Am I capable yet?

…where I argue we need better evaluation and better representations to improve cognitive architectures for HRI

Séverin Lemaignan
Computer Human Interaction for Learning and Instruction, École Polytechnique Fédérale de Lausanne
CH-1015 Lausanne, Switzerland
severin.lemaignan@epfl.ch

I suggest in [1] that the cognitive foundations that are required to further advance research on cognitive architectures for human-robot interaction must answer these three questions:

1. How can I represent my environment and my beliefs in such a way that perception modalities as diverse as my own sensors, verbal descriptions of distant objects or events, or symbolic common-sense knowledge extracted from online bases could be mixed?
2. How do I identify the current situation under a complex combination of diverse, changing circumstances? How do I project myself in other situations to predict, remember, explain, learn?
3. How do I access and represent mental models of the agents I interact with?

These questions can be reformulated into explicit research objects: *a*) real-time situation assessment that builds on semantic mapping and supports perspective-taking and affordances analysis, *b*) interleaved geometric, temporal and symbolic reasoning that supports in particular identification of situations and actions, *c*) management and exploitation of independent, possibly contradictory, belief models for each agent the robot interacts with, *d*) identification and representation of overlapping and multi-scale interaction contexts: temporal, spatial, but also social and cultural, *e*) natural multi-modal communication, also including *backchannel* communication like nodding and facial expressions.

As complex as it may appear, I believe the global challenge formed by these items to be actually tractable by adopting an holistic approach to the design of cognitive architectures for interactive robots.

I believe however that two important pre-requisites must first be addressed: How do we *evaluate* the (socio-)cognitive skills of the robots? And can we rethink the acquisition and representation of the robot's environment as a fundamental cognitive building block that requires special attention? I propose to discuss those two questions in this position paper.

HRI'14, March 3–6, 2014, Bielefeld, Germany.
ACM X-XXXXX-XX-X/XX/XX.

Metrics for robotic cognition

Am I capable yet? Assessing the cognitive skills of robots is difficult because tools for quantitative measurement of such skills remain mostly to be devised.

The robotic community relies mostly on qualitative assessment. Langley et al. [2] propose five such dimensions of evaluation: the *generality* of the system (can it adapt easily to new tasks?), the *rationality* or relevant of the inference/reasoning/decisions the system take, the *reactivity* and *persistence* that evaluates if the behaviour of a cognitive system is appropriate under unpredicted changes, the *improvability* of the system as a function of the knowledge added to it, and finally, the resulting *autonomy* of the system.

Recent work from Zhang et al. [3] surveys operational frameworks and metrics for performance evaluation of cognitive robots. They however underline that most of these benchmarks are focused on physical capabilities that do not necessarily require advanced knowledge representation and manipulation.

Assessment of cognitive performances can also benefit from the support of tools developed in cognitive psychology. Several classical tests, like the False-Belief experiment, related to the Theory of Mind, or the Token test, have been used to assess the cognitive abilities of robots ([4, 5] amongst others). Much remains to be done, however, to draw a complete picture of the know-how in cognitive psychology when projected onto robots: what are the existing metrics, and how suitable and applicable to embodied artificial agents they are.

My hope is that our community researches and builds a rigorous, operational framework for the assessment of cognitive skills of robots, that would address both individual facets of cognition (performance for dealing with underspecified tasks, theory of mind, language tests, etc.) and *global measurements of the cognitive activity* (I propose in [6] a first idea to explore this last point: By plotting the frequency of interactions between the software modules of the robot and a central knowledge, I build a measurement for the *cognitive load* of the robot).

I believe that designing such an evaluation standard for assessment of cognitive skills would provide us with an important baseline to further research cognitive architectures for HRI.

Amodal representations and contexts

In the stack of software components required for an autonomous robot, the layer that provides a uniform representation of the robot's environment not only suitable, but even convenient for decision making, is crucial.

As expected, a large body of literature discusses approaches and techniques to build such representations, collectively designated as *situation assessment* techniques.

Service robots, and even more companion robot, have to handle more perceptual modalities than many other groups of robots: beside proprioception and perception of their environment for navigation and manipulation, these robots need to account for humans. Not only their (dynamic) physical features (location, posture, gestures...), but also mental features: cognitive capabilities, beliefs, desires. These features are not directly observable and usually require inference based on other cues like gestures or backchannel communication (nodding for instance). The correct interpretation of these cues requires building new modality-independent (thus, *amodal*) representations that support parallel and hybrid (continuous/symbolic) perspectives, including temporal and spatial models, models of the (grounded) beliefs of each of the agents, cultural/social contexts.

