Algorithmic Social Sciences Research Unit

ASSRU

Department of Economics
University of Trento
Via Inama 5
381 22 Trento Italy

Discussion Paper Series

3 – 2011/II

HERBERT A. SIMON[♠]

15 JUNE 1915 (MILWAUKEE WI) – 9 FEBRUARY 2001 (PITTSBURGH, PA)

K. Vela Velupillai & Ying Fang Kao

June 2011

'The central task of a natural science is to make the wonderful commonplace: to show that complexity, correctly viewed, is only a mask for simplicity; to find pattern hidden in apparent chaos. …
…
This is the task of natural science: to show that the wonderful is not incomprehensible, to show how it can be comprehended – but not to destroy wonder. For when we have explained the wonderful, unmasked the hidden pattern, a new wonder arises at how complexity was woven out of simplicity. The aesthetics of natural science and mathematics is at one with the aesthetics of music and painting – both inhere in the discovery of a partially concealed pattern.'

Herbert Simon, 1996, pp. 1-2.

[♠] An earlier, more concise, version of this paper was prepared for the **Handbook of the History of Economic Analysis**, edited by *Gilbert Faccarello* & *Heinz Kurz*, Edward Elgar Publishing.

ADVENTURES OF AN ARTIFICED LIFE

No other single person has won the *Nobel Memorial Prize for Economics* (in 1978), the *Turing Award* of the *Association for Computing Machinery* (shared with Allen Newell, in 1975), the Orsa/ Tims John von Neumann Theory Prize in 1988), the Distinguished Scientific Contribution Award of the American Psychological Association in 1969 and the *National Medal of Science* in 1986.

In a virtuoso academic life and career spanning more than six decades, Simon managed, almost single-handedly, to create the wholly new disciplines of behavioural economics and the cognitive sciences and nurture through to growth and prosperity one of the great academic institutions, the *Graduate School of Industrial Administration* (GSIA) at the Carnegie Institute of Technology (now the *Carnegie Mellon University*) in Pittsburgh, where these disciplines challenged orthodoxies with the pugnacious visions and solid theoretical and experimental foundations that had been provided by their courageous and visionary creator[1].

Herbert Simon was the second son of Arthur Carl Simon, a German émigré electrical engineer, inventor and patent lawyer, and Edna Marguerite Merkel, pianist and a second-generation descendant of immigrants from Prague and Cologne. He was wholly Jewish on his father's side; partly Lutheran on his mother's. His brother Clarence was five years older. He was introduced to Dorothea Isabel Pye by William W Cooper[2] and they married on Christmas

---

[1] It was not only the GSIA that he nurtured and fostered. The eminent proof theorist, metamathematician and philosopher, Wilfreed Sieg, currently the Patrick Suppes Professor of Philosophy, recalled, during a memorial event , that Simon was instrumental in the creation of a department of philosophy at CMU and went on to observe:

> How intimately theoretical issues and practical affairs were intertwined for Herb! Having discussed some difficult administrative problems with him, he remarked without further explanation: *Proceed, as if you were proving a theorem*! I followed his advice and, lo and behold, it worked: proving a theorem requires, after all, to look at the problem from a variety of perspectives and to expend lots of patience.

Of course, Simon would not have suggested this metaphor to a Bourbakist or an unadulterated Hilbertian Formalist, but only to a proof-theorist like Sieg, who would have thought of a proof as a procedure, in the same sense in which Simon extolled the virtues of procedural decision making in all aspects of administrative and economic life. Incidentally, Sieg's colleague in the philosophy department at CMU, Teddy Seidenfeld, is the current holder of the *Herbert Simon Professorship of Philosophy and Statistics*.

[2] Of 'Charnes-Cooper' fame and coincidentally also the man who helped establish the School of Urban & Public Affairs at CMU, serving also as its Dean from 1969 to 1975. Simon had met 'Bill' Cooper in his very early period as an undergraduate at the University of Chicago, where they lived in the same dormitory.

day in 1937. Their daughter Kathie was born in 1942, followed by son Peter in 1944 and another daughter Barbara in 1946.

From the public elementary and high schools in Milwaukee, he won a full scholarship ($300 per year) to the University of Chicago, taking the exam in Physics, Mathematics and English. An early intellectual influence was his maternal uncle Harold Merkel, "an ardent formal debater [whom] I followed in that activity too", who had died young, at the age of 30, in 1922. Uncle Harold had graduated with distinction in Law from the University of Wisconsin, having also studied economics under the legendary John R. Commons and leaving behind copies of The Federalist Papers and William James's Psychology in the family library, both of which were devoured by the young Herbert, leaving indelible impressions on the future civil libertarian, behavioural economist, computer scientist and cognitive psychologist. His first publication, whilst still in grade school, was a letter to the Editor of the Milwaukee Journal, defending atheism.

To buttress his debating skills he began to read widely in the social sciences. Two books in particular were decisively influential: Richard T. Ely's *Outlines of Economics* (1893) and Henry George's *Progress and Poverty* (1882). By the time he was ready to embark upon a university career, he had developed a clear sense of the general direction he intended to take in his studies. He would devote himself to becoming a "*mathematical social scientist*". We cannot imagine anyone, not even Kenneth Arrow, other than the Herbert Simon he became, encapsulating and wearing this mantle with more grace and justification.

He obtained his BA in Political Science from Chicago in 1936 and a PhD in 1943. He decided to major in Political Science because his first choice of major, Economics, required him to take an obligatory course in accounting, which he detested.

The undergraduate-term paper, written for graduation, led to a research assistantship at the Milwaukee City Government in the field of Municipal Administration, which in turn led to a Directorship at the Bureau of Public Measurement in the University of California at Berkeley, from 1939 to 1942. For Milwaukee he undertook a study of how the municipal employees made budget decisions, for example, when deciding between planting trees and hiring a recreation director. From this work grew his PhD thesis that, subsequently, became one of the fountainheads for the whole field of organisation theory: **Administrative Behavior**, published first in 1946 and still in print.

Simon's main intellectual impulses during the Chicago years came from Henry Schultz in mathematical economics and econometrics (still in their early years), who was also a mentor, from Rudolf Carnap, in the philosophy of science, from Nicholas Rashevsky, in mathematical biophysics and from Harold Lasswell and Charles Merriam, in political science.

