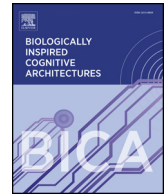




Contents lists available at ScienceDirect

Biologically Inspired Cognitive Architectures

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

Research article

Knowledge acquisition through introspection in Human-Robot Cooperation

Antonio Chella^{a,b}, Francesco Lanza^a, Arianna Pipitone^{a,*}, Valeria Seidita^{a,b,*}^a Dipartimento dell' Innovazione Industriale e Digitale, Università degli Studi di Palermo, Italy^b C.N.R., Istituto di Calcolo e Reti ad Alte Prestazioni, Palermo, Italy

ARTICLE INFO

Keywords:

Cognitive agent
 Knowledge acquisition
 Ontology
 Cognitive architecture
 Introspection

ABSTRACT

When cooperating with a team including humans, robots have to understand and update semantic information concerning the state of the environment. The run-time evaluation and acquisition of new concepts fall in the critical mass learning. It is a cognitive skill that enables the robot to show environmental awareness to complete its tasks successfully. A kind of self-consciousness emerges: the robot activates the introspective mental processes inferring if it owns a domain concept or not, and correctly blends the conceptual meaning of new entities. Many works attempt to simulate human brain functions leading to neural network implementation of consciousness; regrettably, some of these produce accurate model that however do not provide means for creating virtual agents able to interact with a human in a teamwork in a human-like fashion, hence including aspects such as self-conscious abilities, trust, emotions and motivations. We propose a method that, based on a cognitive architecture for human-robot teaming interaction, endows a robot with the ability to model its knowledge about the environment it is interacting with and to acquire new knowledge when it occurs.

Introduction

In human-robot teaming cooperation scenario, artificial agents interact and cooperate with humans and with the environment to reach a shared and common objective. The societal level established in this scenario includes human features the artificial agents have to possess. The knowledge is one of the most crucial key tools for enabling a robot to cooperate with human partners.

The ability to update semantic information is necessary given the dynamicity of the context, which naturally evolves during cooperation; such an ability improves the robot's reasoning and allows more in-depth inferences about the context. It is a human-like ability.

Studying solutions to this kind of problems concerns the field of cognitive architectures and mostly of biologically inspired cognitive architectures. A *cognitive architecture* is a *computational framework for designing intelligent agents, and we call it "biologically inspired" when it aims to reproduce functional properties of the human mind* (Gray et al., 2007, 1990).

The principal aim of research in the field of biologically inspired cognitive architecture is to create an artificial equivalent of the human mind based on biologically inspired architectures so to understand, and then report in the artificial agent, how a human develops his learning

competencies during daily activities. The proposal of researchers in this filed is to create an intelligent agent that implements the essential behavior of the human and his mind in a way that let the artificial agent be perceived as a human. Thus the artificial agent might be able to learn from humans, to cooperate in a team and to properly behave in a society ruled by known and shared laws and rules.

Albus et al. (2007) underlined five fundamental lines in the roadmap for creating a human-level artificial intelligence. Then, others extended this roadmap by adding other interesting elements such as¹:

- accept as a partner – the possibility for a human being to accept the virtual agent as if it were a partner of a team, this brings with it aspects related to trust in the virtual agent, the ability to understand each other delegating when possible and if possible;
- attention awareness, intentionality – agent has to show knowledge and awareness of current situations, voluntary and intentional behavior and so on;
- believable behavior – the agent must be trustworthy then it has to show a behavior like a conscious being's one,
- creativity – the agent must be able to design, generate and evaluate new concepts, rules, strategies, goals, etc.,
- emotional intelligence – the agent must be able to identify and

* Corresponding authors at: Dipartimento dell' Innovazione Industriale e Digitale, Università degli Studi di Palermo, Italy.

E-mail addresses: antonio.chella@unipa.it (A. Chella), francesco.lanza@unipa.it (F. Lanza), arianna.pipitone@unipa.it (A. Pipitone), valeria.seidita@unipa.it (V. Seidita).

¹ This is an excerpt of the whole roadmap (Albus et al., 2007) we consider fundamental for our work.

<https://doi.org/10.1016/j.bica.2018.07.016>

Received 4 July 2018; Received in revised form 11 July 2018; Accepted 11 July 2018

2212-683X/ © 2018 Elsevier B.V. All rights reserved.

understand emotional motivations in the other agents, reason in terms of emotions and generate emotional responses;

- imagery – the agent has been able to generate simulations of different worlds and scenarios;
- learner critical mass – the agent has to be able to learn as a student in the same situation, it should own the critical mass

Many of these elements are related to each other and together form a single model, a cognitive architecture, of what are the characteristics of the human mind and therefore the behavior of the human being. Existing cognitive architectures in the literature (Anderson, Matessa, & Lebiere, 1997; Benjamin, Lyons, & Lonsdale, 2004; Christensen & Hooker, 2004; Franklin, Madl, D’Mello, & Snaider, 2014; Laird, Newell, & Rosenbloom, 1987; Shanahan & Baars, 2005) have addressed all these aspects in isolation, i.e. many of them address basic cognitive functions. It is desirable to have a cognitive model, a cognitive architecture, able to take into account all the aspects mentioned above.

