Cogn Comput (2017) 9:1–4
DOI 10.1007/s12559-016-9444-2

Anatomy of the Mind: a Quick Overview

Ron Sun [1]

Received: 22 November 2016 / Accepted: 9 December 2016 / Published online: 27 December 2016

Abstract The recently published book, "Anatomy of the Mind," explains psychological (cognitive) mechanisms, processes, and functionalities through a comprehensive computational theory of the human mind—that is, a cognitive architecture. The goal of the work has been to develop a unified framework and then to develop process-based mechanistic understanding of psychological phenomena within the unified framework. In this article, I will provide a quick overview of the work.

Keywords Cognition · Psychology · Cognitive architecture · Simulation · Motivation · Subsystem · Implicit

Introduction

The book, *Anatomy of the mind: exploring psychological mechanisms and processes with the Clarion cognitive architecture* [1], has recently been published by the Oxford University Press, as part of its *Oxford series on cognitive models and architectures*. This book aims to explain psychological (cognitive) mechanisms, processes, and functionalities through a comprehensive computational theory of the human mind, namely, a cognitive architecture. More specifically, the book does so through the Clarion cognitive architecture [2–6]. The goal of this work has been to develop a unified and comprehensive framework, and within a unified framework, to develop process-based mechanistic understanding of a large variety of psychological phenomena across domains and functionalities.

✉ Ron Sun
rsun@rpi.edu

[1] Department of Cognitive Science, Rensselaer Polytechnic Institute, Troy, NY, USA

The book describes the essential Clarion framework, its cognitive-psychological justifications, its computational instantiations, and its applications to capturing, simulating, and explaining various psychological phenomena and empirical psychological data. The book shows how models and simulations shed light on psychological mechanisms and processes through the lens of the unified framework (namely, Clarion). Below, I will provide some details of the Clarion project.

Some Details

This work has had, relatively speaking, a long history. In the summer of 1994, the ONR cognitive science basic research program issued a call for proposals, which prompted me to put together a set of emerging ideas. That was the beginning of Clarion. The grant from the ONR program enabled the development and the validation of the initial version of Clarion. During the 1998–1999 academic year, I spent my sabbatical leave at the NEC Research Institute. A theoretically oriented book on Clarion took shape during that period, which was subsequently published [3]. Starting in 2000, research grants from ARI enabled the further development of a number of subsystems within Clarion [7]. Then, from 2008 on, new grants from ONR enabled the extension of the work to social simulation and other related topics [1].

One question that naturally arises is: among many other cognitive architectures, why should one pay attention to, adopt, or even further develop Clarion? One might choose Clarion for the totality of the following reasons:

- Clarion is a cognitive architecture that is more comprehensive in scope than most other cognitive architectures [1].

Table 1 Issues and desiderata for Clarion

Fundamental issues relevant to Clarion (see [1], chapter 1 for details):
Ecological-functional perspective
Modularity
Multiplicity of representation
Dynamic interaction
Essential desiderata for Clarion (see [1], chapter 2 for details):
Sequentiality
Routineness
Trial-and-error adaptation
Implicit versus explicit processes
Synergistic interaction
Bottom-up and top-down learning
Procedural versus declarative processes
Motivational and metacognitive control

- Clarion is psychologically realistic to the extent that it has been validated through simulating and explaining a wide variety of psychological tasks, data, and phenomena (see [1], chapters 5–7).
- Its basic principles and assumptions have been extensively argued for and justified, in relation to a variety of different types of evidence (see [1], chapters 2–4).
- It has major theoretical implications, as well as practical relevance. It has provided useful explanations for a variety of empirical data, leading to a number of significant theories regarding psychological phenomena (more on this point later).
- In addition to addressing problems at the psychological level, it has also taken into account higher levels, for example, regarding social processes and phenomena, as well as lower levels [6].

