Hindawi Publishing Corporation
Advances in Decision Sciences
Volume 2010, Article ID 798030, 16 pages
doi:10.1155/2010/798030

Research Article

Foundations of Boundedly Rational Choice and Satisficing Decisions

K. Vela Velupillai

Department of Economics, University of Trento, Via Inama 5, 381 00 Trento, Italy

Correspondence should be addressed to K. Vela Velupillai, kvelupillai@gmail.com

Received 25 December 2009; Accepted 15 April 2010

Academic Editor: James Moffat

Formally, the orthodox rational agent's "Olympian" choices, as Simon has called orthodox rational choice, are made in a static framework. However, a formalization of consistent choice, underpinned by computability, suggests by, *satisficing* in a *boundedly* rational framework is not only more general than the model of "Olympian" rationality, it is also consistently dynamic. This kind of naturally process-oriented approach to the formalization of consistent choice can be interpreted and encapsulated within the framework of *decision problems*—in the formal sense of *metamathematics* and mathematical logic—which, in turn, is the natural way of formalizing the notion of *Human Problem Solving* in the Newell-Simon sense.

1. Introduction[1]

No one person better combined and encapsulated, in an intrinsically *dynamic, decision-theoretic* framework, a *computationally founded* [2] system of *choice* and *decision*, both entirely rational in a broad sense, than Herbert Simon. In this paper, I try, by standing on the shoulders of Herbert Simon, in fairly precise and formal ways, to suggest *computable foundations* for *boundedly rational* choice and *satisficing* decisions. In a nutshell, the aim is to reformulate, with textual support from Herbert Simon's characterizations and suggestions, bounded rationality and satisficing *in a computable framework* so that their intrinsic (complex) dynamics is made explicit in as straightforward a way as possible. To achieve this aim, in the tradition of Simon, I start from orthodox underpinnings of rational choice theory and extract its inherent procedural content, which is usually submerged in the inappropriate mathematics of standard real analysis.

Before proceeding with an outline of the contents and structure of this paper, it may be useful and apposite to remark on recent resurgences of interests in resurrecting Simon's original definition of—in particular—bounded rationality, especially in the important, interesting, and influential writings of Gigerenzer and Selten [1] and Smith [2]. They have all made it clear, in these cited writings and in their more recent, and not-so-recent, articles and books, that the way bounded rationality has been interpreted by what I have come to call *Modern Behavioural Economics* [3] is not faithful to the letter and spirit of Simon's original definitions. Where I part ways with these giants of game theory and behavioural and experimental economics is in the consistent, almost relentless, way in which I cast the Simonian world of behavioural economics—not just bounded rationality—within a framework of *computability theory*.

In his fascinating and, indeed, provocative and challenging chapter titled "*What is bounded rationality*" (cf., [1, op.cit., Chapter 2, page 35]), Reinhard Selten first wonders what bounded rationality is,and then goes on to state that an answer to the question "cannot be given" now:

> "What is bounded rationality? A complete answer to this question cannot be given at the present state of the art. However, empirical findings put limits to the concept and indicate in which direction further inquiry should go."

In a definitive sense—entirely consistent with the computational underpinnings Simon always sought—I try to give a "complete answer" to Selten's finessed question. I go further and would like to claim that the "limits to the concept" derived from current "empirical findings" cannot point the direction Simon would have endorsed for "further inquiry" to go—simply because current frameworks are devoid of the computable underpinnings that were the hallmark of Simon's behavioural economics.

It may well be apposite, in this particular context of a reference to Selten's challenging remark, to also comment—in a very general way—on "*heuristics*".[4] In Herbert Simon's overall vision and work, the place of *heuristics* is crucial. It appears from almost the very beginning of his work on *HumanProblem Sloving*[5](cf., [3]) as *procedures* that are devised to search, in a structured way, in spaces that are *computationally complex.* As always in Simon's work, the notion of "*computationally complex*" is underpinned by a *model of computation*. Almost without exception, the *model of computation* underpinning all of Simon's *procedural economics*—whether of problem solving or of any other aspect of decision making by reasonable man or in organisations—is the *Turing Machine*[6].Essentially, of course, a heuristic is a procedure which is, more precisely, an *algorithm*. The mathematical foundations of algorithms are provided by either *recursion theory* or *constructive mathematics*. In this paper I confine myself to recursion theoretic foundations for Simon's vision on some aspects of his work. I am not particularly interested in the secondary literature on heuristics—whether of the *fast and frugal variety* or any other variety—mainly because *none* of them are based on the mathematical foundations of the theory of algorithms.

In the next section, some substantiation for "standing on Simon's shoulders" will be outlined. On the basis of Simon suggestion's given in Section 2, I go on, in Section 3, to outline the kind of formalism that provides computable foundations for a complexity approach to decision theory and choice, both rationally conceived. In Section 4, suggestions on the formal machinery that can be built, to make explicit the kind of dynamic and computational complexities intrinsic to the computable foundations of decision and choice, are given. A brief concluding Section 5, summarizes the results and ends with brief signposts towards the care that must be taken in assertions about bounded rationality and satisficing as special cases of, or constrained versions of, the orthodox formalisms.

