Available online at www.sciencedirect.com

ScienceDirect

Cognitive Systems Research 29–30 (2014) 40–52

Cognitive Systems RESEARCH

www.elsevier.com/locate/cogsys

Computable and computational complexity theoretic bases for Herbert Simon's *cognitive behavioral economics* ☆

Action Editor: Ron Sun

K.Vela Velupillai [a,b,*], Ying-Fang Kao [c]

[a] *Department of Economics, The New School for Social Research, 6 East 16th Avenue, New York, NY 10003, USA*
[b] *Algorithmic Social Sciences Research Unit, Department of Economics, University of Trento, Trento, Italy*
[c] *Algorithmic Social Sciences Research Unit & the School of Social Sciences, University of Trento, Via inama, 5, Trento, Italy*

Received 11 June 2013; accepted 17 July 2013
Available online 31 July 2013

Abstract

This paper aims to interpret and formalize Herbert Simon's *cognitive* notions of bounded rationality, satisficing and heuristics in terms of computability theory and computational complexity theory. Simon's theory of *human problem solving* is analyzed 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 orthodox, non-cognitive, Olympian rationality and optimization, respectively, and not the other way about.

Keywords: Bounded rationality; Satisficing; Heuristics; Computability; Computational complexity

1. Introduction

> 'What a person *cannot* do he or she *will not* do, no matter how strong the urge to do it.'Simon, 1996, p. 28; italics in the original

Herbert Simon's *cognitive behavioral economics*, is underpinned by *a model of computation*, highlighting the complexity of behavioral decision processes on the basis of *computational complexity theory*. Simon was aware, from the outset, of the theoretical possibilities of interpreting the emerging field of computer science providing a foundational anchor to his conviction that the best way to study decision problems in the behavioral 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 (Simon, 1983b) optimizing framework.

☆ We are both immensely grateful to our ASSRU colleagues, Stefano Zambelli and Ragu Ragupathy for many enlightening conversations. In addition, the extremely supportive and constructively critical comments and suggestions by two anonymous referees have helped improve this revised version quite considerably. We, as authors, are fully responsible for all remaining infelicities.
* Corresponding author at: Department of Economics, The New School for Social Research, 6 East 16th Avenue, New York, NY 10003, USA.
E-mail addresses: velupilk@newschool.edu (K.V. Velupillai), seldakao@gmail.com (Y.-F. Kao).

1389-0417/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cogsys.2013.07.005

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 research program in the refocusing of the social sciences as cognitive behavioral 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 behavioral 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.[1]

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, 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.

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 behaviorally 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.

We may point out Simon's cognitive behavioral economics, as characterized above, is to be distinguished sharply from what we have in other writings (Kao & Velupillai, 2012 and Kao & Velupillai, 2013) called *Modern Behavioral Economics*.[2] The latter is simply a 'revisionistic' form of traditional neoclassical, subjective expected utility maximization economics, with no anchoring whatsoever in a model of computation. This was strongly underlined by Simon himself, when he – presciently, as always – pointed out (Simon, 1958 (1982), p. 384, footnote 4):

> "There is a great danger at the present moment that the economists and statisticians will carry the day even within the territories of psychology and sociology. As can be seen from the recent review article, 'The Theory of Decision Making,' by Ward Edwards, *Psychological Bulletin*, July, 1954, and *Decision Processes*, edited by Robert Thrall et.al., John Wiley & Sons, 1954, the

[1] 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, 1949). 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. 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. The notion of computational complexity that underpins our visions of Simon's cognitive behavioural economics is formally equivalent to Solomonoff's definition in that classic, later to be called algorithmic complexity (or Kolmogorov complexity). It was this latter notion that Simon used in his classic on *The Architecture of Complexity* (Simon, 1962, p. 478).

[2] We have in the papers referred to above gone into greater substantiation for this distinction, where Simon's cognitive behavioural economics is referred to as *Classical Behavioural Economics*.