Algorithmic Social Sciences Research Unit

ASSRU

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

DISCUSSION PAPER SERIES

11 – 2012/II

RECONSTRUCTING A COMPUTABLE AND COMPUTATIONALLY COMPLEX THEORETIC PATH TOWARDS SIMON'S *BEHAVIOURAL ECONOMICS*[♣]

Ying-Fang Kao[1] & K. Vela Velupillai[♥]

DECEMBER 2012

[♣] We refer to Simon's work in this field, of which he is, arguably, the founding father, as *Classical Behavioural Economics* (see footnote 1, below). We are both immensely grateful to our ASSRU colleagues, Stefano Zambelli and Ragu Ragupathy for many enlightening conversations and also to Shu-Heng Chen for inspiring support. John Davis and Cassey Lee, in their unobtrusive and learned ways, provided important inspiration, too. We are particularly indebted to our friends, Brian Hayes and John Davis, the former for bringing to our attention the important contribution of Kline (2011) and, thereby, also Crowther-Heyck (2006) and Heyck (2008a & 2008b). John introduced us to the related work of Heukelom (2012). These friends are living examples of an *interdisciplinary community*, instanced with masterly narrative skills by Crowther-Heyck (*op.cit*), for the life and times of Herbert Simon's unparalleled scientific career. None of them are remotely responsible for the remaining infelicities in this paper.

[♦] Algorithimc Social Sciences Research Unit, Department of Economics & the School of Social Sciences, University of Trento.

[♥] *Algorithmic Social Sciences Research Unit*, Department of Economics, University of Trento & Department of Economics, The New School of Social Research, New York.

Abstract

This paper aims to interpret and formalize Herbert Simon's notions of bounded rationality, satisficing and heuristics in terms of computability theory and computational complexity theory. Simon's theory of *human problem solving* is analysed in the light of Turing's work on *Solvable and Unsolvable Problems*. It is suggested here that bounded rationality results from the fact that the deliberations required for searching computationally complex spaces exceed the actual complexity that human beings can handle. The immediate consequence is that satisficing becomes the general criterion of decision makers and heuristics are the procedures used for achieving their goals. In such decision problems, it is demonstrated that bounded rationality and satisficing are more general than Olympian rationality and optimization, respectively, and not the other way about.

Keywords: Bounded Rationality, Satisficing, Heuristics, Computability, Computational Complexity

§ 1. Introduction

Crowther-Heyck's comprehensive, even masterly, series of studies on Herbert Simon's institution-building activities, coupled to a series of clearly substantiated and eminently reasonable claims on the various transitions in Simon's research focus, over a period of over a half-a-century, provide the touchstone for any future studies of this modern renaissance scientist's intellectual contributions, their institutional backdrops and their impact on *at least* three different disciplines: the cognitive sciences, economics (including management and administrative sciences) and computer science (Crowther-Heyck, 2005, 2006; Heyck, 2008a, 2008b[1]).

The well-defined theme of understanding the background mechanisms – human and institutional – that went into establishing the Graduate School of Industrial Administration (GSIA) at the Carnegie Institute of Technology, later Carnegie Mellon University (CMU), especially in terms of *funding agencies* and their priorities is narrated with admirable completeness in the above series of works by C-H and Heyck. The emergence of an active field of frontier research, interdisciplinary in essence, challenging the 'orthodox', perhaps 'monolithic', framework of economics as the dominant quantitative and mathematical social science, called the science of decision making in the behavioural sciences, had its successes when, in the immediate postwar years, internal priorities within funding agencies came to a serendipitous alignment with the intellectually fierce determination of a supremely all round scientist, equipped with vision, technique and a mastery of the fundamental conceptual foundations of a range of fields.

The theses convincingly put forward by C-H and Heyck have been updated, on the basis of equally convincing and serious archival research, by Heukelom (2012)[2] – except that this time the older funding agencies of the earlier story have been replaced by the Alfred P. Sloan and Russell Sage

---

[1] We assume that **Crowther-Heyck** (henceforth referred to as C-H) and **Heyck** refer to the same author. Our only – entirely trivial – point of disagreement is with the suggestion that Simon was awarded the *Nobel Memorial Prize in Economic Sciences* for 'work in econometrics', (C-H, 2006, p. 311) or 'for his contribution to mathematical economic analysis and to the theory of the firm' (Heyck, 2008b, p. 49). In fact the award citation states, quite explicitly, that it was for Simon's '*pioneering research into the decision-making process within economic organizations*'. As a matter of fact, the second author was a research assistant in the department of economics at the University of Lund at the time Simon received the award, but it was the department of management economics (*Företagsekonomiska Institutionen*) that strongly sponsored the award to Simon and also hosted him, on the occasion of his sojourn in Sweden in 1978.The department of economics (*Nationalekonomiska Institutionen*) showed absolutely no interest whatsoever in even co-hosting Simon's visit to Lund.

[2] We disagree strongly, however, with Heukelom's 'loose' characterization of Simon's critique of the orthodox rationality framework (loc.cit, p. 273) as being 'random limits on rational decision making' (and, in this disagreement, we believe we have the support of Crowther-Heyck's studied interpretation of Simon's critique of orthodox rationality, see, for example, Crowther-Heyck, 2006, p. 317). As a matter of fact, one of the key premises of our own paper is that the formal notion of bounded rationality, characterized in terms of computability theory, is more general, the orthodox concept being a special and trivial case, applied to non-empirical, non-falsifiable and non-experimental decision problems.

Foundations[3], and the support is for what we have called *Modern Behavioural Economics*, originating in the work of Ward Edwards (1954, 1961) and continuing with the research program of Kahneman-Tversky and Thaler (see Kao & Velupillai, 2012a, 2012b for a fairly detailed discussion and full references). The fundamental methodological – and, thereby, also epistemological – difference between the behavioural decision sciences pioneered by Herbert Simon, *Classical Behavioural Economics*, and that by Ward Edwards, *Modern Behavioural Economics*, is that the former is underpinned by a model of computation, highlighting the complexity of behavioural decision processes on the basis of computational complexity theory; the latter, instead, is squarely within the orthodox tradition of subjective expected utility maximization.

As also C-H has emphasized most cogently, boundedly rational decisions have nothing whatsoever to do with mistakes – deliberate or inadvertent; nor are decisions that lead to satisficing due to any kind of formalisable 'irrational' behavior. This is in complete contrast to modern behavioural economics.

One of the theses in C-H, but more particularly in Heyck (2008b), is that *there was a marked shift in Simon's research focus from about the mid-1950s*, now aware of the theoretical possibilities of interpreting the emerging field of computer science providing a foundational anchor to Simon's increasing conviction that the best way to study decision problems in the behavioural sciences – particularly in economics – was to view *the rational agent as an information processer faced with problem solving*.

It is in this context that *problem solving* was formalized, by Simon, as that which a *boundedly rational agent*, facing a *complex environment*, and invoking the powers of a *computationally constrained cognitive mind*, satisficed – in contrast to the mathematical economist's paradigmatic Olympian rational agent's[4] optimizing framework.

Our thesis, on the other hand, is that there was no such sharp delineation in Simon's work on formalizing decision making within a behavioural tradition underpinned by varieties of computationally constrained mechanisms, both internal and external. Problem solving, heuristics, computation and computational complexity, in the specific context of human decision processes, underpinned by a sustained vision of rationality as a process, was the main foundation of Simon's

[3] However, funding the new orientation in behavioural economics was also an initiative of those, like the Ford Foundation, who had supported aspects of Simon's research program, in the 1950s.

