subjective probabilities can only be assigned to events not governed by rational decision makers; for the latter, one must substitute an equilibrium (or other game-theoretic) notion. The above formulation synthesizes the two viewpoints: Equilibrium is viewed as the *result* of Bayesian rationality, condition (*) appearing as a simple maximization of utility, given player j's subjective probability distribution over the states of the world.

This definition of Equilibrium Point encompasses the notion of a Nash Equilibrium Point (i.e. one in mixed strategies) as a special case. But it also encompasses significantly more; cf. [1].

References for Appendix 4

[1] R. Aumann (1974). "Subjectivity and Correlation in Randomized Strategies", *Journal of Mathematical Economics 1*, 67–95 [Chapter 31].

[2] J. Harsanyi (1967–68). "Games of Incomplete Information Played by Bayesian Players", Parts I–III, *Management Science 14*, 159–182; 320–334; 486–502.

Appendix 5

Repeated Prisoner's Dilemma with Memory Zero

Strategies F and D stand for "Friendly" and "Doublecross" respectively. The letter outside the brace indicates the initial move. Inside the brace the two letters indicate the steady-state response to the opponent's playing F and D respectively on the previous move. Each pair of strategies generates a payoff stream that is periodic in the steady state, and whose average appears in the matrix. For the matrix of the one-shot game and other explanatory material, see the body of the article.

		F				D			
		FF	FD	DF	DD	FF	FD	DF	DD
F	FF	3,3	3,3	0,4	0,4	3,3	3,3	0,4	0,4
	FD	3,3	3,3	2,2	1,1	3,3	2,2	2,2	1,1
	DF	4,0	2,2	2,2	0,4	4,0	2,2	0,4	0,4
	DD	4,0	1,1	4,0	1,1	4,0	1,1	4,0	1,1
D	FF	3,3	3,3	0,4	0,4	3,3	3,3	0,4	0,4
	FD	3,3	2,2	2,2	1,1	3,3	1,1	2,2	1,1
	DF	4,0	2,2	4,0	0,4	4,0	2,2	2,2	0,4
	DD	4,0	1,1	4,0	1,1	4,0	1,1	4,0	1,1

For each player, all strategies except FFD are successively eliminated as follows: First FDF (by DDF) and DFD; then FFF and DFF; then DDF; then DDD; and finally FDD. Five stages are needed in all. Except that of FDF in the first stage, all eliminations are by FFD.

Appendix 6

A Partial Bibliography of Repeated Games of Incomplete Information (up to 1979)

Aumann, R. and Maschler, M. (1966). "Game Theoretic Aspects of Gradual Disarmament", *Report of the U.S. Arms Control and Disarmament Agency/ST-80,* Chapter V, Washington, D.C.

Aumann, R. and Maschler, M. (1967)."Repeated Games with Incomplete Information: A Survey of Recent Results", *Report of the U.S. Arms Control and Disarmament Agency/ST-116,* Chapter III, 287, Washington, D.C.

Aumann, R. and Maschler, M. (1968). "Repeated Games of Incomplete Information: The Zero-Sum Extensive Case", *Report of the U.S. Arms Control and Disarmament Agency/ST-143,* Chapter III, 37, Washington, D.C.

Aumann, R., Maschler, M. and Stearns, R. E. (1968). "Repeated Games of Incomplete Information: An Approach to the Non-Zero-Sum Case", *Report of the U.S. Arms Control and Disarmament Agency/ST-143,* Chapter IV, 117, Washington, D.C.

Kohlberg, E. (1974). "Repeated Games with Absorbing States", *Annals of Statistics 2,* 724.

Kohlberg, E. (1975a). "Optimal Strategies in Repeated Games of Incomplete Information", *International Journal of Game Theory 4,* 7.

Kohlberg, E. (1975b). "The Information Revealed in Infinitely-Repeated Games of Incomplete Information", *International Journal of Game Theory 4,* 57.

Kohlberg, E. and Zamir, S. (1974). "Repeated Games of Incomplete Information: The Symmetric Case", *Annals of Statistics 2,* 1040.

Mayberry, J. (1967)."Discounted Repeated Games with Incomplete Information", *Report of the U.S. Arms Control and Disarmament Agency/ST-116,* Chapter V, 435, Washington, D.C.

Mertens, J. F. (1971–72). "The Value of Two-Person Zero-Sum Repeated Games – The Extensive Case", *International Journal of Game Theory 1,* 217.

Mertens, J. F. (1973). "A Note on 'The Value of Two-Person Zero-Sum Repeated Games – The Extensive Case'", *International Journal of Game Theory 2,* 231.

Mertens, J. F. and Zamir, S. (1971–72). "The Value of Two-Person Zero-Sum Repeated Games with Lack of Information on Both Sides", *International Journal of Game Theory 1,* 39.

Mertens, J. F. and Zamir, S. (1975). "The Maximal Variation of a Bounded Martingale", *Israel Journal of Mathematics 27,* 252.

Mertens, J. F. and Zamir, S. (1976a). "On a Repeated Game Without a Recursive Structure", *International Journal of Game Theory 5,* 1973.

Mertens, J. F. and Zamir, S. (1976b). "The Normal Distribution and Repeated Games", *International Journal of Game Theory 5,* 187.

Mertens, J. F. and Zamir, S. (1980). "Minmax and Maxmin of Repeated Games with incomplete Information", *International Journal of Game Theory 9,* 201.

Ponssard, J. P. (1975a). "Zero-Sum Games with Almost Perfect Information", *Management Science 21,* 794.

Ponssard, J. P. (1975b). "A Note on the LP Formulation of Zero-Sum Sequential Games", *International Journal of Game Theory 4,* 1.

Ponssard, J. P. (1976). "On the Subject of Nonoptimal Play in Zero-Sum Extensive Games: The Trap Phenomenon", *International Journal of Game Theory 5,* 107.

Ponssard, J. P. and Sorin, S. (1980). "The LP Formulation of Finite Zero-Sum Games with Incomplete Information", *International Journal of Game Theory 9,* 99.

Ponssard, J. P. and Zamir, S. (1973). "Zero-Sum Sequential Games with Incomplete Information", *International Journal of Game Theory 2,* 99.

Sorin, S. (1979). "A Note on the Value of Zero-Sum Sequential Repeated Games with Incomplete Information", *International Journal of Game Theory 8,* 217.

Stearns, R. E. (1967). "A Formal Information Concept for Games with Incomplete Information", *Report of the U.S. Arms Control and Disarmament Agency/ST-116,* Chapter IV, 405, Washington, D.C.

Zamir, S. (1971–72). "On the Relation Between Finitely and Infinitely-Repeated Games with Incomplete Information", *International Journal of Game Theory 1,* 179.

Zamir, S. (1973a). "On Repeated Games with General Information Function", *International Journal of Game Theory 2,* 215.

Zamir, S. (1973b). "On the Notion of Value for Games with Infinitely Many Stages", *Annals of Statistics 1,* 791.

Notes:

1. Stochastic Games are not included in this bibliography, although they have important applications to repeated games with incomplete information.
2. The help of Jonathan Cave and Sylvain Sorin in researching this bibliography is gratefully acknowledged.

Note added in proof A fuller account of the incomplete information case was published as *Repeated Games with Incomplete Information,* by R. J. Aumann and M. B. Maschler, Cambridge: MIT Press, 1995. This book also contains a fairly complete bibliography of the subject up to 1995.