Several important background caveats on the mathematical underpinnings of the computable methodology with which I approach the issues tackled in this paper must be pointed out, at the very outset—lest the unwary or unhoned (in algorithmic mathematics) reader concentrates on inessentials. The main complexity concept I will ultimately be interested in, for rationally conceived decisions and choices, is *computational complexity* (although the kind of *dynamic complexity*, associated with formal dynamical systems, that also will be discussed, can be "*reduced*" to formal computational complexity).

Computational complexity theory is doubly related to mathematical economics and economic theory: firstly, as a theory of the *efficiency of computations* it is best viewed as the *economic theory of computations*; secondly, in having at its central core the paradigmatic combinatorial, intractable, *NP-Complete*, *Travelling Salesperson's Problem* (TSP). In the former case, it must first be remembered that the pure theory of computations abstracts away from all kinds of *resource constraints*. Computational complexity theory, the "applied" theory of computation, is its finessing, taking explicit account of resource constraints, typically time and space constraints. One of the modern pioneers of computational complexity theory, Karp, perceptively noted [4, page 464], Italics added:

> "[I] do think there are some very worthwhile and interesting analogies between complexity issues in computer science and in economics. For example, economics traditionally assumes that the agents within an economy have universal computing power and instantaneous knowledge of what's going on throughout the rest of the economy. Computer scientists deny that an algorithm can have infinite computing power. *They're in fact studying the limitations that have arisen because of computational complexity. So, there's a clear link with economics*."

Unfortunately, where even this generous analogy is misleading is in assuming that "economics traditionally assumes that the agents within an economy have universal computing power." In fact, not even this fantastic assumption is explicitly made "in economics" (unless it is of the Simonian variety of *behavioural economics*). This is why it is important to be aware that, in computational complexity theory, the characterizing framework is one of *problem solving*, with *a model of computation* explicitly underpinning it, as *decision problem.*

Now, a *decision problem* asks whether there exists an *algorithm* to *decide* whether a mathematical assertion does or does not have a proof; or whether a formal problem does or does not have a solution. Thus the characterization makes clear the crucial role of an underpinning model of computation; secondly, the answer is in the form of a *yes/no* response. Of course, there is the third alternative of "*undecidable*", too, but that is a vast issue outside the scope of this paper. It is in this sense of *decision problems* that I will interpret the word "decisions" in this paper.

As for "problem solving", I will assume that this is to be interpreted in the sense in which it is defined and used in the monumental classic by Newell and Simon [3].

Decisions, in the computational and *problem solving* tradition of Herbert Simon, have this same general and fundamental characterization in *computable economics*.

Finally, the *model of computation*, in the above senses and contexts, is the *Turing model*, subject to the *Church-Turing Thesis*. I will adhere to this tradition, but—at least for my results and propositions—this is only for convenience; I believe that all my formal results can also be derived without assuming the Church-Turing Thesis, hence within the formalism of constructive mathematics.

2. Standing on Simon's Shoulders[7]

In this section I will try to provide a "Simonian context" for the way I aim to tackle the problem of a "computable approach" to "decisions and choice". This is provided by means of two extensive "quotations"—one, from a long letter Herbert Simon wrote me, in May, 2000; and the other, from one of his classic pieces. They make explicit his visions of *complexity*, based on the *Turing model of computation* and the nature of the way internal and external constraints determine *satisficing* in a *boundedly rational* context. I proceed in this unconventional way simply to make it clear, from the outset, that my own vision is that a boundedly rational agent satisficing by implementing (rational) decisions *is the general case*; the Olympian model of rational choice—the orthodox model—is the special case.

On May 25th, 2000, Herbert Simon wrote me as follows (referring to having read my book to in [5]; emphases added):

> I want to share some first impressions on my reading of "Computable Economics."... I was delighted and impressed by the mileage you could make with Turing Computability in showing how nonsensical the Arrow/Debreu formulation, and others like it, are as bases for notions of human rationality. Perhaps this will persuade some of the formalists, where empirical evidence has not persuaded them, of what kinds of thinking humans can and cannot do—especially when dealing with the normative aspects of rationality....
>
> As the book makes clear, my own journey through bounded rationality has taken a somewhat different path. Let me put it this way. There are many levels of complexity in problems, and corresponding boundaries between them. Turing computability is an outer boundary, and as you show, any theory that requires more power than that surely is irrelevant to any useful definition of human rationality. A slightly *stricter boundary is posed by computational complexity*, especially in its common "worst case" form. We *cannot expect people (and/or computers) to find exact solutions for large problems in computationally complex domains*. This still leaves us far beyond what people and computers actually can do. The next boundary, but one for which we have few results except some of Rabin's work, is *computational complexity for the "average case"*, sometimes with an "almost everywhere" loophole. That begins to bring us closer to the realities of real-world and real-time computation. Finally, we get to the empirical boundary, measured by laboratory experiments on humans and by observation, of the level of complexity that humans actually can handle, with and without their computers, and—perhaps more important—what they actually do *to solve problems that lie beyond this strict boundary* even though they are within some of the broader limits....
>
> The latter is an important point for economics, because *we humans* spend most of our lives *making decisions that are far beyond any of the levels of complexity we can handle exactly*; and *this is where satisficing*, floating aspiration levels, recognition and heuristic search, and similar *devices for arriving at good-enough decisions* [8]*take over*. A parsimonious economic theory, and an empirically verifiable one, shows how *human beings, using very simple procedures, reach decisions that lie far beyond their capacity for finding exact solutions by the usual maximizing criteria* ...