Simon observed, when writing an appreciation for his almost lifelong collaborator and co-Turing prizewinner Allen Newell, that the four great questions of human intellectual endeavour are those on the *nature of matter*, the *origins of the universe*, the *nature of life* and the *working of the mind*. There is little doubt that he himself devoted the whole of his professional life to various aspects of the problem of *the working of the mind*. How does the mind perceive the external world? How does perception link up with memory? How does memory act as a reservoir of information and knowledge in interacting with the processes that are activated in human decision-making, in individual and social settings? In short, *human problem solving* (the title of his 1972 book with Newell): in the face of *internal constraints* emanating from the working of the mind; and *constraints*, imposed on its workings, by the *external*, perceived, world.

Much is made these days of 'ecological rationality' and 'ecological cognitive computing' (cf. Vernon Smith, 2008 and Andrew Wells, 2006), by which is meant that any internal constraints on the working of the mind should be taken in conjunction with the 'external constraints of the environment' in which the mind is situated for its interaction with the external world. Simon's definition of *Bounded Rationality*[3], from the very outset, was to encapsulate both of these aspects of the workings of the mind, in its rational, decision making, incarnation. This kind of shoddy scholarship, reinventing the square wheel, was even 'better' displayed by the purveyors of current versions of behavioural economics, who attribute the genesis of this field to a 1980 paper by Richard Thaler (1980)[4].

[3] The term – although not the concept – was introduced in the Introduction to Part IV of Simon (1957), p. 198. The analytical content, together with its conceptual underpinnings, were fully developed in Simon (1955 & 1956).

[4] Three of the undisputed frontier researches in '*modern*' behavioural economics, Colin Camerer, George Lowenstein and Matthew Rabin (2004), in the *Preface* to **Advances in Behavioral Economics** state:

> "Twenty years ago [i.e., 1984], behavioural economics did not exist as a field……Richard Thaler's 1980 article '*Toward a Theory of Consumer Choice*', of the remarkably open-minded (for its time) ***Journal of Economic Behavior and Organization***, is considered by many to be the first genuine article in ***modern*** behavioural economics." (pp.xxi-xxii; underlining added).

Nothing and no one in the burgeoning field of modern behavioural economics seem to have ever underpinned any of their theories on a model of computation. Thus they are unable to comprehend the nature of the decision problem framework, intrinsically framed with an underlying algorithmic basis, within which Simon first advanced, and then developed, *boundely rational* and *satisficing* decisions.

George Polya's influential little book *How to Solve It?* (Polya, 1945) introduced generations of students to *heuristics - the art of guided search*. Simon and Newell felt that the Polya framework provided a starting point for investigating, *experimentally*[5], the creative aspects of the workings of the mind in two formally and rigorously definable areas - *human problem solving* and in *the art of discovery*. From lessons that could be learnt in understanding the formal aspects of *human problem solving* they felt they could move on to more ambitious tasks: to an understanding of *human thinking*, in general. From there it would, then, be a natural step, even if not an easy one, to a formal understanding of the underpinnings of *human decision-making* in general.

As for the *art of discovery*, based on *heuristics* as guided search, that path led Simon to develop *Models of Discovery*, resurrecting the Charles Peirce-Norwood Russell Hanson[6] emphasis on *Retroduction* (or *Abduction*; and, thus, circumventing the worn out dichotomy between induction and deduction in scientific discovery) and, simultaneously, taking a well

---

This claim cannot be described as anything other than 'scandalous', especially since the variant of behavioural economics they refer to as modern – which has no computational underpinnings whatsoever – had, in fact, its noble origins in the work of Ward Edwards (1954, 1961), himself a teacher of Kahneman and Twersky. To avoid confusion with this kind of non-computationally underpinned behavioural economics we refer to the Simon-inspired variant as *Classical* Behavioural Economics.

[5] In the sense of exploring by computer simulation, guided by programmed heuristics. It may be useful to point out that heuristics, for Simon, were algorithms that were not necessarily constrained by the *Church-Turing Thesis* of Computability Theory. In this sense his lifelong adherence to *heuristical methods* in human problem solving, models of discovery, design of organisations and evolutionary dynamics, had more in common with the notion of *algorithms as proofs in constructive mathematics*. A lack of appreciation of this subtle difference may have been the reason for even the great Hao Wang to be critical of the way Newell, Shaw and Simon automated proofs in **Principia Mathematica**, in contrast to the way he had done the same (Wang, 1970, p. 227; italics added):

> 'There is no need to kill a chicken with a butcher's knife. Yet the net impression is that Newell-Shaw-Simon failed even to kill the chicken with their butcher's knife. … To argue the superiority of '*heuristic*' over *algorithmic methods* by choosing a particularly *inefficient algorithm* seems hardly just."

Even the 'gods nod', from time to time, and it may be that Hao Wang was a bit hasty in his appeal to the analogy of the butcher's knife! What is a heuristic, if not an algorithm?

[6] In particular, *Patterns of Discovery* by Norwood Russell Hanson (1958).

aimed attack on the nihilism in Karl Popper's stance that there was no scientific basis for a '*logic of scientific discovery*' (see, in particular, Simon, 1977, chapter 5.4, & Langley, Simon, Bradshaw and Zytkow, 1987).

Simon was a member of the Cowles Foundation for Economic Research in its early Chicago days, before its decisive move to Yale in New Haven; he was also a member of the Rand Corporation in its glory days, the early 1950s. The former nurtured, in Simon's own words, the econometric "mafia"; the latter fostered the mathematical economics "mafia". To his considerable distinction he preserved the courage of his convictions and remained a gadfly inside these citadels of orthodoxy whilst enjoying the respect, perhaps even the envy, of his distinguished and eminent peers.

His contributions to formal and traditional economic theory - both to micro and macro variants - and to econometric theory were fundamental and path-breaking. At a very early stage in the mathematisation of economics he deduced, in joint work with the mathematician David Hawkins (Hawkins & Simon, 1949) conditions for stability, which came to be known as the *Hawkins-Simon conditions* in the folklore of the subject, for linear multisectoral models of the economy[7]. This led to an amusing episode with the House Un-American Committee hearings, during the "McCarthy era", because Hawkins - whom Simon had never met and with whom he had written the famous paper entirely by correspondence - was a paid-up member of the Communist Party.