[4] In C-H (2006, p. 324), this is referred to an "*omniscient*" rationality (used also, for example, by Simon, 1963, p. 737). We rely, as referred to earlier, on Simon (1983) and the use of '*Olympian* rationality' for the neoclassical agent, capable of optimizing anything, at any time, all the time. C-H captures the difference succinctly and accurately (*loc.cit*):

> "In particular, his vigorous advocacy of a behavioral approach to economics and management, and his criticism of the assumption of 'omniscient' rationality and maximizing behavior, concepts central to neoclassical economics, did not endear him to his colleagues in that field."

research program in the refocusing of the social sciences as behavioural decision sciences. It began with Simon's early familiarity with Polya's classic '*How to Solve It*' (Polya, 1945) and continued with a serious study of Turing's pioneering studies on the *triptych* of *computability* (1936), *mechanical intelligence* (1951) and formal *problem solving* (1954). It continued with the refinement of the notion of *complexity* that had been a recurring theme in Simon's work, even in organization theory, administrative behavior and hierarchical systems – not only in formal human problem solving in a behavioural decision making context. The study of *causal structures,* initially inspired by his work in these fields, displayed possibilities for *simplification* – but a measure of complexity (and its 'dual', simplicity) had to wait till formal computational complexity theory, nascent in the mid-1950s, became central in the core research area of computability theory[5].

Against the backdrop provided above, in a sense it can be said that the main aims of this paper are twofold: first, to clarify, interpret and reformulate bounded rationality, remaining faithful to the definitions and vision of Herbert Simon; Second, to emphasize that bounded rationality ought to be placed and studied within a well-structured *algorithmic* context, which Simon had been advocating all his life (even if, in the early years, still only implicit).

This paper elaborates the computability theoretic underpinnings of the concept of bounded rationality and discusses the modelling philosophy involved in characterising economic agents. The discussion proceeds along the lines of Turing computability, computational complexity and heuristics, in the belief that this hypothetical reconstruction of the intellectual path traversed by Simon (see the reference mentioned in footnote 6, below), is fruitful to explore – and see where it may lead. In other words, we aim to be able to reconstruct a coherent theoretical narrative for underpinning Simon's path from bounded rationality as a basis for consistent behavior in decision contexts, to its finessing via satisficing, in human problem solving by agents as information processing systems. For this we underpin our narrative in terms of computable and computational complexity theory. This theoretical underpinning allows us to put the pieces of the fascinating mosaic that is the intellectual vision and life of one of twentieth century's most versatile thinkers.

[5] The two works that inspired and influenced him most decisively in these two respects – the study of causal structures and the linking of inductive inference with a formal notion of algorithmic complexity – were Goodwin (1947) and Solomonoff (1964a, 1964b). The former remained a central inspiration in Simon's sustained vision on problem simplification, search space decomposition and evolution, via an interpretation of the notion of unilateral coupling in Goodwin (*loc.cit*) with the formalism of *semi* (or *nearly*)-*decomposable* matrices, which would not have been alien to the author of the famous 'Hawkins-Simon' results (Hawkins & Simon, 1945). An early version of Solomonoff (*op.cit*) was, in fact, the only document 'submitted' by John McCarthy, in lieu of a post-conference report on the famous Dartmouth conference, in which both Simon and Solomonoff were two of the ten official participants (cf., Kline, 2011, pp. 11-12). In fact, one of Simon's last writings, published, indeed, after his death (Simon, 2001), returned to the framework of Solomonoff's *Dartmouth contribution* to tackle issues that had been central to his scientific philosophy and outlook, from his Chicago university days, on linking scientific discovery with pattern recognition by means of a focus on parsimony in modelling.

For example, while viewing bounded rationality in the context of human problem solving, three aspects of *problem solving* become relevant: the *existence of a method*, the *construction of a method*, and *the complexity of the constructed method*. A message that this paper hopes to convey, then, is that the bounds to human rationality will also be dictated by the complexity of different problems that the problem solver encounters and the research program on *Human Problem Solving* initiated by Herbert Simon becomes a natural path along this direction.

In section 2, therefore, an analysis of such a definition of *bounded rationality* and discussions on *satisficing*, *procedural rationality* and *heuristics* can be found. In section 3, the meeting ground between *Turing's computability* and *problem solving* on the one hand, and Simon's work on *Human Problem Solving* and *Information Processing Systems*, on the other, is explored. Section 4 contains a discussion and interpretation of Simon's empirical grounding of behaviourally rational behavior via computational complexity theory. A brief concluding section summarises the vision aimed at, in our computable and computational complexity theoretic interpretation of Simon's research program – in addition to our own vision of how me may go 'beyond Simon', standing on his giant intellectual shoulders.

§ 2. A Boundedly Rational Reconstruction of Simon's Computable Behavioural Economics

> "In my 1956 paper, 'Rational Choice and the Structure of the Environment,' I wove around the metaphor of the maze a formal model of how an organism (a person?) could meet a multiplicity of needs and wants at a satisfactory level and survive without drawing upon superhuman powers of intelligence and computation. The model provided a practicable design for a creature of *bounded rationality*, as all we creatures are."
> Simon, 1991, p. 175; italics added.

Bounded rationality – together with *satisficing* - is the central theme of the *Classical Behavioural Economics* of Herbert Simon. It appears, now, to be ubiquitously accepted as a replacement for the otherwise *computably infeasible* notion of *Olympian rationality*, which was strongly disapproved by Herbert Simon. Contrary to popular understanding, Simon perceived bounded rationality as *the more general notion* compared to Olympian rationality. Orthodox economics literature (for example, Radner, 1975, Sargent, 1993) tends to promote the opposite view, that bounded rationality is a formally constrained version of Olympian rationality.

The theory of *Human Problem Solving* (Newell & Simon, 1972) incorporates these two essential notions in interpreting the decision making processes of the rational agent, now interpreted as an information processor. Although Simon almost never phrased his theories and concepts in terms of computability and computational complexity theories *explicitly*, he devoted himself to construct more realistic boundaries of human rationality, however, always *implicitly* within the framework of rationality as being procedural (algorithmic) and in turn, encapsulated by Turing computability and constrained by theories of computational complexity[6]. This is amply evident in the two pioneering classics of classical behavioural economics, Simon (1955, 1956), where, in fact, even the notion of *computability*, in its strict recursion theoretic senses, and *complexity,* in its computational complexity theoretic senses – although the theory was still in its very nascent stage – are copiously invoked and used.

§ 2. 1 Computable Bounded Rationality

> "Theories that incorporate constraints on the information-processing capacity of the actor may be called *theories of bounded rationality*.
> Simon, 1972, p. 162; italics added

The term "bounded rationality" was coined by Herbert Simon in his introduction to the fourth part of his collected works **- Models of Man**. He wrote:

> "The alternative approach employed in these papers is based on what I shall call *the principle of bounded rationalit*y:
> *The capacity of the human mind for formulating and solving complex problems is very small compared with the size of the problems whose solution is required for objectively rational behavior in the real world* **-** *or even for a reasonable approximation to such objective rationality.*
> If the principle is correct, then the goal of classical economic theory **-** to predict the behavior of rational man without making an empirical investigation of his psychological properties **-** is unattainable."
> Simon, 1957, p.198-199, italics in the original

Although the term appeared in 1957, the original idea of bounded rationality can be found in both Simon (1955) and Simon (1956), and can eventually be traced back to Simon (1947).

After Simon proposed his initial models of rational behaviour, successive models of bounded rationality that were developed had an early and reliable report in March (1978), showing the many different directions in which it was developed by Simon and others, its mild extension and reinterpretations.Often, any *inconsistent* behaviour with respect to orthodox, Olympian, rationality is perceived as a *mistake* on the part of the agent. Consequently, bounded rationality has been explained as the mistake or short coming of human beings that arises due to a variety of factors (largely psychological) for about 50 years. Modern behavioural economics is not the only field that considers bounded rationality as a compromised concept from normative rationality. As March – surely with Simon's endorsement – noted very early on:

> "Alternatively, one can recall all of the deviations from normative specifications as stupidity, errors that should be corrected; and undertake to transform the style of exciting humans into the styles anticipated by the theory. This has, for the most part, been the strategy of operations

[6] We append the letter Simon write to the second author as a partial substantiation of this view – but also as an important doctrine-historical record of a significant episode in the history of the behavioural sciences.