> So I think we will continue to proceed on parallel, but somewhat distinct, paths for examining *the implications of computational limits for rationality*—you the path of mathematical theories of computation, I the path of *learning how people* in fact *cope with their computational limits*. I will not be disappointed however if, in the part of your lives that you devote to experimental economics, you observe phenomena that seduce you into incorporating in your theories some of these less general but very real departures from the rationality of computational theory. This seems to me especially important if we are to deal with the mutual outguessing phenomena (will we call them the Cournot effects?) that are the core of game theory.
>
> I am sure that you will be able to interpret these very sketchy remarks, and I hope you will find reflected in them my pleasure in your book. While I am fighting on a somewhat different front, I find it greatly comforting that these outer ramparts of Turing computability are strongly manned, greatly cushioning the assault on the inner lines of empirical computability.

Several important issues are clarified by Simon in these elegant observations. First of all, the defining—and decisive—role played by the *Turing model of computation* as the benchmark for his own fundamental work on *computationally underpinned work on rationality*—that is, bounded rationality—and satisficing decisions. Secondly, it is also unambiguously clear that the various boundaries delineated and defined by *computational complexity theory*—based, of course, on the Turing model of computation—are with reference to *the problems* that boundedly rational agents try to solve—that is, the level of complexity is that which is defined by the nature of the problem to be solved, not determined *solely* by the complexity of the computational architecture of the boundedly rational agent. Thirdly, boundedly rational agents actually do solve "problems that lie beyond the strict boundary" of formally feasible, computationally solvable, problems. The hint may well be that boundedly rational agents do discover, by heuristic means, methods to satisfactorily solve problems that computational complexity theory places beyond the empirically feasible range.[9] To the extent that computational complexity theory is underpinned by a model of computation, formal complexity boundaries are defined for the degrees of solvability of computable problems; uncomputable problems are beyond the "outer boundary". Fourthly, and perhaps most importantly, boundedly rational agents actually solve decision problems, in a satisficing framework, that lie beyond the orthodox domains of solvability—perhaps the best way to state this is that *Olympian means and aims* are not capable of solving the problems framed within the *Olympian model of rational choice*. The key to interpret this important observation by Simon is to note that the traditional, *half-naked*, framework of "optimization" is replaced by the fully-clothed one of *decision problems*. The *half-naked* nature of the Olympian framework is due to the absence of a "model of computation" to underpin its formalization—and that, in turn, is almost entirely due to the unfortunate reliance of the mathematics of real analysis of a very truncated sort. This is the sort that is founded on set theory, with its uncomputable and nonconstructive handmaiden, the axiom of choice.

The above characterisations and comments are further strengthened by the following, even more explicit, commentaries by Simon, on the distinction between the internal and external constraints going into the definition of a boundedly rational agent's confrontation with a decision problem in a satisficing framework.

> "Now if an organism is confronted with the problem of behaving approximately rationally, or adaptively, in a particular environment, the kinds of simplifications

that are suitable may depend not only on the characteristics—sensory, neural, and other—of the organism, but equally upon the structure of the environment. Hence, we might hope to discover, by a careful examination of some of the fundamental structural characteristics of the environment, some further clues as to the nature of the approximating mechanisms used in decision making. . . .

[T]he term environment is ambiguous. I am not interested in describing some physically objective world in its totality, but only those aspects of the totality that have relevance as the "life space" of the organism considered. Hence, what I call the "environment" will depend upon the "needs," "drives," or "goals" of the organism and upon its perceptual apparatus." (see [6, page 21].)

The point, again, is *not* that the theoretical analyst is concerned with "absolute" constraints—either of the internal structure of the decision making entity, or of the external environment of which a problem is a part—and in which it is embedded. The relevant architecture of the decision making entity, in this case that of a *computationally conceived rational economic agent*, solves a decision problem embedded, and emerging from, an environment, also computationally underpinned. The approximations are two-pronged: one, on the architecture of the computationally conceived rational agent—that is, the boundedly rational agent; the other, on the computationally underpinned environment, now conceived within the satisficing framework of a decision problem. This does not entail, in any way at all, that the approximations of a computationally conceived agent is a special case of the orthodox rational agent in the Olympian mode of choice. Nor does it imply at all that the approximation of the decision problem in the satisficing framework is a special case of the Olympian model of indiscriminate optimization. The numerous attempts, claiming to be within a behavioural economics setting, because, for example, the agents are supposed to be boundedly rational *fail in the former sense*; that is, assuming that the agent in such allegedly behavioural settings are boundedly rational because they are assumed to be constrained—for example by having only "limited" memory, modelled as finite automata, rather than as Turing machines—versions of the Olympian agent. As for an example of the failure from the point of view of the second "vision"—regarding the approximations on, and of, the environment, the canonical example is, of course *the folly of considering an integer linear programming problem as a special case of the standard linear programming problem*.

In fact, this will be the illustrative example I will choose for my formal description and discussion of these distinctions, so as to find a way to state and define the case for the vision that places the boundedly rational agent in a satisficing setting to solve a decision problem as the general one—and the Olympian model as *a special, and uninteresting, case*.

3. Brief Remarks on Decision Problems

"By a *decision procedure* for a given formalized theory T we understand a method which permits us to decide in each particular case whether a given sentence formulated in the symbolism of T can be proved by means of the devices available in T (or, more generally, can be recognized as valid in T). The *decision problem* for T is the problem of determining whether a decision procedure for T exists (and possibly for exhibiting such procedure). A theory T is called *decidable* or *undecidable* according as the solution of the decision problem is positive or negative." (see [7, page 3]; italics in the original.)

