Biologically Inspired Cognitive Architectures (2013) 5, 52–63

Available at www.sciencedirect.com

SciVerse ScienceDirect

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

RESEARCH ARTICLE

An architecture for observational learning and decision making based on internal models

Haris Dindo [a,*], Eric Nivel [b], Giuseppe La Tona [a], Antonio Chella [a], Kristinn R. Thórisson [b,c]

[a] RoboticsLab, DICGIM, University of Palermo, Viale delle Scienze, 90128 Palermo, Italy
[b] CADIA/School of CS, Reykjavik University, Menntavegur 1, 101 Reykjavik, Iceland
[c] Icelandic Institute for Intelligent Machines, 2.h. Uranus, Menntavegur 1, 101 Reykjavik, Iceland

Received 23 March 2013; received in revised form 20 May 2013; accepted 22 May 2013

KEYWORDS
Cognitive architecture;
Internal models;
Imitation learning;
Simulation;
Anticipation

Abstract
We present a cognitive architecture whose main constituents are allowed to grow through a situated experience in the world. Such an architectural growth is bootstrapped from a minimal initial knowledge and the architecture itself is built around the biologically-inspired notion of *internal models*. The key idea, supported by findings in cognitive neuroscience, is that the same internal models used in overt goal-directed action execution can be covertly re-enacted in simulation to provide a unifying explanation to a number of apparently unrelated individual and social phenomena, such as state estimation, action and intention understanding, imitation learning and mindreading. Thus, rather than reasoning over abstract symbols, we rely on the biologically plausible processes firmly grounded in the actual sensorimotor experience of the agent. The article describes how such internal models are learned in the first place, either through individual experience or by observing and imitating other skilled agents, and how they are used in action planning and execution. Furthermore, we explain how the architecture continuously adapts its internal agency and how increasingly complex cognitive phenomena, such as continuous learning, prediction and anticipation, result from an interplay of simpler principles. We describe an early evaluation of our approach in a classical AI problem-solving domain: the Sokoban puzzle.

* Corresponding author.
E-mail addresses: haris.dindo@unipa.it (H. Dindo), eric.nivel@gmail.com (E. Nivel), giuseppe.latona10@unipa.it (G. La Tona), antonio.chella@unipa.it (A. Chella), thorisson@ru.is (K.R. Thórisson).

2212-683X/$ - see front matter
http://dx.doi.org/10.1016/j.bica.2013.05.007

1. Introduction

In the early days of AI, optimism about replicating a broad set of human-level cognitive skills in artificial agents was relatively common. As the decades went by the goal of building a *general* intelligence was substituted by a less ambitious one of building agents whose ''intelligence'' is measured against their capability to solve a well-defined (and narrow) set of problems, usually in relatively structured environments. While many approaches have proven to outperform humans in specific tasks, they still lack some of the most remarkable features of the human intelligence such as, for example, adaptiveness and robustness. Notable exceptions are provided by efforts to create cognitive architectures that take a holistic approach to intelligence, by integrating under the same theoretical umbrella various processes whose interoperation is aimed at giving rise to more complex forms of intelligence.

The architecture we are developing is of this kind: its capabilities are intended to grow through the system's situated experience in the world (Nivel et al., 2013). Such an architectural growth is bootstrapped from a minimal initial knowledge that the system uses as its first principles to build upon. The overall architectural pattern is inspired by a biologically-inspired notion of *internal models* which constitute a computational implementation of the *mirror neuron system* (Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). Indeed, findings in cognitive neuroscience have pushed forward the idea that the same internal models used in overt goal-directed action execution can be covertly re-enacted through a process of *motor simulation* to provide a unifying explanation to a number of apparently unrelated individual and social phenomena, such as motor control and state estimation, action and intention understanding, imitation learning, joint action and theory-of-mind (Wolpert, Doya, & Kawato, 2003; Pezzulo, Candidi, Dindo, & Barca, 2013) just to name a few; see also (Thórisson, 2012) for key underlying assumptions of this work.

Although limited in scope with respect to state of the art cognitive architectures as ACT-R (Anderson & Lebiere, 1998) or SOAR (Laird, 2012), our architecture is meant to merge ideas from the above cited approaches and less classical architectures based on situated sensorimotor loops such as MOSAIC (Wolpert & Kawato, 1998) or HAMMER (Demiris & Khadhouri, 2006). Indeed, executable knowledge is encoded into *internal models* and associated forward and inverse operators are used to implement higher cognitive functions such as learning and reasoning. In addition, we explicitly target the ability to *simulate* alternative curses of actions (Grush, 2004; Hesslow, 2002) and to *anticipate* actions that might prove to be useful in the future (Pezzulo, 2008). Unlike other cognitive architectures, our approach tightly integrates these capabilities into the overall decision making process. For example, the cognitive architecture Polyscheme uses simulation to integrate different representations and algorithms, but not as a support for action selection (Cassimatis, Trafton, Bugajska, & Schultz, 2004). Other studies use simulated sensory input to blindly control robot navigation but do not use simulation mechanisms to process abstract goals (Gigliotta, Pezzulo, & Nolfi, 2011; Ziemke, Jirenhed, & Hesslow, 2005).

This article is organized along two deeply intertwined dimensions: (1) how internal models are learned in the first place (we refer to this process as knowledge acquisition or learning interchangeably throughout the paper) and (2) how are they used in action planning, simulation and execution. Furthermore, we explain how a system based on these principles is able to continuously adapt its internal agency, and how continual learning and anticipation result from an interplay of simpler processes. These features have been recognized as one of the major ingredients of advanced intelligences and inserted in the roadmap for building biologically-inspired architectures (Chella, Lebiere, Noelle, & Samsonovich, 2011). In the next section we provide an overview of our architecture, its processes and their interaction. Then we discuss mechanisms of learning and action selection within our framework. Finally, we provide an early evaluation of our approach in a classical AI problem-solving domain: the Sokoban puzzle.

2. Overview of the architecture

The architecture we propose in this paper is characterized by two main features: it continuously expands its skills either through direct experience or by observing and imitating others (*learning*), and it provides tools for planning, executing and monitoring its own goal-directed actions (*reaction*). Before going into computational details of how learning and reaction processes are efficiently implemented, we first describe the fundamental building blocks of our architecture.

The system continuously gathers data from the environment and from its own inner activity. Every single sample of such an activity is represented as a key-value *message*, where the semantics of a key is to be interpreted by models. A message could represent, for instance, the sensed angle of a joint or a motor command. All such messages are shared by all the processes in a way similar to a blackboard architecture (Hayes-Roth, 1985). The set of messages available at a certain time t constitutes the *state* S_t of the system.

On the other hand, operational knowledge is encoded via internal models. We distinguish between two types of models: *forward* (known as predictors) and *inverse* (known as controllers) (Wolpert & Kawato, 1998). Each model possesses a list of patterns on messages, meaning that we restrict the applicability of a particular model only to states that match the pattern of that model. A model also possesses a *production* that can be executed when its patterns are satisfied. A production is a block of code that the system executes. It can encode predictions of future states, in case of forward models, or controls, in case of the inverse ones. Thus, a forward model is defined as $M_f = \{Precondition, Command, Production\}$ and an inverse model as $M_i = \{Precondition, Goal, Production\}$. The execution of a model may suppress or enable or trigger the execution of other models; in this sense we say that the architecture is model-driven.

Our architecture is composed of concurrent processes that cooperate to achieve system goals. The execution of these processes is data-driven (the availability of new data triggers their execution). Fig. 1 shows the components of