THE WORK OF JOHN NASH IN GAME THEORY

Nobel Seminar, December 8, 1994

(The document that follows is edited version of a seminar devoted to the contributions to game theory of John Nash. The participants, in the order of their appearance, were:

HAROLD W. KUHN
Department of Mathematics, Princeton University, Princeton, NJ 08544, USA

JOHN C. HARSANYI
The Walter A. Haas School of Business, University of California at Berkeley, Berkeley, CA 94720, USA

REINHARD SELTEN
Department of Economics, University of Bonn, Adenauerallee 24 - 42, D-53113 Bonn, Germany

JÖRGEN W. WEIBULL,
Department of Economics, Stockholm University, S-10691 Stockholm, Sweden

ERIC VAN DAMME
Center for Economic Research, Tilburg University, 5037 AB Tilburg, The Netherlands

JOHN F. NASH, JR
Department of Mathematics, Princeton University, Princeton, NJ 08544, USA

PETER HAMMERSTEIN, Max-Plank-Institute für Verhaltens Physiologie, 82319 Seewiesen, Germany, is the co-author of the intervention of Professor Selten.
A unified bibliography for the seminar is provided at the end.)

KUHN

It gives me great pleasure to chair this seminar on the importance of Nash's work on the occasion of the first Nobel award that recognizes the central importance of game theory in current economic theory. I shall be joined by two colleagues whom I've known for over thirty years, John Harsanyi and Reinhard Selten, two new friends, Jörgen Weibull and Eric van Damme, and John Nash, whom I've known since we were graduate students together in Princeton forty-six years ago.

The timing of these awards has historical significance, since this year is the fiftieth anniversary of the publication of "The Theory of Games and Economic Behavior" [52] by the Princeton University Press. Although von

Neumann had laid the mathematical foundation of the theory of games in his paper entitled “Zur Theorie der Gesellshaftsspiele” [51], published in the Mathematische Annalen in 1928, it was largely through the collaboration of von Neumann and Morgenstern that economists learned of this new tool for analyzing economic problems.

Some of you may have read Morgenstern’s own account [33] of this collaboration. There is a new historical study [29] by Robert Leonard of the University of Quebec at Montreal that points out that “understandably, but regrettably, Morgenstern’s reminiscence sacrifices some of the historical complexity of the run up to 1944.” Leonard’s study gives most of the credit for the creation of game theory to von Neumann who had written essentially all of the mathematical manuscript nine months before Morgenstern ever saw it. Nevertheless, had von Neumann and Morgenstern never met, it seems unlikely that we would be here today celebrating the central role of the theory of games in economics.

This leads to a natural question which has been asked repeatedly by journalists in the last two months: ‘Why did it take fifty years for these new ideas to be recognized?” To give a partial answer to this question, we must look more closely at the developments in the late forties and early fifties. A crucial fact is that von Neumann’s theory was too highly mathematical for economists. Therefore, the theory of games was developed almost exclusively by mathematicians during this period. To describe the spirit of the time, allow me to quote from Robert J. Aumann’s magnificent article [3] on game theory in the New Palgrave Dictionary “The period of the late 40’s and early 50’s was a period of excitement in game theory. The discipline had broken out of its cocoon and was testing its wings. Giants walked the earth. At Princeton, John Nash laid the groundwork for the general non-cooperative theory and for cooperative bargaining theory. Lloyd Shapley defined the value for coalitional games, initiated the theory of stochastic games, coinvented the core with D.B. Gillies, and together with John Milnor developed the first game models with a continuum of players. Harold Kuhn reformulated the extensive form of a game, and worked on behavior strategies and perfect recall. Al Tucker discovered the Prisoner’s Dilemma, and supported a number of young game theorists through the Office of Naval Research.”

HARSANYI

When did Tucker discover the Prisoner’s Dilemma?

KUHN

Al Tucker was on leave at Stanford in the Spring of 1950 and, because of the shortage of offices, he was housed in the Psychology Department. One day a psychologist knocked on his door and asked what he was doing. Tucker replied: “I’m working on game theory.“, and the psychologist asked if he would give a seminar on his work. For that seminar, Al Tucker invented the Prisoner’s Dilemma as an example of game theory, Nash equilibria, and the attendant paradoxes of non-socially-desirable equilibria. A truly seminal example, it inspired dozens of research papers and several entire books.

It is important to recognize that the results that I have enumerated did not respond to some suggestion of von Neumann, nor did they follow work that he had outlined or proposed; rather they were revolutionary new ideas that ran counter to von Neumann's theory. In almost every instance, it was a repair of some inadequacy of the theory as outlined by von Neumann and Morgenstern, and indeed in the case of Nash's cooperative and general non-cooperative theory, von Neumann and Morgenstern criticized it publicly on several occasions. In the case of the extensive form, von Neumann claimed that it was impossible to give a practical geometric extensive form. All of the results that Aumann cited were obtained by members of the Mathematics Department at Princeton University. At the same time, the RAND Corporation, funded by the US Air Force, which was to be for many years the other major center of game-theoretic research, had just opened its doors in Santa Monica.

This suggests a second part of our answer to the question: "Why did it take so long for economists to recognize game theory as crucial to their held?" It is a historical fact that initially the main financial support for research in this area came from military agencies in the United States. Quoting Aumann again, "The major applications were to tactical military problems: defense from missiles, Colonel Blotto assignment problems, fighter-fighter duels, etc. Later the emphasis shifted to deterrence and cold war strategy with contributions by political scientists like Herman Kahn, Kissinger, and Schelling."

In any event, it was into this environment at Princeton of research ferment that the twenty-year old John Nash came in September of 1948. He came to the Mathematics Department with a one sentence letter of recommendation from R.L. Duffin of Carnegie Institute of Technology. This letter said, simply: "This man is a genius." As his thesis advisor, Professor A.W. Tucker was to write several years later: "At times I have thought this recommendation was extravagant, but the longer I've known Nash the more I am inclined to agree that Duffin was right." If we do the arithmetic of subtracting the date of Nash's arrival in Princeton, which was September 1948, from the date of submission by Solomon Lefschetz to the Proceedings of the National Academy of Sciences of the main result of Nash's thesis, November 1949, we find the results for which he is being honored this week were obtained in his first fourteen months of graduate study. It is a fine goal to set before the graduate students who are in the audience today. We shall return to the thesis later.

Although the speed with which Nash obtained these results is surprising, equally surprising and certainly less widely known is that Nash had already completed an important piece of work on bargaining while still an undergraduate at the Carnegie Institute of Technology. This work, a paper for an elective course in international economics, possibly the only formal course in economics he has ever had, was done in complete ignorance of the work of von Neumann and Morgenstern. In short, when he did this work he didn't know that game theory existed. This result, which is a model of theoretical elegance, posits four reasonable requirements or axioms: (1), that any solution should be invariant under positive linear affine transformations of the utility functions, (2), that the solution should be efficient in the sense of Pareto optimality, (3), that irrelevant alternatives should not change the outcome of the solution, and (4), that bargaining problems with symmetric outcome sets should have symmetric solutions. If these four reasonable conditions are satisfied then there is a unique solution, namely, the outcome that maximizes the product of the players' utilities. There is evidence in the published form of this paper, [37], that, before it appeared in Econometrica in 1950, he had met von Neumann and Morgenstern. This evidence is a reference to Cournot, Bowley, Tintner, and Fellner. It is almost certain that these were added at the suggestion of Morgenstern, because I don't think John has even read these papers as of now.

If it is clear that Nash had not read those writers, it is equally clear that this paper was written by a teenager. The evidence is that the objects in the example to be bargained over are a bat, a ball, a toy, and a knife. No amount of urging by his colleagues, or by the editor of Econometrica, persuaded John to change this example.

I should now like to discuss the thesis itself and show you some sections of John's work from the actual document. We already know that the main result, the definition of a Nash equilibrium, and a proof of existence had been completed prior to November 1949, the date of submission by Lefschetz to the National Academy of Sciences. The thesis itself was completed and submitted after the persistent urging and counsel of Professor Tucker. John always wanted to add more material, and Tucker had the wisdom to say "get the result out early." It was submitted and accepted by the Mathematics Department in May of 1950.

The formal rules at Princeton require that the thesis must be read by two professors, who prepare a report evaluating the work. In this case, the readers were Tucker and the statistician, John Tukey; the evaluation was written by Tucker himself. He wrote, "This is a highly original and important contribution to the Theory of Games. It develops notions and properties of "non-cooperative games," finite n-person games which are very interesting in themselves and which may open up many hitherto untouched problems that lie beyond the zero-sum two-person case. Both in conception and in execution this thesis is entirely the author's own."

In my discussion of the thesis itself, I shall try not to duplicate observations that will be made by later speakers. Some overlap is inevitable. For example, the abstract begins boldly: "This paper introduces the concept of a non-cooperative game and develops methods for the mathematical analysis of such games." Take careful note, there had been no general theory of non-cooperative games before this thesis. Although he was using the same strategic form as had been developed by von Neumann, the theory which occupies fully half of the von Neumann and Morgenstern book deals with cooperative theory envisaging coalitions, side-payments, and binding agreements. In addition, they proposed as a solution concept a notion we now call a "stable set", which need not exist for every game. By contrast, Nash proved by page 6 of his thesis that every n-person finite non-cooperative game has at least one (Nash) equilibrium point. This is a profile of mixed strategies, one for each player, which is such that no player can improve his payoff by changing his mixed strategy unilaterally.