During the Great Depression Simon had seen a chart on the walls of his father's study, tracking the dismal progress of a faltering American economy. This chart was constructed on the basis of a model of the macroeconomy and its flows built on the principles of *servomechanism theory*, using hydrodynamic analogies. It had been devised by an imaginative engineer, with a doctorate in sociology, A.O. Dahlberg[8]. Simon had begun, inspired by the memories of the Dahlberg chart in his father's study, to look at the economy,

---

[7] Incidentally, the *Hawkins-Simon stability conditions*, when invoked as *productivity* conditions in linear, multisectoral, production models are a kind of 'complement' to the *Perron-Frobenius theorem(s)* for indecomposable matrices. These 'complementarities' appear as obvious constructs in the context of Leontief matrices and, more particularly, in the orthodox formalisation of aspects of Sraffian models. In the latter case, in particular, in the construction of the *Standard System* for an *indecomposable* production system.

[8] We think it likely that the chart on the walls of Simon's father's study was one of the many in Dahlberg's unusual book, *Jobs, Machines and Capitalism* (1932; 1969). We also think that Dahlberg updated and revised this approach in his later, much more explicitly 'graphical' work, *National Income Visualized: A Graphic Portrayal of How Economic Activity is Measured* (Dahlberg, 1956).

particularly the macroeconomy, from the point of view of the theory of servomechanisms and feedback control. This line of research led him to his celebrated results on *certainty equivalence* in the devising of optimal policy in decisions on production scheduling in firms. He did not pursue the servomechanism metaphors for too long, because he did not feel that they gave additional insights or empirical levers that could not be got by the mathematical analysis underlying their structures. He also felt, by then, that *analogue simulations* were a distinct second best to the *digital* possibilities he was pioneering.

The origins of the inspiration that led to the influential work with his eminent Japanese student Yuji Ijiri on the *size distributions* of the growth and decay of business firms and organisations are narrated with humour and candour in his charming autobiography, *Models of My Life* (1991). It is also a tale of academic bloody-mindedness, recounted without rancour, and revisited with nostalgia and regrets on the fallibility of memory.

In 1946 Richard Goodwin (Goodwin, 1947) had begun his own lifelong research programme of interpreting economic agents, markets and the economic system as (nonlinear) oscillators. His earliest paper on this subject analysed *markets as coupled oscillators* with *hierarchies* of coupling strengths: some markets weakly coupled; others strongly coupled. All of them linked by economy-wide, common, expenditure impulses. In a remarkable series of papers, extending over half a century, Simon exploited this simple idea in all sorts of fertile ways: to study *causality* in economic models; to formalise causality and link it with i*dentifiability* in econometric models; to theorise about *aggregation* in economic models; to formalise the idea of the *hierarchy of complexity* utilising *near-decomposability* in hierarchical organizations[9]; to study counterfactuals in scientific theorizing, and on using the idea of near-decomposability to study evolutionary aspects of mind, thought, organizations and nature[10].

[9] This is most clearly argued and presented in the supremely pedagogical essay on *The Architecture of Complexity* in Simon, 1996 (chapter 8, pp. 173-216)

[10] Serendipities galore in this story! The idea of 'near decomposability' that Simon extracted from Goodwin's notion of unilateral coupling in markets with production lags was made mathematically rigorous by the idea of approximately decomposable (or indecomposable) matrices. The link between indecomposability and the Peroon-Frobenius theorem(s) was mentioned, above, in footnote 7. It was in Richard Goodwin's own pioneering work on multisectoral dynamic models that both the link and, indeed, the use and introduction of the Perron-Frobenius theorem(s) first appeared in his contributions of the late 1940s and early 1950s (cf., Velupillai, 1998). The Goodwin-Simon analytical nexus straddled both interindustrial and macroeconomic dynamics, in that Simon's work on applying servomechanism theory to the modelling of macroeocnomic systems was also inspired by the former's classics on aggregate nonlinear macrodynamics (cf., Goodwin, 1951 & Simon, 1952).

In *Models of My Life*, he takes this particular example of the inspiration he got from Goodwin's remarkable attempt to represent markets interacting with delayed responses as hierarchically coupled oscillators *to wonder about the kinds of representations scientists use in thinking about research problems*: where do the metaphors for scientific representations come from? How are they represented and retained in the human mind? How are they recalled - and when and why at that particular juncture? What are the triggering mechanisms and the catalysts? He has answers, tentative, testable and, as always, interesting and provocative.

In his *Raffaele Mattioli Lectures* of 1993 (published as *An Empirically Based Microeconomics*, 1997), he made the case for study of economies in terms of organisations as the basic unit rather than the traditional device of markets. His case is many-pronged and based on solid empirical and theoretical results. He felt - but, again, justified by empirical and experimental data and results - that the reliance on markets led to the unnecessary and false claims for their optimality properties (true only in one of many possible mathematical worlds and false in all others) as well as the propagation of the false dichotomy between the virtues of decentralisation and the vices of centralisation, without forgetting the merits of the former and the disadvantages of the latter.

But he was optimistic about the future of economics and even more so of computer science and the interaction between the two and psychology. In a letter to a former colleague, written after reading Velupillai (2000), he wrote that he thought that "the battle has been won, at least the first part, although it will take a couple of academic generations to clear the field and get some sensible textbooks written and the next generations trained". Much to our regret, we feel his optimism was premature.

Herbert Simon did not accept criticism willingly - not often even gracefully. He felt that he had reached his convictions, beliefs and the stands he was taking at any time after deep theoretical studies, and seriously and carefully designed experiments to harness empirical data to substantiate those theories on which he was betting. The least he expected from his critics and detractors was an equal commitment to the scientific enterprise. On the other hand he was exceptionally generous to junior colleagues and graduate students - with his time, with his advice and with his patronage. He was a passionate liberal and debunked any and every kind of pretence - in scholarship, in manners, in attitude and in interpersonal relationships.

