Social Robots, are they Emotional?

Mauricio Iza[1]

[1] Department Basic Psychology, University of Málaga,
29071 Málaga, Spain
iza@uma.es

During the last few decades many different cognitive architectures such as SOAR (Laird, 2012), ACT-R (Lebiere et al. 2013), CELTS (Faghihi et al. 2011b) CLARION (Sun, 2006, 2016), the iCub (Sandini et al. 2007), LIDA (Faghihi and Franklin 2012), etc.; have been proposed. And, agents based on such infrastructures have been widely tested in several cognitive tasks involving reasoning, learning, perception, action execution, selective attention, recognition, etc. (Anderson et al., 2004; Boden, 2016; Langley et al., 2009; Vernon et al. 2007; Mealier et al., 2017).

Cognitive architectures are a part of research in general AI, which began in the 1950s with the goal of creating projects that could reason about problems across different domains, develop insights, and adapt to new situations. Similarly, the main goal of research in cognitive architectures is to model the human mind, eventually enabling us to build human-level artificial intelligence. That is, cognitive architectures attempt to provide evidence what particular mechanisms succeed in producing intelligent behaviour and thus contribute to cognitive science.

Given the multitude of approaches that may lead to human-level AI, and in the absence of a clear definition and general theory of cognition, each cognitive architecture is built on a particular set of premises and assumptions, making comparison and evaluation of progress across architectures difficult. Several papers were published to resolve these issues, the most prominent being Sun's desiderata for cognitive architectures (Sun 2004) and Newell's functional criteria, and later restated by Anderson & Lebiere (2003). Newell's criteria include flexible behaviour, real-time operation, rationality, large knowledge base, learning, development, linguistic abilities, self-awareness and brain realization. Sun's desiderata are broader and include ecological, cognitive and bio-evolutionary realism, adaptation, modularity, routineness and synergistic interaction. Besides defining these criteria and applying them to a range of cognitive architectures, Sun also pointed out the lack of clearly defined cognitive assumptions and methodological approaches, which hinder progress in studying intelligence. He also noted an uncertainty regarding essential dichotomies (implicit/explicit, procedural/declarative, etc.), modularity of cognition and structure of memory. However, a quick look at the existing cognitive architectures reveals persisting

---

disagreements in terms of their research goals, structure, operation and application.

The goal of achieving Artificial General Intelligence (AGI) is explicitly pursued by a small number of architectures, among which are Soar, ACTR, NARS ((Sun 2004) LIDA (Faghihi and Franklin 2012), and several recent projects, such as SiMA (formerly ARS) (Schaat et al. 2015), and CogPrime (Goertzel & Yu 2014). Majority of architectures study particular aspects of cognition, e.g. attention [ARCADIA (Bridewell and Bello 2015), STAR (Tsotsos 2017)], emotions [CELTS (Faghihi et al. 2011b)], and perception of symmetry [Cognitive Symmetry Engine (Henderson and Joshi 2013) or problem solving [FORR (Epstein 2004), PRODIGY (Epstein 2004)].

All of them have memory storage, control components, data representation, and input/output devices. Moreover, cognitive architectures must change through development and efficiently use knowledge to perform new tasks. According to Sun, psychologically based cognitive architectures should facilitate the study of human mind by modeling not only the human behaviour but also the underlying cognitive processes. Such models, unlike software engineering-oriented "cognitive" architectures, are explicit representations of the general human cognitive mechanisms, which are essential for understanding of the mind.

Most of the surveys define cognitive architectures as a proposal about the mental representations and computational procedures that operate on these representations enabling a range of intelligent behaviours (Butt et al. 2013; Duch et al. 2008; Langley et al. 2009; Profanter 2012; Thagard 2012).

In this work, cognitive architectures are selected on the following criteria: self-evaluation as cognitive, robotic or agent architecture, existing implementation, and mechanisms for perception, attention, action selection, memory and learning. In particular, we will focus on the structure of memory.

A crucial aspect of these architectures is the memory system. Its main function is to perform some cognitive operations. Even more, it must coevolve with the external world in order to generate expectations that can be used in decision-making, to adapt to changes in the external world, and to direct attention to relevant features of the situation. By this way, it can perform chains of inferences based on semantic and episodic relations. Such chains of inference can generate new relations of previous experiences. In social robots, this memory system allows more robust goal directed action, visual search, communication, and planning.

These types of architectures lend themselves to model both perception and semantic memory. These properties are central to the model that we will develop below and are used to process both semantic and episodic memories and to form associations that binds stimuli to situations.