The entire thesis is 27 pages of typescript, very generously double-spaced. Frankly, I have always considered the most important sections to be the first 6 pages summarized above and the last pages (from page 21 to 26) on motivation, interpretation, and applications. For many years, I have accused John of padding the thesis in the middle section (15 pages in all).

The two interpretations which form the essential motivation of work to be described by later speakers occur in the last pages of the thesis. On page 21, we find: "We shall now take up the mass action interpretation of equilibrium points." This interpretation will be discussed in detail by Selten and Weibull. The second interpretation is found on page 23, where we read: "We now sketch another interpretation . . . investigating the question: what would be a "rational" prediction of the behavior to be expected of rational playing the game in question." This interpretation will be discussed by van Damme. It is important to recognize that, although these very influential interpretations are explicitly in the thesis, they appear in no other publication by Nash.

To conclude my introduction to this seminar, I shall quote Aumann [3] again:

> "[The Nash] equilibrium is without doubt the single game theoretic solution concept that is most frequently applied in economics. Economic applications include oligopoly, entry and exit, market equilibrium, search, location, bargaining, product quality, auctions, insurance, principal-agent [problems], higher education, discrimination, public goods, what have you. On the political front, applications include voting, arms control and inspection, as well as most international political models (deterrence, etc.). Biological applications all deal with forms of strategic equilibrium; they suggest an interpretation of equilibrium quite different from the usual overt rationalism. We cannot even begin to survey all of this literature here."

It is now my pleasure to introduce an economist whom I have known since we were co-directors of a Summer Institute on Bargaining and Conflict in Princeton in 1962: John Harsanyi.

HARSANYI

In the short period of 1950 - 53, John Nash published four brilliant papers ([35], [37], [38], [39]), in which he made at least *three* fundamentally important contributions to game theory:

(1) He introduced the distinction between *cooperative* and *non-cooperative* games. The former are games in which the players can make *enforceable* agreements and can also make *irrevocable* threats to other players. That is to say, they can fully *commit* themselves to specific strategies. In contrast, in non-cooperative games, such self-commitment is not possible.1

(2) As a natural solution concept for non-cooperative games, he introduced the concept of *equilibrium points* ([35], [38]), now usually described as *Nash equilibria.* He also established their existence in all finite games.[2]

(3) As a solution concept for two-person cooperative games, he proposed the *Nash bargaining solution,* first for games with *fixed threats* [37], and later also for games with *variable threats* [39]. He also showed that, in the latter case, the two players' optimal strategies will have *maximin* and *minimax* properties.

The best way to understand the importance of Nash's contributions is by comparing the state of game theory just after publication of von Neumann and Morgenstern's book in 1944 with its state after publication of Nash's four papers in 1953.

von Neumann and Morgenstern's book contains an excellent mathematical analysis of one class of *non-cooperative* games, viz. of *two-person zero-sum* games and of the *minimax* solution for such games. It contains also an excellent mathematical discussion of one *cooperative* solution concept, that of *stable sets,* for many specific games.

Yet, it so happens that the concept of two-person zero-sum games has *very few* real-life applications outside of the military field. The concept of stable sets has *even fewer* empirical applications

Had these two distinguished authors had Nash's notions of cooperative and non-cooperative games available to them, then presumably they would have asked the question of how to act rationally in a two-person *nonzero-sum* game or in a *more-than-twoperson* game if this is played as a *non-cooperative* game, permitting no enforceable agreements and no irrevocable threats. Perhaps they would have asked also whether one could not find for *cooperative* games a more convincing solution concept than *stable sets* are. For instance, whether one could not find a solution concept yielding *sharper-predictions* about the players' actual payoffs than the concept of stable sets does.

Of course, in actual fact, they did not have these two notions available to them and therefore *did not* ask these two questions. But I merely want to point out how much our ability to ask important game theoretic questions was enhanced by Nash's work.

Nash's contributions described above under (1), (2), and (3) had an almost immediate effect on game-theoretic research. At first their effect was to encourage game theorists to develop the theories of cooperative games and of non-cooperative games as virtually *separate* disciplines, and for some time to devote much more effort to devise alternative *cooperative* solution concepts than to further development on *non-cooperative* game theory.

Then, in about the decade 1970 - 80, the focus of game theoretic research shifted once more. Interest in cooperative solution concepts decreased whereas interest in non-cooperative games and in non-cooperative-game models of cooperative games substantially increased.

This shift was due to a number of different factors. But one of these factors was what came to be known as *Nash's program* . One of Nash's papers ([38], p. 295) contains the following interesting passage:

> "The writer has developed a "dynamical" approach to the study of cooperative games based on reduction to non-cooperative form. One proceeds by constructing a model of the pre-play negotiation so that the steps of [this] negotiation become moves in a larger non-cooperative game...describing the total situation.
>
> This larger game is then treated in terms of the theory of this paper[3]...and if values are obtained [then] they are taken as the values of the cooperative game. Thus, the problem of analyzing a cooperative game becomes the problem of obtaining a suitable, and convincing, non-cooperative model for the negotiation."

When game theorists speak of "Nash's program," it is this two-paragraph passage they have in mind. That is to say, they are talking about the program of trying to reduce *cooperative* games to *non-cooperative* games by means of suitable non-cooperative models of the bargaining process among the players.

It is an interesting fact of intellectual history (if I am right in my reading of this history) that Nash's papers in the early 1950's at first encouraged game theorists to cultivate *cooperative* and *non-cooperative* game theory as largely independent disciplines, with a concentration on *cooperative* theory. But twenty-five years later they encouraged a shift to *non-cooperative* game theory and to *non-cooperative* models of the negotiations among the players.

Both Reinhard Selten and I were very pleased indeed when we learned that we received our Nobel Memorial Prizes in Economics together with John Nash. Not only do we both have the highest regard for his work, but our own work in game theory has been to an important extent based on his.

One of Reinhard's important contributions was his distinction between *perfect* and *imperfect* Nash equilibria. It was based on his realization that even strategy combinations fully satisfying Nash's definition of Nash equilibria might very well contain some *irrational* strategies. To exclude such imperfect Nash equilibria containing such irrational strategies, at first he proposed what now are called *subgame-perfect* equilibria (Selten, [45]). Later he proposed the even more demanding concept of *trembling-hand perfect* equilibria (Selten[4] [46]).

Reinhard's work on *evolutionarily stable strategies* was likewise based on the concept of Nash equilibria.

In my own case, an important part of my own work was likewise based on Nash's results. Thus, in my first game-theoretic paper [17], my main point was to show the *mathematical equivalence* of Nash's and of Zeuthen's bargaining models.

In the same paper (pp. 152 - 53), I pointed out an interesting corollary to Nash's theory of optimal threats: Suppose we measure the costs of a conflict to either party in terms of von Neumann-Morgenstern utilities. Suppose also that one bargainer makes a threat against the other. Then this will *strengthen* his own bargaining position only if carrying out his threat would increase the costs of a conflict for his opponent in a *higher proportion* than it would increase the costs of a conflict for him.

In a later paper [18], I extended the Shapley value to games *without* transferable utility and showed that my new solution concept was not only a generalization of the Shapley value, but also a direct generalization of Nash's two-person bargaining solution with variable threats.

A Nash equilibrium is defined as a strategy combination with the property that every player's strategy is a *best reply* to the other players' strategies. This of course is true also for Nash equilibria in *mixed* strategies. But in the latter case, besides his mixed *equilibrium strategy,* each player will also have infinitely many *alternative* strategies that are his *best replies* to the other players' strategies. This will make such equilibria potentially unstable.

In view of this fact, I felt it was desirable to show [20], that "almost all" Nash equilibria can be interpreted as strict equilibria in *pure* strategies of a suitably chosen game with randomly fluctuating payoff functions.

KUHN

In the early sixties, I had the great good fortune to hire both John Harsanyi and our next speaker as consultants to a project that I initiated for a research company in Princeton, called MATHEMATICA. The project was funded by the Arms Control and Disarmament Agency and a major topic was games with incomplete information. Our speaker has written about this experience in his autobiographical note [47]: Reinhard Selten.

SELTEN

Let me first tell you that this intervention has been prepared by Peter Hammerstein and myself. When John Nash published his basic papers on 'equilibrium points in n-person games' [35], and 'non-cooperative games' [38], nobody would have foretold the great impact of Nash equilibrium on economics and social science in general. It was even less expected that Nash's equilibrium point concept would ever have any significance for biological theory. To most game theorists it came as a complete surprise that beginning with the pioneering paper by Maynard Smith and Price [31] non-cooperative game theory, as it was founded by Nash, became one of the central tools for understanding the evolutionary logic of animal and plant interaction.

