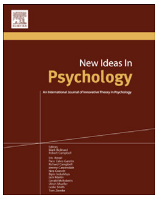
New Ideas in Psychology 34 (2014) 37–55

Contents lists available at ScienceDirect

New Ideas in Psychology

journal homepage: www.elsevier.com/locate/newideapsych



Autonomous learning in psychologically-oriented cognitive architectures: A survey

Sébastien Hélie [a], [*], Ron Sun [b]

[a] Department of Psychological Sciences, Purdue University, 703 Third Street, West Lafayette, IN 47907-2081, USA
[b] Rensselaer Polytechnic Institute, USA

Keywords:
Cognitive architectures
Autonomous learning
Psychology
ACT-R
Clarion
Soar

ABSTRACT

This survey paper discusses the topic of autonomous learning in psychologically-oriented cognitive architectures and reviews some of the most popular cognitive architectures used in psychology, namely ACT-R, Soar, and Clarion. Autonomous learning is critical in the development of cognitive agents, and several learning-related desiderata useful for 'psychological' cognitive architectures are proposed. This article shows that all the reviewed cognitive architectures include some form of explicit ('symbolic') and implicit ('subsymbolic') learning. Additionally, ACT-R and Clarion are shown to include a top–down learning algorithm (from explicit to implicit), and Clarion also includes a bottom–up learning process (from implicit to explicit). Two simulation examples are presented with each cognitive architecture to illustrate the autonomous learning capacities of each modeling paradigm. While Clarion is more autonomous (requiring less a priori knowledge), Soar and ACT-R have so far been used in more complex tasks. The presentation concludes with some general considerations for future work.

1. Introduction

This article surveys existing cognitive architectures in relation with autonomous learning for psychologically-realistic applications. A cognitive architecture is the essential structures and processes of a domain-generic computational cognitive model used for a broad, multiple-level, multiple-domain, analysis of cognition and behavior (Sun, 2004). Specifically, cognitive architectures deal with componential processes of cognition in a structurally and mechanistically well-defined way. Its function is to provide an essential framework to facilitate more detailed exploration and understanding of various components and processes of the mind. In this way, a cognitive architecture serves as an initial set of assumptions to be used for further development. These assumptions may be based on available empirical data (e.g., psychological or biological), philosophical thoughts and arguments, or computationally-inspired hypotheses concerning psychological processes (Sun, 2002). A cognitive architecture is useful and important precisely because it provides a comprehensive initial framework for further modeling and simulation in many task domains.

In order to achieve generality in a psychologically-realistic way, cognitive architectures should include only minimal initial structures and independently learn from their own experiences – that is, autonomous learning (Sun, 2000). Autonomous learning is an important way of developing additional structure, bootstrapping all the way to a full-fledged cognitive model (Sun, 2004). However, it is important to be careful and devise only minimal initial learning capabilities that are capable of "bootstrapping" relevant capacities for whatever phenomenon is modeled (e.g., Sun, Merrill, & Peterson, 2001). By doing so, many

---

[*] Corresponding author. Tel.: +1 765 496 2692.
E-mail address: shelie@purdue.edu (S. Hélie).

http://dx.doi.org/10.1016/j.newideapsych.2014.03.002
0732-118X/© 2014 Elsevier Ltd. All rights reserved.

structures can be placed back into the world, instead of placing them in the head of the agent (Bickard, 1993; Hutchins, 1995; Sun, 2000). The avoidance of overly complicated initial structures, and thus the inevitable use of autonomous learning, may often help to avoid overly representational models that are designed specifically for the task to be achieved (Sun, 2000). This flexibility is essential in achieving general intelligence (Newell, 1990).

The remainder of this article is organized as follows. Section 2 discusses the importance of psychologically-oriented cognitive architectures for cognitive science. Section 3 discusses some learning-related considerations (i.e., desiderata) in the development of psychologically-realistic cognitive architectures. Section 4 reviews some of the most popular cognitive architectures that are currently used in psychology and cognitive science. Section 5 presents a discussion of autonomous learning in the cognitive architectures surveyed in Section 4. Section 6 describes some simulation examples to illustrate the learning capabilities of the architectures reviewed in Section 4. Section 7 presents a short summary and prescriptions for future work.

