Biologically Inspired Cognitive Architectures xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Biologically Inspired Cognitive Architectures

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

Research article

A social practice oriented signs detection for human-humanoid interaction

Agnese Augello[a,*], Frank Dignum[b], Manuel Gentile[c], Ignazio Infantino[a], Umberto Maniscalco[a], Giovanni Pilato[a], Filippo Vella[a]

[a] Institute of High Performance and Networking (ICAR), National Research Council (CNR), Italy
[b] Utrecht University, Netherlands
[c] Institute for Educational Technology, National Research Council of Italy, Italy

ARTICLE INFO

Keywords:
Social practice
Cognitive architecture
Human robot interaction
Social signs classification

ABSTRACT

In this work we propose a cognitive architecture, based on the Social Practice (SP) theory, aimed at the modeling of socially adaptive robots, able to interact with people, recognizing and interpreting the specific social context where it is acting. The proposed social robot is able to recognize and interpret social signs during ongoing social practices. The cognitive architecture is inspired by the well-known Psi model, and it is equipped with a Social Practice Engine that manages the whole conduct of the robot. The use of such an architecture simplifies and makes more natural the interaction between human beings and a robot. In particular, the scenario of a robot fulfilling the tasks of welcoming people in an office environment is being considered.

Introduction

Every time we interact each other within the society, we follow, even without realizing it, *social practices*, i.e. routinely performed behaviors like *going to work, have a meeting*, and so on. Social practices drive our behavior by integrating physical and mental activities, competences, knowledge, emotions (Dignum & Dignum, 2014; Reckwitz, 2002). Our behavior is conditioned by our role within the society and the practices that are commonly used within it. Social practices are shared and created by mutual, explicit or implicit, mutual agreement, within the society. They trigger our attention on what is happening around us, they affect the importance we give to our needs as well as our expectations about the behavior of the other participants in the practice.

When we talk about social robots often we focus on their communication abilities without considering the context in which the interaction takes place. The communication skills of the last

generation of robots do not automatically make them *social*, and optimizing their interaction for specific tasks do not make their behavior more *social* either (Dautenhahn, 2007; Dignum, Prada, & Hofstede, 2014; Kaminka, 2013).

Instead, coping with the issue of modeling a form of social intelligence in a robot, means properly modeling its identity (i.e. its knowledge and its point of view about users and the surrounding environment), endowing its knowledge with socio-cultural practices, providing the robot with the ability of understanding the current social situation and, as a consequence, properly planning and carrying on the interaction.

As discussed in Dignum and Dignum (2014), the social context must be considered as the foundation from which deliberation starts, instead of adding extra conditions to be analyzed from time to time, which complicates the formalization of agents and robots (Augello, Gentile, & Dignum, 2016).

However, there is a growing attention towards socially adaptive robots (Breazeal, 2004; Dautenhahn, 2007; Fong, Nourbakhsh, & Dautenhahn, 2003; Hegel, Muhl, Wrede, Hielscher-Fastabend, & Sagerer, 2009; Williams, 2012) and different proposal have been made to model social abilities in cognitive architectures, as an example in Malfaz, Castro-González, Barber, and Salichs (2011), Ho (2016), Adam, Johal, Pellier, Fiorino, and Pesty (2016). In such proposal, a key role in robots decisional processes emerges from their needs, emotions and motivations. As social beings, among the needs that drive and motivate our behavior, we give an important role to the so-called *Affiliation* (Bartl & Dorner, 1998), i.e. the need to gain the recognition or acceptance from people in a social group. We are able to understand and evaluate this acceptance by analyzing and interpreting, either consciously or unconsciously, the information conveyed in the several multi-modal signals coming from people interacting with us.

Therefore, the processing of verbal and not verbal signals can help a robot to recognize and understand the intention of people (Lobato,

* Corresponding author.
E-mail addresses: agnese.augello@icar.cnr.it (A. Augello), F.P.M.Dignum@uu.nl (F. Dignum), manuel.gentile@itd.cnr.it (M. Gentile), ignazio.infantino@icar.cnr.it (I. Infantino), umberto.maniscalco@icar.cnr.it (U. Maniscalco), giovanni.pilato@cnr.it (G. Pilato), filippo.vella@icar.cnr.it (F. Vella).

https://doi.org/10.1016/j.bica.2018.07.013
Received 15 June 2018; Received in revised form 10 July 2018; Accepted 10 July 2018

Warta, Wiltshire, & Fiore, 2015; Loth & De Ruiter, 2016), their emotions, attitudes and relationships (Poggi & Francesca, 2010; Vinciarelli, Pantic, & Bourlard, 2009). In particular, non verbal behavior, such as face expressions, postures and gestures, are particularly meaningful in a social interaction and reveal the attitude of people towards others and their affective states (Vinciarelli et al., 2009). Another important form of not verbal communication is the eye gaze. It plays a crucial role in social engagement: as an example, the work (Ehrlich, Wykowska, Ramirez-Amaro, & Cheng, 2014) focuses on the prediction of both intention and the detection of the initiator of a gaze contact during a human-robot social engagement. Providing a robot with the ability of interpreting eye gaze can improve the interaction and the performance of collaborative tasks with human partners (Palinko, Rea, Sandini, & Sciutti, 2016; Palinko, Sciutti, Wakita, Matsumoto, & Sandini, 2016).