> and management analysis for the past twenty years; and it has had its successes. But it has also had failures."
> March (1978), p. 597

On the surface of it, Simon's descriptions of bounded rationality might seem that the existence of a ‘bound' is simply due to the limitations of a psychological – more precisely, a cognitive, computationally underpinned - nature in human decision making. When we consider *both* the decision maker *and* the associated, or ‘coupled’, environment, bounded rationality emerges naturally within such a setting. Given the characteristics of the environment and the decision maker, "*satisficing*" (first used in Simon, 1956, p.129, contrasting it, explicitly, with ‘optimize’) is the reasonable action to be pursued in a procedurally rational decision making setting, and *heuristics* are the means through which boundedly rational satisficing behaviour becomes algorithmically implementable.

Later, through the interpretation of the principle of bounded rationality, with computable foundations, it is made clear that bounded rationality is neither *irrationality* (Simon, 1957, p. 200) nor *approximate* optimality (Simon, 1972, p.170).

The models of *rational* decision making suggested in Simon (1955, 1956) do not require utility functions to be defined over (even countably finite) alternatives. They provide some important ideas regarding how a boundedly rational entity could be – indeed, should be - modelled. In, for example, Simon (1955), a simplified value function $V(\cdot)$ which takes only two values (1,0) was introduced[7]. The binary values can be associated with "satisfactory and unsatisfactory", "accept and reject", etc. The domain of function V is S, the set of all possible outcomes which is mapped to A, a set of all *behavioural alternatives*. This is in order to distinguish the means from the ends.

The ensuing rational decision-process is defined as

1. Search for a set of possible outcomes $S' \subseteq S$ such that the pay-off function is satisfactory $(V(s)=1), \forall s \in S'$
2. Search for behavioural alternatives $a \in A'$, whose possible outcomes are all in S' through the mapping.

This process does not guarantee the existence and uniqueness of a solution, until the sequence in which the alternatives arrive, and the dynamics of aspiration levels (a psychological concept), are *formally* incorporated into it.

In real life, the alternatives are often examined sequentially and the first satisfactory alternative evaluated is the one that is selected. The difficulty of discovering a satisfactory choice depends on the cost of obtaining better information regarding the mapping of A on S. Thus, *if* the aspiration level

[7] It is precisely here that Simon first used the notion of *computability* explicitly (loc.cit, footnote 2, p.247, in the reprinted version in Simon, 1957).

grows when the cost of search is low and declines when the cost of search is high, then this dynamic can lead to near-uniqueness and existence of a solution in the long run[8].

In Simon (1955), the focus is on suggesting a dynamic process for decision makers, without going into the details of the mapping between A and S. However, in Simon (1956), the focus is more on the other important aspect - the environment. Here, the organism is assumed to have a single and a fixed aspiration level - it needs only food. But, the food heaps are located in such a way that the organism has to walk in a maze, where there are branches at each node. Each node is a possible location for food. This is combined with the constraint that its vision is limited and therefore it cannot see as far as it would like. However, if it sees a food heap in the range of its vision, it knows the way to reach the food. It has to eat the food before it dies of starvation and there is a maximal number of moves it can make after eating before its energy runs out. These are some of the parameters[9] regarding the environment that the organism faces and the "physical" constrains that the organism has:

- p: $0<p<1$, is the percentage of branch points, randomly[10] distributed, at which food is found.
- d: is the average number of paths diverging from each branch point.
- ν: is the number of moves ahead the organism can see.
- H: is the maximum number of moves the organism can make between meals without starving.

The first two parameters concern the environment (exogenously given *problem space*), on how the targets are distributed and (conjectures) on the size of the problem space is. The last two parameters are regarding the organism on how far it can search and its capacity for searching. This setting can be applied to a much broader class of problems. The parameters do not have to be limited only by physical needs and constraints. Especially, *probability* is *not* central in many realistic cases of decision making for Simon. For example, in chess, a game that was studied intensively by Simon, the goals (some particular patterns) that a player might seek are *not* randomly distributed in the problem space.

Integrating the models in the two papers mentioned above, we can summarize the situation of rational decision making postulated by Simon as the following: There are always two aspects of

[8] A similar modelling logic can be found, *not surprisingly*, in a simple job search model in McCall (1970); ‘not surprisingly’ in view of Simon’s prescient observation (loc.cit, footnote 2, p. 247) of the role of aspiration levels in Wald’s decision theory which was one fulcrum upon which McCall developed his pioneering work in Search Theory. Wald’s formalization via Markov Decision Processes, introduced by McCall, was later adopted as one of the three ‘recursive’ bedrocks of Newclassical *Recursive Macroeconomics* (Ljungqvist & Sargent, 2004) – the two other two recursive tools being (Kalman) filtering and (Bellman’s) dynamic programming. None of these so-called ‘recursive tools’ have anything to do with recursion theory.

[9] These parameters are algebraic, rational, numbers or other computable numbers.

[10] It is an easy mathematical task to consider this in notion of randomness in algorithmic terms, which Simon must have known from the work of Solomonoff, one of the three acknowledged pioneers of algorithmic probability or algorithmic information theory (the other two being Kolmogorov and Chaitin).

decision making - the environment and the mechanism of the decision maker. The two aspects are highly interrelated.

The characteristics of the environment or the problem space are the following:

- The alternatives are assumed to be formalizable as discrete values – i.e., can be *effectively enumerated*;
- The alternatives or the offers arrive in a *sequence*, while the order is not necessarily known;

The characteristics of decision makers are

- Satisficing (influenced by aspiration levels – hence, eventually, coupled to the ***SAT*** problem of computational complexity theory)
- Limited computational capacities (such as time and memory – hence, eventually, underpinned by computational complexity theory)
- Use of *heuristics* to search (hence, algorithmic search)
- Some knowledge or clue regarding the *stopping rule* for searching (starting from any node – the halting state for a Turing Machine)
- Adapting aspiration levels (the eventual notion of learning machines is built from the basics here)
- Knowledge of what to choose and what not to choose (depending on the 'state' of a Turing Machine)

The problem space, viewed formally, as a tree, is 'explored' by the decision maker, often initial conjectures are obtained by prior theoretical understanding and some elementary, preliminary, computer simulations. This problem step-up of 'searching in a tree' is, surely, inspired by the means-end schema proposed in chapter IV of Simon (1947). In Simon's view, the description of the environment depends on the needs, drive and goals of the decision maker. This seems to underpin the '*maze*' metaphor that Simon frequently and fruitfully invokes and suggests for many decision problems in real life. Therefore, human decision making, which is part of human thinking activity, can be associated in fertile ways to many deep areas, such as, computer science, graph theory, formal logic, etc., as Savitch, whose important result will be referred to later, observed, in a similar context:

> "Informally, a maze is a set of rooms connected by one way corridors. Certain rooms are designated goal rooms and one room is designated the start room. Thus, a maze is a directed graph with certain nodes or rooms distinguished. The maze is threadable if there is a path from the start room to some goal room"
> Savitch (1970), p.187

It is important to note that the probabilities that are used to calculate the likelihood of failing to survive or finding a solution in the model discussed earlier are *trivial or meaningless* in many real life problems:

> "From still a third standpoint, the chess player's difficulty in behaving rationally has nothing to do with *uncertainty* - whether of consequences or alternatives - but is a matter of *complexity*. For there is no risk or uncertainty, in the sense in which those terms are used in economics or statistical decision theory, in the game of chess. As von Neumann and Morgenstern observe, it is a game of perfect information. No probabilities of future events need enter the calculations, and no contingencies, in a statistical sense, arise.
> From a game-theoretical standpoint, the presence of the opponent does not introduce contingencies. The opponent can always be counted on to do his worst. The point becomes clear if we replace the task of playing chess with the task of proving theorems. In the latter

task, there is no opponent. Nor are there contingencies: the true and the derivable theorems reside eternally in Plato's heaven. *Rationality in theorem proving is a problem only because the maze of possible proof paths is vast and complex*."
Simon (1972), pp. 169-170, italics added

However, it is debatable whether the opponent can be counted on to do the worst in games like Chess. The search spaces of games like Chess or Go are *certain*, known to both players, but only waiting to be *discovered* (as in 'Plato's heaven'). We can also imagine the opponent[11] as using *heuristics* in her own mind, in order to decide what the possible reacting moves of his opponent will be, where outguessing as infinite regress indeterminacy has plenty of scope for indeterminacies, but can be tamed by constructive or computable formalisms with convincing simplicity. Facing such indeterminacies, and interpreting them as complexity, the decision maker has to incorporate some mechanisms for her to terminate the searching process. This leads to satisficing[12].