In particular, Clarion has been successful in computationally modeling, simulating, accounting for, and explaining a wide variety of psychological data and phenomena. These simulations not only reproduced empirical data, but more importantly, they provided insight that led to some major new theories concerning important psychological functionalities. Some new theories that resulted from work on Clarion include:

- The theory of "bottom-up" learning (from implicit learning to explicit learning), as developed in Sun et al. [4]
- The theory of the implicit-explicit interaction and their synergistic effects on skill acquisition, as developed in Sun et al. [5]
- The theory of creative problem solving, as described in Helie and Sun [2]
- The theory of human motivation and its interaction with cognition, as described in Sun [8], as well as in related simulation papers (e.g., [9, 10]; and others)
- The theory of human reasoning (based on implicit and explicit interaction), as developed in Sun [11, 12] and Sun and Zhang [13]

Some general outlines, as well as some details, of the Clarion cognitive architecture are now in order here. First, a number of essential (philosophical and psychological) desiderata have been central to the conception of the Clarion framework, including those listed in Table 1. Together, they present a situated/embodied view of the mind in a generalized sense. Moreover, on top of that, motivational and metacognitive control of behavior has been emphasized in Clarion. For one thing, a full account of behavior must address why one does what one does; hence, motivational processes need to be understood [8]. In a way, cognition has evolved to serve the

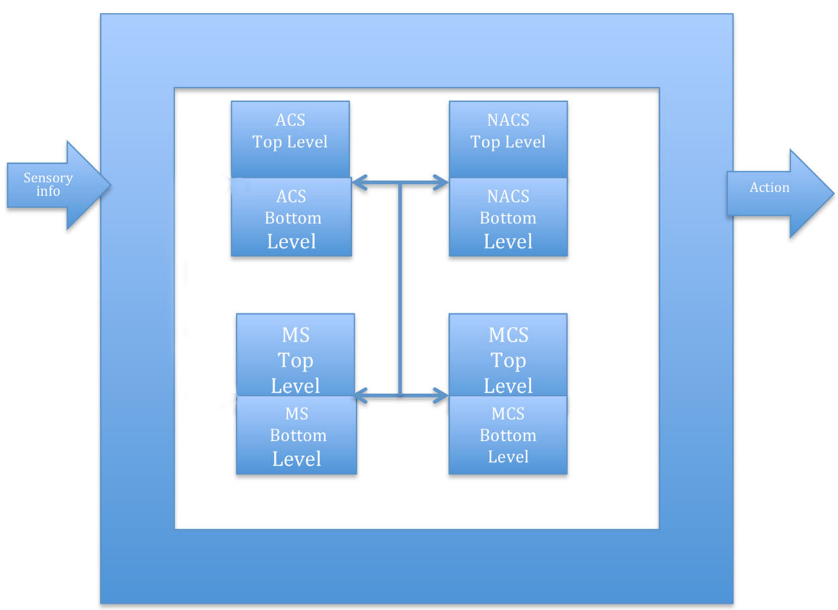

Fig. 1 The subsystems of the Clarion cognitive architecture. The major information flows are shown with arrows. See Sun [1] for technical details of the subsystems

Table 2 Some essential constituting ideas of Clarion

- A distinction exists between implicit and explicit processes (though they tend to interact), with different forms of representation and learning.
- Knowledge is often redundantly represented in both explicit and implicit forms. Implicit and explicit processes are often simultaneously involved in a task. Results of explicit and implicit processing are integrated (leading to possible synergy).
- A distinction exists between procedural and declarative processes, in terms of their contents and forms (including learning), orthogonal to the distinction between implicit and explicit processes.
- These processes are motivationally driven and modulated. Motivations consist of drives and goals, with goals determined mostly based on drives.
- Metacognition regulates these processes in a number of specific ways based (in part) on motivations.

essential needs (motives) of an individual and bridge the needs (motives) of an individual and his/her environments.

Based on these desiderata, Clarion was developed as an integrative cognitive architecture consisting of a number of distinct (but interacting) subsystems. These subsystems capture distinct types of representational contents (addressed by these desiderata above), in a functionally somewhat separate but mutually dependent and dynamically interacting fashion. See Fig. 1 for a sketch of the major subsystems of Clarion. In the figure, "ACS" stands for the action-centered subsystem; "NACS" stands for the non-action-centered subsystem; "MS" stands for the motivational subsystem; "MCS" stands for the metacognitive subsystem. Note that, among them, "ACS" captures procedural processes, whereas "NACS" captures declarative processes. Also, within each subsystem, "top level" captures explicit (consciously accessible) processes, whereas "bottom level" captures implicit processes.

