

# LEARNING GAME THEORY FROM JOHN HARSANYI

by Roger Myerson

August 31, 2000

When a scholar publishes a paper, it is a letter sent to unknown recipients. If the job has been well done, then with luck it may be found and read, perhaps years later, by people who may take it into their lives and let it change their understanding of the world. It is in this way that I was touched and changed by John Harsanyi, before I ever met him.

In 1972, when I was a college student trying to learn about game theory, I found a reference to John Harsanyi's 1963 paper on a general solution concept for cooperative games without transferable utility. This paper was not so easy to read, and I pored over it in my college library for 3 days. But when I finally understood what he had done, I knew that this was what I wanted to do too. I continued studying all of John Harsanyi's papers on game theory, and I found in them a dissertation topic and, beyond that, a research career. My debt to him is immense.

This month, I have taken some time to reread those papers that influenced me so much in my early studies, and also to read some of his other early papers that I never knew before. I want to tell you something here about what I have learned from this reading, to try to share with you some understanding of why his contributions have been so important.

First and foremost, what made John Harsanyi's work so special was his consistent focus on analytical generality. The key to understanding his work may have been best expressed in a short paper on theoretical analysis in social science that he published in 1961 in the *Australian Journal of Politics and History*. There he wrote:

The unusually rapid social, economic, and political changes of our period, the conflicts between rival economic and political systems, the hot and cold wars, revolutions and other social crises that we have been experiencing, represent an intellectual and practical challenge that cannot be met by a mere description of the facts but only by theoretical analysis, by explaining and evaluating these facts in terms of a general theoretical framework. (Harsanyi, 1961a, p 60.)

The madness of the Nazi and Communist systems which had persecuted him cried out for explanation. He responded in his own way, not by dwelling on the social evils that he had experienced, but by working systematically to reconstruct the logical foundations of social theory, so that people in the future might better understand the dynamics of social institutions. In this labor, he drew deeply on the mathematical and philosophical skills that he developed in Hungary, during the years when scholarship was not forbidden to him.

In this 1961 paper he sketched the logical progression that led him to game theory. He began with some references to other general structures that he had studied in social science, including those of Max Weber and Talcott Parsons. But he went on to explain carefully why he found the most effective social theories to be those which explain people's behavior in terms of their individual preferences, and this led him to game theory.

The game theory that he found in the 1950s was a new subfield where a small but extraordinary group of mathematicians and economists wrestled with basic questions of methodology. There was a proliferation of different solution concepts, but most of them could be applied only to special types of games, either to games where players' payoffs were transferable like money, or to games where there were only two players. Many theories could yield multiple solutions or no solutions at all for a game.

John Harsanyi's view of the field was clarified by his insistence that a good solution concept should yield one well-defined solution to any game. And in the mid 1950s, if you looked for game theoretic solution concepts that yielded unique solutions to broad classes of games, you would find two. Lloyd Shapley's (1953) solution yielded a unique value with any number of players, but was only defined for games where payoffs were transferable. John Nash's (1950, 1953) bargaining solution did not require transferable payoffs, but was only defined for games with two players. In 1956, Harsanyi showed that Nash's bargaining solution could be derived from an earlier theory of bargaining by the Danish economist Frederik Zeuthen (1930). Then in 1959 and 1963, John Harsanyi published papers to show how the Nash bargaining solution and the Shapley value could be unified into one general solution concept that could be applied to any cooperative game with complete information. This 1963 paper, with its remarkable unification of the two nicest solution concepts in game theory, was the paper that got me started in game theory. (I still treasure the photocopy of it that I made in 1972.)

But this early work was all in cooperative game theory. John Nash had introduced the basic concept of noncooperative equilibrium in 1951, but in the decade thereafter there was very little work in noncooperative game theory. The most important advance in this decade was Thomas Schelling's (1957, 1958) theory of the focal-point effect in games with multiple equilibria, which he advocated as a better way to understand bargaining in the real world. John Harsanyi responded to Schelling in a 1961 paper in the *Journal of Conflict Resolution*. In this paper, John Harsanyi argued that the distribution of power that is measured by a cooperative solution concept can become the focal factor that selects among the many noncooperative equilibria of a bargaining game. To me, this paper still stands as the definitive defense of cooperative game theory against the arguments of a noncooperative game theorist. But, in 1961, virtually all the active game theorists except Schelling still wanted to do cooperative game theory. The advance of noncooperative game theory really began in footnote 7 of this paper, where John Harsanyi promised to show, in a subsequent paper, how the axioms of his general bargaining theory could be extended to noncooperative games.