EVOLUTIONARY STABILITY

Maynard Smith and Price [35] introduced the concept of an evolutionarily stable strategy (ESS). Initially they were not aware of the relationship between the concept of an ESS and that of a Nash equilibrium. Rational game theory looked at mixed strategies as produced by conscious randomization. Nash's interpretation of a mixed equilibrium as a mass action phenomenon was buried in his unpublished dissertation and not found in textbooks on game theory. In biology the mass action interpretation is very natural and guided the work on evolutionary stability already from its beginning.

In their original paper, Maynard Smith and Price [35] restricted their attention to two-person games in normal form. They defined an ESS as a strategy prescribed by a symmetric equilibrium point and imposed on this strategy an additional stability requirement. This requirement had its roots in the idea that a population in evolutionary equilibrium should be stable against the invasion of mutants. There is no problem of instability if the mutant does not play a best reply to the symmetric equilibrium. However, if it plays an alternative best reply, i.e., a best reply different from the equilibrium strategy, it may spread by genetic drift. Maynard Smith and Price argued that this is excluded if against the alternative best reply the equilibrium strategy achieves a higher payoff than the alternative best reply itself does. This is the additional stability requirement in the definition of an ESS.

Nowadays it almost seems to be obvious that the correct application of Darwinism to problems of social interaction among animals requires the use of non-cooperative game theory, but when this idea was first conceived it was a revolutionary great insight. Of course the strategies of animals and plants are not the result of conscious deliberation. They are thought of as behavioral programs transferred by genetical inheritance from generation to generation. Game equilibrium is achieved by the process of natural selection which drives organisms towards the maximization of fitness. Roughly speaking, Darwinian fitness is the expected number of surviving offspring.

The original restriction to symmetric two-person games was soon removed and much more general definitions of an ESS were elaborated. The fruitfulness of game-theoretic thinking in biology is revealed by a multitude of intriguing applications. References to the theoretical and empirical literature can be found in our review paper (Hammerstein and Selten [16]).

THE CRISIS OF DARWINIAN ADAPTATION THEORY

In the early forties, biological thought on evolution reached a consensus often referred to as the 'new synthesis'. The apparent contradiction between Mendelian inheritance and gradual adaptation had been resolved by the population genetic work of Fisher [11], Haldane [14], and Wright [55]. Fisher's famous 'fundamental theorem of natural selection' had shown that under appropriate assumptions about the genetical system, the mean fitness of a population rises until a maximum is reached. However, in the sixties a new generation of population geneticists became aware of the fact that plausible genetic systems are very unlikely to satisfy the assumptions of Fisher's theorem. In the framework of a standard selection model, Moran [32] found examples in which mean fitness decreases over time until an equilibrium is reached. He looked at a two-locus model in which an evolving trait is coded for by two genes. Later Karlin [26] showed that these examples are not just degenerate cases.

The curious phenomenon of decreasing mean fitness becomes understandable if one looks at the nature of the resulting equilibrium. In this equilibrium one finds genotypes of high and low fitness but the offsprings of high fitness genotypes can have the same mean fitness as those of low fitness genotypes. This is an effect of recombination which tears genes at different loci apart. The phenomenon of decreasing mean fitness is a serious challenge to the theory of Darwinian adaptation. Some population geneticists came to the conclusion that the whole idea of fitness maximization has to be discarded as the main explanatory principle of biological evolution. The difficulties arise in the context of what is called 'frequency-independent selection' or, in other words, when there is no game interaction. In the presence of game interaction, the situation is even less favorable for the maximization of fitness. Of course, mean fitness is not maximized by game theoretic equilibrium, but Moran's and Karlin's results mean that game equilibrium is not necessarily reached.

By these developments, Darwinian adaptation theory was thrown into a true crisis. However, very few empirically oriented biologists were really disturbed by this problem. To them the crisis seemed to be one of mathematical theory rather than a failure of Darwinism as an explanation of biological facts. They continued to be impressed by the overwhelming empirical evidence for adaptation by natural selection. Nevertheless, the problem posed a serious challenge to theory.

THE STREETCAR THEORY

The process which generates the phenomenon of decreasing mean fitness governs the adjustment of genotype frequencies in the absence of mutations. Eshel and Feldman [10] were the first to ask the question under what conditions a stable equilibrium reached by this process is also stable against the invasion of mutants. As has been shown by Moran [32] and Karlin [26], the 'internal stability' with respect to the process without mutation does not necessarily lead to fitness maximization of game equilibrium. However, they succeeded to show that for an internally stable state in a sufficiently small vicinity of an ESS, the inflow of a destablizing mutation has a tendency to initially move the system in the direction of the ESS.

This opens the possibility that the notion of an ESS has more significance for the analysis of genetic systems than one might think if one looks only at internal stability and not also at external stability against the invasion of mutants. Admittedly, the results of Eshel and Feldman do not yet go very far in this direction but they were an ingenious step towards a new genetic interpretation of Darwinian adaptation. In the process of writing our review paper for the handbook of game theory (Hammerstein and Selten [16]) , we became intrigued by the possibility of providing a better foundation for the application of non-cooperative game theory in biology along the lines of Eshel and Feldman. We ended up in proving two theorems whose biological implications we like to describe by an analogy elaborated by one of us (Hammerstein [15]). The analogy involves a streetcar whose stops correspond to internally stable states. Only at the terminal stop the population state is phenotypically stable in the sense that the probabilities of pure strategies cannot be changed any more by the invasion of a mutant.

The first theorem shows that only a Nash equilibrium can be phenotypitally stable in a standard two-locus model of population genetics with game interaction. This means that in the long run the process of natural selection and mutation if it converges at all, must converge to Nash equilibrium. It therefore turns out that Nash equilibrium is of central importance for evolutionary biology. Of course, the streetcar may often stay for a while at a temporary stop at which some passengers exit and others enter before, finally, the terminal stop is reached at which it stays much longer.

The second theorem shows that a phenotypically monomorphic population state can be a terminal stop if and only if it is an ESS in the sense of Maynard Smith and Price [31]. Wherever one has reason to suppose that a trait is phenotypically monomorphic, this result establishes a firm foundation for the concept of an ESS. However, polymorphism is often observed in nature and in this respect Nash equilibrium is of more far reaching significance.

CONCLUDING REMARK

Originally, von Neumann and Morgenstern [52] developed game theory as a mathematical method especially adapted to economics and social science in general. In the introduction of their book, they emphasized their view that methods taken over from the natural sciences are inadequate for their purpose. They succeeded in creating a new method of mathematical analysis not borrowed from physics. In the case of game theory the flow of methodological innovation did not go in the usual direction from the natural to the social sciences but rather in the opposite one. The basis for this extremely successful transfer is the concept of Nash equilibrium.

KUHN
About five years ago, the Economics Department at Princeton University was fortunate to have the next speaker as a visiting professor. He has been in the forefront of recognizing the importance of Nash's mass action interpretation: Jörgen Weibull.

WEIBULL

THE MASS ACTION INTERPRETATION

In his unpublished Ph.D. dissertation, John Nash provided two interpretations of his equilibrium concept for non-cooperative games, one rationalistic and one population-statistic. In the first, which became the standard interpretation, one imagines that the game in question is played only once, that the participants are "rational," and that they know the full structure of the game. However, Nash comments: "It is quite strongly a rationalistic and idealizing interpretation" ([36], p. 23). The second interpretation, which Nash calls the *mass-action* interpretation, was until recently largely unknown (Leonard [28], Weibull [53], Björnerstedt and Weibull [6]). Here Nash imagines that the game in question is played over and over again by participants who are not necessarily "rational" and who need not know the structure of the game:

> "It is unnecessary to assume that the participants have full knowledge of the total structure of the game, or the ability and inclination to go through any complex reasoning processes. But the participants are supposed to accumulate empirical information on the relative advantages of the various pure strategies at their disposal.
>
> To be more detailed, we assume that there is a population (in the sense of statistics) of participants for each position of the game. Let us also assume that the 'average playing' of the game involves *n* participants selected at random from the *n* populations, and that there is a stable average frequency with which each pure strategy is employed by the 'average member' of the appropriate population.
>
> Since there is to be no collaboration between individuals playing in different positions of the game, the probability that a particular n-tuple of pure strategies will be employed in a playing of the game should be the product of the probabilities indicating the chance of each of the *n* pure strategies to be employed in a random playing" ([36], pp. 21 - 22.)

Nash notes that if s_i is a population distribution over the pure strategies $\alpha \in A_i$ available to the i'th player position, then $s = (s_i)_{i \in I}$ is formally identical with a mixed strategy profile, and the expected payoff to any pure strategy α in a random matching between an n-tuple of individuals, one from each player population, is identical with the expected payoff $\pi_{i\alpha}(s)$ to this strategy when played against the mixed strategy profile s:

> "Now let us consider what effects the experience of the participants will produce. To assume, as we did, that they accumulate empirical evidence on the pure strategies at their disposal is to assume that those playing in position i learn the numbers $\pi_{i\alpha}(s)$. But if they know these they will employ only optimal pure strategies [...]. Consequently, since s_i expresses their behavior, s_i attaches positive coefficients only to optimal pure strategies, [...]. But this is simply a condition for s to be an equilibrium point.
>
> Thus the assumption we made in this 'mass-action' interpretation lead to the conclusion that the mixed strategies representing the average behavior in each of the populations form an equilibrium point." (op cit., p. 22)[5]

These remarks suggest that Nash equilibria could be identified as stationary, or perhaps dynamically stable, population states in dynamic models of boundedly rational strategy adaptation in large strategically interacting populations. In spirit, this interpretation is not far from Friedman's [13] subsequent "as if" defense of microeconomic axioms. For just as Nash argued that boundedly rational players will adapt toward strategic optimality, Friedman argued that only profit maximizing firms will survive in the long run under (non-strategic) market competition. Moreover, the view that games are played over and over again by individuals who are randomly drawn from large populations was later independently taken up by evolutionary biologists (Maynard Smith and Price [31], Taylor and Jonker [49]).

