The World of Game Theory and Game Theory of the World: A Personal Journey

Robert J. Aumann*

Text of the author's talk at the colloquium held in his honor on October 11, 2021, at the University Paris-Panthéon-Assas, to mark his receipt on the next day of a doctorate honoris causa granted by the same university. After thanking the initiators and organizers of the event, and paying homage to the great figures of game theory who have passed away, the author describes his "personal journey" in Science. The leitmotif, illustrated by his undergraduate studies, his doctoral dissertation, and his subsequent work in game theory, is that pure and applied science are ultimately one and the same. This is illustrated, inter alia, by the relationship between game theory (GT) and behavioral economics (BE): it is argued that BE's supposedly irrational heuristics and biases almost always lead to behavior that accords well with GT's rational analysis.

LE MONDE DE LA THÉORIE DES JEUX ET THÉORIE DES JEUX DU MONDE : UN PARCOURS PERSONNEL

Ceci est le texte de l'intervention de l'auteur au colloque organisé en son honneur le 11 octobre 2021 à l'Université Paris-Panthéon-Assas, à l'occasion de la réception le lendemain d'un doctorat honoris causa décerné par cette même université. Après avoir remercié les initiateurs et les organisateurs de l'événement, et rendu hommage aux grandes figures disparues de la théorie des jeux, l'auteur décrit son « parcours personnel » dans la science. Le leitmotiv, mis en lumière par ses études universitaires, sa thèse de doctorat et ses travaux ultérieurs en théorie des jeux, est que la science pure et la science appliquée ne sont finalement qu'une seule et même chose. Cela est illustré, entre autres, par la relation entre la théorie des jeux (GT) et l'économie comportementale (BE) : l'idée défendue ici est que les heuristiques et les biais supposés irrationnels par la seconde (BE) conduisent presque toujours à définir un comportement qui s'accorde bien avec l'analyse rationnelle proposée par la première (GT).

Keywords: game theory, philosophy of science, rationality, evolution

Mots clés : théorie des jeux, philosophie des sciences, rationalité, évolution

JEL Codes: B20, B31, B40, C00, C70.

* The Hebrew University of Jerusalem. *Correspondence:* Federmann Center for the Study of Rationality, The Hebrew University, 91904 Jerusalem, Israel. *Email:* raumann@math.huji.ac.il

To start with, MERCI BEAUCOUP to Professor Antoine Billot and Dr. Christina Pawlowitsch, who provided both the impetus for, and the organization of, this moving event; to the University of Paris II Pantheon-Assas, who sponsored the event; to the speakers—the President of the University Stéphane Braconnier, and the game theorists Enrico Minelli, Itzhak Gilboa, Herakles Polemarchakis, Dov Samet, Françoise Forges, Sylvain Sorin, and Jorgen Weibull; to the entire audience; and to the *Revue Économique*, the publisher of these proceedings. To each and every one—thank you very very much.

And thanks also to the giants of game theory with whom I had the privilege of interacting and working closely who are no longer with us. Inter alia, Oskar Morgenstern, Lloyd Shapley, Michael Maschler, Jean-François Mertens, Bezalel Peleg, Jacques Drèze, Kenneth Arrow, Harold Kuhn, John Nash, John Harsanyi, Reinhard Selten, Martin Shubik, and Herb Scarf.

Finally, allow me to pay tribute to the French school of mathematical game theory. The Talmud says that a baker does not testify as to his own dough, so I will say nothing about the Israeli school of mathematical game theory; but that aside, the French school, led by Sylvain Sorin, is without doubt the deepest and most profound in the world.

Now to my personal journey.

In the mid-20th century, when I started studying math, the mathematical vogue was: The less useful, the better. "Pure" mathematics ruled the roost; applied mathematics was considered somehow sullied, second-rate, not befitting a real mathematician and gentleman. The great number theorist Godfrey Harold Hardy reportedly said that the only thing he regretted doing in his life was his work on mathematical genetics. Edmund Landau, another famous number theorist, mocked applied math with his dictum "*Die Zahlentheorie ist etwa nützlich weil man mit ihr promovieren kann*" (Number theory may be considered useful because one can do a doctorate with it).

Young and impressionable, I was swept up by the vogue. When studying for my first degree at New York's City College, I read number theory voraciously, in particular the books of Landau and Hardy. What attracted me was that (i) the *results* are easily formulated and natural, a schoolchild can understand them; in contrast, (ii) the *proofs* are often deep and difficult; and finally, (iii) the whole subject seemed absolutely useless.

