Obituary

Herbert A. Simon (1916–2001)

Herbert A. Simon died on February 9, 2001, in Pittsburgh, at the age of 84. He was the Richard King Mellon University Professor of Computer Science and Psychology at Carnegie Mellon University, where he taught since 1949. He won the Nobel Prize in economics in 1978. He has influenced management scientists, decision theorists and artificial intelligence specialists all over the world, probably more than any other person in those fields.

Simon also received A.M. Turing Award (1975), National medal of Science (1986), high honors and awards from the American Psychological Association, Chinese Academy of Sciences (1994) and the International Joint Conference on Artificial Intelligence and the American Society of Public Administration (1995).

Herbert Alexander Simon was born on June 15, 1916, the son of German immigrants, in Milwaukee. Strong European values, combined with exacting interdisciplinary interests in systems, brought forth concepts, thoughts and writings of the highest caliber. How many of us spent hours of intellectual joy with the 'Sciences of the Artificial', 'Administrative Behavior' or 'Models of Bounded Rationality'? How many students were lead on the path of surprising and unraveling discovery in studying his papers and articles, one leading to another, often in obscure publications and journals, many never acknowledged by the 'establishment'. Simon's thought was like an iceberg: there is still so much more 'under the water', unseen, unknown – potentially dangerous to many cozy fellow travelers of decision-making and artificial-intelligence seas. How many students and researchers know his 27 published books?

Those of us who knew him personally were perhaps influenced the most. Richard M. Cyert, past President of Carnegie-Mellon and first HSM Editorial Board member, as well as Jared Cohon, the current President of the university, have nurtured Simon with links to both Human Systems Management and Multiple Criteria Decision Making (MCDM) communities. The theories of 'satisficing' and 'bounded rationality' have in fact become integral part of both areas. Satisficing (satisfactory, doing well enough) has been incorporated explicitly into Goal programming, in contrast to optimization (doing one's best under constraints). Concepts of bounded rationality became even more influential, rejecting the so-called 'economic man'.

There is very little of 'economic man', complete information, rational behavior, utility function, and other formal constructs in human affairs, especially in economics. People do not behave according to formal axioms of rationality, yet they try to behave 'rationally' in different contexts. Because different decision contexts are missing in economics, as is production and valuation of knowledge, multiple decision criteria, added values, and so on, economic man – and especially economic woman – are just formalistic fictions.

The notion that 'all available information' somehow enters into 'rational' decision making and behavior of humans is part of the previous century thinking. The experiments of Tversky, Kahneman and Slovic have shown that only a small part of information input goes into decision quality (or 'rationality'); its larger part goes into building up human *confidence* in decisions already taken. That of course is quite rational, because decision-making confidence enhances decision implementation. However, traditional economic theory does not do confidence or contexts and must declare anybody who does not use all of information (that is all of us) an exception, an island of irrationality in the sea of axiomatic rationality.

Simon paid serious attention to so called anomalies, exceptions and deviations in human decision-making behavior. Mechanistic contrivance of *homo oeconomicus* has been successfully attacked by Amos Tversky (also deceased now), Paul Slovic and Daniel Kahneman. More recently, also by Richard Thaler (a schoolmate from my University of Rochester days), who ended up teaching at the same school where H.A. Simon got his doctorate: University of Chicago. Thaler's 'mental accounting' theory has pointed out that irrational exuberance can indeed be rational – in humans, not in machines.

Human Systems Management 20 (2001) 3–4
ISSN 0167-2533 / $8.00

Human rationality is indeed bounded because humans place their own boundaries on themselves, differentially in their own different problem-solving contexts. Individuals are certainly not axiomatically rational, but can populations of irrational investors behave rationally on average, as in the stock market? Large populations do tend to revert to average and this law of averages has very little to do with rationality or perfect rationality. Ignoring the tendency towards average would be irrational in itself.

There is an apparent paradox in Simon's teachings. On one hand, he passionately reduced, scaled down and bounded the myths of full rationality and optimum decision making in humans, on the other hand, he predicted, upgraded, and actually believed – with irrational exuberance – that the abilities of machines and artificial contrivances can achieve virtually unbounded and unlimited performance. 'Within the visible future, machines that think, that learn and that create will be able to handle challenges coextensive with the range to which the human mind has been applied', he wrote in 1957. His rationality turned out to be truly unbounded. His 'Over the Christmas holiday, Al Newell and I invented a thinking machine', became one of the most famous and enduring folklore around Carnegie-Mellon in Pittsburgh.

Can machines think? Of course they cannot and never will. Machines can carry out instructions and commands and follow rules, often in much more complex patterns than humans ever can or will. They 'calculate' and 'simulate', but do not think. They cannot create, invent rules and impose purposes.

Humans can act as machines, carry out instructions and follow rules, but not as well as machines. Humans think. They formulate rules, give orders and issue commands; they can formulate purposes, autonomously and from within. Playing chess better than Kasparov is not proof of thinking, but of better and more efficient application of rules to complex combinatorial spaces. Humans will ultimately lose in tasks suitable for machines, but machines cannot surpass humans in their autonomous tasks of goal, rule and command settings. Humans and machines are collaborative and complementary human systems, not separate adversarial entities.

Quite illuminating is the anecdote related to Bertrand Russell. Simon and colleagues wrote a computer program (1955) that could prove mathematical theorems from Russell and Whitehead's *Principia Mathematica*. Then he wrote a letter to Russell explaining the feat. Lord Russell replied: 'I am delighted to know that *Principia Mathematica* can now be done by machinery. I wish Whitehead and I had known of this possibility before we wasted 10 years doing it by hand'. Through an exquisite irony, Russell pointed out that machines can only reproduce or mimic what humans thought of or 'did by hand' earlier. 'Thinking machines' would be lost and 'dead' without Russells and Whiteheads – and also without Simons.

Why is Simon's view of rationality in humans bounded, while the 'rationality' of human-created machines can be virtually unbounded? Simon did not offer any answers to his paradox. In 'Models of My Life' (1991) he confirmed that he felt as good about his predictions concerning machines as he did about those related to humans. It was the duality of his love, for both machines and humans, that lead to paradoxes, dilemmas and controversial pronouncements. Only in their integration, as human systems, can machines and humans be reconciled at a higher level of their mutually self-producing interaction. But that is already the stuff of the 21st century.

Milan Zeleny
Graduate School of Business
Fordham University at Lincoln Center
New York, NY 10023
USA