NOTATION AND PRELIMINARIES

Consider a finite n-player game G in normal (or strategic) form. Let A_i be the pure-strategy set of player position $i \in I = \{1, ..., n\}$, S_i its mixed-strategy simplex, and $S = \Pi_{i \in I} S_i$ the polyhedron of mixed-strategy profiles. For any player position i, pure strategy $\alpha \in A_i$ and mixed strategy $s_i \in S_i$, let $s_{i\alpha}$ denote the probability assigned to α. A strategy profile s is called *interior* if *all* pure strategies are used with positive probability. The expected payoff to player position i when a profile $s \in S$ is played will be denoted $\pi_i(s)$, while $\pi_{i\alpha}(s)$ denotes the payoff to player i when he uses pure strategy $\alpha \in A_i$ against the profile

$s \in S$. A strategy profile $s \in S$ is a *Nash equilibrium* if and only if $s_{i\alpha} > 0$ implies $\pi_{i\alpha}(s) = \max_{\beta \in A_i} \pi_{i\beta}(s)$

In the spirit of the "mass-action" interpretation, imagine that the game is played over and over again by individuals who are randomly drawn from (infinitely) large populations, one population for each player position i in the game. A *population state* is then formally identical with a mixed-strategy profile $s \in S$, but now each component $s_i \in S_i$ represents the distribution of pure strategies in player population i, i.e., $s_{i\alpha}$ is the probability that a randomly selected individual in population i will use pure strategy $\alpha \in A$, when drawn to play the game. In this interpretation $\pi_{i\alpha}(s)$ is the (expected) payoff to an individual in player population i who uses pure strategy α – an "a-strategist" – and $\pi_i(s)\, \Sigma_\beta s_{i\beta} \pi_{i\beta}(s)$ is the average (expected) payoff in player population i, both quantities being evaluated in population state s.

Suppose that every now and then, say, according to a statistically independent Poisson process, each individual reviews her strategy choice. By the law of large numbers the aggregate process of strategy adaptation may then be approximated by deterministic flows, and these may be described in terms of ordinary differential equations.

INNOVATIVE ADAPTATION

We first consider the case when strategy adaptation is memory-less in the sense that the time rate of strategy revision and the choice probabilities of strategy-reviewing individuals are functions of the current state s (only):

$$\dot{s}_{ia}(t) = f_{i\alpha}[s(t)] \tag{1}$$

for some functions $f_{ia} : S \rightarrow R$. The quantity f_{ia} (s) thus represents the net increase per time unit of the population share of a-strategists in player population i when the overall population state is s. The (composite) function f is assumed to be Lipschitz continuous and such that all solution trajectories starting in S remain forever in S. Such a function f will be called a *vector field* for (1).

The class of population dynamics (1) clearly allows for an innovative element; some individuals may begin using earlier unused strategies, either intentionally, by way of experimentation or calculation, or unintentionally, by way of mistakes or mutations. Indeed, a certain degree of inventiveness in this sense is easily seen to imply that only those population states that constitute Nash equilibria can be stationary.[6] The requirement is simple: if there is some (used or unused) pure strategy which results in a payoff above the current average payoff in the player population in question, then some such pure strategy will grow in population share. Formally, for any population state $s \in S$ and player position $i \in I$, let $B_i(s)$ denote the (possibly empty) subset of better-than-average pure strategies, $B_i(s) = \{\alpha \in A_i : \pi_{ia}(s) > \pi_i(s)\}$.

Inventiveness can then be formalized as

$$[\mathbf{IN}]: \text{ If } B_i(s) \neq \emptyset, \text{ then } f_{ia}(s) > 0 \text{ for some } \alpha \in B_i(s).$$

This condition is, for instance, met if reviewing individuals move toward the best replies to the current population state. Note that [*IN*] requires no knowledge about payoffs to other player positions, nor is any detailed knowledge of the payoffs to one's own strategy set necessary. It is sufficient that individuals on average tend to twitch toward *some* of the better-than-average performing strategies.

Proposition 1 *Suppose f meets [IN]. If a population state s is stationary under the associated dynamics (1), then s constitutes a Nash equilibrium of G.*

An example of innovative adaptation is given by

$$f_{ia}^{+}(s) = \pi^{+}_{ia}(s) - s_{ia}\sum_{\beta\in A_i}\pi^{+}_{i\beta}(s) \tag{2}$$

where $\pi^{+}_{ia}(s) = \max\{\pi_{i\alpha}(s - \pi_i(s), 0\}$; the excess payoff to pure strategy α over the average payoff in its player population. It is not difficult to verify that f^+ meets [*IN*]. The associated population dynamics (1) is nothing else than the continuous-time version of the iteration mapping introduced in Nash's [38] influential existence proof for equilibrium points – later adopted in general equilibrium theory, see Arrow and Debreu [1].[7]

In order to incorporate memory in the dynamic process of strategy adaptation, one may introduce real variables $P_{i\alpha}$, one for each player position i and pure strategy $\alpha \in A$, that represent the i'th population's recollection of earlier payoffs to pure strategy a. Assume that the recalled payoff to any pure strategy $\alpha \in A$, changes with time according to

$$P_{i\alpha}(t) = h_{i\alpha}(\pi_{i\alpha}[s(t)], p_{i\alpha}(t), t), \tag{3}$$

where $h_{i\alpha}$ is a Lipschitz continuous function such that the recalled payoff changes only if the current payoff differs from the recalled payoff $(h_{i\alpha}(\pi_{i\alpha}, p_{ia}, t) = 0 \Rightarrow \pi_{i\alpha} = p_{i\alpha})$.

The full adaptation dynamics with memory is then a system of differential equations in the state vector x = (s, p) , where p moves according to (3) and s according to

$$\dot{s}_{i\alpha}(t) = f_{i\alpha}[s(t), p(t)] \tag{4}$$

A counterpart to the earlier requirement [*IN*] of inventiveness is: if there is some (used or unused) pure strategy which is *recalled* to result in a payoff above the average of the currently recalled payoffs in the player population in question, then some such pure strategy will increase its population share. Formally, for any state (s, p) and player position $i \in I$, let $B_i(s, p) = \{\alpha \in A_i : p_{i\alpha} > \Sigma_{\beta\in A_i} s_{i\beta} p_{i\beta}\}$. Inventiveness can then be formalized as

[**IN'**]: If $B_i(s, p), \neq \emptyset$ then $f_{i\alpha}(s, p) > 0$ for some $\alpha \in B_i(s, p)$.

The following extension of proposition 1 is straightforward:

Proposition 2 *Suppose f meets [IN']. If* (s, p) is stationary under (3) and (4), then s is a Nash equilibrium of G.[8]

A special case of memory processes of the above type is when the recalled payoff to a pure strategy is the *time average* of its earlier payoffs: $p_{i\alpha}(t) = \int^t_0 \pi_{i\alpha}[s(\tau)]d\tau$. This is the memory process in fictitious play (Robinson [40]), written in continuous time. (In that model players always use best replies against the time average of past play.) Differentiation of $p_{i\alpha}(t)$ with respect to t gives

$$h_{i\alpha}(\pi_{\iota\alpha}, p_{i\alpha}, t) = \frac{1}{t}(\pi_{i\alpha} - p_{i\alpha}), \tag{5}$$

a function which can be slightly modified so that it belongs to the above-discussed class (add an arbitrarily short time interval before $t = 0$).