Given the basic framework outlined above, Clarion naturally embodies a set of essential ideas concerning human cognition-psychology. Table 2 outlines briefly these basic ideas. They correspond to the desiderata on which Clarion was based as listed in Table 1. See Sun [1, 3] for further details of these basic ideas.

Back to the book in question ("Anatomy of the Mind"; [1]), it extensively argues for and applies these ideas in exploring psychological mechanisms and processes in different domains and functionalities. Within the book, the first chapter provides necessary background (such as those fundamental issues covered above). The rest of the book, divided into eight more chapters, describes various details of the work on Clarion. They include three chapters for presenting various theoretical, conceptual, and technical aspects of Clarion (chapters 2, 3, and 4), three chapters on various simulations using Clarion (chapters 5, 6, and 7), and additional materials (issues, questions, and comparisons) in the remaining two chapters (chapters 8 and 9).

Final Remarks

In all, Clarion is grounded in empirical research, is reasonably compact (given its broad scope), and captures a wide range of empirical data (as discussed in detail in the book [1]).

It is worth noting that comprehensive, integrative models, such as cognitive architectures, serve as important antidotes to the increasing specialization of scientific research, especially in the social and behavioral sciences. Cognitive architectures that integrate a broad range of functionalities go against over-specialization and help to fit pieces together again.

Thus, in fields ranging from cognitive science (especially cognitive modeling), to psychology, to artificial intelligence, and even to philosophy, academic researchers, graduate and undergraduate students, and practitioners of various kinds may have interest in topics covered by this book. The book may also be suitable for graduate-level seminars or courses on cognitive architectures, but might also be extended to use at the advanced undergraduate level.

Detailed information about the book can be found at https://global.oup.com/academic/product/anatomy-of-the-mind-9780199794553?cc=us&lang=en&.

Acknowledgements I would like to thank Dr. Amir Hussain for his invitation to write this overview. Thanks are also due to the many students, collaborators, and colleagues of mine, past and present, for their contributions to the Clarion project.

Compliance with Ethical Standards

Funding No funding for this article.

Conflict of Interest The author declares that he has no conflict of interest.

Ethical Approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Sun R. Anatomy of the mind: exploring psychological mechanisms and processes with the Clarion cognitive architecture. New York: Oxford University Press; 2016.
2. Helie S, Sun R. Incubation, insight, and creative problem solving: a unified theory and a connectionist model. Psychol Rev. 2010;117(3):994–1024.
3. Sun R. Duality of the mind. Mahwah, NJ: Lawrence Erlbaum Associates; 2002.
4. Sun R, Merrill E, Petersen T. From implicit skills to explicit knowledge: a bottom-up model of skill learning. Cogn Sci. 2001;25:203–44.
5. Sun R, Slusarz P, Terry C. The interaction of the explicit and the implicit in skill learning: a dual-process approach. Psychol Rev. 2005a;112(1):159–92.

6. Sun R, Coward LA, Zenzen MJ. On levels of cognitive modeling. Philos Psychol. 2005b;18(5):613–37.
7. Sun R. A tutorial on Clarion 5.0. Technical report, Cognitive Sciences Department, Rensselaer Polytechnic Institute, Troy, NY, 2003. http://www.cogsci.rpi.edu/~rsun/sun.tutorial.pdf.
8. Sun R. Motivational representations within a computational cognitive architecture. Cogn Comput. 2009;1(1):91–103.
9. Sun R, Wilson N. Roles of implicit processes: instinct, intuition, and personality. Mind and Society. 2014;13(1):109–34.
10. Wilson N, Sun R, Mathews R. Performance degradation under pressure. Neural Netw. 2009;22:502–8.
11. Sun R. Integrating rules and connectionism for robust commonsense reasoning. New York: John Wiley and Sons; 1994.
12. Sun R. Robust reasoning: integrating rule-based and similarity-based reasoning. Artif Intell. 1995;75(2):241–96.
13. Sun R, Zhang X. Accounting for a variety of reasoning data within a cognitive architecture. J Exp Theor Artif Intell. 2006;18(2):169–91.