2. Why are psychologically-oriented cognitive architectures important?

While there are all kinds of cognitive architectures in existence, in this survey we are concerned specifically with psychologically-oriented cognitive architectures.[1] Psychologically-oriented cognitive architectures are particularly important because they shed new light on human cognition and therefore they are useful tools for advancing the understanding of cognition (Newell, 1990). In understanding cognitive phenomena, the use of computational simulation on the basis of cognitive architectures forces one to think in terms of processes, and in terms of details (Sun, 2002). Instead of using vague, purely conceptual theories, cognitive architectures force theoreticians to think clearly. They are therefore critical tools in the study of the mind (Newell, 1990). Cognitive psychologists who use cognitive architectures must specify a cognitive mechanism in sufficient detail to allow the resulting models to be implemented on computers and run as simulations. This approach requires that important elements of the models be spelled out explicitly, thus aiding in developing better, conceptually clearer theories. It is certainly true that more specialized, narrowly-scoped models may also serve this purpose, but they are not as generic and as comprehensive and thus may not be as useful to the goal of producing general intelligence (Sun, 2002, 2004).

It is also worth noting that psychologically-oriented cognitive architectures are the antithesis of “expert systems”: Instead of focusing on capturing performance in narrow domains, they are aimed to provide broad coverage of a wide variety of domains in a way that mimics human performance (Langley, Laird, & Rogers, 2009). While they may not always perform as well as expert systems, business and industrial applications of intelligent systems increasingly require broadly-scoped systems that are capable of a wide range of intelligent behaviors, not just isolated systems of narrow functionalities. For example, one application may require the inclusion of capabilities for raw image processing, pattern recognition, categorization, reasoning, decision-making, and natural language communications. It may even require planning, control of robotic devices, and interactions with other systems and devices. Such requirements accentuate the importance of research on broadly-scoped cognitive architectures that perform a wide range of cognitive functionalities across a variety of task domains.

3. Learning-related desiderata of psychologically-realistic cognitive architectures

This section presents a short list of desiderata for the development of psychologically realistic cognitive architectures (extended from Sun, 2004). Such desiderata are considered here because they are crucially important for understanding and modeling autonomous learning in a psychologically-realistic way (Sun, 2000, 2002, 2004; Sun et al., 2001). Here, we focus on desiderata of learning processes as well as on what needs to be learned. Other non-learning-related desiderata (e.g., compositionality, universal computation) can be found in Anderson and Lebiere (2003).

3.1. *Dichotomy of implicit and explicit processes*

Implicit processes are often described as inaccessible and imprecise, while explicit processes are contrasted as accessible and precise (Dreyfus & Dreyfus, 1987; Hélie & Sun, 2010; Reber, 1989; Smolensky, 1988; Sun, 1994; Sun, Slusarz, & Terry, 2005). These differences are related to some other well-known dichotomies such as ‘symbolic’ vs. ‘subsymbolic’ processing (Sun, 1994), conceptual vs. subconceptual processing (Smolensky, 1988), and conscious vs. unconscious processing (Sun, 2002). It can also be justified psychologically by the voluminous empirical studies of implicit and explicit learning (Reber, 1989), implicit and explicit memory (Schacter, Wagner, & Buckner, 2000), implicit and explicit perception (Bornstein & Pittman, 1992), and so on. These empirical dichotomies are closely related, and thus they can all serve as justification for a more general distinction between implicit and explicit cognitive processes.

3.2. *Explicit learning and memory*

Explicit memories are those accessible to conscious awareness (Eichenbaum, 1997). Typically this includes working memory, episodic memory, and often also (explicit) semantic memory (Anderson et al., 2004). There are several types of explicit learning, such as memorization and hypothesis testing (Ashby & O’Brien, 2005). Memorization is useful mostly in situations where no further elaboration is required (Craik & Tulving, 1975) whereas hypothesis testing (Evans, 2006) can be used not only to

---

[1] As opposed to software engineering-oriented cognitive architectures (e.g., LIDA; Franklin & Patterson, 2006), cognitive robotic architectures (e.g., SS-RICS; Kelley & Avery, 2010), or neurally-oriented cognitive architectures (e.g., ART; Carpenter & Grossberg, 1987).