At the time of his death he was the Richard King Mellon Professor of Computer Science and Psychology at Carnegie-Mellon University, a post he has held since 1966.

The precept that may have guided his astonishingly fertile scientific life may well have been, in his own words (Simon, 1996, p. 28; italics in the original):

> 'What a person *cannot* do he or she *will not* do, no matter how strong the urge to do it.'

REFERENCES

Camerer, Colin. F, George Loewenstein and Matthew Rabin (2004; editors), **Advances in Behavioral Economics**, Princeton University Press, Princeton, New Jersey.

Dahlberg , Arthur (1932; 1969), **Jobs, Machines, and Capitalism**, AMS Press, New York.

Dahlberg, Arthur (1956), **National Income Visualized: A Graphic Portrayal of How Economic Activity is Measured**, Columbia University Press, New York.

Edwards, Ward (1954), *The Theory of Decision Making*, **Psychological Bulletin**, Vol. 51, No. 4, pp. 380-417.

Edwards, Ward (1961), *Behavioral Decision Theory*, **Annual Review of Psychology**, Vol. 12, pp. 473-498.

Goodwin, Richard. M (1947), *Dynamical Coupling with Especial Reference to Markets Having Production Lags*, **Econometrica**, Vol. 15, No. 3, July, pp. 181-204.

Goodwin, Richard M (1951), *The Nonlinear Accelerator and the Business Cycle*, **Econometrica**, Vol. 19, #1, January, pp. 1-17.

Hanson, Norwood Russell (1958), **Patterns of Discovery**, Cambridge University Press, Cambridge.

Hawkins, David & Herbert A. Simon (1949), *Note: Some Conditions of Macroeconomic Stability*, **Econometrica**, Vol. 17, No. 3/4, July-October, pp. 245-8.

Langley, Pat, Herbert A. Simon, Gary L. Bradshaw & Jan M. Zytkow (1987), **Scientific Discovery: Computational Explorations of he Creative Process**, The MIT Press, Cambridge, Massachusetts.

Polya, George (1945), **How to Solve It: A New Aspect of Mathematical Method**, Princeton University Press, Princeton, NJ.

Polya, George (1945, 2nd edition, 1957): **How to Solve It – A New Aspect of Mathematical Method**, Princeton University Press, Princeton, New Jersey.

Simon, Herbert A (1952), *On the Application of Servomechanism Theory in the Study of Production Control*, **Econometrica**, Vol. 20, No. 2, April, pp. 247-268

Simon, Herbert. A (1955): *A Behavioral Model of Rational Choice*, **Quarterly Journal of Economics**, Vol. 69, No.1, pp. 99-118.

Simon, Herbert. A (1956): *Rational Choice and the Structure of the Environment*, **Psychological Review**, Vol. 63, pp. 129-38

Simon, Herbert. A (1957): **Models of Man: Social and Rational**, John Wiley & Sons, Inc., New York.

Simon, Herbert. A. (1977): **Models of Discovery**, D.Reidel Publishing Company, Dordrecht, Holland.

Simon, Herbert. A (1991): **Models of My Life**, The MIT Press, Cambridge, Massachusetts.

Simon, Herbert A (1996), **The Sciences of the Artificial** (Third Edition), The MIT Press, Cambridge, Massachusetts.

Simon, Herbert. A (1997): An Empirically Based Microeconomics - *The Raffaele Mattioli Lectures* (Delivered on 18th and 19th March, 1993), Cambridge University Press, Cambridge.

Smith, Vernon. L (2008), **Rationality in Economics: Constructivist and Ecological Forms**, Cambridge University Press, Cambridge.

Thaler, Richard (1980), *Toward a Positive Theory of Consumer Choice*, **Journal of Economic Behavior and Organization**, Vol. 1, #1, pp. 39-60.

Velupillai, Kumaraswamy (2000), **Computable Economics**, Oxford University Press, Oxford.

Velupillai, Kumaraswamy (1998), *Richard Goodwin: 1913-1996*, **The Economic Journal**, Vol. 108, September, 1998, pp.1436-1449.

Wang, Hao (1970), **Logic, Computers and Sets**, Chelsea Publishing Company, New York.

Wells, Andrew (2006), **Rethinking Cognitive Computation: Turing and the Science of the Mind**

APPENDIX

K. Vela Velupillai

The Logic of Discovery, Problem Solving and Retroduction♠

> "If we are willing to regard all human complex problem solving as *creative*, then .. successful programs for problem solving mechanisms that simulate human problem solvers already exist, and a number of their general characteristics are known. If we reserve the term '*creative*' for activities like the discovery of the special theory of relativity or the composition of Beethoven's Seventh Symphony, then *no example of a creative mechanism exists at the present time*."
>
> Newell, Shaw and Simon, 1962 [Simon, 1979, p.145; italics added].

Herbert Simon did not mince his words or his views on Popper's nihilistic vision for a *Logic of Scientific Discovery*:

> "It is unusual for an author, less than one-tenth of the way through his work, to disclaim the existence of the subject matter that the title of his treatise announces. Yet that is exactly what Karl Popper does in his classic, The Logic of Scientific Discovery, announcing [it] in no uncertain terms on p. 31…."
>
> Simon, 1977, p. 326.

Simon countered Popper's provocative claim that there was no 'logic of (or 'to') scientific discovery', by providing a *computational theory of scientific discovery*. He did this by resurrecting a tradition that came down from Plato's *Meno* and Archimedes's essay on *Method*, through Bernard Bolzano's *Wissenschftslehre* and Peirce's innovative ideas, via George Polya's multifaceted visions and theories of problem solving and Hanson's *Patterns of Discovery*.

Popper's Challenge, taken up and resolved - in my opinion decisively - by Simon via his *Models of Discovery* (Simon, 1977, Langley et. al., 1987), was:

> "[M]y view of the matter, .. , is that there is no such thing as a logical method of having new ideas, or a logical reconstruction of this process. My view may be expressed by saying that every discovery contains an 'irrational element', or 'a creative intuition', in Bergson's sense."
>
> Popper, 1959, p.32[11].