Many problems have relatively closed and a pre-defined problem spaces, though the problem space (tree) may be – in fact, *are* - massive. There are many other problems which are far more complex, for example, finding a particular quotation amongst the books in a library. However, most of the time, the material that one is looking for is in the formalizable neighbourhood of initially conjectured structure of the complex problem space, but it is hard to find a good *heuristic* – i.e., a practicable algorithm - to reach it.

§ 2. 2 Algorithmic Satisficing and Olympian Optimizing

"A decision maker who chooses the best available alternative according to some criterion is said to *optimize*; one who chooses an alternative that meets or exceeds specified criteria, but that is not guaranteed to be either unique or in any sense the best, is said to *satisfice*. The term 'satisfice' which appears in the *Oxford English Dictionary* as a Northumbrian synonym for 'satisfy', was borrowed for this new use [in Simon, 1956]."
Simon, 1997, p. 295; first two italicized words, added.

Satisficing is the other pillar, the first being bounded rationality, on which Simon's behavioural economics stands on. Here, the decision maker does not look for an optimal choice, the search `procedure' will itself, if constructed adaptively, lead the decision maker and problem solver to choose a satisfactory outcome as and when one encounters it. This would mean that even though there might be an outcome that could yield a higher level of satisfaction, the choice process stops once a `good enough' alternative that matches the aspiration level is met. Simon also comments on the relation between satisficing and optimizing and that the latter is a special case of the former.

"A satisficing decision procedure can often be turned into a procedure for optimizing by introducing a rule for optimal amount of search, or, what amounts to the same thing, a rule for fixing the aspiration level optimally....

[11] Who may well be typically fallible Greek or Hindu Gods!
[12] *Minmax* could, therefore, be such a *satisficier*.

> Although such a translation is formally possible, to carry it out in practice requires additional information and assumptions beyond those needed for satisficing"
> Simon, 1972, p.170

For a decision maker, the act of optimization would require *a priori* knowledge of ***all*** the available options and the associated outcomes. Moreover, she also requires ***a method*** (i.e., an algorithm) for listing all the options and to compare each of them. When the decision maker is confronted with multiple goals, then association between choice and outcomes gets even more complex. As Simon remarks, this is both unrealistic and excessively demanding (Simon, 1956, p.136). Simon further emphasizes that the optimizing approach facing real-life complexity is indeed approximate optimization. The statisficing approach , on the other hand, is linked with the dynamics of aspiration levels and tackles the problem very differently. We have developed two formal results, in Velupillai (2000, chapter 3) that conceptualize, as theorems, the above considerations by Simon:

Theorem 1:

There is no *effective procedure* to generate preference orderings.

Theorem 2:

Given a class of choice functions that do generate preference orderings (pick out the set of maximal alternatives) for any agent, there is no *effective procedure* to *decide* (algorithmically) whether or not any arbitrary choice function is a member of the given class.

In Velupillai (op.cit), 2010a and 2010b), we have recursion theoretically formalized decision making as an act of algorithmically choosing a subset from a finite, non-empty, countable set, as opposed to uncountably infinite sets, by using a computable choice function. *Solving* optimally an instance of the latter class can be shown to be equivalent to *solving* linear integer programming problems. We, then, transformed the linear integer programming problem into the optimization problem of a combinatorial system, and then constructed abstract Turing machines to study the characteristics of problem solving. The key here is the demonstrable formal double equivalence between bounded rationality and the behavior of a Turing Machine during a formal (symbolic) computation and 'algorithmically choosing' and 'satisficing'.

§ 2. 2 Procedural and Substantive Rationality

> "The search for *computational efficiency* is a search for *procedural rationality*, and computational mathematics is a normative theory of such rationality. In this normative theory, there is no point in prescribing a particular substantively rational solution if there exists no procedure for finding that solution with an acceptable amount of computing effort."
> Simon, 1976, p.133

The theory of computational efficiency is computational complexity theory. The part of computational mathematics that formalizes the notion of a normative theory of rationality is the computing activity of a Turing Machine, subject to the Church-Turing Thesis. The dichotomy between procedural rationality, underpinning Simon's kind of (classical) behavioural economics and the substantive rationality of orthodox behavioural economics is highlighted by the divide between the behavior of an ideal computing machine – the Turing Machine – and that between one that is constrained by time and space constraints, encapsulating economic costs, formalized by computational complexity theory (reversing the two former dichotomies).

When we begin to claim that "This decision maker is satisficing.", the next question we may ask primarily becomes "What are the procedures the decision maker uses?", instead of, "What does the decision maker choose?". The fundamental distinction between Simon's approach and the other theories that invoke behavioural traits, such as in modern behavioural economics, is the insistence on `methods' or `procedures' involved in choosing and their centrality in the theory of decision making. The link between a procedurally rational choice and computation is present from the very outset in Simon's scheme. The insistence here is on the complexity of this decision process, in terms of the effort devoted in doing it.

§ 2. 3 Heuristics

> "Most weak methods require larger or smaller amounts of search before problem solutions are found, but the search need not be blind trial-and-error--in fact, usually cannot be, for the search spaces are generally far too vast to allow unselective trial and error to be effective. Weak methods generally incorporate Polya's idea of "*heuristics*"- *rules of thumb* that allow search generators to be highly selective, instead of searching the entire space."
> Simon,1983a, p.4570; italics added.

Simon's use, refinement and explorations of formal decision making and formalizing problem solving were decisively influenced by Polya's 'little' classic of 1945 (Polya, 1945), from even before he completed Simon (1947). There is still no better description of the nature and scope of 'heuristics' than given in that Polya classic:

> "**Heuristic** or heuretic, or 'ars inveniendi' was the name of a certain branch of study, not very clearly circumscribed, belonging to *logic*, or to *philosophy*, or to *psychology,* often outlined, seldom presented in detail, and as good as forgotten today. The aim of heuristic is to study the *methods and rules of discovery and invention*. …
> …
> Heuristic, as an adjective, means '*serving to discover*'."
> Polya, op.cit., pp. 112-3; bold emphasis in the original; italics, added

No one person made the notion of 'heuristic' to underpin (mathematical) logic, philosophy (of science) and (cognitive) psychology of the *sciences of the artificial* (Simon, 1996) and for 'serving to discover' (Simon, 1977; Langley, et.al., 1987) than Herbert Simon. From an early reference to 'rules of thumb', surely derived from 'rules of anticipated reactions' in Simon (1947) and 'rule of action'

(March & Simon, 1958) to fully fledged formal algorithms, via heuristics in Newell & Simon (1956) was one of the computable and computational complexity theoretic paths Simon took.

Newell and Simon initiated the project on constructing a program – an algorithm - which learns to play good Chess, in 1954, while they pinned down the investigation of *heuristics* on proving theorems in *Principia Mathematica* to start with. The program they designed was hand-simulated first and interpreted into machine language which gave birth to the *Logic Theorist* which was their first example of human problem solving (Newell, et.al., 1958). However, it was subject to an uncharacteristically strong criticism from the otherwise enlightened Hao Wang:

> "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."
> Wang (1970), p.227

It was unfortunate that Wang had misunderstood Simon's stated priorities (Simon,1991, p.209-210) and the path to it, from the concept of 'rules of thumb'. Newell and Simon were involved in finding procedures used by human beings in solving problems and used this information to construct the program *Logic Theorist*. What are hueristics, if not algorithms? A perceptive remark on heuristics is that "a method is simply a plan that you use twice" (Newell & Simon, 1972, p.835), which, on the 'third' occasion approaches the formal status of an algorithm. That is, heuristics represents the methods that human beings actually use to search in a problem space. They are nothing but algorithms.