IMITATIVE ADAPTATION

It may be argued that the above classes of population dynamics go somewhat beyond the spirit of the mass-action interpretation since they presume that individuals perform a certain amount of calculations. Therefore, now assume no memory and no inventiveness as defined above. Thus, individuals now switch only between strategies already in use, and they do so only on the basis of these strategies' current performance. Technically, this means that the population dynamics (1) has a vector field f of the form

$$f_{i\alpha}(s) = g_{i\alpha}(s)\, s_{i\alpha}. \tag{6}$$

The involved functions $g_{i\alpha}$ will be called *growth-rate functions*—$g_{i\alpha}(s)$ being the growth rate of the population share of pure strategy α in player population i when the population state is s. No vector field of the form (6) is innovative in the sense of condition [IN], because if all individuals in a player population initially use only one (or a few) pure strategy then they will continue doing so forever, irrespective of whether some unused strategy yields a high payoff or not. Consequently, stationarity does *not* imply Nash equilibrium for the present class of dynamics, which will be called *imitative.*

A prime example of such dynamics is the so-called replicator dynamics used in evolutionary biology (Taylor and Jonker [49], Taylor [48]). In this strand of literature, pure strategies represent genetically programmed behaviors, reproduction is asexual, each offspring inherits its parent's strategy, and payoffs represent reproductive fitness. Thus $\pi_{i\alpha}(s)$ is the number of (surviving) offspring to an α-strategist in population i, and $\pi_i(s)$ is the average number of (surviving) offspring per individual in the same population. In the standard version of this population model, each pure strategy's growth rate is proportional to its current payoff:[9]

$$g_{i\alpha}(s) = \pi_{i\alpha}(s) - \pi i(s). \tag{7}$$

We will here consider a broad class of vector fields which contains the replicator vector field as a special case. The defining requirement is close in spirit to that in the previous section: If there exists a pure strategy which results in a payoff above average in its player population (whether this pure strategy be currently used or not), then some such pure strategy has a positive *growth rate.* Hence, if all such strategies are present in the population, then some such population share will grow. Formally:

[**POS**]: If $B_i(s) \neq \emptyset$, then $g_{i\alpha}(s) > 0$ for some $\alpha \in B_i(s)$.

The next proposition establishes the following implications under payoff positive imitation: (a) if all strategies are present in a stationary population state, then this constitutes a Nash equilibrium, (b) A dynamically stable population state constitutes a Nash equilibrium, (c) If a dynamic solution trajectory starts from a population state in which all pure strategies are present and the trajectory converges over time, then the limit state is a Nash equilibrium.[10] Claim (b) is a generalization of a result due to Bomze [7] for the single-population version of the replicator dynamics as applied to symmetric two-player games, and (c) generalizes a result due to Nachbar [34]. (See Weibull [54] for a proof.)

Proposition 3 *Suppose g meets [POS], and consider the associated population dynamics (1) where f is defined in (6).*

(a) If s is interior and stationary, then s is a Nash equilibrium.

(b) If s is dynamically stable, then s is a Nash equilibrium.

(c) If s is the limit of some interior solution trajectory, then s is a Nash equilibrium.

Note that claims (a) and (c) involve hypotheses that no pure strategies are extinct. Indeed, these claims are otherwise not generally valid. Implication (b), however, allows for the possibility that some pure strategy is extinct. This is permitted because dynamic stability by definition asks what happens when the population state is slightly perturbed - in particular, when currently extinct strategies enter the population in small population shares.

CONCLUSION

The mass-action interpretation of Nash equilibria is in stark contrast with the usual rationalistic interpretation, but is closely related to ideas in evolutionary game theory. It opens new avenues for equilibrium and stability analysis of social and economic processes, and suggests new ways to combine insights in the social and behavior sciences with economic theory.

KUHN
In Reinhard Selten's talk, he did not mention his major discovery of two refinements of the concept of Nash equilibria, the so-called subgame perfect equilibria [45] and trembling-hand perfect equilibria [46]. A large body of research followed these discoveries; it has been summarized in a magnificent manner in a book [50] by our next speaker: Eric van Damme.

VAN DAMME
Ideas, concepts and tools that were introduced by John Nash [36] have been extremely important in shaping modern economic theory. He introduced the fundamental solution concept for non-cooperative games, one of the main solution concepts for cooperative games and he proposed the Nash program for providing non-cooperative foundations of cooperative concepts. In his analysis he introduced seminal tools and techniques that served as essential building blocks in the later development of the theory and that contributed to its successful application. Below we provide a perspective on Nash's work and trace its influence on modern economic theory.

NASH EQUILIBRIUM: THE RATIONALISTIC INTERPRETATION

A non-cooperative game is given by a set of players, each having a set of strategies and a payoff function. A strategy vector is a *Nash equilibrium* if each player's strategy maximizes his pay-off if the strategies of the others are held fixed. In his Ph.D. thesis, Nash introduces this concept and he derives several properties of it, the most important one being existence of at least one equilibrium for every finite game. In published work ([35], [38]), Nash provides two alternative, elegant proofs, one based on Kakutani's fixed point theorem, the other based directly on Brouwer's theorem. These techniques have inspired many other existence proofs, for example, in the area of general equilibrium theory (see [9]).

In the section "Motivation and Interpretation" of his thesis, Nash discusses two interpretations of his equilibrium concept. The first, "mass-action" interpretation is discussed in Jorgen Weibull's contribution to this seminar. Here, we restrict ourselves to the "rationalistic and idealizing interpretation" which is applicable to a game played just once, but which requires that the players are rational and know the full structure of the game. Nash's motivation runs as follows:

> "We proceed by investigating the question: What would be a "rational" prediction of the behavior to be expected of rational playing the game in question? By using the principles that a rational prediction should be unique, that the players should be able to deduce and make use of it, and that such knowledge on the part of each player of what to expect the others to do should not lead him to act out of

> conformity with the prediction, one is led to the concept of a solution defined before.” [36]

In other words, a theory of rational behavior has to prescribe the play of a Nash equilibrium since otherwise the theory is self-destroying. Note that the argument invokes three assumptions: (i) players actively randomize in choosing their actions, (ii) players know the game and the solution, and (iii) the solution is unique. Later work has scrutinized and clarified the role of each of these assumptions. Harsanyi ([20]) showed that a mixed strategy of one player can be interpreted as the beliefs (conjectures) of the other players concerning his behavior. This reinterpretation provides a “Bayesian” foundation for mixed strategy equilibria and eliminates the intuitive difficulties associated with them. Aunrann developed the concepts of an interactive belief system, which provides a formal framework for addressing the epistemic conditions underlying Nash equilibrium, i.e., it allows one to formalize players’ knowledge and to investigate how much knowledge is needed to justify Nash equilibrium. In 2-player games less stringent conditions are sufficient than in general n-player games. (Aumann and Brandenberger [4])

Since the rationalistic justification of equilibria relies on uniqueness, multiplicity of equilibria is problematic. Nash remarks that it sometimes happens that good heuristic reasons can be found for narrowing down the set of equilibria. One simple example that Nash provides (Ex. 5 [38]) is the game that is reproduced here in Figure 1. This game has equilibria at (a, α) and (b, β), as well as a mixed equilibrium. Nash writes that “empirical tests show a tendency toward (a,α),” but he does not provide further details. One heuristic argument is that (a, α) is less risky than (b, β), an argument that is formalized by Harsanyi and Selten’s [23] concept of risk dominance. It figures prominently both in the literature that builds on the “rationalistic interpretation” as well as in the literature that builds on the “mass-action” interpretation of Nash equilibrium. We will return to it in EQUILIBRIUM SELECTION.

	α	β
a	1,2	-1, -4
b	-4, -1	2,1

Fig. 1

THE NASH PROGRAM

It can be said that the "rationalistic argument" leading to Nash equilibrium was already discussed in von Neumann and Morgenstern [52] cf. their "indirect argument" (pp. 147 - 148). They advocate the (equilibrium) solution implied for the 2-person zero-sum case, however, they argue that this solution is not satisfactory for games outside this class, since in these one cannot neglect coalitions nor the possibility that players will want to make payments outside the formal rules of the game (p. 44). They argue that for these games "there seems no escape from the necessity of considering agreements concluded outside the game" (p. 223) and they see themselves forced to assume that coalitions and agreements concluded outside of the game are respected by the contracting parties (p. 224). Hence, they end up with having two distinct theories.

Nash proposes to distinguish between cooperative and non-cooperative games. In games of the latter type, players are unable to conclude enforceable agreements outside the formal rules of the game. Cooperative games allow such agreements. Nash suggests that non-cooperative games are more basic, that cooperative games may fruitfully be analyzed by reformulating them as non-cooperative ones and by solving for the Nash equilibria. This approach has come to be known as the *Nash program*. It allows unification of the theory and enables a better understanding of the different solution concepts that have been proposed in cooperative theory. (See Harsanyi [21] for an example dealing with the von Neumann and Morgenstern stable set concept; also see BARGAINING THEORY.) By following the Nash program, an abstract discussion about the "reasonableness" of certain outcomes or principles can be replaced by a more mundane discussion about the appropriateness of the rules of the game.

The non-cooperative approach also has its limitations. First, values can be obtained only if the game has a unique solution, i.e., one has to address the equilibrium selection problem. Secondly, the non-cooperative model must at the same time be relevant - similar to reality in its essential aspects, and mathematically tractable. Consequently, the axiomatic approach - which aims to derive the outcome directly from a set of 'convincing' principles, is not redundant. On the contrary, if a solution can be obtained from a convincing set of axioms, this indicates that the solution is appropriate for a wider variety of situations than those captured by the specific non-cooperative model. As Nash concludes "The two approaches to the problem, via the negotiation model or via the axioms, are complementary; each helps to justify and clarify the other." ([39], p. 129)

BARGAINING THEORY

According to orthodox economic theory, the bargaining problem is indeterminate: The division of the surplus will depend on the parties' bargaining skills. Nash breaks radically with this tradition. He assumes that bargaining between rational players leads to a unique outcome and he seeks to determine it. He solves the problem in the 2-person case and he derives his solution both by means of the axiomatic approach and as the outcome of a non-cooperative model.