♠ Norwood Russell Hanson notes, in Patterns of Discovery (Hanson, 1958, p.85):

> "Aristotle lists the type of inferences. These are deductive, inductive and one other called [*apagoge*] … Peirce translates is as 'abduction' or 'retroduction'."

[11] The above paragraph goes on as follows:

> "In a similar way Einstein speaks of the 'search for those highly universal laws .. from which a picture of the world can be obtained by pure deduction. There is no logical path', he says, 'leading to these ...

We have the apparent paradox of Simon, conventionally associated with psychological theories of human rational behaviour, in the 'red corner', defending the possibility of a *Logical Theory of Scientific Discovery*. In the 'blue corner' we find, surprisingly, that implacable enemy of all kinds of psychologism, claiming that there can be no logic of discovery.

Are these paradoxes? I think not. Simon was the supremely rational artificer; it is just that he never forgot to take into account psychological constraints on behaviour. Logic did not develop independently of psychology and its mechanisms – except in Vienna in those heady interwar years, when the two lived side by side and did not communicate with each other. Schlick, Neurath and others, who set the themes that Popper thought he was crucifying, on one side; Freud and his followers and patients on all the other sides. Simon's visions were not totally blinded by Carnap, his teacher at Chicago. Like all students who went beyond their masters, Simon took the grounding Carnap gave him and turned into a quest for a resurrection of that noble tradition of a search for a theory of the *logic of discovery* that began with Plato's *Meno*.

Simon noted, in his fascinating collaborative work on *Scientific Discovery*[12] that:

> "A hypothesis that will be central to our inquiry is that the *mechanisms of scientific discovery* are not peculiar to that activity but can be subsumed as special cases of the general *mechanism of problem solving*."
>
> Langley, et.al, op.cit., p.5; italics added.

Simon came to adopt the problem solving vision – that vision which placed *heuristics* at the center of the process of analyzing discoveries, problems, creativity, experimentation etc. - as his scientific credo after the initial impetus garnered from a reading of Polya's wonderful little book, *How to Solve It* (Polya, 1945)[13]. The work that went into *Administrative Behavior* had been completed by the time Polya captivated him. But the way he had approached the problems of administering an organization had sown the seeds of a problem solving approach to the logic of human decision making in individual, social and organizational contexts. Hence the immediate impact and, subsequently, a lifelong commitment to, a problem solving approach to behaviour, discovery, hierarchies and much else.

> laws. They can only be reached by intuitions, based upon something like an intellectual love of the objects of experience."

Clark Glymour, in his commemorative essay in honour of Alan Turing (Glymour, 1966, p.269), noted, with admirable irreverence and impeccable accuracy, that:

> For Popper – who quite confused a psychological question with a mathematical issue – it sufficed to quote Einstein to disprove the possibility of a discovery algorithm; for Carnap it sufficed to quote Popper quoting Einstein."

Simon, as noted above, had been Carnap's pupil in his Chicago student days.

[12] With the unambiguous subtitle (*pace* Popper!): *Computational Explorations of the Creative Process*.

[13] In his '*Biographical Memoir*' on Allen Newell for the National Academy of Sciences, Simon recalled that (italics added):

> "Polya's widely read book, *How to Solve It*, published in 1945, had introduced many people (*including me*) to heuristic, the art of discovery."

So, what is *problem solving*? Indeed, *what is a problem*? In all my own unstructured search for a 'definition' of this crucial concept I have never found anything to better Polya's direct, intuitively clear and logically impeccable definition:

> "[T]o have a problem means: *to search consciously for some action appropriate to attain a clearly conceived, but not immediately attainable, aim.* To solve a problem means to find such action."
>
> Polya, 1981, Vol.1, p.117; italics in original.

'Action' is what is usually referred to as a 'method', an 'algorithm', a 'mechanism' to solve a problem. Now, I could ask, what other way can we define a method than via a Turing Machine? I shall go even further and assert the following, challenging, in the process, the ghost of Popper and any of his living acolytes: give me whatever you call a method and I will show you that it is either a *Turing Machine* or, what I shall call, a *Constructive Machine*. Any other device, designed to solve a problem, will have to rely on magic, the mind of God, metaphysics or some such appeal to an oracle.

Why don't I confine the '*computational explorations of the creative process*' to the Turing Machine model? Not because I believe Simon went beyond the Turing Machine model. But because Simon enquired, seriously and systematically, whether there was a need, in theory or dictated by the actual, observable, activity of the practicing scientist, for a special logic of scientific discovery and problem solving. By 'a special logic', in his enquiries, he meant, whether it was necessary to go beyond the standard logic of the declarative, by appeals to an imperative, modal or deontic logic. His answers were emphatically and clearly in the negative in every such case (cf., in particular, chapters 3.1, 3.2 and 4.3, in Simon, 1977)[14]. His conclusion in every such enquiry and examination was that the ordinary declarative logic was sufficient and, at most, there was a need to transform a well specified collection of imperatives to a declarative form during any action (i.e, implementation of a method to solve a problem).

The reason I adjoin *Constructive Machines* to the *Turing Machines* as an underpinning for Simon's *computational explorations of the creative process* is that Kolmogorov showed, in a beautiful demonstration, the validity of the following remarkable proposition:

> "The *calculus of problems* is formally identical with the Brouwerian intuitionistic logic …|
>
> Kolmogorov, 1932 [1998, p.328]; italics added.

And, therefore:

---

[14] I had myself felt that von Wright's way of using *Deontic Logic* was a promising avenue for a theory of problem solving and, hence, also human decision making in social contexts. But I had found it almost intractably difficult to make the formalisms of *Deontic Logic* numerically operational via the Turing Machine model and was about to give up the endeavour when a reading of Simon pushed me decisively over the edge! As Simon pointed out (ibid, ch.3.2, p.154):

> "Approaches to the theory of action have often taken their starting point in the modalities of common language: in the meanings of 'ought' and 'must' and 'can'. But because common language is complex, flexible, and imprecise, travellers on this path have encountered formidable difficulties."

'The theory of action' simply means 'methods for a solution of a problem'.

> "[O]ne should consider the solution of problems as the independent goal of mathematics (in addition to the proofs of theoretical propositions)."
>
> ibid, p. 333; italics added.