Algorithms are formally connected to symbolic structures which are underpinned by computability theory. Physical symbolic systems such as human beings and digital computers, as Newell and Simon pointed out, can process only a finite number of steps in any given interval of time. However, the finiteness to which the most general model of algorithms - Turing Machine - appeals is in many cases not strong enough to show the severe limitation that human minds have to confront (Newell & Simon, 1976, p.120). Empirical boundaries correspond to the level of complexity that the human minds can actually handle (see also the contents of Simon's letter, referred to above, in footnote 6). A study of heuristics is crucial in order to understand how human beings handle problems whose complexity is beyond the empirical boundary. In other words, heuristics act as procedures that help reduce the problem to a level of complexity which can be handled algorithmically

The approach described in *Human Problem Solving* (Newell & Simon, 1972) that encompasses heuristics is underpinned by computational complexity theory, which in turn is based on computability theory. Moreover, the significance of heuristics is not revealed until some algorithmic impossibilities concerned with procedural decision making are formally proved. To further explore this connection, we need to examine the interconnections between the approaches of Turing and Herbert Simon to Machine and Human Problem Solving.

§3. From Turing to Simon: Problem Solving and Decision Problems

> "Given a Diophantine equation with any number of unknown quantities and with rational integral numerical coefficients: To devise a process according to which it can be determined by a finite number of operations whether the equation is solvable in rational integers."
> Hilbert, 1902, p. 451.

That's where it all began! This came to be known, famously, as *Hilbert's Tenth Problem* and, as a decision problem, could not be answered without a formal definition of "a process according to which it can be determined by a finite number of operations"; in other words, a formal definition of the intuitive notion of an algorithm. Until about 1936-1937, when the definitions of such a finite procedure was defined formally, the mathematicians had only an *intuitive* notion of an algorithm.

Church defined algorithm (effective calculability) with the λ-calculus, and Turing defined it in terms of Turing Machines. They were proved to be equivalent definitions and the intuitive notion of algorithm captured by these definitions imply the so-called Church-Turing Thesis. *If this thesis is true*, then the halting problem for Turing machines is **unsolvable**. Church (1938), mentioned that the intuitive notion of an effective procedure can be formalized in three different ways: Turing Machines (Turing (1936), λ-definability (Church (1936)) and the general recursive functions (of Herbrand and Gödel, Kleene, 1936a). The equivalence of the three notions had been proved in Kleene (1936b) and Turing (1937). The equivalence of recursiveness and computability enables one to apply the definition of recursive function to prove more classes of computable functions. (Davis, 1982, Ch. 3). The consensus on the notion of effective calculability was reached in the late 1930s and this led to the development of computability theory.

Herbert Simon was fully aware of all these developments at a very early stage. Both Herbert Simon and Allen Newell were recipients of the ACM *Turing Award* in 1975 for their contribution to the human problem solving approach, which they initiated in the mid-1950s together with Cliff Shaw. It is evident that their work, where computation plays an important role, was grounded on Turing's contributions:

> "Concurrently with Turing's work appeared the work of the logicians Emil Post and (independently) Alonzo Church. Starting from independent notions of logistic systems (Post productions and recursive function, respectively), they arrived at analogous results on *undecidability* and *universality* - results that were soon shown to imply that all three systems were equivalent. Indeed, the convergence of all these attempts to define the most general class of *information processing systems* provides some of the force of our conviction that *we have captured the essentials of information processing in these models*."
> Newell & Simon, 1976, p.117; italics added.

If there were any lingering doubts that Simon's boundedly rational agent, encapsulated in the formalism of an information processing system, is processing symbol sequences in the same way as a Turing Machine, subject to the same computational complexity norms, then the above observation should dispel them once and for all, at least we hope so. The *Human* Problem Solving *information* *processing system* of **Herbert Simon** was exactly equivalent to the Mathematical Problem Solving *Machine* constructed by **Alan Turing**.

Although Turing does not seem to have made an attempt to solve Diophantine decision problems in the above Hilbertian sense, he ‘constructed’ the hypothetical Turing Machine for solving another decision problem posed by Hilbert in 1928 (Turing,1936) and theoretical developments based on Turing machines contributed eventually to the negative solution of *Hilbert's 10th problem*, 70 years after the question was posed (see Matiyasevich, 1994 for details). The decision problem that concerned Turing, in the specific context of *problem solving*, was general: Is there a systematic procedure to decide whether a given problem (puzzle) is solvable or not? This decision problem regards all those problems which can be transformed into what he called ‘substitution puzzles’. The answer to this was proved to be negative by Turing. The negative solution to this decision problem indicates that we need to develop *specific* procedures in order to decide *specific* problems. There is no *general* solution - i.e., algorithmic procedure - to any given problem. This has a direct bearing for the theories of decision making that rely on optimization, without addressing the procedural aspects. This is also why Simon, understanding the need for specific studies, indulged in simulation by computation to explore the structure of problem spaces – not only, indeed, not even, for popular reasons of cognitive computational constraints.

In order to understand how he arrived at the negative solution, it could be useful to take a close look at the problem setting – which also elucidates why, again even if in an *ex-post* sense, Simon may have concentrated so much in understanding *Human Problem Solving* by studying games and puzzles:

> "Given any puzzle [problem] we can find a corresponding substitution puzzle which is equivalent to it in the sense that given a solution of the one we can easily use it to find a solution of the other. If the original puzzle is concerned with rows of pieces of a finite number of different kinds, then the substitution may be applied as an alternative set of rules to the pieces of the original puzzle. A transformation can be carried out by the rules of the original puzzle if and only if it can be carried out by the substitutions and leads to a final position from which all marker symbols have disappeared."
> Turing1954, p.15, italics in the original.

He further wrote, "In effect there is no opposition to the view that every puzzle is equivalent to a substitution puzzle."[13] The production rules[14] that are introduced in Turing's example follow type 0 grammar[15], though, the time at which Turing proposed it was before the Chomsky hierarchy was defined (Chomsky, 1956, 1959).

[13] See Turing (1954), p.13 for an example of a substitution puzzle.
[14] Its formalism can be traced back to Post (1947).
[15]Type 0 grammar is the superset of the hierarchy and includes all recursively enumerable languages. Turing Machines are the most general kind of symbol operators which are capable of recognizing the languages generated from all types of grammar.

It is assumed that there exists a systematic procedure for deciding whether a puzzle is solvable or not. At the same time, this systematic procedure can be transformed into a substitution puzzle whose set of rules is K. Naturally, K has unambiguous moves and it always comes out with final result no matter what R, the puzzle of interest, is. In particular, it will come out with, say B(lack), when R belongs to class I, and W(hite) when it belongs to class II. Then, when we look at the puzzle P(K,K)[16] to be investigated, we will find inconsistent results. That is, we should be able to classify that P(K,K) belongs to class I or II. But according to the substitution puzzle K, it has the potential to result in both possibilities, as a result, we could not classify it into either one or the other of the two classes. This leads to a contradiction!

This demonstration towards showing that there is no general algorithm for deciding whether a puzzle is solvable or not suggests that we need to seek for separate algorithms in order to decide whether a kind of problem is solvable or not, given the initial puzzle and the desired outcome.

It was here where the formalized, machine-based, problem solving methods of Turing synchronized perfectly with the information processing system that was the human problem solver of Simon. Thus, it was also natural that the ground where bounded rationality and computability meet is clearly presented in Newell & Simon, 1972, where symbolic systems can be adopted to understand human thinking, especially in the activities of information processing. In the theory of *Human Problem Solving*, the *vague* idea of an "environment" is precisely formulated into a "problem space" and a problem solver into an *Information Processing System*. In addition, in Simon's approach, the notion of "complex problems" needs to be given a *precise* definition, and it is done within the context of Turing's computability theory and computational complexity theory.

