Cognitive Architectures for Social Human-Robot Interaction

Paul Baxter, Séverin Lemaignan
Centre for Robotics and Neural Systems
Plymouth University, Plymouth, U.K.
{paul.baxter, severin.lemaignan}@plymouth.ac.uk

J. Gregory Trafton
Naval Research Laboratory
Washington DC, USA
greg.trafton@nrl.navy.mil

***Abstract*—Social HRI requires robots able to use appropriate, adaptive and contingent behaviours to form and maintain engaging social interactions with people. Cognitive Architectures emphasise a generality of mechanism and application, making them an ideal basis for such technical developments. Following the successful first workshop on Cognitive Architectures for HRI at the 2014 HRI conference, this second edition of the workshop focusses specifically on applications to social interaction. The full-day workshop is centred on participant contributions, and structured around a set of questions to provide a common basis of comparison between different assumptions, approaches, mechanisms, and architectures. These contributions will be used to support extensive and structured discussions, with the aim of facilitating the development and application of cognitive architectures to social HRI systems. By attending, we envisage that participants will gain insight into how the consideration of cognitive architectures complements the development of autonomous social robots.**

***Index Terms*—Cognitive Architectures; Cognitive Robotics; Social Human-Robot Interaction**

I. Introduction

Achieving social interactions between humans and robots is a complex task that has yet to be attained, but which is necessary for the increasing range of real-world applications for social robots. It requires an understanding of human social behaviour, and it requires the robots to use appropriate, adaptive and contingent behaviours to form and maintain these social interactions. Given that pre-programmed approaches are clearly insufficient for this problem, Cognitive Architectures provide a good alternative as they propose general mechanisms of 'intelligence' and behaviour generation.

Following the successful First Workshop on Cognitive Architectures for HRI held at the HRI conference in 2014 (Bielefeld) [1], we we are running a second edition, in which we focus specifically on cognitive architectures for *social* human-robot interaction[1]. As previously, the intention is to have the workshop be as inclusive as possible, catering both for experienced researchers in the area, but also for those for whom this may be a new topic. For all, we intend the workshop to provide a forum for discussion and the exchange of ideas. To facilitate this discussion and to provide a basis for a concrete contribution to the research community, we request short position paper contributions, and will organise a special issue (based on extended version of the position papers) after the workshop to consolidate the progress made, and provide a reference point for the community.

[1]https://sites.google.com/site/cogarch4socialhri2016/

Fig. 1. Workshop logo: cogs are typically used to represent cognition in an individual agent, this has been adapted to acknowledge the central role that interaction must play in social human-robot interactions in addition to the 'internal' cognition of the individuals.

II. Background

Cognitive Architectures are constructs (encompassing both theory and models) that seek to account for cognition (over multiple timescales) using a set of domain-general structures, mechanisms and/or processes [2]. Typically (but not necessarily) inspired by human cognition [3], the emphasis is on deriving a set of general principles of operation not constrained to a specific task or context. Despite the multitude of implementations used [4], they encourage the system designer to initially take a broader perspective than the computational mechanisms to be used and consider what sort of functionality needs to be present for the type of application, and how this relates to other cognitive competencies that are required.

For HRI, such an approach to building autonomous systems based on Cognitive Architecture would emphasise first those aspects of behaviour that are common across domains, before applying these to specific interaction contexts for evaluation. In the case of social interaction, the problems are numerous, encompassing the coordination of multiple sensory and motor modalities for the robot, the timing of proactive and reactive actions, and the recognition of interacting human states (cognitive, affective, physical, etc). Indeed, recent theoretical developments have emphasised the complex temporal coordination dynamics of human social behaviour, rather than the internal state of any individual agent [5]. This leads to questions

978-1-4673-8370-7/16/$31.00 © 2016 IEEE

regarding how the human should be taken into account in the action preparation/selection for the robot: explicit and individual models of performance, theory-of-mind, and/or generalised statistical models of human behaviour? It also gives rise to the question of whether and how the robot 'cognition' and actions should be directly informed by (or indeed constrained by) human psychology and physiology, with the complexity and 'non-optimal' behaviours that may result, e.g. [6]. Should our cognitive architectures for social robots be based directly on models of human behaviour, or is there no need for this? These, and related, questions are outstanding in the field and require addressing if the utility and efficacy of social robots in the real world is to be realised.