A *decision problem* asks whether there exists an algorithm to decide whether a mathematical assertion does or does not have a proof; or a formal problem does or does not have a solution.

Thus the characterization must make clear the crucial role of an underpinning model of computation; secondly, the answer is in the form of a yes/no response.

Of course, there is the third alternative of "undecidable", too, but that is a vast issue outside the scope of this paper.

Remark 3.1. Decidable-Undecidable, Solvable-Unsolvable, Computable-Uncomputable, and so forth, are concepts that are given content algorithmically.

The three most important classes of decision problems that almost characterise the subject of computational complexity theory, *underpinned by a model of computation*,[10] are the **P**, and **NP-Complete** classes.

Concisely, but not quite precisely, they can be described as follows:

(1) **P** defines the class of computable problems that are solvable in time bounded by a polynomial function of the size of the input.

(2) **NP** is the class of computable problems for which a solution can be *verified* in polynomial time.

(3) A computable problem lies in the class called **NP-Complete** if every problem that is in **NP** can be *reduced to* it in polynomial time.

Consider the following three-variable Boolean formula:

$$\neg x_3 \wedge (x_1 \vee \neg x_2 \vee x_3). \tag{3.1}$$

Just as in the case of equations with integer (or rational) values, given a truth assignment $t(x_i) = 1$ or 0 for each of the variables x_i $(i = 1, \ldots, 3)$, the above Boolean formula can be evaluated to be true or false, globally. For example the following assignments gives it the value *true*: $t(x_1) = 1;\ t(x_2) = 1;\ t(x_3) = 0$. Boolean formulas which can be made true by some truth assignments are said to be *satisfiable*.

Now consider the Boolean formula:

$$(x_1 \vee x_2 \vee x_3) \wedge (x_1 \vee \{\neg x_2\}) \wedge (x_2 \vee \{\neg x_3\}) \wedge (x_3 \vee \{\neg x_1\}) \wedge (\{\neg x_1\} \vee \{\neg x_2\} \vee \{\neg x_3\}). \tag{3.2}$$

Remark 3.2. Each subformula within parenthesis is called a clause; The variables and their negations that constitute clauses are called literals; It is "easy" to "see" that for the truth value of the above Boolean formula to be $t(x_i) = 1$, all the subformulas within each of the parenthesis will have to be true. It is equally "easy" to see that no truth assignments whatsoever can satisfy the formula such that its global value is true. This Boolean formula is unsatisfiable.

Problem 1. SAT—*The Satisfiability Problem*

Given m clauses, C_i $(i = 1, \ldots, m)$, containing the literals (of) x_j $(j = 1, \ldots, n)$, *determine* if the formula $C_1 \wedge C_2 \wedge \cdots \wedge C_m$ is *satisfiable*.

Determine means "find an (efficient) algorithm". To date it is not known whether there is an *efficient* algorithm to solve the satisfiability problem—that is, to determine the truth value of a Boolean formula. In other words, it is not known whether SAT $\in$ **P**. But the following id considered.

Theorem 3.3. *SAT* $\in$ **NP**

Now to go from here to an optimization framework is a purely mechanical affair. Denoting the union operator as ordinary addition and the negation operator related to arithmetic operators as: $\neg x = (1 - x)$ and noting that it is necessary, for each clause C, there should, at least, be one true literal, we have, for any formula

$$\sum_{x \in C} x + \sum_{x \in C} (1 - x) \geq 1. \tag{3.3}$$

With these conventions, the previous Boolean formula becomes the following *integer linear programming* (*ILP*) problem:

$$\begin{gathered} x_1 + x_2 + x_3 \geq 1, \\ x_1 + (1 - x_2) \geq 1, \\ x_2 + (1 - x_3) \geq 1, \\ x_3 + (1 - x_1) \geq 1, \\ (1 - x_1) + (1 - x_2) + (1 - x_3) \geq 1, \\ 0 \leq x_1, x_2, x_3 \leq 1, \quad \text{and integer.} \end{gathered} \tag{3.4}$$

Definition 3.4. A Boolean formula consisting of many clauses connected by conjunction (i.e., $\wedge$) is said to be in Conjunctive Normal Form (CNF).

Remark 3.5. A CNF is satisfiable if and only if the equivalent ILP has a feasible point.

Clearly, the above system of equations and inequalities do not, as yet, represent an ILP since there is no "optimization". However, it can be turned into a complete ILP in the ordinary sense by, for example, replacing the first of the above inequalities into:

$$\text{Max } y, \quad \text{s.t.} \quad x_1 + x_2 + x_3 \geq y \tag{3.5}$$

Remark 3.6. The formula is satisfiable if and only if the optimal value of y, say $\hat{y}$ exists and satisfies $\hat{y} \geq 1$.

Finally, we have Cook's famous theorem, rounding off all these connections and bringing into the fold of computational complexity theory, the quintessential combinatorial economic optimization problem.

Theorem 3.7 (Cook's theorem). *SAT is* **NP**-*Complete*

It is the above kind of context and framework within which I am interpreting Simon's vision of behavioural economics. In this framework optimization is a very special case of the more general decision problem approach. The real mathematical content of *satisficing* [11] is best interpreted in terms of the satisfiability problem of computational complexity theory, the framework used by Simon consistently and persistently—and a framework to which he himself made pioneering contributions.