In the context of the roadmap we are studying a cognitive architecture that can serve as a basis for the creation of virtual agents that interact with a human being in a team scenario. This allows us to explore everything related to the metacognitive level that includes episodic memory, knowledge, Theory of Mind, self-regulation, the ability to evaluate complex and social emotions and to generate goals, the ability to reflect about itself and introspection.

In this paper, we focus on the role of knowledge acquisition by cognitive agents/robots. In order to exploit metacognitive abilities and to realize the elements of the roadmap, we consider the basic principles of cognitive semantics, according to which the concept of *mental schema* (McCarthy, 1995) roughly represents mental structures for humans that recur in construals of environment; the mental schema plays a fundamental role for conceptualization of a new percept by robot, and hence for knowledge discovering. The robot becomes capable of the mental situations about the single percept, acquiring the *imagery* skill. Moreover the robot will own the *critical mass* (Albus et al., 2007) because it becomes able to grow in environmental awareness as a human could do under same condition. Briefly, the method we investigate attempts to implement environmental awareness and self-consciousness for knowledge acquisition.

We claim that the ontological structures represent the mental schemas the robot owns; we argue to define a distribution over the space of possible mental schemas from the ontology, that allows us to estimate the probability of each possible schema given a percept. Both mental schema and percept are represented by plain text.

Once the plausibly schema for a percept is inferred, the percept is acquired by blending the correspondent new concept and the related instance (if the concept does not already exist in the knowledge base) or only the instance (if the concept already exists), and the knowledge grows. Each time a new percept is detected, the knowledge incrementally grows if it does not yet include such a percept; the acquisition is at run-time. The method integrates two kinds of approaches, that are sub-symbolic and symbolic. In this way, the important problems of transparency and transfer learning are addressed, as shown in the final discussions.

The rest of the paper is organized as follows: in Section “A cognitive architecture for human-robot teaming interaction” we briefly discuss the cognitive architecture we identified for modeling human-robot teaming interaction and we highlight the parts developed in this paper. In Section “Problem description: the example of a robot working in a partially known environment” we anticipate the discussion about the method we propose by starting with a simple example, then in Section “Modeling the mapping and merging processes” we cover the proposed method for updating knowledge at run-time. Finally, in Sections “Discussions” and “Conclusions” we draw some discussions and conclusions.

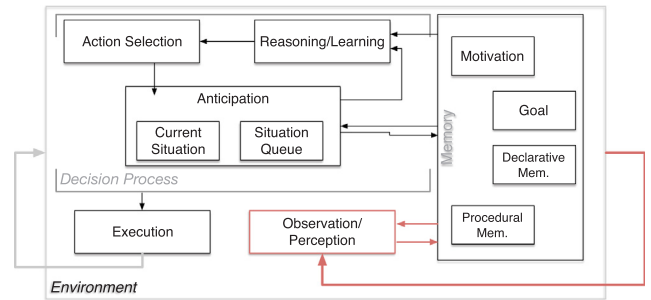


Fig. 1. The cognitive architecture for human-robot teaming interaction.

A cognitive architecture for Human-Robot Teaming Interaction

In recent years, various cognitive architectures have been studied and developed, which have served to explore the main characteristics of the human mind. Some of these architectures have been used as the basis for the implementation of robotic systems that behave like humans. However, most existing architectures take into account only one of the aspects outlined in the Roadmap. Some focus on learning, many on memory, both short and long-term. All of these aspects are essential but they are never dealt with considering a system that works in a dynamic environment and that has to make decisions at run-time.

Our work is part of the BICA roadmap; our aim is developing a cognitive model that includes all the necessary modules for a robot to cooperate in a team to achieve a common goal. The robot has to apply a decision-making process that takes its cue from the objective situation of the environment but also from the knowledge it has of itself and the other members of the team.

The shown architecture in Fig. 1 contains useful modules for self-modeling, for representing the surrounding (physical) world, including the other and including all the mental states that this involves. In a cognitive agent, it is the mental state that really triggers actions, memory in all its forms serves only as a support for the mental state processing phase rather than for data, which comes from sensory activity, processing.

To date, most architectures base their decision making and learning process only on the concept of stored data or facts and not on the idea of a mental state. Our contribution lies in the creation of memory modules containing all the information about the mental state in the world so that the perceive-act cycle becomes what we call the perceive-proact cycle. We identified four main modules: the module devoted to the observation of the environment, the one realizing the decision process (including reasoning, learning and actions anticipation), the execution and the memory. A cognitive agent knows its goal and the state of affairs around and within itself, perceives objects and everything relevant for a mission in order to trigger a decision about which action to perform. Before performing actions it produces the anticipation of the action results, in order to check if that action brings to acceptable situation of the world, and alternatively a queue of situations to be selected if necessary. This process is performed iteratively, each time interacting with all the elements in the environment through perception and observation.

In this paper, we detail the path highlighted in red that is related to the process of knowledge management, above all its acquisition through introspection as it is illustrated in the next sections.

Problem description: the example of a robot working in a partially known environment

This work aims to make a robot *aware* of objects in dynamic environments and *self-conscious* of its knowledge by automatically updating it when a new entity is perceived. If the robot cannot “understand” an object, then it would not be able to use and refer to it during task

Download English Version:

<https://daneshyari.com/en/article/10150918>

Download Persian Version:

<https://daneshyari.com/article/10150918>

[Daneshyari.com](https://daneshyari.com)