Kolomogorov did not attempt to define a 'problem' or what was entailed in 'solving a problem' (ibid, p.328; italics in original):

> "We do not define what a *problem* is; rather we explain this by some examples."[15]

Kolmogorov's examples satisfy Polya's definition given above. The reason Kolmogorov was able to identify Brouwerian *Intuitionistic Logic*[16] in the formalism of the calculus of problems is simple. The formal mathematics that is erected on the foundations of intuitionistc logic is some version of constructive mathematics, where there is a special status for existential propositions: they must come with a method of constructing the object. Problem solving, as Polya has defined it, a definition that Simon embraced, is exactly such an activity: to *find* in a particular set an object with prescribed properties. It is not enough to state that such an object exists. This *activity of finding* is the equivalent of the method of solution, the constructive process or the Turing Machine. It is also fundamental and pervasive in *Administrative Behavior* and its subsequent codification in behavioral economics.

Thus, when Simon quite correctly states, at the end of his didactic survey of *The Logic of Heuristic Decision Making* (Simon, 1977, ch.3.2, p.174):

> "Our brief survey … demonstrates again the central thesis of this paper: ordinary mathematical reasoning, hence the ordinary logic of declarative statements, is all that is required.."

But he was like M. Jourdain in Molière's *Le Bourgeois Gentilhomme*, speaking prose all his life without realising it. Simon was a natural practitioner of a constructive mathematics and the 'ordinary logic' of constructive mathematics could have been given an intuitionistic underpinning, if required, but like Errett Bishop, he did not see the need for it – and he was absolutely right. Simon simply practised computable or constructive mathematics and neither of them requires anything more than a refined and nuanced form of declarative statements for their practice to be made numerically operative in the form of algorithms. I would even be bold enough to suggest that a careful reading of all the examples of problem solving that Simon and his collaborators analyzed would reveal that they did not necessarily accept the restrictions of the Church-Turing Thesis. But that is another story.

The examples of the way in which the *rules* behind discoveries and problem solving are learned or revealed should be viewed against a particular backdrop of traditions which were

[15] Perhaps Courant (and Robbins) were following in Kolmogorov's footsteps in their elegant and influential book titled <u>*What is Mathematics*</u>. There is no attempt to define mathematics in this book; the reader is expected to infer what it is that may well be mathematics by examples (Courant-Robbins, 1958).

[16] This may be the appropriate place to refer to Popper's failed attempt at a '*criticism of Brouwer's epistemology*' (cf. Popper, 1972, pp. 128-140). I have been surprised that serious intuitionistic mathematicians have not taken Popper to task for the distortions on, and of, Brouwer that permeate this essay, full of false assertions, misleading and out of context references and plain incorrect mathematics.

the explicit or implicit settings for much of Simon's work in these areas. Plato's Meno, the essay on Method by Archimedes, Bolzano's Theory of Science, the works of Peirce and Polya in almost their entirety and Norwood Russell Hanson's Patterns of Discovery are the precursors. They, each of them, provide an impetus or a background for a tradition that Simon himself nurtured and refined.

The Socratic dialogue in Meno initiates a noble Western tradition[17] of inquiry into the *processes* of arriving at knowledge and understanding of concepts independently and intuitively perceived. The dialogue tries to *discover* the *rule*, if any, that may characterize concepts in such a way that it will also be recognized as such. In other words, the *rule*, will not only define the concept but also provide a *method* to assure the inquirer that the application of the method will also provide a means to *recognize*, at its termination, that it results in an object with the attributed characteristics[18].

Such a non-standard interpretation puts Simon squarely in a particular Socratic epistemological tradition of *learning*, *heuristics* and *problem solving*. Traditionaly, Meno has been interpreted to have provided a theory of learning – knowledge as recollection and remembering; the paradox of recognizing something one knows, the obliteration of problem of *problem solving*. According to the traditional interpretation, to which Polanyi adheres, there is no question of problem solving if one knows - meaning, if one *knows* that the object one is looking for *exists* - what one is looking for. Neither *how* one comes to *know*, nor by what *means* one locates where it exists, is relevant. Simon challenges this interpretation in Simon (1977, ch.5.5) and in demonstrating its fallaciousness gives the method that he came to adopt, generally, in all problem solving contexts: *the model of computational exploration of the creative process* – the Turing Machine Model.

No careful reading of the relevant parts of *Meno* could even remotely substantiate Polanyi's interpretations. Just for the record, and to make my case that the *Meno* supports Simon's interpretations, even of learning as problem solving, I quote from Jowett's translations (Jowett, 1892; the sources used by Simon himself). In perusing the parts of the dialogue between Meno and Socrates that I select to quote here, a few pointers might be of use. First of all, note that Socrates first castigates Meno for confusing an enumeration with a definition; but, then, immediately goes on to define *the limit of an enumeration* as the definition that is sought. This is precisely a case of learning as somewhere between induction – the Gold, Putnam view – and retroduction, the Simon view[19]. The definition sought for is determined

---

[17] I believe this qualification is essential for a variety of reasons, not least because there are other, equally rich traditions of inquiry and epistemology that have equal claim on scholarship. In particular, within my own cultural and educational background, Buddhist epistemology provides a rich source of epistemological precepts of learning via inquiry dialogues as thought experiments (cf. Jayatilleke, 1963, and the dialogue in the Tamil Epic, *Manimekalai*).

[18] Kolmogorov made this point quite unambiguously (ibid, p.330; second emphasis added; cf also f.n 8 to this citation on p.334):

> "If I have solved a logical or a mathematical problem, then I can present this solution in a way that is intelligible to all and it is *necessary* that it be *recognized* as a correct solution although this necessity has to a certain extent an ideal character, for it presupposes a sufficient intelligence on the part of the listener."