4. Bounded Rationality as a Superset of Olympian Rationality

> Linear Programming problems are solvable in polynomial time.... in "Integer Linear Programming", we come to a field where the problems in general are less tractable, and are NP-Complete. It is a general belief that these problems are not solvable in polynomial time. The problems in question are
>
> (i) solving systems of linear diophantine inequalities, that is, solving linear inequalities in integers,
>
> (ii) solving systems of linear equations in nonnegative integer variables,
>
> (iii) solving *integer linear programming* problems.
>
> "[T]hese three problems are equivalent in the sense that any method for one of them yields also methods for the other two. Geometrically, the problems correspond to the intersection of a lattice and a polyhedron." (see the study by Schrijver in [8, page 2-3]; italics in the original.)

The simple analogy I wish to appeal to, for substantiating the case that the Boundedly Rational Agent is the general case and the Olympian Agent is the special case, is in terms of the classic difference between Integer Linear Programming and Linear Programming. *From the point of view of problem solving, underpinned by a model of computation*, the former is unambiguously the more general and the more *complex* case; the latter is the less general, *simple* case. It must also be emphasized that "more complex" refers to the precise sense of computational complexity—as made clear by reference to NP-Complete in the above quote.

Consider the following abstract version of a formalization of what may be called the standard economic optimization problem (SEP):

$$\begin{aligned} &\text{Minimize} && f(x) \\ &\text{subject to} && g_i(x) \geq 0, \quad i = 1, 2, \ldots, m, \\ &\text{and} && h_j(x) = 0, \quad h_j = 1, 2, \ldots, p \end{aligned} \tag{4.1}$$

(Naturally, with standard—i.e., "convenient but irrelevant"—assumptions on f, g, and h).

Now, consider the following variant of SEP.

Definition 4.1 (SEP*). An optimization problem is a pair $\{F, c\}$, where:

F is the set—the domain—of possible alternatives,

$c : F \to \mathfrak{R}$ (e.g., the criterion function), Then the *problem to solve*, associated with SEP* *is*: Find $f \in F$, such that $c(f) \leq c(g)$, for all $g \in F$.

Now, make explicit the computational content of an SEP* as follows.

Definition 4.2 (SEP$^{\text{TM}}$). (i) Given a *combinatorial object* (i.e., a number-theoretically specified object) f and a set of parameters, S, *decide* whether $f \in F$ (where F is characterized by S).

(ii) Assume that this *decision procedure* is executed by algorithm T_f (standing for the *Turing Machine* indexed by f, which has been *effectively* encoded, number-theoretically).

(iii) After the decision implemented by T_f, use another (*algorithmic*) decision procedure to *compute* the value $c(f)$, where c is characterised by the set of parameters Q. Call this latter decision procedure T_c.

(iv) Note that S and Q are to be represented number-theoretically—for example, *Gödel-numbered*.

Remark 4.3. Firstly, to start with a "given combinatorial object" ab initio is part of the claim to generality of the decision problem approach to problem solving in the satisficing, boundedly rational, vision. Secondly, the combinatorial object is encoded number theoretically to be processed by a model of computation. Simon does not always assume that the human problem solver is endowed with the full facilities of the most powerful model of computation (subject to the Church-Turing Thesis), but limited by various psychological and neurologically determined and informed factors. It is in this step that the qualification *limited* or *bounded* gets its full significance in a problem solving context. Satisficing, however, comes together with the *decision problem* approach to problem solving, that is, in the third of the above four-step scheme. Finally, approximating the combinatorial object suitably, by the agent or the problem solver, is the step where the structure of the environment [6] comes into play.

Now, consider the standard integer linear programming (SLIP) problem as an example of SEP$^{\text{TM}}$ as follows.

Minimize $c'x$ such that $Ax = b$ and $x \geq 0$, and possibly also c, b, and $A \in \mathbb{N}$ (the variables are, naturally, vectorial of suitable dimensions).

According to the SEP$^{\text{TM}}$ interpretation this means the following.

(i) The parameters S, for the decision procedure T_f, are given by A, b.

(ii) Given any integer (vector) x, T_f *decides* whether $Ax = b$ and $x \geq 0$ are simultaneously satisfied.

(iii) "Then", T_c is implemented, which has c for Q to evaluate $c'x$ for each x decided by T_f.

Remark 4.4. "Then", in the third step above, does not necessarily imply sequential actions by TMs. More complex decision tasks, employing richer varieties of SEP$^{\text{TM}}$ could imply a set of TMs operating on a parallel architecture and executing decisions both synchronously and asynchronously. However, Simon almost invariably worked within a sequential, synchronous, framework—although he was, of course, quite familiar with the richer relative possibilities of parallel architectures.

The two main conclusions of this section are the following. Firstly, given the computational underpinning of a problem solving approach to rational decision making and, therefore, the necessity of a model of computation to implement a decision problem, every such process has an intrinsic complexity measure in terms of computational complexity theory—in general in the form of *NP-Completeness*. Secondly, the whole setup is naturally more general than the setting in which the Olympian Model is framed and formalized.