That subsequent paper became a series of papers (Harsanyi 1962a, 1964, 1966, 1975), and ultimately led to his great book with Reinhard Selten (1988). Applying ideas like risk dominance to identify a unique rational equilibrium for any noncooperative game proved to be a very hard problem. But as soon as he began to work on it, John Harsanyi realized the force of Nash's early arguments for the greater generality of noncooperative game theory over the cooperative approach. Thus Harsanyi switched camps during the 1960s, and he became the leading advocate of noncooperative game theory. In this role, he began the search for theoretical criteria to select among multiple equilibria, because he understood first that the noncooperative approach could not become the standard methodology of game theoretic analysis without some refinements of Nash's equilibrium concept.

In 1962(b), John Harsanyi published a short speculative piece about bargaining when players are ignorant about each others' preferences. He found some fundamental difficulties here, because the tools of game-theoretic analysis could only be applied under the assumption that all players know the complete structure of the game. So he began to think systematically about how to model incomplete information in games, and he developed his ideas in discussions with Reinhard Selten and the other great game theorists in the now-legendary *Mathematica* arms-control project. The result was a

monumental three-part series of papers on games with incomplete information, which John Harsanyi published in 1967 and 1968. Here he carefully and systematically showed how any kind of uncertainty about a game should be modeled by bringing the uncertainty into the game model itself. He defined a Bayesian game model in which each player has a set of possible types that characterize the player's possible preferences and beliefs at the beginning of the game. Then each player is assumed to understand the whole structure of this Bayesian model and, in addition, to know his own actual type.

If this approach seemed counterintuitive at first, it was because the initial problem of uncertainty about payoff functions in a game was solved by assuming instead that payoff functions were known in a much bigger and more complicated game. But John Harsanyi showed that, with an additional consistency assumption, these Bayesian games could themselves be understood as conventional complete-information games where the beginning of the game has just been moved backwards in time, so that a historical chance move could be introduced to account for the differences in players' information.

In the later 1970s and 1980s, John Harsanyi's consistent Bayesian model became the standard conceptual structure that economists of my generation used to understand informational problems in economics. We studied his 1967-1968 papers, and they gave us a common framework in which Vickrey's (1961) auctions, Akerlof's (1970) market for lemons, and Spence's (1973) labor markets, and Rothschild and Stiglitz's (1976) insurance markets could all be seen as interesting examples of people playing Harsanyi's Bayesian games. Having this common framework for such informational problems enabled us to apply insights from the study of any one of them to all the others, and thus the new economics of information was developed. Harsanyi's Bayesian framework even transformed our understanding of the older analysis of complete-information games, following his 1973 paper which taught us to interpret randomized equilibria as indicators of the decisive impact of minor private information in rational competitive behavior.

So when I first met John Harsanyi at Berkeley in 1976, I had already learned game theory from him. I told him about my thesis and about the typo on page 210 in his 1963 paper. He smiled and thanked me. I got to know him better, along with Anne and Tom, when we were together in Bielefeld during the 1978-79 academic year. I also had the privilege of accompanying him at Northwestern's

graduation in 1989 when he received an honorary degree. And I got to congratulate him on the day that he won the Nobel prize. But whenever we were together, I wanted him to know how much his work meant to me and others like me, in the next generation of game theorists and economists.

For above all, John Harsanyi was right about the importance of a general theoretical framework for social science. Good applied work in any academic field must stay within the scope of the discipline's basic theoretical framework, because beyond these limits we have only ad-hoc intuition. By working consistently to develop the broadest scope for game-theoretic analysis, John Harsanyi helped to liberate economists from the older limits of price theory. With the general analytical structures that he developed, economists could boldly expand the scope of their analysis to study problems of information and incentives in any social, economic, or political institution. With his contributions, we now have an analytical methodology that can be applied more broadly to help understand the changes of our period, the conflicts between rival economic and political systems, the hot and cold wars, the revolutions and other social crises such as he experienced. Perhaps we can even learn to keep the world a better place than it may have seemed in Hungary in 1944 or 1950. We have lost a great intellectual leader, but we will keep working at it.

