Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 123 (2018) 579–584

Procedia Computer Science

www.elsevier.com/locate/procedia

8th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA 2017

Joint Goal Human Robot collaboration-From Remembering to Inferring

Vishwanathan Mohan[1] and Ajaz Ahmad Bhat[2]

[1] *CSEE Department, University of Essex, U.K.*

[2] *School of Psychology, University of East Anglia, U.K.*

Vishwanathan.mohan@essex.ac.uk, A.Bhat@uea.ac.uk

Abstract

The ability to infer goals, consequences of one's own and others' actions is a critical desirable feature for robots to truly become our companions-thereby opening up applications in several domains. This article proposes the viewpoint that the ability to remember our own *past* experiences based on *present* context enables us to infer *future* consequences of both our actions/goals and observed actions/goals of the other (by analogy). In this context, a biomimetic episodic memory architecture to encode diverse learning experiences of iCub humanoid is presented. The critical feature is that partial cues from the present environment like objects perceived or observed actions of a human triggers a recall of context relevant past experiences thereby enabling the robot to infer rewarding future states and engage in cooperative goal-oriented behaviors. An assembly task jointly done by human and the iCub humanoid is used to illustrate the framework. Link between the proposed framework and emerging results from neurosciences related to shared cortical basis for 'remembering, imagining and perspective taking' is discussed.

Peer-review under responsibility of the scientific committee of the 8th Annual International Conference on Biologically Inspired Cognitive Architectures

Keywords: Human robot collaboration, Robot episodic memory, iCub humanoid, internal simulation, Cumulative learning

1 Introduction

From dining together to jointly assembling an IKEA table from constituent parts we are acting and anticipating consequences of potential actions, goals of both, oneself and the other. This seemingly effortless "real-time" inference of others actions/goals and prediction of ensuing future events enables us to both plan our actions accordingly or engage in cooperative goal oriented behaviors. Undoubtedly, as articulated in several recent robotics roadmaps [1, 2] this is a critical desirable feature for robots working alongside humans in unstructured environments- from industry to homes. Even today, most existing approaches in industrial manufacturing involving human robot coexistence is based on each agent performing isolated steps independently with minimal communication exchanged as necessary

Peer-review under responsibility of the scientific committee of the 8th Annual International Conference on Biologically Inspired Cognitive Architectures
10.1016/j.procs.2018.01.089

[3]. In this context, recent advances in the design and availability of safe, compliant robots like Baxter, iCub humanoid, Kuka LBR, Universal Robots is gradually enabling humans and robots to share workspaces. This opens up the scope for "joint-goal" human robot scenarios- where both agents act/perceive/plan collaboratively in a continuously evolving unstructured environment (as a simple example, assembling something from constituent parts). Given that, successful collaboration with another agent in a joint goal task requires a complex integration of multiple subsystems like perception, action, goal directed reasoning, we are only beginning to scratch the surface of understanding the computational basis of social intelligence in autonomous robots (see [4] for recent reviews).

This article is an exploration into this topic with the working hypothesis that the ability to remember our own *past* experiences based on *present* context enables us to infer *future* consequences of both our actions/goals and observed actions/goals of the other (by analogy). Emerging trends from neurosciences importantly the discovery of the Default Mode Network (DMN) in the brain [5] is providing converging evidence in this direction. In particular, studies on DMN indicate that there is an extensive overlap in cortical networks activated while remembering the past and those engaged during simulation of the future and adopting the perspective of the other [6-7]. At the core of DMN are the brain areas in the Medial Temporal Lobe known to be involved in episodic memory. Disruption to the DMN also indicates suppressed social behavior as observed in cognitive disorders like ASD [7].

In the context of cognitive robotics and from a computational/functional perspective, presently there is consensus that the central function of DMN is to generate self-referential episodic simulations- that include recall of past experiences, prediction of potential future states and inferring the perspective of the other. Given the trends in neuroscience of memory, computational modelling and implementation of biomimetic robot episodic memory has been a topic of emerging interest in cognitive robotics ([8-9] see, Vernon, Beetz and Sandini, 2015 for a review). Robot episodic memory systems have been instantiated both sub-symbolically using ANNs and symbolically using content-addressable image databases with traditional image indexing and recall algorithms. Importantly, unlike in synthetic systems where memory is usually treated as a passive storage device, this viewpoint looks at memory as an active process involved in forward simulation and presepctive taking.

In this context, we present a growing, multimodal episodic memory framework to encode diverse experiences of the robot acquired cumulatively by interacting with the environment. The central idea is that the episodic memory network is activated autonomously based on diverse partial cues emerging from the environment mainly- a) vision-objects perceived in the present scene; b) linguistic words, for example the word "red ball" or "assemble fuse box"; c) Actions performed by a human counterpart. Partial cues trigger the retrieval dynamics enabling the robot to recall its own past experiences in relation to the present context. We then demonstrate how context specific recall of past episodic experiences based on observation of the actions of a human counterpart enables the robot to simulate future states and engage in cooperative goal directed behaviours. A playful scenario where the robot learns cumulatively through multiple experiences to assemble the tallest possible tower with random object's and then exploits such knowledge to co-operate with the human to jointly assemble the tallest tower is used to illustrate the results. In sum, the architecture offers shared computational basis for "remembering, imagining and perspective taking" in cognitive robots. As a side effect, since such episodic memories are derived from direct experiences (of the robot), also finesses the symbol grounding problem [10].

The rest of the article is organized as follows. Section 2 describes the central building blocks related to perception, action, robot episodic memory system in the proposed framework. An example iCub humanoid robot learning to assemble a tower suing different objects presented randomly, the ensuing representation in the episodic memory network, the encoding and retrieval dynamics is presented to illustrate the computational model. Section 3 presents results where the iCub humanoid exploits its past experiences to creatively collaborate with a human counterpart assembling a tower. A discussion concludes.