5. Computable Rational Agents and Satisficing[12]

> "The theory proclaims *man to be an information processing system*, at least *when he is solving problems* . . . .
>
> An information processing theory is dynamic,. . ., in the sense of describing the change in a system through time. Such a theory describes the time course of behavior, characterizing each new act as a function of the immediately preceding state of the organism and of its environment. . . .
>
> The natural formalism of the theory is the program, which plays a role directly analogous to systems of differential equations in theories with continuous state spaces. . .
>
> All dynamic theories pose problems of similar sorts for the theorist. Fundamentally, he wants to infer the behavior of the system over long periods of time, given only the differential laws of motion. Several strategies of analysis are used, in the scientific work on dynamic theory. The most basic is taking a completely specific initial state and tracing out the time course of the system by applying iteratively the given laws that say what happens in the next instant of time. *This is often*, but not always, *called simulation*, and is one of the chief uses of computers throughout engineering and science. It is also the mainstay of the present work." (see the study by Newell and Herbert [3, pages 9–12]; italics added.)

The point here is that a (rational) problem solving entity is considered to be an information processing system, which is intrinsically dynamic, encapsulated in the "program" and, hence, naturally analogous to the role played by, say, "differential equations", in classical dynamics.[13] With this in mind, and against the backdrop provided by the discussion in the previous section, the strategy for my formalization exercise can be summarized in the following sequence of steps.

(i) Extract the procedural content of orthodox rational choices (in theory).

(ii) Formalize such a procedural content as a process of computation.

(iii) The formalized procedural content is Given as a process of computation, to be able to discuss its computational complexity.

(iv) Show the equivalence between a process of computation and a suitable dynamical system.

(v) The possibility of nonmaximum rational choice.

(vi) Then, that such behaviour is that which is manifested by a boundedly rational, satisficing, agent.

5.1. Rational Choice as a Computation by a Universal Turing Machine

> "In situations that are complex and in which information is very incomplete (i.e., virtually all real world situations), the behavioral theories deny that there is any magic for producing behavior even approximating an objective maximizing of profits and utilities. They therefore seek to determine what the actual frame of the decision is, *how that frame arises from the decision situation*, and *how, within that frame, reason operates*.
>
> In this kind of complexity, there is no single sovereign principle for deductive prediction. The *emerging laws of procedural rationality* have much more the complexity of molecular biology that the simplicity of classical mechanics." See the study by Simon in [9, page S223]; italics added.

The following result encapsulates, formally, the content of the first three steps of the above six-step scheme.

Theorem 5.1. *The process of rational choice by an economic agent is formally equivalent to the computing activity of a suitably programmed (Universal) Turing machine.*

Proof. By construction. See, [5, Section 3.2, pages 29–36] □

Remark 5.2. The important caveat is "process" of rational choice, which Simon—more than anyone else—tirelessly emphasized by characterizing the difference between "procedural" and "substantive" rationality; the latter being the defining basis for Olympian rationality, the former that of the computationally underpinned problem solver facing decision problems. Any decision—rational or not—has a time dimension and, hence, a content in terms of some process. In the Olympian model the "process" aspect is submerged and dominated by the static optimization operator, By transforming the agent into a problem solver, constrained by computational formalisms to determine a decision problem, Simon was able to extract the procedural content in any rational choice. The above result is a summary of such an approach.

Definition 5.3 (Computation Universality of a Dynamical System). A dynamical system—discrete or continuous—is said to be capable of computation universality if, using its initial conditions, it can be programmed to simulate the activities of any arbitrary Turing Machine, in particular, the activities of a Universal Turing Machine.

Lemma 5.4. *Dynamical Systems capable of Computation Universality can be constructed from Turing Machines.*

Proof. See [5, 10]. □

Theorem 5.5. *Only dynamical systems capable of computation universality are consistent with rationality in the sense that economists use that term in the Olympian Model.*

Proof. See, [5, page 49-50]. □

Remark 5.6. This result, and its proof, depend on the first theorem in this subsection and, therefore, its background basis, as explained in the Remark following it, given above. In this way, following the Simon's vision as outlined in the opening quote of this section, the definition of rationality is divorced from optimization and coupled to the decision problems of an information processing problem solver, emphasizing the procedural acts of choice.

Theorem 5.7 (nonmaximum rational choice). *No trajectory of a dynamical system capable of universal computation can, in any "useful sense" (see Samuelson's Nobel Prize lecture, [11]), be related to optimization in the Olympian model of rationality.*

Proof. See [12]. □

Theorem 5.8. *Boundedly rational choice by an information processing agent within the framework of a decision problem is capable of computation universality.*

Proof. An immediate consequence of the definitions and theorems of this subsection. □

Remark 5.9. From this result, in particular, it is clear that the Boundedly Rational Agent, satisficing in the context of a decision problem, encapsulates the only notion of rationality that can "in any useful sense" be defined procedurally.

The above definitions, theorems and lemma give formal content to the six-point formalization strategy outlined at the beginning of this section.