§4. Herbert Simon’s Empirically Rational Problem Solver and Computational Complexity Theory

> "Moreover, since *Homo sapiens* shares some important psychological invariants with certain nonbiological systems - the computers - I shall want to make frequent reference to them also. One could even say that my account will cover the topic *human* and *computer psychology*."
> Simon (1990), p.3; italics added.

Simon stressed the infeasibility of a procedure of optimization by showing that digital computers which overpower human beings in terms of their physical computational capacity find their strength insignificant when confronted with the *complexity of real world problems* (Simon,1976, p.135). Understanding our own limits is one of the lessons we obtain from digital computers. The same logic should be applied to economic decision makers, as Simon suggested in the following:

> "The human mind is programmable: it can acquire an enormous variety of different skills, behaviour patterns, problem solving repertoires, and perceptual habits.....There seems to be no

[16] To derive a contradiction via a diagonal argument.

escape. If economics is to deal with uncertainty, it will have to understand how human beings in fact behave in the face of uncertainty, and by what *limits of* information and *computability* they are bound."
Simon, 1976, p. 144; italics added.

For empirical reasons, the unsolvability of a problem does not really stop people from looking for a solution for it, particularly not Herbert Simon. Sipser admirably summarises the *pros* and *cons* of proving the unsolvability of a problem, and then coming to terms with it:

> "After all, showing that a problem is *unsolvable* doesn't appear to be any use if you have to solve it. You need to study this phenomenon for two reasons. First, knowing when a problem is *algorithmically unsolvable* is useful because then you realize that the problem must be simplified or altered before you can find an algorithmic solution. Like any tool, computers have capabilities and limitations that must be appreciated if they are to be used well. The second reason is *cultural*. Even if you deal with problems that clearly are solvable, a glimpse of the unsolvable can stimulate your imagination and help you gain an important perspective on computation."
> Sipser1997, p.151, italics added.

As we have stressed, Simon was concerned with both empirical and theoretical questions of decision making and he grounded himself, for tackling these questions *effectively*, on Turing Computability. He later looked for computational complexity in average cases or empirical complexity that is relevant for human problem solving. Now that we have seen that there exist algorithmically unsolvable decision problems, we can appreciate how Turing computability should be the outer limit for human rationality or machine computability. Anything that goes beyond Turing computability is clearly meaningless (especially for procedural decision making) since even the most powerful abstract computing machine cannot solve such a problem, even in principle. But this can only be an outer boundary of how far procedural rationality can go in theory because the notion of pure computability in theory does not take into account time and space limitations, which are essential to solve a problem or compute a function. They become particularly important in the case of human decision making. This only reinforces the conclusions and strengthens the concepts that Simon advocated. This was why Simon spent so much time in understanding the nature and structure of a problem – against the backdrop of a deep knowledge of the limits of computability – so that essentials need not be sacrificed in the simplifications that had to be achieved to make the unsolvable approach empirical solvability:

> "How complex or simple a structure is depends critically upon the way in which we describe it. Most of the complex structures found in the world are enormously redundant, and we can use this redundancy to simplify their description. But to use it, to achieve the simplification, we must find the right representation."
> Simon (1962), p.481

The above observation can be interpreted in terms of computational complexity theory, where the complexity of a problem is determined by the time and space requirements of an algorithm that solves the problem[17]. There are three aspects of problem solving: the inherent solvability of a problem, the procedure to solve a problem and the complexity of the procedure. Provided we have Turing's abstract model of computation, we can use this idea to construct an abstract machine for solving a particular problem. We can then analyse the number of steps or memory that the algorithm would require, approximately, without going about to count the precise time and space required by the problems of the same kind. This helps us to have an idea of the associated difficulty of a problem we are dealing with before we really start to solve it. Computational complexity provides a more solid, inner boundary of bounded rationality with Turing computability as its outer boundary. Although the scale of time steps and space (memory) that computational complexity theory regards is normally pretty large, it is important to have a general idea of tackling a problem by knowing the complexity of the algorithm which solves it. In theory, the reducibility among problems is also used to study the complexity without actually constructing a real algorithm.

As far as problem solving is concerned, according to Turing's interpretation, a decision problem is to decide whether one can change a string of symbols to the desired string of symbols, by only using a set of rules that are given in advance. Knowing that a decision problem is unsolvable leads us to ask different questions and try to solve them, otherwise, it provides no practical help when we try to solve a problem. We need to find a set of rules for a substitution puzzle, that is an "algorithm", to solve our problem. However, even if we have an algorithm to solve a certain kind of problem, it does not guarantee that we can solve the problem within the desirable period of time. If the problem involved is complex, it can demand immense amount of computation by the problem solver. Time complexity and space complexity are very useful and standard tools for providing measures of quantitative ideas on how much effort is needed for solving a problem.

When we have an algorithm for solving a problem, we can look at its general behaviour and analyse how many time steps and the space or memory it would require. Time complexity tells the number of steps needed for running an algorithm, and space complexity takes care of the memory needed. Time and space complexity are the functions of size of input, for example, playing 3-disk *Tower of Hanoi*[18] needs much less time steps than 10-disk Tower of Hanoi. In many cases, it is very difficult to obtain the exact reduced form of time and space complexity of an algorithm. Therefore, in computational complexity, asymptotic notations, such as $O(n)$, are used to present the asymptotic behaviour of an algorithm as the asymptotic approximation of the true function behind it. When large input sizes are involved, exponential time grows drastically faster than polynomial time, and the problem becomes unmanageable, in precisely definable ways, very quickly.

[17] The rigorous definitions of time and space complexity and those of different complexity classes can be found, for example, in chapter 7 & 8 in Sipser (1997).
[18] A favourite example in Simon's pedagogy (see, for example, Simon, 1975).

It should be remembered that there exist always more than one method to solve a problem, therefore, the complexity of a problem is determined by the method that solves it[19].

Space complexity has attracted relatively less attention and effort compared to time complexity, despite the powerful result that PSPACE=NPSPACE[20]. By default, when the complexity of a problem is discussed, time complexity is the one that is referred to. Arguably, it is because whether **P=NP** is one of most popular unsolved problems. We, however, would like to emphasize, for the domain of *human* problem solving, *space complexity is at least as important as time complexity*. Although, there is no doubt that the architecture of human brain has the potential to store huge amount of knowledge, the amount of information that minds can process at a given moment is severely limited.

For example, it is very tough to calculate 4593*3274 in the mind for an ordinary person, unless this person has pencil and paper at hand or he/she is an expert in arithmetic calculations. Such calculation requires a certain amount of temporary memory which is a function of input size. In terms of time limitation, minds are constrained by attention span, apart from other externally imposed time constraints, e.g., a chess player has to make a move in 5 minutes. How minds are constrained by time and memory varies with different contexts and structure of the problems and among different persons. Furthermore, these two dimensions should not be completely independent, i.e. the memory constraints affects the time which is needed for solving a problem and vice versa. Therefore, it is important to investigate the time complexity of a problem (or an algorithm) together with the space complexity; consequently, we will be able to know what kinds of heuristics are needed based on these two dimensions. Space complexity is even more crucial when the problem concerned requires no aid of external memory.

In spite of the fact that the time and space complexity of an algorithm can be analysed, human beings are constrained very differently from (digital) computers. We normally have only a certain amount of time to make a decision, and we have very limited working memory (no matter expert or layman) to process this task, regardless of the presumably unlimited long-term memory. We are forced to use those algorithms which will be able to halt within certain amount of time, by applying the knowledge and experience we have in the long-term memory. Although, we are often assigned to a task like "find the best person for this job", we are not able to solve it as an optimization problem. At best, we will have the criteria for appropriate candidates and consider only a small group of people. Depending on the time and memory we are supplied with and the procedure we should go through, we have to be selective to different degrees.