[19] Glymour (op.cit, p.271) is the clearest articulation of this view, explicitly within the context of a discussion of the place of Plato's *Meno* in Western epistemological traditions:

> "The epistemological idea about knowledge in the limit is implicit in many contexts in the twentieth century. .. But the articulation of the idea came almost simultaneously in the 1960s from two independent sources. Hilary Putnam and E.Mark Gold. It seems likely that Putnam took the idea from

by a limit process; without the latter, the knowledge claimed of the thing defined is incomplete. Devising a (constructive) limit process defines the object. Thus:

> "*Socrates*: How fortunate I am, Meno! When I ask you for one virtue, you present me with a swarm of them, which are in your keeping. Suppose that I carry on the figure of the swarm and ask of you, What is the nature of the bee? And you answer that there are many kinds of bees, and I reply: But do bees differ as bees, because there are many and different kinds of them; or are they not rather to be distinguished by some other quality, as for example beauty, size, or shape? How would you answer me?
>
> *Meno:* I should answer that bees do not differ from one another, as bees.
>
> *Socrates*: And if I went on to say: That is what I desire to know Meno; tell me what is the quality in which they do not differ, but are alike; - would you be able to answer?
>
> *Meno*: I should.
>
> *Socrates*: And so of the virtues, however many and different they may be, they have all a common nature which makes them virtues; and on this he who would answer the question, 'What is virtue?' would do well to have his eye fixed: Do you understand?"
>
> [Jowett, 1892, p.29][20].
> "Socrates: Do you not understand that I am looking for the 'simile in multis'"
>
> [ibid, p.32].
>
> "Socrates: Well then, you are now in a condition to understand my definition of figure. I define figure to be that in which the solid ends; or more concisely, *the limit of solid*."
>
> [ibid, pp.33-4].

For Polanyi, on the other hand, we know, *tacitly*, like Meno, what *virtue* is. We *know* what we are looking for; hence there is no problem situation. This is a crippled definition of a problem.

But it is in that supreme empirical scientist, the true precursor of Simon, Archimedes and his practices that one finds the later *Epicurean Adventurer*'s[21] method articulated most

---

> Hans Reichenbach and combined it with reflections on Turing's conventions for output of a computing machine."

My own views on this problem had been formed before I was motivated to read *Meno* for the preparation of this essay. In Velupillai (2000, chapters 5 & 6) I did advocate Gold's method and an earlier Putnam vision, that which was from his remarkable and almost totally unknown doctoral dissertation written under Reichenbach (cf. Putnam, 1951[1990]). I was, unfortunately, totally unaware of the possibilities of the *retroductive* method at the time I was preparing *Computable Economics* and, hence, relied almost entirely on inductive methods for learning and discovery problems.

[20] On this part of the dialogue, Jowett, in his introduction (p.3; italics added) perceptively observed:

> "Socrates reminds Meno that *this is only an enumeration* of the virtues and not a definition of the notion which is common to them all."

[21] i.e., Simon!

persuasively. Archimedes 'divided and conquered' when faced with problem solving: 'first a *heuristic* mechanical method to *discover* the answer to a problem before looking for a demonstration [i.e., proof] of it' (Crombie, 1994, p.176; italics added). As described in *Method* his strategy of *computationally exploring the discovery process* was as follows:

> "I thought it fit to write out … the peculiarity of a certain method, by which it will be possible .. to get a start to enable [one] to investigate some of the problems in mathematics by means of mechanics. This procedure is, I am persuaded, no less useful even for the proof of the theorems themselves; for certain things first become clear .. by a mechanical method, although they had to be demonstrated [proved] by geometry [the axiomatic method] afterwards because their investigation by the said method did not furnish the actual demonstration [proof]. But it is of course easier, when we have previously acquired by the method some knowledge of the questions, to supply the proof, than it is to find it without any previous knowledge."
>
> Quoted in Crombie, op.cit., p.176

Simon elevated this tradition to a new level of rigour and routine, which enabled the determined novice to traverse a path towards the status of an expert, so far as discovering rules were concerned in a learnable way and, therefore, also a teachable way. This latter point, the need to learn and teach methods of inquiry into the nature of knowing concepts, was also crucial in the dialogue between Socrates and Meno.

Archimedes felicitously separated the question of *discovery* from that of *proof*; the former could be done experimentally – both in the classical sense and in the sense of thought experiments. For the latter, he used both the technique of proof by *reductio ad absurdum* and *ad contradictionem*. It was only with the formalism of problem solving that this separation was healed. But that had to wait till Brouwer challenged the latter day Platonists in the early 20^{th} century, Kolmogorov re-integrated solvability with proof and Church and Turing put it all together in one fell swoop as a computational paradigm.

There were many who lit the path between Archimedes and Simon, none more so than that father of the *Bohemian Enlightenment*, Bernard Bolzano. He resurrected the Aristotelian triple division of methods of inquiry into induction, deduction and retroduction. The latter as the *apagogic procedure* identical with *reductio ad absurdum* in the context of what he called *HEURETIC* (cf. Bolzano, 1972, Books Four and Five). This became, in the hands of Polya and Simon, *heuristics* – the art of guided search in a complex space.

One part of Bolzano's methodology that is reflected in Simon's practice comes via the influence the former had on Polya's approach to problem solving. The other part of the influence, perhaps via Norwood Russell Hanson, was in Simon identifying the *logic of discovery* with the *logic of retroduction* in his forceful and uncompromising critique of Popper's nihilism regarding the feasibility of a theory and logic of discovery:

> "This mystical view [i.e., Popper's view] towards discovery …. has not gone without challenge. Peirce coined the term 'retroduction' as a label for the systematic processes leading to discovery; while Norwood Hanson, in his *Patterns of Discovery*, revived that term and gave us a careful account of the retroductive path that led Kepler to the elliptical orbits of the planets. …

> Hanson made his case for retroduction by examining historical examples of scientific discovery. He did not propose an explicit formal theory of the retroductive process ..
> It is the aim of this paper to clarify the nature of retroduction, and to explain in what sense one can speak of a 'logic of discovery' or 'logic of retroduction'. Like Hanson, I shall proceed from examples of retroductive processes…"
>
> Simon, 1977, ch. 5.4, pp.326-7.