6. Concluding Notes

> "In your opening chapter, you are very generous in crediting me with a major role in the attention of the economics profession to the need to introduce limits on human knowledge and computational ability into their models of rationality ... But you seem to think that little has happened beyond the issuance of a manifesto, in the best tradition of a Mexican revolution". (Simon to Rubinstein [13, page 189])

To give a rigorous mathematical foundation for bounded rationality and satisficing, as decision problems, it is necessary to underpin them in a dynamic model of choice in a computable framework. However, these are not two separate problems. Any formalization underpinned by a model of computation in the sense of computability theory is, dually, intrinsically dynamic. I believe—and hope—this has been demonstrated in this paper, in a setting that is entirely faithful to Herbert Simon's precepts and lifelong decision-theoretic research program. A by-product of the results in this paper is the exploitation of the mentioned duality between dynamical systems and computability. With this duality it was possible to show in what sense bounded rationality is the more general case, in the case of an information processing problem solver, set in the context of a decision problem, and the Olympian model is the special case.

A rational choice framework that is entirely underpinned by computability and dynamical systems theory is naturally amenable to complexity theoretic analysis—both in terms of standard computational complexity theories and the more vague dynamic complexity theories. In other, companion writings (see [10, 14]), I have developed these two themes in much greater detail and I will have to refer the interested reader to them for further developments and the many felicitous connections.

Most importantly, I hope the message in this paper disabuses unscholarly assertions about bounded rational behaviour being a case of approximations to, or constrained versions of, the Olympian Model and satisficing, concomitantly, a case of suboptimal decision process. These kinds of unscholarly assertions permeate every strand of modern behavioural economics and behavioural game theory and in *some* varieties of experimental economics. For example, in the case of games played by automata, bounded rationality is modelled in terms of finite automata, ostensibly to take into account "limited" memory as one case of constrained Olympian rationality. Nothing in the Olympian model has anything to do with any kind of model of computation. How, then, can a special case of that become a model for computation by a finite automaton? A similar series of misguided examples can be cited from modern behavioural economics and behavioural game theory—not to mention orthodox choice theory.

Simon's intuitive perception of the importance of computability theoretic underpinnings for choice theory had *dynamic, computable* and *empirical* origins. In this paper I have tried to make precise the dynamic– computable duality theoretically explicit. Together—the dynamic and the computable—they combine to produce a ferociously complex framework, when implemented *purely theoretically*. But, mercifully, serious *empirical* investigations—of the kind Simon routinely practised all his life—entails judicious simplifications, as indicated in that letter from Simon to me, from which I quoted at the beginning of this paper.

Acknowledgments

An earlier version of this paper was given as part of one of the two Herbert Simon Lectures the author delivered at the National Chengchi University in Taipei, on 23 March, 2010. The early inspiration to write this paper, as also a way of emphasizing the distinction between Classical and Modern Behavioural Economics, was provided by lectures on Behavioural Economics given by Professor Shu-Heng Chen, the authors's host in Taipei, at Trento, a few months ago. Naturally, he is not responsible for the remaining infelicities in this version of the paper.

Endnotes

1. My main motivation and justification for making the case I am outlining in this paper is the fact that Herbert Simon himself seemed to have endorsed my interpretation of his vision. This is substantiated by appeal to his detailed letter to me, just before his untimely demise. Computability and computational complexity were the defining bases for the behavioural economics he pioneered. This is quite different from current fashions in behavioural economics, even—or, perhaps, especially—those claiming adherence to the traditions broached by Herbert Simon.
2. "Computational" has always meant "computable" in the Turing sense, at least in my reading of Simon's magisterial writings. In particular, in the context of bounded rationality, satisficing and their underpinnings in the architecture of human thinking, it was the path broached by Turing that guided Simon's pathbreaking contributions. In a volume celebrating "*The Legacy of Turing*" (see [15, pages 81 and 101]), Simon's paper, *Machine as Mind*, began and ended as follows:

   > "The title of my talk is broad enough to cover nearly anything that might be relevant to a collection memorializing A.M. Turing. . . . If we hurry, we can catch up to Turing on the path he pointed out to us so many years ago."

3. I have been trying to make a clear distinction between *Modern* and *Classical Behavioural Economics* for many years, mostly in my lectures to graduate students. I identify the latter with the pioneering works of Herbert Simon, James March, Richard Nelson, Richard Day and Sidney Winter. The beginning of the former kind of *Behavioural Economics* is generally identified as Thaler (see [16]), for example by Camerer et al. (cf., [17, p.xxii]), although my own view would be to begin with the pioneering inspirations Ward Edwards (see [18, 19]). But this is not a paper on the history of the origins of behavioural economics and, therefore, I will not go into further details on this matter, at this juncture.
4. A slightly more detailed discussion of the logic of heuristics, in the context of problem solving, is appended to Part IV of [20].
5. In the context of the approach taken in this paper, the most interesting and relevant discussion of heuristics by Simon is in [21].
6. Naturally, within classical recursion theory, by the *Church-Turing Thesis*, such a model is formally equivalent to any of the other models of computation, such as partial recursive functions, Post Machines, λ-functions, and so forth. Simon was, of course, well aware of these results.
7. My first attempts at trying to make the case for boundedly rational, adaptive behaviour and satisficing, in solving decision problems in a computable framework, were made in chapter 4 (see [5]). To the best of my knowledge, no other work makes this point—whether in a computable framework or not.
8. *The famous Voltaire precept comes to mind: "The perfect is the enemy of the good"*!
9. I hope knowledgeable readers do not try to read into this sentence even the mildest of hints that "heuristics" make it possible to go beyond the "Turing Limits". Nothing of the sort is intended here—or, indeed, can be meaningfully intended in any theoretically rigorous computable context. No engineer in his or her right mind would try to build a machine that violates the second law of thermodynamics. Economists constantly build models of rationality that imply mechanisms of decision making that go beyond even the ideal. Simon spent more than half a lifetime pointing out this absurdity—one that underpins, by the way, models of decision making in *modern behavioural economics*.
10. In general, the model of computation in this context is the *Nondeterministic Turing Machine*.
11. In [22, page 295], Simon clarified the semantic sense of the word *satisfice*, by revealing the way he came to choose the word:

    > "The term "satisfice", which appears in the *Oxford English Dictionary* as a Northumbrian synonym for "satisfy", was borrowed for this new use by Simon (1956) in "Rational Choice and the Structure of the Environment" (i.e., [6])"

12. The "classic" attempt at formalizing Satisficing, from an orthodox point of view was the elegant paper by Radner [23]. However, within the formalism of decisions problems, in the sense defined in this paper, Radner's formalization is contrary to Simon's vision. Gigerenzer and Selten (op.cit.) have resurrected this "classic" vision, in terms of "aspiration levels", within the context of modern behavioural economics.
13. Indeed, even more so in modern dynamical systems theory, particularly in its post-Smale varieties.

References

[1] G. Gigerenzer and R. Selten, Eds., *Bounded Rationality: The Adaptive Toolbox*, The MIT Press, Cambridge, Mass, USA, 2001.

[2] V. L. Smith, *Rationality in Economics: Constructivist and Ecological Forms*, Cambridge University Press, Cambridge, UK, 2008.

[3] A. Newell and H. A. Simon, *Human Problem Solving*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1972.

[4] R. Karp, "Complexity and parallel processing: an interview with Richard Karp, a turing award interview with Karen A. Frenkel," in *ACM Turing Award Lectures—The First Twenty Years, 1966–1985*, ACM Press/ Addison-Wesley, Reading, Mass, USA, 1987.

[5] K. Velupillai, *Computable Economics*, Oxford University Press, Oxford, UK, 2000.

[6] H. A. Simon, "Rational choice and the structure of the environment, psychological review," *Psychological Review*, vol. 63, pp. 129–138, 1956, Reprinted as chapter 1.2, in: Models of Thought, Vol. 1, pp. 20–28, Yale University Press, New Haven, Conn, USA.

[7] A. Tarski, A. Mostowski, and R. M. Robinson, *Undecidable Theories*, Studies in Logic and the Foundations of Mathematics, North-Holland, Amsterdam, The Netherlands, 1953.

[8] A. Schrijver, *Theory of Linear and Integer Programming*, Wiley-Interscience Series in Discrete Mathematics, John Wiley & Sons, Chichester, UK, 1986.

[9] H. A. Simon, "Rationality in psychology and economics," *The Journal of Business*, vol. 59, no. 4, part 2, pp. S209–S224, 1986.

[10] N. Dharmaraj and K. V. Velupillai, "Nonlinear dynamics, complexity and randomness: algorithmic foundations," to appear in *The Journal of Economic Surveys*.

[11] P. A. Samuelson, "Maximum principles in analytical economics," in *Les Prix Nobel en 1970*, pp. 273–288, The Nobel Foundation, Stockholm, Sweden, 1971.

[12] K. V. Velupillai, "The impossibility of an effective theory of policy in a complex economy," in *Complexity Hints for Economic Policy*, S. Massimo and C. David, Eds., pp. 273–290, Springer, Milano, Italy, 2007.

[13] A. Rubinstein, *Modelling Bounded Rationality*, The MIT Press, Cambridge, Mass, USA, 1998.

[14] K. V. Velupillai, "A computable economist's perspective on computational complexity," in *The Handbook of Complexity Research*, J. B. Rosser Jr., Ed., chapter 4, pp. 36–83, Edward Elgar, Cheltenham, UK, 2009.

[15] H. A. Simon, "Machine as mind," in *Machines and Thought—The Legacy of Alan Turing*, P. Macmillan and A. Clark, Eds., chapter 5, pp. 81–101, Oxford University Press, Oxford, UK, 1996.

[16] R. Thaler, "Toward a positive theory of consumer choice," *Journal of Economic Behavior and Organization*, vol. 1, no. 1, pp. 39–60, 1980.

[17] C. F. Camerer, G. Loewenstein, and M. Rabin, Eds., *Advances in Behavioral Economics*, Princeton University Press, Princeton, NJ, USA, 2004.

[18] W. Edwards, "The theory of decision making," *Psychological Bulletin*, vol. 51, no. 4, pp. 380–417, 1954.

[19] W. Edwards, "Behavioral decision theory," *Annual Review of Psychology*, vol. 12, pp. 473–498, 1961.

[20] K. V. Velupillai, *Computable Foundations for Economics*, Routledge, London, UK, 2010.

[21] H. A. Simon, "The logic of heuristic decision making," in *Models of Discovery—and Other Topics in the Methods of Science*, chapter 3.2, pp. 154–175, H. A. Simon, D. Reidel, Dordrecht, The Netherlands, 1967–1977.

[22] H. A. Simon, "Satisficing," in *Models of Bounded Rationality—Empirically Grounded Economic Reason*, vol. 3, chapter IV.4, pp. 295–298, The MIT Press, Cambridge, Mass, USA, 1997.

[23] R. Radner, "Satisficing," *Journal of Mathematical Economics*, vol. 2, no. 2, pp. 253–262, 1975.