As a matter of fact, the communicated signals must be interpreted in the specific socio-cultural context of the interaction. Some signals are proper in a specific socio cultural context, while they are inappropriate in another. As an example, let us consider a social practice of welcoming and scheduling the appointment in a office: people observing the behavior of the receptionist should consider unacceptable a rude gesture towards a customer, while the same gesture is considered appropriate if the customer did not respect the social norms and showed an aggressive behavior. In that case, the receptionist stops the current practice and starts another one aimed at the maintenance of the security in the office.

Various machine learning approaches exist in literature to process social verbal and not verbal signals (see for example the review in Rudovic, Nicolaou, and Pavlovic (2017)) by extracting and analyzing features that are considered useful to evaluate the human behavior. However, while different gestures recognition modules have been proposed in literature (Barros, Parisi, Jirak, & Wermter, 2014; Canal, Escalera, & Angulo, 2016), few of them suggest heuristics of processing according to specific situations (Loth, Jettka, Giuliani, & de Ruiter, 2015).

In this work, we model the behavior of a social robot, by relying on a proper and well known social theory. We introduce a social practice engine at the heart of the cognitive architecture of a robot. Based on what affirmed by the social practice theory, the engine allows the robot, according to current social practice, to focus the attention towards specific perceptions, weight its urges, determine the meaning of what is perceived by the environment, and to deliberate according to the specific social situation. In particular, in this work we focus the attention to the use of the engine to interpret the verbal and gestural social signs in an interaction. We use the term “social sign” in order to generalize and extend the meaning of *social signal:* from the social practice perspective, a social sign is context dependent, and it is part of a given behavioral schema; moreover, it could be executed by any actor of the social interaction (in our case either human or artificial embodied agent) depending on their specific role at a given instant. According to our interpretation of social sign, we highlight here a “social sign” is more linked to the meaning associated, while the signal is more bound to the mere perception process. The SP engine allows the robot to engage a social interaction with people, recognizing and interpreting from time to time the communicative signs according to the specific social context where it is acting. The results obtained by an actions classification module (Augello et al., 2018) are analyzed in the context of the current social practice in order to understand if they are conform to the social expectations.

As a case study we present the abilities of a robot concierge to interpret and react (both verbally and with gestures) to the social signs communicated by the users visiting a research laboratory. In particular we consider the behavior in a social practice, that we called *ReceptionSP*, of user’s reception in a laboratory. The understanding of the social signs inside the practice affects the *Affiliation* need of the robot and therefore its *Motivation*. This determines the behavior of the robot and the conversation, that is managed by a dialogue engine supported by a reasoner.

Social practice model

Social Practice theory is a sociological theory that studies the behavior of groups of people into the society (Reckwitz, 2002). According to this theory, social practices are routinized behaviors typically and habitually performed by people. The actors of a practice play a certain role and have some predefined purposes and expectations about the behaviour of other agents in the practice and social norms to respect. This theory inspired the formalization of a social practice model for cognitive agents, that is used as heuristic to determine their plans and behavior in social contexts (Dignum & Dignum, 2014). Often, to consider the social context in the formalization of intelligent agents, extra conditions must be added from time to time and the complexity increases considerably. This happens for example by considering goal driven approaches, such as BDI, as shown with some examples in Dignum and Dignum (2014). Recently there is a growing interest in employing normative models (developed by the multi-agent systems community) to design robots behaviour in social environments, by introducing norms that are narrowed to social and cultural conventions. Some interesting works are presented for example in Carlucci, Nardi, and Nardi (2015), Malle, Scheutz, and Austerweil (2017). At the same time, however, while a normative approach can be enough when the scenario is well constrained, moving out robot from constrained environments and introducing them in society requires a socially adaptable behaviour. The social behavior, the interpretation and reasoning processes of an agent are determined by the different social situations that, moreover, from time to time, arise different social expectations, and require the compliance of social norms (Can, Seibt, et al., 2016). Social practice are not rigid, shared structures of knowledge which enable the agents to interpret what happens, analyzing from time to time the different social signs and the behaviour of the other agents, and to act in certain ways, choosing the goal to pursue, and by respecting the social norms. Social practices allow a cognitive agent to manage the social context in a more efficient way. The context is considered at the basis of the deliberation process and not as extra norms to manage after the deliberation process started when the agent tries to reach a goal (Dignum & Dignum, 2014; Dignum, Dignum, Prada, & Jonker, 2015).

The model, summarized in Fig. 1, formalizes the elements of a social practice: 1) a *Physical Context* that describes the element of the environment, i.e. the physical objects (*Resources*), the individuals involved in the practice (*Actors*) and their locations (*Places*); 2) a *Social Context*, consisting in a *Social Interpretation* of the context, the *Roles* of the individuals in the practice and the *Norms* expected inside the practice; 3) the possible *Activities* an agent could perform; 4) the *Meaning* of the agents activities and their social effects; 5) the *Competences* an agent should have to perform the activities in the social practice; 6) the *Plan Patterns* to reach the goals. The knowledge formalized in social practices is used as heuristics to guide a context-oriented plan identification for agents. Social practices model give a means to choose between reactive behavior in standard circumstances and pro-active behavior that is necessary for an agent to be socially intelligent (Dignum & Dignum,

Fig. 1. Social practice model.