In Nash [39] the axiomatic method is described in the following way:

> "One states as axioms several properties that it would seem natural for the solution to have and then one discovers that the axioms actually determine the solution uniquely." ([39], p. 129)

In the case of the fixed-threats, Nash's basic axioms are that rational players are characterized by von Neumann Morgenstern utility functions, and that the bargaining situation is fully represented by its representation, Bin utility space. Three axioms specify the relation which should hold between the solution and the set B: (i) Pareto efficiency, (ii) symmetry and (iii) independence of irrelevant alternatives. These axioms determine the solution to be that point on the north-east boundary of *B* where the product of the utilities, u_1u_2, is maximized.

Axion (iii) states that, if the set of feasible utility pairs shrinks but the solution remains available, then this should remain the solution. It is more difficult to defend than the others and there has been considerable discussion of it in the literature. Nash writes that it is equivalent to an axiom of "localization", specifically "Thinking in terms of bargaining, it is as if a proposed deal is to compete with small modifications of itself and that ultimately the negotiation will be understood to be restricted to a narrow range of alternative deals and to be unconcerned with more remote alternatives." ([39], p 139.) Recent developments in non-cooperative bargaining theory (which build on the seminal paper [42]) have confirmed this interpretation. Namely, assume players alternate in proposing points from *B* until agreement is reached. Assume that if an offer is rejected there is a small but positive probability that negotiations break down irrevocably. This game admits a unique subgame perfect Nash equilibrium (see EQUILIBRIUM REFINEMENT) and agreement is reached immediately. The equilibrium condition states that each time each responder is indifferent between accepting the current proposal and rejecting it. Consequently, the equilibrium proposals are close together when the stopping probability is small, hence, we obtain "localization" property. Indeed the equilibrium conditions imply that both equilibrium proposals have the same Nash product, hence, since they have the same limit, they converge to the Nash solution.

Of course, it is gratifying to find that this natural bargaining model implements Nash's bargaining solution. However, even more important is that this application of the Nash program may clarify some ambiguities concerning the choice of the threat point in applications of the Nash bargaining model. (See [5] for further discussion.)

In the variable threat case, each party has a choice how much pressure to put on the other. The theory now has to determine both the threats that the players will use in case they don't agree, as well as the agreement that will be concluded. Two additional axioms allow reduction of the problem to the case with fixed threats and, hence, determine the theory. The first is equivalent to assuming that each player has an optimal threat strategy, i.e., it is postulated that the problem admits a solution. The second says that a player cannot improve his payoff by eliminating some of his strategies. In the non-cooperative approach, Nash assumes that players first commit themselves to their threats. Players will be forced to execute their threats if they cannot come to an agreement in the second stage. Each pair of threats induces a (fixed-threat bargaining) subgame in which the distinguished equilibrium that maximizes the product of the utility gains is selected. Applying backwards induction and using this selection (i.e., by replacing each subgame with its unique solution), the choice of threat strategy in the first stage essentially reduces to a strictly competitive game, i.e., this reduced first stage game has an equilibrium with minmax properties. Consequently, the overall game has a value and optimal threat strategies. Needless to say, the solution obtained by the non-cooperative approach coincides with that obtained by means of the axioms.

EQUILIBRIUM REFINEMENT

Nash equilibrium expresses the requirement that a theory of rational behavior recommends to each player a strategy that is optimal in case all of the other players play in accordance with the theory. It imposes no conditions concerning behavior after a deviation from the theory has occurred. von Neumann and Morgenstern, however, already argued that a solution concept should also address the question of how to play when the others do not conform and that, in the presence of "non-conformists", a believer in the theory should still be willing to follow the theory's advice. It turns out that not all Nash equilibria have this desirable property: After a deviation from the equilibrium has occurred, a believer in this equilibrium may prefer to deviate from it as well. As an example, modify the game from Figure 1 such that player I (the "row player") makes his choice before player 2, with this choice being revealed before the latter makes his decision. The strategy pair in which player 1 chooses *a* and player 2 responds with α no matter what player 1 chooses, is a Nash equilibrium. However, the equilibrium action of player 2 is not optimal if player I deviates and chooses *b*: In that case player 2 prefers to deviate to β. The equilibrium relies on a non-credible threat of-player 2.

A clear discussion of the credibility issue can already be found in Nash's work on variable threat bargaining. Nash's paper is appropriately called "Two-person Cooperative Games" since it relies essentially on the existence of an umpire who enforces contracts and commitments, Nash writes "it is essential for the success of the threat that *A* be *compelled* to carry out his threat if *B* fails to comply. Otherwise it will have little meaning. For, in general, to execute the threat will not be something *A* would want to do, just of itself." (Nash [39], p. 130).

To eliminate equilibria that rely on non-credible threats, various refinements of the Nash equilibrium concept have been proposed, which will not be surveyed here see [50]. Let us just note that two papers of Reinhard Selten were fundamental. Selten [45] argues that a theory of rational behavior has to prescribe an equilibrium in every subgame since otherwise at least one player would have an incentive to deviate once the subgame is reached. He calls equilibria having this property subgame perfect. They can be found by a backwards induction procedure. Unfortunately, this procedure generally does not eliminate all "questionable" equilibria. Selten [46] suggests a further refinement that takes the possibility of irrational behavior explicitly into account, i.e., he suggests viewing perfect rationality as a limiting case of incomplete rationality. Formally, he considers slightly perturbed versions of the original game in which players with small probabilities make mistakes and he defines a (trembling hand) perfect equilibrium as one that is a limit of equilibrium points of perturbed games. It is interesting to note that Nash already discussed a game with an imperfect equilibrium (see Ex. 6 in [39]).

This suggestion to discriminate between equilibria by studying their relative stabilities had already been made in Nash's work on bargaining (see EQUILIBRIUM SELECTION). An important difference between Selten's approach and that of Nash, however, is that Selten requires stability only with respect to *some* perturbation, while Nash insisted on stability against all perturbations in a certain class. Consequently, a game typically allows multiple perfect equilibria. Kohlberg and Mertens *[27]* have argued that Selten's perfectness requirement is not restrictive enough and they have proposed various refinements that require stability, of sets of equilibria, with respect to all perturbations in a certain class. At present, the debate is still going on of whether these strong stability requirements indeed capture necessary requirements of rational behavior. What can be said, however, that Nash's ideas were fundamental in shaping this research program.

EQUILIBRIUM SELECTION

We have already argued that, since the rationalistic interpretation of Nash equilibrium relies essentially on the uniqueness assumption, the fact that a game frequently has multiple equilibria makes the equilibrium selection problem prominent. Nash already encountered this problem and in his study of the fixed-threat bargaining problem. In Nash's non-cooperative model both players simultaneously state their demands and if the pair of demands is feasible then each player gets his demand; otherwise disagreement results. Clearly, any pair of demands that is Pareto efficient constitutes a pure equilibrium of the game. The following quote describes the multiplicity problem as well as Nash's solution of it

> "Thus the equilibrium points do not lead us immediately to a solution of the game. But if we discriminate between them by studying their relative stabilities we can escape from this troublesome non-uniqueness. To do this we "smooth" the game to obtain a continuous pay-off function and then study the limiting behavior of the equilibrium points of the smoothed game as the amount of smoothing approaches zero." ([39], pp. 131 - 132)

The smoothed game is determined by a continuous, strictly positive, function h, where $h(d)$ can be interpreted as a probability that the demand vector d is compatible. (It is assumed that $h(d) = 1$ if d is feasible in the unperturbed problem, i.e., $d \in B$, and that h tapers off very rapidly towards zero as d moves away from B.) The smoothed game, in which players i's payoff function is $u_i^h(d) = dh_i(d)$ can be thought of as representing uncertainties in the information structure of the game, the utility scales, etc. ([39], p. 132). Any maximizer of the function $d_1d_2h(d)$ is an equilibrium of this perturbed game and all these maximizers converge to the unique maximizer of the function u_1u_2 on B as the noise vanishes. Furthermore, if h varies regularly, the perturbed game will have the unique maximizer of $d_1d_2h(d)$ as its unique equilibrium. It follows that the Nash bargaining solution is the unique necessary limit of the equilibrium points of the smoothed games. Consequently, the original game has only one "robust" equilibrium, which may be taken as the solution of the game.

Building on Nash's ideas, and motivated by the attempt to generalize Nash's bargaining solution to games with incomplete information, Harsanyi and Selten [22] construct a coherent theory of equilibrium selection for general games. A crucial concept in this theory is that of risk dominance, and the influence of Nash's ideas on the theory is demonstrated by the following quote:

> "Our attempts to define risk dominance in a satisfactory way have been guided by the idea that it is desirable to reproduce the result of Nash's cooperative bargaining theory with fixed threats. The Nash-property is not an unintended by-product of our theory." ([22], p. 215)

The Nash-property that is referred to in this quote is the property that in certain classes of games (such as unanimity games and P-person 2 x 2 games) the selected equilibrium is the one for which the product of the losses associated with deviating from the equilibrium is largest. For example, in the game of Figure 1, the equilibrium (a, a) has a Nash product of 30, while the Nash-product of (b, β) is 6. Hence, (a, α) risk-dominates (b, β) For the special case of 2 x 2 games, Harsanyi and Selten derive the risk-dominance relation from a convincing set of axioms that resembles those with which Nash justifies his bargaining solution. For more general games, risk dominance cannot be based on a simple comparison of Nash products and it is not clear that Harsanyi and Selten's definition (which is based on the tracing procedure and which will not be given here) is the most appropriate one. Carlsson and van Damme [8] compare several concepts that all derive their inspiration from Nash's work and that coincide with Nash's solution for 2×2 games, but that yield different outcomes outside of this class. In any case it is clear that Nash's ideas figure prominently in the theory of equilibrium selection.