An example is Fermat's last theorem. We know that the sum of two squares can itself be a square; for example, $3^2 + 4^2 = 5^2$. But what about cubes? Can two cubes add up to a cube? The answer is negative; the equation $a^3 + b^3 = c^3$ has no solution (in positive integers).

This already is not easy to prove. Next, we go to fourth, fifth, sixth powers, and so on. Is there *any w* other than 2 for which the equation $a^w + b^w = c^w$ has a solution in positive integers? In 1637, the French mathematician Pierre de Fermat conjectured that the answer is negative. But it took 257 years until this conjecture was proved—by very deep, complex, and ingenious methods.

After City College, I did a doctorate at MIT (the Massachusetts Institute of Technology), with a thesis in the theory of knots (those tied in ropes). Again, what attracted me to the subject was that (i) the *results* are easily formulated and natural, even more so than in number theory; a kindergarten child can understand them. In contrast, (ii) the *proofs* are deep and difficult; and finally, (iii) the whole subject seemed absolutely useless.

Specifically, my thesis concerns knots that *alternate,* in the sense that each strand alternately passes over and under the other strands (or itself); I showed that such knots cannot come apart. The proof, though of course nowhere near as deep or profound as that of Fermat's last theorem, nevertheless depends on fairly deep results of algebraic topology. At one point in October of 1954, the proof was almost done, but then a particularly stubborn hurdle presented itself. I remember standing in the shower in October of 1954, thinking and thinking about the problem. And then, right there in the shower, the solution hit me. The thesis was complete.

After the doctorate, I left knot theory, turning to the research that has taken up the rest of my life; more on that forthwith. But first, fast forward to October of 2004. It is ten in the evening, and the phone in my flat rings. On the line is my grandson Yakov Rosen, who is in the second year of medical school.

Yakov: Grandpa, what are linking numbers?

Me: You mean like in knot theory?

Yakov: Yes.

Me: Why are you interested in knot theory?

Yakov: Well, linking numbers played an important role in one of today's lectures; I didn't understand the lecturer's explanation of what they are, and think that he himself didn't understand it.

Me: Why are you studying knots? Does the University require you to take non-medical courses? Or are you taking this course to broaden horizons?

Yakov: You don't understand. This course is in the medical school.

Me: Why is knot theory taught in the medical school?

Yakov: Well, sometimes the DNA in a cell gets knotted up. This can lead to cancer or other problems, depending on the linking numbers. So we have to understand linking numbers.

Overwhelmed with emotion, I had to sit down. The work I'd been doing half a century earlier *because* it seemed absolutely useless was now being taught in the second year of medical school. The world of knot theory has given birth to knot theory of the world.

Also number theory, which I'd read as an undergraduate *because* of its apparent uselessness, suffered a similar fate: it became highly useful. Starting in the 1970s, the theory of prime numbers became the basis of computer cryptology. The world of number theory gave birth to number theory of the world.

Back to the 1950s: After completing the doctorate in 1955, I did a post-doctoral stint at a unit of Princeton University's math department called the Analytical Research Group (ARG). Most of the department did highly abstract "pure" mathematics, but ARG did highly practical operations research consulting for the US government and for industrial organizations like Bell Telephone Laboratories. One of the matters about which Bell Labs consulted ARG concerned a ground-to-air missile they were developing for the US defense department. Specifically, how to program the missile to deal with an attack by a squadron of aircraft some of which are carrying nuclear bombs while most are decoys. This was of great relevance in the 1950s, when there were few nuclear bombs altogether. The problem was referred to me.

John Nash, fresh from his groundbreaking Ph. D. in game theory, had come to teach at MIT at the beginning of my last year there. We got to know each other fairly well, talked math a lot. Inter alia, game theory. At the time, wrapped up as I was with other areas of math, game theory did not really seize me, though it did interest me mildly. But when starting to think about the missile and the decoys a year later, I realized that this was what Nash had been talking about. So I read up on the subject, did what I could with the problem, and wrote a report. In the process, I became hooked on game theory.

In 1957, Duncan Luce and Howard Raiffa published their book *Games and Decisions*, which I read from cover to cover. Among the many important topics they discuss is that of repeated games. Their discussion, though informal and inconclusive, is suggestive; I started to think about how one could formulate and prove a general result on the topic. What emerged was the following principle: *Repeated games are conducive to cooperation*.