Turing (1951) suggested that a machine should be programmed to learn to play the games like Chess, GO and Bridge. Chess is the recurring example and an important one for Simon, it is also one

[19] See Sipser (1997), p.229-231, for an example.
[20] This is implied by the first theorem in Savitch (1970).

of the examples of **complex combinatorial problems**[21] which make brute-force algorithms infeasible. Even though the problem space of chess is closed and certain, the massive size of the game tree prevents human beings or even supercomputers to use brute search algorithm. Let us take the number of the possible continuations of Chess, which is approximated in Shannon (1950), as an example. If we want to know, from the beginning of the game, whether Black or White has a winning strategy, we have roughly 10^{120} variations to calculate; when each branch reaches its end, we can see whether that branches leads to a win, loss or a draw. Suppose we have a high-speed computer which uses only one microsecond (10^{-6} a second) for one variation, for searching the whole problem space, it will take 10^{100} million years! Obviously, in real life, different actions should be taken, that is why we can always find plenty of Chess tactical guides in the bookshops.

Finally, some notes on *satisficing* as an instance of the *satisfiability* problem[22] of computational complexity theory. Simon showed that Olympian rationality is a special case of bounded rationality by appealing to the act of *satisficing*. He suggested descriptively that a model of satisficing can be turned into optimizing by setting the aspiration level at an optimal level. This lucid point can be substantiated mathematically by applying the results in combinatorial complexity. We noted that (Velupillai, 2010b, p.9):

> "The real mathematical content of satisficing is best interpreted in terms of satisfiability problem of computational complexity theory, the framework used by Simon consistently and persistently - and a framework to which himself made pioneering contributions."

The formal decision problem framework for a boundedly rational information processing system can be constructed in one of the following ways: systems of linear Diophantine inequalities, systems of linear equations in non-negative integer variables, integer programming. Solving the former three problems are equivalent in the sense that the method of solving one problem provides a method to solve the other two as well. The Integer Linear Programming (ILP) problem and SAT can be translated both ways, i.e, one can be transformed into another.

The satisfiability problem is one of the important problems in modern computer science. In Velupillai (2010b), it is demonstrated that SAT problem is the meeting ground of Diophantine problems and satisficing, in turn this connection leads to the conclusion that bounded rationality is the superset of Olympian rationality, which Simon had been advocating.

§ 5 Concluding Remarks: Can We Go Beyond Simon?

> “So when we reach a bifurcation in the road. Of the labyrinth. ‘something’ chooses which branch to take. And the reason for my researches, and the reason why labyrinths have

[21] See Guy (ed.), 1991.

[22] The problem *Satisfiability* (**SAT**) is defined as follows: Given a Boolean formula (The Boolean formula is, itself, in the *Conjunctive Normal Form*) φ, determine whether there is an assignment that satisfies it (i.e., more formally, SAT is the set of all *satisfiable* Boolean formulas).

fascinated me, has been my desire to observe people as they encounter bifurcations and try to understand why they take the road to the right or to the left."
Simon, 1991, p. 179[23]

Now we cannot observe and be guided by the lessons Simon learnt, formalized and experimented with, as he observed boundedly rational agents negotiating the maze of life and its problems. We have to go (sic!) beyond, standing on his mighty shoulders. How might we do it?

There may well be many ways to go 'beyond Simon'. We mention just one possibility, only because we think it encapsulates every precept we think Simon's vision worked with: The Game of GO – as a Game, as a domain for the articulation of formal problem solving in both his and Turing's senses and, of course, as a means of playing the game with information processing systems!

The game of GO, thus, can be one of the possible problem solving paradigms to go beyond Simon, yet remaining faithful to the research program he developed. In a mathematical sense, GO is more flexible and general than Chess, because its board size can be, in principle, unlimitedly enlarged. The philosophy and heuristics of GO are not necessarily consistent with those of Chess, which makes human problem solving more interesting. GO was also in the choice list of Simon, but Newell and Simon finally chose chess as their paradigmatic example of human problem solving[24]. We think they made the right decision for the pioneers they were.

This game has already been studied by combinatorial game theory. GO has been shown to be PSPACE-hard *and* EXP-time complete, which means GO is in the exponential time class and it is proved that there can be no PSPACE algorithm for solving GO (Lichtenstein, 1980, Robson, 1983). Problems in exponential time are considered to be among the most difficult problems. This result suggests that deciding whether Black or White has winning strategy, at an arbitrary position, is practically infeasible. Clearly, any expectation for finding an optimal strategy to win in such a complex setting is both meaningless and futile. On the other hand, it also provides a perfect setting for studying 'actual' modes of decision making without being tied to the search for optimal strategies.

The theme of this paper is that Simon's Behavioural Economics should be formally understood within the framework of computability theory and computational complexity theory. By extracting the procedural content of decision making, heuristics are considered as algorithms. In computational complexity theory, the complexity of a problem is analysed not only through the structure of task environments but also through the heuristics that problem solvers used to solve the problems. In this framework, we are able to show that bounded rationality via satisficing is the general notion of rationality.

[23] From the 'dialogue' with Jorge Luis Borges, held in Buenos Aires, in December, 1970.
[24] The reasons for which are stated in (Newell & Simon, 1972, pp.664-665). Given these reasons, we can only wonder whether they would have chosen the game of GO as their paradigm if they had been from the Orient!

If we view a piece of knowledge as an articulated paragraph, composed by words, which are in turn composed by symbols from a finite set, it becomes clear that it is never straightforward to understand or learn the knowledge by reading or memorizing it. We might have to follow the author with the same or different path on how the attained knowledge is reached. What matters is the knowledge generated in the mind, not the knowledge written. This is partly because a segment of words might have been a compression or definition of bigger segment of words. Tacit knowledge, too, played an important part in Simon's behavioural economics – as important a part as heuristics, and for the same reasons.

Simon's vision and definitions regarding bounded rationality were always intuitive and straightforward. He thus left a large canvas for others to build models based on bounded rationality. Simon's notions concerning bounded rationality can be interpreted more clearly in the light of alternative mathematical formalisms, those which are faithful to the notion of procedural decisions. Also, models should be constructed according to different situations and the actors who handle those situations. In this paper, it is argued that the two aspects of human problem solving - the task environment (problem space) and problem solver (algorithm) should be distinguished and then studied.

By appealing to computability theory, it is shown that bounded rationality is a superset of Olympian rationality. Subsequently, the empirical boundary of rationality is further narrowed down to an inner boundary - the one established by computational complexity. Finally, it is suggested that Simon's empirical boundaries can be further approached - along the same methodology - by investigating the heuristics which are the algorithms (methods) that are used by human beings in problem solving circumstances.

References

Chomsky, Noam (1956), *There Models for the Description of Language*, **IRE Transaction on Information Theory**, 2(3), 113-124.

Chomsky, Noam (1959), *On Certain Formal Properties of Grammars*, **Information and Control**, 2, 137-167.

Church, Alonzo (1936), *An Unsolvable Problem of Elementary Number Theory*, **American Journal of Mathematics**, 58(2), 345-363.

Church, Alonzo (1938), *The Constructive Second Number Class*, **Bulletin of the American Mathematical Society**, 44, 224-232.

Crowther-Heyck, Hunter (2006), *Herbert Simon and the GSIA: Building an Interdisciplinary Community*, **Journal of the History of the Behavioral Sciences**, Vol. 42 (4), pp. 311-334, Fall.

Heyck, Hunter (2008a), *Defining the Computer: Herbert Simon and the Bureaucratic Mind – Part 1*, **IEEE Annals of the History of Computing**, April-June, pp. 42-51.

Heyck, Hunter (2008b), *Defining the Computer: Herbert Simon and the Bureaucratic Mind – Part 2*, **IEEE Annals of the History of Computing**, April-June, pp. 52-63.