EXPERIMENTAL GAMES

In the previous sections, we have documented the influence of Nash's ideas on the development of normative, rationalistic game theory. This paper would be incomplete if it would not also mention the pioneering work of Nash, together with Kalisch, Milnor and Nering [25] in experimental economics. That paper reports on a series of experiments concerning *n*-person games and aims to compare various theoretical solution concepts with the results of actual play, i.e., it deals with the behavioral relevance of the rationalistic theory. The authors find mixed support for various theoretical solution concepts and they discuss several reasons for the discrepancy between theoretical and empirical results. Among others, the role of personality differences, the fact that utility need not be linear in money and the importance of apparent fairness considerations are mentioned. In addition, several regularities are documented, such as framing effects, the influence of the number of players on the competitiveness of play, the fact that repetition of the game may lead to more cooperative play, and the possibility of inducing a more competitive environment by using stooges. As documented by the importance of the above mentioned concepts in current experimental economics, the paper is an important milestone in the development of descriptive game theory. (See [41]).

A second important contribution of Nash to the experimental economics literature is his discussion of the repeated prisoners' dilemma experiment conducted by Melvin Dresher and Merrill Flood. In this experiment, two players played 100 repetitions of a prisoners' dilemma. They did not constantly play the one-shot equilibrium, but they also did not succeed in reaching an efficient outcome either. The experimenters view their experiment as a test of the predictive relevance of the one-shot equilibrium and they interpret the evidence as refuting this hypothesis. Nash, however, argues that the experimental design is flawed, that the repeated game cannot be thought of as a sequence of independent games and he suggests that the results would have been very different if the interaction between the trials had been removed. He concedes that constant repetition of the static equilibrium is the unique equilibrium of the overall game, but he argues that a pair of trigger strategies ("Cooperate as long as the other Cooperates, Defect forever as soon as the other has Defected once") is nearly in equilibrium and that this pair is an exact equilibrium in the infinitely repeated game. Furthermore, he suggests that the situation might actually be better represented by the latter game "since 100 trials are so long that the Hangman's Paradox cannot possibly be well reasoned through on it". (Nash in [12]) Hence, he not only specifies an appropriate design for testing static equilibrium predictions; he also describes the essential insight in the theory of repeated games and he points to a specific form of bounded rationality as an explanation for observed discrepancies between theoretical predictions and empirical outcomes.

CONCLUSION

Aumann [2] has forcefully argued that a game theoretic solution concept should be judged primarily by the insights that it yields in applications, by "its success in establishing relationships and providing insights into the workings of the social processes to which it is applied" (pp. 28 - 29). On this score, "Nash equilibrium is without a doubt the most "successful" - i.e., widely used and applied - solution concept of game theory" (p. 48). Indeed, much of the modern literature in economics (and related disciplines) takes the following form: A social situation is modeled as a non-cooperative game, the Nash equilibria of the game are computed and its properties are translated into insights into the original problem.

The Nash bargaining solution can also be considered a very successful solution concept since it has also been applied frequently. Of course, its scope is much more limited than that of the equilibrium concepts. Furthermore, because of its more abstract nature, it is associated with ambiguities, which might inhibit successful applications. Such ambiguities may be resolved by application of the Nash program, i.e., by making explicit the bargaining process by means of which agreements are reached and by solving the resulting game for its equilibria.

The problems associated with multiplicity of equilibria and with the fact that not all equilibria need correspond to rational behavior, have hampered successful application of the Nash program. Nash resolved these difficulties for the special case of 2-person bargaining games. Inspired by his ideas and building on his techniques an important literature dealing with these issues has been developed, which enables the analysis and solution of more complicated, more realistic games. Hence, the domain of applicability of non-cooperative game theory has been extended considerably. It is to be expected that, as a result, our insights into the workings of society will be enhanced as well.

KUHN

I would now like to open the floor to questions for any of the participants including John Nash. I shall invoke a privilege of the chair to pose a question to him. Why did you not publish the interpretations which are in your thesis when it was made the basis of the article [38], in the Annals of Mathematics?

NASH
I am afraid I can't, simply can't answer that question. I really don't remember, but I do know that Tucker changed the form of what was to be my thesis and that the paper "Two person cooperative games", which might have been incorporated originally, if it had been allowed to, was not included. So that became a separate publication in Econometrica, differentiated from the rest of it, while that which could be presented more as a mathematical paper went into the Annals of Mathematics. So I don't know whether it was just pruned down in style for the Annals of Mathematics.

KUHN
It is certainly the case that the Annals of Mathematics has different standards than economics journals, and it may well have pruned down by an editor or a reviewer there, but I think it is a great shame, because the delay in recognizing these interpretations has been marked, I know that Jörgen Weibull was especially prominent in bringing forward the mass interpretation, and I think Eric has shown today that the reexamination, having the thesis available, has been very fruitful for a number of people. The meeting is now open to questions from anyone.

WERNER GÜTH (Humboldt University of Berlin)
I just want to ask one question, because I think John Nash proved a generic result by showing that for every finite game there exists a Nash equilibrium point. I found the assumptions that there are only finitely many strategies very intuitive, very natural, but of course to prove it you have to assume that you can vary the mixed strategies continuously. And if I now think that having only finitely many actions available is very natural. I also have to assume that only finitely many options in randomizing are available. Would you agree that this should be viewed as an assumption for the definition of rational players to justify that a player can continuously vary probabilities in choosing pure actions? How do you justify it? Otherwise I would have the conceptual philosophical problem. I think I can live with this finitely many actions, but the Nash theorem somehow has to rely on continuous variation of probabilities. Would you also see it as an assumption of rational players, so it is more philosophical. Thank you.

NASH
That's really a philosophical question. Mathematically of course it is clear that you must have the continuity. You can get quite odd numbers in fact. I think if you have two players and you have the mixed strategies, you have specific numbers that are rational, but if you have more players you get algebraic

numbers. So if there is some philosophical basis on which that type of number cannot be arrived at, describing a mixed strategy, then that's out.

There is something I just wanted to say. When I heard about the Nobel awards, and I heard that the persons were who they were, I wondered how they were connected. Of course I knew that Harsanyi and Selten made some use of the concept of equilibrium points or Nash equilibrium points, but I wondered what else there was of interrelation, and I started reviewing things, because I hadn't been following the field directly. And I discovered this book: "A General Theory of Equilibrium Selection in Games [23] which I think was published in 1988 by joint authors Harsanyi and Selten, and then I discovered also that in relation to this book, from opinions expressed about it, that it is very controversial. It's very interesting, but also somewhat controversial. And talking to some persons I found that impression is sort of confirmed, that there may be specific aspects of it that are not immediately accepted. But something can be more interesting if it is not immediately accepted. So there is the problem; the possibility that all cooperative games could really be given a solution. This could be analogous to the Shapley value. If it were really valid you would be able to say, here is a game, here are the possibilities of cooperation, binding agreements, threats, everything, this is what it is worth to all the players. You have a management-labor situation, you have international politics, the question of membership in the common market; this is what it is worth. That is, if everything could be measured or given in terms of utilities. So the possibility that there could be something like that is very basic, but Shapley would have had that very early if the Shapley value were really the true value. But one example in this book I studied shows how the solution considered there in fact differs from the Shapley value, and so it is a very interesting comparison. In principle, experiments might distinguish between different theories, so I think that's a very interesting area. I think there will be further work. I had better not say too much, because of course Harsanyi and Selten will be speaking tomorrow and I don't exactly know that they, Harsanyi and Selten, what they will say.

NOTES

[1] Actually, Nash also assumed that in a non-cooperative game, the players will be unable to *communicate* with each other. Yet, in my own view, this would be a needlesssly restrictive assumption. For if the players cannot enter into enforceable agreements, then their ability to communicate will be of no real help toward a cooperative outcome.

[2] Note that Nash equilibria seem to be the only solution concept applying *both* to games in *normal* form and in *extensive* form.

[3] Nash here refers to his theory of non-cooperative games based on the concept of Nash equilibria.

[4] He does not use the term "trembling-hand perfect" equilibria. But this is the term used by many other game theorists to describe this class of Nash equilibria.

5 Nash denotes payoffs with a Roman p instead of, as here, a Greek π.

6 A population state $s \in S$ is *stationary* if no change over time takes place once the population is in state s, i.e., if $f_{i\alpha}(s) = \forall i, \alpha$.