Up to now, there have only been limited and relatively isolated attempts to addressing these questions, particularly within the HRI community, with few examples of direct applications, e.g. [7]. Building on the first iteration of this workshop [1], we seek to bring together researchers who are attempting to formalise knowledge of appropriate robot behaviours for naturalistic interaction with people, typically emphasising generally applicable, holistic perspectives (i.e. striving to consider the full gamut of socially interactive behaviour rather than only individual aspects).

III. Outline of the Workshop

This workshop is aimed at researchers from a wide range of backgrounds who may be interested in applying concepts from Cognitive Architectures to their work, specifically Social HRI. Participation in this workshop is open to all interested researchers.

Prospective participants are requested to submit a 2-4 page position paper on (preferably) their work involving cognitive architectures (including the development and/or application thereof). In order to facilitate interactions and discussions at the workshop (by providing a basis for comparison), we ask that all authors additionally use their position papers to provide an answer to six guiding questions. These are as follows:

1) Why should you use cognitive architectures - how would they benefit your research as a theoretical framework, a tool and/or a methodology?
2) Should cognitive architectures for social interaction be inspired by and/or limited by models of human cognition?
3) What are the basic requirements for social interaction for a cognitive architecture?
4) How the requirements for social interaction would inform your choice of the fundamental computational structures of the architecture (e.g. symbolic, sub-symbolic, hybrid, ...)?
5) What is the primary outstanding challenge in developing and/or applying cognitive architectures to social HRI systems?
6) Can you devise a social interaction scenario that current cognitive architectures would likely fail, and why would this be the case?

Submission of a position paper is not a pre-requisite for attendance, and we encourage researchers to attend the workshop even if not willing/able to submit a position paper, in order to maximise community engagement and the uptake of these concepts within the field of HRI. By attending, we envisage that participants will gain insight into how the consideration of cognitive architectures complements the development of autonomous social robots.

IV. Organisers

Paul Baxter is a researcher at Plymouth University (UK) in the Centre for Robotics and Neural Systems, and the Cognition Institute. After obtaining a PhD in Developmental Cognitive Robotics (University of Reading, UK), he worked on the EU FP7 ALIZ-E project to apply and evaluate a memory-centred perspective on cognition to social child-robot interaction. His current research work involves the development of supervised autonomous therapy robots for children with ASD (EU FP7 DREAM project), with a specific focus on cognitive robot control.

Greg Trafton is a Cognitive Scientist at the Naval Research Laboratory in Washington, DC, USA. He has degrees in both Computer Science (Trinity University) and Psychology (Princeton University) and works on Human-Robot Interaction from a cognitive modeling / architectures perspective.

Séverin Lemaignan is a researcher at Plymouth University (UK) in the Centre for Robotics and Neural Systems, and the Cognition Institute, focusing on the cognitive pre-requisites of social interaction between humans and robots. He conducts both basic work on mechanisms like the Theory of Mind, and technical realisations on interactive robots.

Acknowledgement

This work was partially supported by the EU FP7 project DREAM (grant number 611391, http://dream2020.eu/), and the Marie Sklodowska-Curie Actions of the EU H2020 project DoRoThy (grant number 657227).

References

[1] P. Baxter and J. G. Trafton, "Cognitive Architectures for Human-Robot Interaction," in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction - HRI '14*. Bielefeld, Germany: ACM Press, 2014, pp. 504–505.

[2] P. Langley, J. E. Laird, and S. Rogers, "Cognitive architectures: Research issues and challenges," *Cognitive Systems Research*, vol. 10, no. 2, pp. 141–160, jun 2009.

[3] R. Sun, L. A. Coward, and M. J. Zenzen, "On levels of cognitive modeling," *Philosophical Psychology*, vol. 18, no. 5, pp. 613–637, oct 2005.

[4] D. Vernon, *Artificial Cognitive Systems - A Primer*. MIT Press, 2014.

[5] E. Di Paolo and H. De Jaegher, "The Interactive Brain Hypothesis," *Frontiers in Human Neuroscience*, vol. 6, no. June, pp. 1–16, 2012.

[6] J. L. McClelland, "The Place of Modeling in Cognitive Science," *Topics in Cognitive Science*, vol. 1, no. 1, pp. 11–38, jan 2009.

[7] J. G. Trafton, L. M. Hiatt, A. M. Harrison, P. Tamborello, S. S. Khemlani, and A. C. Schultz, "ACT-R/E : An Embodied Cognitive Architecture for Human-Robot Interaction," *Journal of Human-Robot Interaction*, vol. 2, no. 1, pp. 30–54, 2013.