Contexts, in particular, seem both critical and under-studied in our community. Proper context management should allow the robot to mentally *move around its own experiences* to place itself in the mental situation where the interpretation of an event, an interaction or a situation makes sense.

The role and importance of context identification for correct interpretation of a situation is well understood in cognitive science. A classical example considers two series of words: `FOX; OWL; SNAKE; TURKEY; SWAN; D?CK` and `BOB; RAY; DAVE; BILL; HENRY; D?CK`

If you read through these lines, you are likely to have guessed the last words of each row, `DUCK` and `DICK`, only from the context induced by the others words.

Applied to service robotics in households, an example of context-dependent interpretation of two similar situations could be: *A cat walks in the living room* versus *A baby crawls towards a power socket*.

The example involves perception issues (distinguishing between a cat and a baby), but even if we consider that the scene is perceptually recognised, its interpretation relies on selecting relevant contexts (for instance, the *caregiver* context: What is the role of the robot in presence of a cat/baby? The *baby* context: knowledge about the baby capabilities, predictions of baby intentions, salient features of the room for a baby, etc.)

Cognitive functions like episodic memory, theory of mind, projection, diagnosis amongst others can be seen as special cases of a generic context management capability.

Managing context means at least three things: context identification, context representation, context restitution. Depending on what context we talk about, identifying contexts can be relatively easy (Who is talking to me? Where am I?) to difficult (What past experience does my interactor implicitly refer to?). One of the main problem we see with context identification is that it is a fundamentally *multi-scale* problem: At any moment, several temporal, spatial, social, cultural context co-exist and overlap.

This leads to the second aspect, context representation. Contexts are currently often limited to the current spatial and temporal situation. Some projects offer the possibility to jump in the past or to switch to another agent's perspective, but in current approaches, selecting a context always basically consists in retrieving a set of beliefs corresponding to a situation, and temporarily replacing the current beliefs by those other ones. This misses the fact that at a given moment, many contexts co-exist at different scales.

The ability to explicitly manage contexts and context switches would endow the robot with a cognitive capability similar to what is known as *context-dependent memory* in cognitive psychology. This is also related to Tulving's *autonoetic consciousness*: the ability to reflect upon its own past or future experiences. Much remains to be done to this regard, starting with a formal analysis of what are the relevant contexts for our robots.

To conclude

This position paper side-steps from usual discussions on cognitive architectures, and puts instead a focus on two issues that, I believe, need to be addressed to support and foster research on cognitive architectures in HRI.

First, we need to define and agree on reproducible *metrics* for cognitive and social skills. Those metrics can draw from existing artificial intelligence tests, cognitive sciences (cognitive and developmental psychology, but also neurosciences), and reproducible field experiments (the *robocup@home* competition may be an interesting starting point).

Then, I believe that progresses on situation assessment are required to advance intelligent autonomous human-robot interaction. Current state-of-the-art mainly focuses on geometric and temporal situation assessment, we need to extend it to hybrid (continuous and symbolic) assessment of the environment at large, including other agents' perspectives and mental models, and modeling of the overlapping interaction contexts that define a situation. This would allow us to feed the upper layers of cognitive architectures with richer representations of the world to support decision making.

1. REFERENCES

[1] S. Lemaignan, M. Warnier, E. Sisbot, and R. Alami, "Human-robot interaction: Tackling the AI challenges," *Artificial Intelligence*, 2014, submitted.

[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, June 2006.

[3] L. Zhang, S. Rockel, A. Saffiotti, F. Pecora, L. Hotz, Z. Lu, D. Klimentjew, and J. Zhang, "Evaluation metrics for an experience-based mobile artificial cognitive system," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013.

[4] N. Mavridis and D. Roy, "Grounded situation models for robots: Where words and percepts meet," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006.

[5] G. Trafton, L. Hiatt, A. Harrison, F. Tamborello, S. Khemlani, and A. Schultz, "ACT-R/E: An embodied cognitive architecture for human-robot interaction," *Journal of Human-Robot Interaction*, vol. 2, no. 1, pp. 30–55, 2013.

[6] S. Lemaignan and R. Alami, "Explicit knowledge and the deliberative layer: Lessons learned," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013, to appear.