Davis, M (1982), **Computability and Unsolvability**, Dover Publications, Inc., New York.

Edwards, Ward (1954), *The Theory of Decision Marking*, **Psychological Bulletin**, 51, 380-417.

Edwards, Ward (1961), *Behavioral Decision Theory*, **Annual Review of Psychology**, 12, 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.

Guy, Richard. K (1991; editor), **Combinatorial Games**, American Mathematical Society, Short Course, August 6-7, 1990, Columbus, Ohio, American Mathematical Society, Providence, Rhode Island.

Hawkins, David and Herbert. A Simon (1949), *Some Conditions on Macroeconomic Stability*, **Econometrica**, Vol. 17

Heukelom, F. (2012), *A Sense of Mission: The Alfred P. Sloan and Russell Sage Foundations' Behavioral Economics Program, 1984-1992*, **Science in Context** 25(2), pp. 263-286.

Hilbert, David (1902), *Mathematical Problems*, **Bulletin of the American Mathematical Society**, Vol. 8, 437- 479.

Kao, Y-F. and K. Vela Velupillai, (2012a), *Origins and Pioneers of Behavioural Economics*, **Interdisciplinary Journal of Economics and Business Law**, 1(3), 47-73.

Kao, Y-F and K. Vela Velupillai (2012b), *Behavioural Economics: Classical and Modern*, Forthcoming: **European Journal of the History of Economic Thought.**

Kleene, S. C (1936a), *General Recursive Functions of Natural Numbers*, **Mathematische Annalen**, 112, pp. 727-242.

Kleene, S. C, (1936b), *λ-Definability and Recursiveness*, **Duke Mathematical Journal**, 2, pp. 340-353.

Kline, Ronald. R (2011), *Cybernetics, Automata Studies, and the Dartmouth Conference on Artificial Intelligence*, **IEEE Annals of the History of Computing**, October-November, pp. 5-16.

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.

Lichtenstein, D and M. Sipser (1980), *Go is Polynomial-Space Hard*, **Journal of the Association for Computing Machinery**, 27(2), 393-401.

Ljungqvist, Lars & Thomas. J Sargent (2004), **Recursive Macroeconomic Theory** (Second Edition), The MIT Press, Cambridge, Massachusetts.

March, James. G (1978), *Bounded Rationality, Ambiguity and the Engineering of Choice*, **The Bell Journal of Economics**, 9(2), 587-608.

March, James. G & Herbert. A Simon (1958; 1993), **Organizations** (Second Edition), Blackwell Publishers, Oxford

Matiyasevich, Yuri (1994), **Hilbert's Tenth Problem**, MIT Press, Cambridge Massachusetts.

McCall, John. J, (1970), *Economics of Information and Job Search*, **The Quarterly Journal of Economics**, 84(1), 113--126.

Newell, Allen and Herbert. A Simon (1956), *The Logic Theory Machine*, **IRE Transactions on Information Theory IT-2**, Vol. 3, pp. 61-79.

Newell, Allen, and Herbert. A Simon, J. Cliff Shaw (1958), *Elements of the Theory of Human Problem Solving*, **Psychological Review,** 65(3), 151--66.

Newell, Allen and Herbert. A Simon (1972), **Human Problem Solving**, Prentice-Hall, INC, Englewood Cliffs, NJ.

Newell, Allen and Herbert. A Simon (1976), *Computer Science as Empirical Inquiry: Symbols and Search*, **Communications of the ACM**, 19(3), 113-126.

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

Post, Emil. L (1947), *Recursive Unsolvability of a Problem of Thue*, **The Journal of Symbolic Logic**, 12(1), 1-11.

Radner, Roy (1975), *Satisficing*, **Journal of Mathematical Economics**, Vol. 2, #2, pp. 253-262.

Robson, J. M. (1983). *The Complexity of Go*, **Information Processing**, 83, 413--417.

Sargent, Thomas J (1993), **Bounded Rationality in Macroeconomics**, Clarendon Press, Oxford.

Savitch,W. J. (1970), *Relationships between Nondeterministic and Deterministic Tape Complexities*, **Journal of Computer and System Sciences** 4, 177-192.

Shannon, C. E. (1950), *Programming a Computer for Playing Chess*, **Philosophical Magazine**, 41(314).

Simon, Herbert. A, (1947), **Administrative Behavior** (Second ed.). The Free Press, New York.

Simon, Herbert. A (1955), *A Behavioral Model of Rational Choice*, **Quarterly Journal of Economics**, 69(1), 99-118.

Simon, Herbert. A (1956), *Rational Choice and the Structure of the Environment*, **Psychological Review**, 63(2), 129-38.

Simon, Herbert. A (1957), **Models of Man**, John Wiley & Sons, Inc.

Simon, Herbert. A (1962), *The Architecture of Complexity*, **Proceedings of the American Philosophical Society**, 106(6), 467- 482.

Simon, Herbert. A (1963), *Economics and Psychology*, in: **Psychology: A Study of a Science**, ed. By S. Koch, Vol. 6, pp. 715-752, McGraw-Hill, New York.

Simon, Herbert. A (1972), *Theories of Bounded Rationality*, In C. B. Radner and R. Radner (Eds.), **Decision and Organization**, pp. 161-176, North-Holland, Amsterdam.

Simon, Herbert. A (1975), *The Functional Equivalence of Problem Solving Skills*, **Cognitive Skills**, Vol. 7, pp. 268-288.

Simon, Herbert. A, (1976), *From Substantive to Procedural Rationality*, In Spiro. J Latsis (Ed.), **Method and Appraisal in Economics**, pp. 129-148. Cambridge University Press, Cambridge.

Simon, Herbert. A (1977), **Models of Discovery – and Other Topics in the Methods of Science**, D. Reidel Publishing Company, Dordrech & Boston.

Simon, H. A. (1983a). *Discovery, Invention, and Development: Human Creative Thinking*. **Proceedings of the National Academy of Sciences**, Vol. 80, 4569-4571.

Simon, Herbert. A, (1983b), **Reason in Human Affairs**, Basil Blackwell, Oxford.

Simon, Herbert. A (1990), *Invariants of Human Behavior*, **Annual Review of Psychology**, 41, 1-19.

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), **Models of Bounded Rationality – Volume 3: Empirically Grounded Economic Reason**, The MIT Press, Cambridge, Massachusetts.

Simon, Herbert. A (2001), *Science Seeks Parsimony, Not Simplicity: Searching for Pattern in Phenomena*, chapter 3, pp. 32-72, in: **Simplicity, Inference and Modelling: Keeping it Sophisticatedly Simple**, edited by Arnold Zellner, Hugo A. Keuzenkamp and Michael McAleer, Cambridge University Press, Cambridge.

Sipser, Michael (1997), **Introduction to the Theory of Computation**, PWS Publishing Company.

Solomonoff, Ray. J (1964a), *A Formal Theory of Inductive Inference: Part I*, **Information and Control**, Vol.7, pp.1-22.

Solomonoff, Ray. J (1964b), *A Formal Theory of Inductive Inference: Part II*, **Information and Control**, Vol.7, pp.224-54.

Turing, Alan. M (1936), *On Computable Numbers, with an Application to the* ***entscheidungsproblem***, **Proceedings of the London Mathematical Society**, 42(2), 230--265.

Turing, Alan. M (1937), *Computability and λ-definability*, **The Journal of Symbolic Logic** 2(4), 153--163.

Turing, Alan. M (1951), *Intelligent Machinery, a Heretical Theory*, **Philosophia Mathematica**, 4(3), 256--260.

Turing, Alan. M (1954), *Solvable and Unsolvable Problems*, **Science News**, 31, 7--23.

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

Velupillai, K. Vela (2010a), **Computable Foundations for Economics**, Routledge, London.

Velupillai, K. Vela (2010b), *Foundations of Boundedly Rational Choice and Satisficing Decisions*, **Advances in Decision Sciences**, 2010, 16 pages.

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