7 Nash [36] used the mapping $T: S \rightarrow S$ defined by $T_i(s) = s_i'$ for

$$s_{i\alpha}' = \frac{s_{i\alpha} + \pi_{i\alpha}^{+}(s)}{1 + \sum_{\beta \in A_i} \pi_{i\beta}^{+}(s)} \quad (\forall\ i \in I, \alpha \in A_i)$$

8 In the present context, a state $x^* = (s^*, p^*)$ is *stationary* if $s(0) = s^*$ and $p(0) = p^*$ together imply $s(t) = s^*$ and $p(t) = p^*$ for all t in some open time interval containing 0.

9 An alternative version (Maynard Smith [30], Hofbauer and Sigmund [24]) presumes $u_i(s) > 0$ and is given by $g_{i\alpha}(s) = u_{i\alpha}(s)/u_i(s) - 1$.

10 A population state s is *dynamically stable* if small perturbations of the state do not lead the population away, i.e., if every neighborhood V of s contains a neighborhood U of s such that no solution curve starting in U leaves V.

REFERENCES

[l] Arrow, K., and G. Debreu. 1954. Existence of an equilibrium for a competitive economy. *Econometrica 3: 265 - 290.*

[2] Aumann, R. J. 1987. What is game theory trying to accomplish? In *Frontiers of Economics* , edited by K. Arrow and S. Honkapohja. Oxford: Basil Blackwell, 28 - 100.

[3] Aumann, R. J. 1987b. Game theory. In *The* new *Palgrave dictionary of economics,* edited by M. Milgate and P. Newman, 460 - 482.

[4] Aumann, R. and A. Brandenburger. 1991. *Epistemic conditions /or Nash equilibrium,* working paper 91 - 042, Harvard Business School.

[5] Binmore, K., A. Rubinstein, and A. Wolinsky. 1986. The Nash bargaining solution in economic modelling. Rand *Journal of Economics* 17: 176 - 188.

[6] Björnerstedt, J., and J. Weibull. 1993. Nash equilibrium and evolution by imitation. In *The rational foundations of economic behavior,* edited by K. Arrow et al. New York: Macmillan, forthcoming.

[7] Bomze, I. 1986. Non-cooperative two-person games in biology: A classification. *International Journal of Game Theory* 15: 31 - 57.

[8] Carlsson, H., and E. Van Damme. 1993. Equilibrium selection in stag hunt games. In *Frontiers in game theory,* edited by K. Binmore and A. Kirman. Cambridge: MIT Press, 237 - 254.

[9] Debreu, G. 1984. Economic theory in the mathematical mode. In Les Prix Nobel 1983. (Reprinted in *American Economic Review 74: 267 - 278.*

[10] Eshel, I., and M. W. Feldman. 1984. Initial increase of new mutants and some continuity properties of ESS in two locus systems. *American Naturalist* 124: 631 - 640.

[11] Fisher, R. A. 1930. *The genetical theory of natural selection.* Oxford: Clarendon Press.

[12] Flood, M. M. 1959, Some experimental games. *Management Science 5: 5 - 26.*

[13] Friedman, M. 1953. *Essays in positive economics.* Chicago: University of Chicago Press.

[14] Haldane, J. B. S. 1932. *The causes of evolution.* London: Longman.

[15] Hammerstein, P. 1994. Darwinian adaptation, population genetics and the streetcar theory of evolution. *Journal of Mathematical Biology,* in press.

[If] Hammerstein, P., and R. Selten. 1994. Game theory and evolutionary biology. In *Handbook of game theory with economic applications,* vol. 2, edited by R. J. Aumann, and S. Hart, 929 - 993. Amsterdam: Elsevier.

[17] Harsanyi, John C. 1956. Approaches to the bargaining problem before and after the theory of games: A critical discussion of Zeuthen's, Hicks', and Nash's theories. *Econometrica 24:* 144 - 157.

[18] - - - 1963. A simplified bargaining model for the n-person cooperative game. *International Economic Review 4:* 194 - 220.

[19] - - -, 1967 - 1968. Games with incomplete information played by Bayesian players. *Management Science* 14: 159 - 182, 320 - 334, 486 - 502.

[20] - - -. 1973. Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points. *International Journal of Game Theory 2:* 1 - 23.

[21] - - -. 1974. An equilibrium-point interpretation of stable sets and a proposed alternative definition. *Management Science 20:* 1472 - 1495.

[22] Harsanyi, John C, and R. Selten. 1972. A generalized Nash solution for two-person bargaining games with incomplete information. *Management Science 18: 80* - 106.

[23] - - -_ 1988. *A general theory of equilibrium selection in games* Cambridge: MIT Press.

[24] Hofbauer J., and K. Sigmund. *1988. The theory of evolution and dynamical systems.* Cambridge: Cambridge University Press.

[25] Kalisch, G. K.,]. W. Milnor, nor, J. F. Nash, and E. D. Nering. 1954. Some experimental n-person games. In *Decision Processes* edited by R. M. Thrall, C. H. Coombs, and R. L. Davis. New York: Wiley, Chapter 21, 301 - 327.

[26] Karlin, S. 1975. General two-locus selection models: Some objectives, results and interpretations. *Theoretical Population Biology* 7: 364 - 398.

[27] Kohlberg, E., and J.-F. Mertens. 1986. On the strategic stability of equilibria. Econometrica 54: 1003 - 1037.

[28] Leonard. R. 1994, Reading Cournot, reading Nash: The creation and stabilization of the Nash equilibrium. *Economic Journal* 104: 492 - 511

[29] Leonard, R. 1995. From parlor games to social science: von Neuman, Morgenstern, and the creation of game theory, 1928 - 1944. *Journal of Economic Literature,* in press.
[30] Maynard Smith, J. 1982 *Evolution and the theory of games.* Cambridge: Cambridge University Press.
[31] Maynard Smith, J., and G. R. Price. 1973. The logic of animal conflict. *Nature 246:* 15 - 18.
[32] Moran, P. A. P. 1964. On the nonexistence of adaptive topographies. *Annals of Human Genetics 27: 338 - 343.*
[33] Morgenstern, Oskar. 1976. The collaboration between Oskar Morgenstern and John von Neumann on the theory of games. *Journal of Economic Literature* 14: 805 - 816.
[34] Nachbar, J. 1990. 'Evolutionary' selection dynamics in games: Convergence and limit properties. *International Journal of Game Theory* 19: 58 - 89.
[35] Nash, Jr., John F. 1950a. Equilibrium points in n-person games. *Proceedings National Academy of Sciences* 36: 48 - 49.
[36] - - -. 1950b. *Non-cooperative games,* Ph.D. thesis, Mathematics Department, Princeton University.
[37] - - -. 1950c. The bargaining problem. *Econometrica 18:* 155 - 162.
[38] - - -. 1951. Non-cooperative games. *Annals of Mathematics 54: 286 - 295.*
[39] - - -. 1953. Two-person cooperative games. *Econometrica* 21: 128 - 140.
[40] Robinson, J. 1951. An iterative method of solving a game. *Annals of Mathematics 54:* 296 - 301.
[41] Roth, A. E. 1993. The early history of experimental economics. *Journal of the History of Economic Thought* 15: 184 - 209.
[42] Rubinstein, A. 1982. Perfect equilibrium in a bargaining model. *Econometrica 50: 97* - 109.
[43] Rubinstein, A., Z. Safra, and W. Thomson. 1992. On the interpretation of the Nash bargaining solution and its extension to non-expected utility preferences. *Economettica 60:* 1171 - 1186.
[44] Schelling, T. *C.* 1960. *The strategy of conflict.* Cambridge: Harvard University Press.
[45] Selten, Reinhard. 1965. Spieltheoretische behandlung eines oligopolmodells mit nachfragetragheit. *Zeitschrift für die Gesamte Staatswissenschuft* 121: 301 - 324 and 667 - 689.
[46] - - -_ 1975. Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of* Game *Theory 4: 25 - 55.*
[47] - - -. 1993. In search of a better understanding of economic behavior. In *The* makers *of modern economics,* edited by A. Heertje. Hertforshire: Harvester Wheatscheaf, Vol. I: 115 - 139.
[48] Taylor, P. 1979. Evolutionarily stable strategies with two types of players. *Journal of Applied Probability* 16: 76 - 83.
[49] Taylor, P., and L. Jonker. 1978. Evolutionary stable strategies and game dynamics. *Mathematical Biosciences 40:* 145 - 156.
[50] van Damme, E. 1987. *Stability and perfection of Nash equilibria* Berlin: Springer-Verlag.
[51] von Neumann, John. 1928. Zur theorie der gesellschaftsspiele. *Mathematische Annalen 100: 295 - 320.*
[52] von Neumann, John, and Oskar Morgenstern. 1944. *Theory of games and economic behavior* Princeton: Princeton University Press.
[53] Weibull, J. 1994. The 'as if' approach to game-theory: Three positive results and four obstacles. European *Economic Review* 38: 868 - 882.
[54] Weibull, J. 1994. *Evolutionary Game Theory.* Cambridge: The MIT Press, forthcoming.
[55] Wright, S. 1931. Evolution in Mendelian populations. *Genetics 16: 97* - 159.