## REFERENCES

- George Akerlof, "The market for lemons: quality uncertainty and the market mechanism," Quarterly Journal of Economics 84 (1970), 488-500.
- John C. Harsanyi, "Approaches to the bargaining problem before and after the theory of games," Econometrica 24 (1956) 144-157.
- John C. Harsanyi "A bargaining model for the cooperative n-person game," in Contributions to the Theory of Games 4, edited by A. W. Tucker and R. D Luce, Princeton University Press (1959), 325-355.
- John C. Harsanyi, "Theoretical analysis in social science and the model of rational behavior," Australian Journal of Politics and History 7 (1961a), 60-74.
- John C. Harsanyi, "On the rationality postulates underlying the theory of cooperative games," Journal of Conflict Resolution 5 (1961b), 179-196.
- John C. Harsanyi, "Rationality postulates in cooperative and noncooperative games," Management Science 9 (1962a), 141-153.
- John C. Harsanyi, "Bargaining in ignorance of the opponent's utility function," Journal of Conflict Resolution 6 (1962b), 29-38.
- John C. Harsanyi, "A simplified bargaining model for the n-person cooperative game," International Economic Review 4 (1963), 194-220.
- John C. Harsanyi, "A general solution for finite non-cooperative games, based on risk-dominance," in Advances in Game Theory, edited by M. Dresher, L. S. Shapley, and A. W. Tucker, Princeton University Press (1964), 651-679.
- John C. Harsanyi, "A general theory of rational behavior in game situations," Econometrica 34 (1966), 613-634.
- John C. Harsanyi "Games with incomplete information played by Bayesian players," Management Science 14 (1967-1968), 159-182, 320-334, 486-502.
- John C. Harsanyi, "Games with randomly disturbed payoffs: a new rationale for mixed strategy equilibrium points," International Journal of Game Theory 2 (1973), 1-23.
- John C. Harsanyi, "The tracing procedure: a Bayesian approach to defining a solution for n-person

- noncooperative games," International Journal of Game Theory 4 (1975), 61-94.
- John C. Harsanyi and Reinhard Selten, A General Theory of Equilibrium Selection in Games, MIT Press (1988).
- John Nash, "The bargaining problem," Econometrica 18 (1950), 155-162.
- John Nash, "Non-cooperative games," Annals of Mathematics 54 (1951), 286-295.
- John Nash, "Two-person cooperative games," Econometrica 21 (1953), 128-140.
- Michael Rothschild and Joseph Stiglitz, "Equilibrium in competitive insurance markets, an essay on the economics of imperfect information," Quarterly Journal of Economics 90 (1976), 629-649.
- Thomas C. Schelling, "Bargaining, communication and limited war," Journal of Conflict Resolution 1 (1957), 19-36.
- Thomas C. Schelling, "The strategy of conflict: prospectus for a reorientation of game theory," Journal of Conflict Resolution 2 (1958), 203-264
- Lloyd S. Shapley, "A value for n-person games," in Contributions to the Theory of Games 2, edited by H. W. Kuhn and A. W. Tucker, Princeton University Press (1953), 307-317.
- Michael Spence, "Job market signaling," Quarterly Journal of Economics 87 (1973), 355-374.
- William Vickrey, "Counterspeculation, auctions, and competitive sealed tenders," Journal of Finance 16 (1961), 8-37.
- Frederik Zeuthen, Problems of Monopoly and Economic Warfare London: Routledge (1930).

*Note: This paper was written for a Memorial Service for John C. Harsanyi, at the University of California at Berkeley, on August 31, 2000.*

*A full list of John Harsanyi's publications up to 1992 can be found in Rational Interaction: Essays in Honor of John C. Harsanyi, edited by Reinhard Selten, Springer-Verlag (1992).*