Bolzano confined *his* application of the retroductive process, i.e., *the apagogic* procedure, to the *mathematical question* of ascertaining the truth of a proposition. However, 'ascertaining the truth of a proposition' was an instance of 'problem solving', which is why it entered Polya's scheme of things seamlessly. Bolzano's criterion for the solution of problems, the successful demonstration of a proposition, is predicated upon two requirements, even before the thought experiment is conceived: firstly, the proposition to be demonstrated must be chosen with care; secondly, it must be tested for its consequences in every conceivable way[22]. Simon adheres to these two strictures almost religiously and his sustained criticism of 'armchair theorizing' reflects the importance he gives to the second of Bolzano's criteria for the consideration of a proposition as worthy of attention. He goes beyond Bolzano, of course, because he goes that extra distance with the testing of propositions: not only as thought experiments; but also as components in real experiments. He also goes the extra distance in suggesting a theory for the careful selection of the relevant proposition, and that theory, too, is computational.

Bolzano's testing proceeds as follows:

> "If the mere clear representation of a proposition M does not lead to a judgement about it, or if this judgement does not appear reliable enough, the next stage in its *testing* is that we attempt to deduce, either from M alone or from M together with other already known premises, several consequences and from these further consequences, etc. … This procedure of showing the truth of proposition M, and thus of *solving the indicated problem*, is generally called the *reduction to absurdity*, or apagogic procedure. Examples are common in the mathematical sciences…"
>
> ibid, p.373; first set of italics added.

In concluding my brief attempt at chronicling and exploring Simon's adventures with the logic of discovery in terms of a theory of problem solving, I cannot summarize any more succinctly or more cogently that path he took in anything like the elegant way Richard Day succeeded in doing it:

> "[Simon] was motivated by the view that empirical, systematic, and logically consistent theory could enhance effective choice, action, and coordination in ongoing human organizations. His commitment to a scientific approach led him to consider the nature of scientific methodology, the logical content of causality, the possibility for identifying causal structure, and the foundations of human reason itself. [Simon espoused] the ancient philosophers' consilient game of crossing disciplines to create unity of knowledge. Moreover, … Simon, as much as any other theorist of his era, *built his work precisely on a close examination of the basic premises upon which*

[22] Not in the preposterous Popperian sense of attempts to falsify it.

economic theory is built; namely rational choice and the firm. He did this in terms of the psychology of individual choice and the social architecture of decision making in organizations."

Day, 2001, p. 2; italics in original.

References for the Appendix

Bolzano, Bernard (1837 [1972]): Theory of Science – Attempt at a Detailed and in the main Novel Exposition of Logic With Constant Attention to Earlier Authors, Edited and Translated by Rolf George, University of California Press, Berkeley and Los Angeles.

A.C.Crombie (1994): Styles of Thinking in the European Tradition - The History of Argument and Explanation Especially in the Mathematical and Biomedical Sciences and Arts, Vol.1., George Duckworth & Co., Ltd., London.

Courant, Richard and Herbert Robbins (1958): What is Mathematics: An Elementary Approach to Ideas and Methods, Oxford University Press, Oxford.

Day, Richard, H (1964): "A Behavioral Theory of the Firm", Econometrica, Vol. 32, No. 3, July, pp. 461-5

Day, Richard, H (2001): "Consilience and the Economic Legacy of Herbert A. Simon", Paper prepared for presentation at the Atlanta Meeting of the AEA, January 6, 2002.

Glymour, Clark (1996): " The Hierarchies of Knowledge and the Mathematics of Discovery", in: Machines and Thought: The Legacy of Alan Turing – Vol. 1, edited by P.J.R.Millican and A.Clark, Oxford University Press, Oxford, 1996.

Hanson, Norwood Russell (1958): Patterns of Discovery- An Inquiry into the Conceptual Foundations of Science, Cambridge University Press, Cambridge.

Jayatilleke, K.N (1963): Early Buddhist Theory of Knowledge, Motilal Banarsidass Publishers, Private Limited, Delhi.

Jowett, Benjamin (Ed.) (1892): The Dialogues of Plato – Translated into English with Analysis and Introductions by B. Jowett, Vol.II, 3rd Edition, OUP, Humphrey, Milford.

Kolmogorov, A.N (1932/1998): "On the Interpretation of Intuitionistic Logic", in P.Mancosu (ed.): From Brouwer to Hilbert: The Debate on the Foundations of Mathematics, Oxford University Press, Oxford, pp. 306-310.

Langley, Pat, Herbert A. Simon, Gary L. Bradshaw and Jan M. Zytkow (1987): Scientific Discovery: Computational Explorations of the Creative Process, The MIT Press, Cambridge, Massachusetts.

Newell, Allen, J.C.Shaw and Herbert.A.Simon (1962): "The Processes of Creative Thinking" in Herbert Simon: Models of Thought, Vol. 1, Chapter 4.1.,pp.144-74. Yale University Press, New Haven, 1979.

Polya, George (1945, 2nd edition, 1957): **How to Solve It – A New Aspect of Mathematical Method**, Princeton University Press, Princeton, New Jersey.

Polya, George (1981): **Mathematical Discovery – On Understanding, Learning and Teaching Problem Solving**, Vol. I and Vol. II, John Wiley & Sons., New York

Popper, Karl. R (1959): **The Logic of Scientific Discovery** Hutchinson & Co., London.

Popper, Karl. R (1972): **Objective Knowledge: An Evolutionary Approach**, Oxford University Press, Oxford, UK.

Popper, Karl. R & John C. Eccles (1977): **The Self and its Brain: An Argument for Interactionism**, Routledge & Kegan Paul, London.

Putnam, Hilary (1951[1990]): The Meaning of the Concept of Probability in Application to Finite Sequences, Garland Publishing, New York and London, 1990.

Simon, Herbert. A. (1977): Models of Discovery, D.Reidel Publishing Company, Dordrecht, Holland.

Simon, Herbert. A (1979): Models of Thought: Volume 1, Yale University Press, New Haven.

Simon, Herbert. A (1996a): The Sciences of the Artificial (3rd Edition), The MIT Press, Cambridge, Massachusetts.

Simon, Herbert. A (1997a): Administrative Behavior – A Study of Decision-Making Processes in Administrative Organizations (4th Edition), The Free Press, New York, NY.

Simon, Herbert, A: "Allen Newell: March 19, 1927 – July 19, 1992", Biographical Memoirs of the National Academy of Sciences (downloaded from the url site: http://bob.nap.edu/html/biomems/anewell.html)

Velupillai, K (2000): Computable Economics, Oxford University Press, Oxford.