In 2005, when I was awarded the Nobel Memorial Prize in Economics for my work in game theory, it was in considerable part for my work on the above principle. And people asked: Is *that* why you got the prize? Isn't that obvious? In a repeated game, people get to know each other, trust each other, work together. It's not about playing games—has nothing to do with rationality; it's about building relationships.

I myself began to wonder about this. And then I hit on the answer: It *is* about game theory, it *is* about rationality. Though people don't necessarily think rationally, they do *behave* rationally. Rational behavior, like all characteristics of living organisms, has evolved (or been learned) because it is beneficial, it increases fitness. Evolution works by trial and error; it stumbles on the right way to do things. If modes of behavior were not rational, they would not survive. The trust, the relationships mentioned above are modes of behavior that have evolved in repeated game situations. We, the analysts, have a way of discovering how to behave rationally that is more efficient than trial and error; namely, through game theory analysis.

This is closely related to evolutionary game theory (EGT). When John Maynard Smith and George Price first came out with EGT, it was considered an interesting curiosity—the same equations that define Nash equilibrium also define evolutionary population equilibrium. But now we realize that the evolutionary viewpoint is the more basic. We behave (not think!) rationally *because* of evolution.

It is also closely related to behavioral economics (BE); in fact, it's all about BE. BE says that people don't think things out rationally; rather, they behave by rules of thumb—what BE calls "heuristics" or "biases." That's exactly what we said above. But *we* say that people behave rationally, whereas BE says that people often behave irrationally. In fact, that's the main point of BE: that mainstream economics—and mainstream game theory—do not describe how people behave in practice.

In fact, BE is right that people behave by rules of thumb, but wrong that that behavior is often irrational. Many of the behaviors that the BE people describe as irrational are in fact rational. All others occur in contrived or highly unusual scenarios. Recall that the rules of thumb—the modes of behavior—are products of evolution; and evolution operates on the usual, the natural, not the contrived.

An example is the ultimatum game, played as follows: Two players, the proposer (*P*) and responder (*R*), must divide €100. If they agree how, each gets his agreed share. If not, both get nothing. They sit at computers in separate rooms and can't communicate directly. *P* starts by making a numerical offer to *R*, without words. *R* responds by clicking "yes" or "no"; no other response is possible. The game is then over; the players get their payoffs (if any) and leave by separate doors. They never see each other nor learn each other's identity. The subjects are students, not particularly long on money.

The observed behavior in this game is that most *P*'s offer around 35. Smaller offers—say 20—are rejected.

The rule that prescribes this behavior is: *Don't let people kick you in the stomach: reject lop-sided offers.* Mechanisms for executing the rule are feelings of wounded pride, insult, desire for revenge, honor. The rule and its mechanisms evolved in natural scenarios, where the negotiators know each other. If in such scenarios you accept lop-sided offers, you'll get a reputation for doing so, and in the future will get *only* such offers; so rejecting is highly rational. In the contrived, artificial ultimatum game, reputational effects don't apply, as the players are totally anonymous; but the rule and its mechanism evolved in natural scenarios, where they do apply. The rule prescribes rational behavior in naturally occurring scenarios, but not in the contrived ultimatum game.

The ultimatum game illustrates our contention that a rule that usually calls for rational behavior may in unnatural, contrived scenarios lead to suboptimal results. But also, many of the behaviors that the BE people describe as irrational are in fact rational. An example is hyperbolic discounting: Offered a choice between €10 on the spot and €11 tomorrow, some experimental subjects choose €10 on the spot; whereas the same subjects, offered a choice between €10 in a year and €11 in a year and a day, choose €11 in a year and a day.

Though BE views this behavior as irrational, it is in fact perfectly rational. The rule that prescribes the behavior is: *A bird in the hand is worth two in the bush.* In our case, if you give me €10 now, I pocket it, and that's the end of the story. €11 tomorrow? Maybe yes, maybe no; there's a qualitative difference between now and later. Between 365 and 366 days, there is no such difference.

To conclude: By and large, people behave rationally, though usually not because they have thought the matter through. That's game theory of the world. To understand their behavior, or to advise them, or just out of curiosity, we think the matter through. That's the world of game theory.

In short, game theory is a beautiful edifice. But not an ivory tower.

Revue économique – vol. 74, n° 4, juillet 2023, p. 1-5