Pervasive Memory: the Future of Long-Term Social HRI Lies in the Past

Paul Baxter[1] and **Tony Belpaeme**[2]

Abstract.
The technologies underlying long-term social Human-Robot Interaction (sHRI) continually advance, with impressive results. However, applications typically remain on relatively short time-scales, or if long-term, are more focussed on the human perspective than that of the robot. What are still missing are the broader theories of temporal organisation of robot social behaviour for long-term interaction, and more particularly, how experience informs the adaptivity of behaviour. Based on this necessity, we propose in this position paper that memory should play a more central role in accounts of sHRI. Specifically, we introduce the concept of "Pervasive Memory", i.e. a broader notion of memory than is typically used: memory as underlying and involved in all aspects of social behaviour, beyond the mere passive storage of (symbolic) semantic information. The working hypothesis is that by committing to such a perspective, an integrated and coherent approach to long-term sHRI can be formulated. A number of examples are described in this position paper, including studies currently under way, to support this approach. We suggest that, in this sense, the future development of social robotics really does lie in the past.

1 Introduction

Having social robots capable of interacting over extended periods of time will be useful for many applications in the medical and educational domains for example, where there is an emphasis on adaptable contingent interactions for success, e.g. [13, 26]. However, this ideal has not yet been achieved, for a range of technical and theoretical reasons. What we propose in this contribution is that in order to achieve this goal, a greater consideration of memory is required: what memory is, what role it plays, and how it is structured and implemented are all questions that need to be addressed. Here, we seek to provide an overview of the beginnings of an effort to tackle these issues, and those that arise during the process.

There remain numerous unanswered questions in the research underlying social Human-Robot Interaction (sHRI), concerned as much with the reactions of the human interactants as with the characteristics of the robot itself. For example, the mere appearance of the robot or the task to be undertaken would influence the context of the interaction and thus the perception of the robot by the human, and indeed the interaction itself. These aspects, and others, would necessarily form part of a complete account of sHRI. However, in this paper we are primarily concerned with the behaviour of the robot – i.e. how the robot organises it's behaviour over the course of both a single interaction and a series of interactions with an individual or a number of individuals. This is not to say that the robot can be considered in isolation from its prospective interaction partner(s) – it has been persuasively argued elsewhere that it is necessary to consider the role of the dynamics of the interaction itself, e.g. [8, 15], beyond the competencies of the individual interactants. What we consider instead are (some of) those functionalities necessary on the part of the robot to engage in such interaction dynamics, and examine the mechanisms underlying such capabilities.

Specifically, starting from the general perspective that memory is "*...an active process that serves current and future adaptive behavior, based on previously acquired information...*" [31], we posit that it is fundamentally involved in all aspects of social interaction, and thus also in social HRI. In this paper, we introduce the concept of 'Pervasive Memory', which emphasises an integrated approach to the organisation of (robot) social behaviour founded on the substrate[3] of memory. This perspective on memory differs fundamentally from the approach typically taken in robot control systems, and indeed in synthetic cognitive architectures. This typical approach decomposes memory into functionally specific modules, each responsible for distinct types of information [28], as occurs in many cognitive architecture for example [19]. Through the application of the Pervasive Memory perspective, we are effective advocating an inversion of this. For our work we do not restrict ourselves here to a particular type of information (since memory is typically taken as only declarative), or to any particular representation (since symbolic information is typically assumed). By taking such a broad concept of memory as the starting point, we commit not only to examining the common aspects of memory across multiple facets of sHRI, but also to examining the commonalities across these multiple facets of sHRI themselves. While memory has previously been considered and used in sHRI, this generally is either a superficial treatment (such as passive storage of timestamped semantic information), or as additional functionality to an existing learning/cognitive algorithm. The purpose of this paper is to set out the start of this inversion process.

We begin with a consideration of a number of important issues in sHRI, illustrating the important role that a (somewhat general) memory must play (section 2). Then we describe how a theoretical framework of memory-based cognition can be applied to sHRI, and from this how we instantiate the notion of Pervasive Memory (section 3). Finally, we explore some of the wider implications of this position by noting how ongoing developments in both cognitive modelling and

[1] Centre for Robotics and Neural Systems; The Cognition Institute; Plymouth University, U.K.; email: paul.baxter@plymouth.ac.uk
[2] Centre for Robotics and Neural Systems; The Cognition Institute; Plymouth University, U.K.; email: tony.belpaeme@plymouth.ac.uk

[3] We use the term 'substrate' as indicative of commonalities of memory in function (what role memory plays), structure (how memory is organised) and mechanism (how memory operates) across the different aspects of sHRI discussed in this paper.

cognitive architectures can be leveraged to improve the efficacy of social robots over the longer term (section 4).

2 Long-Term Social HRI and Memory

At the outset of this paper, we contended that memory is pervasive in all aspects of sHRI. The intention in this section is therefore to provide examples in support of this contention. There are of course many diverse aspects to sHRI; it is naturally not possible to cover all of these in this short contribution. We have therefore chosen a range, from general to more specific, and each of which are by themselves fairly uncontroversial. For these, we have identified a straightforward but plausible mechanism that could be used to fulfil the function. By doing so, we seek to illustrate how applying *ad hoc* solutions to memory in sHRI contexts misses the broader benefits that the Pervasive Memory perspective can bring. The common theme for these examples is of course that memory is required. Recall that the definition of memory taken (in section 1) is a general one, emphasising function over computational instantiation. Based on the following examples, this definition is revisited and refined to form the basis of a computational model that can be applied to sHRI (section 3).

Long-term sHRI. The first, and most general, illustration of this need is the goal of long-term social interaction, which has the capacity to find a range of practical applications, such as in the home, e.g. [18]. It is generally accepted that for this to be feasible, various degrees of adaptivity are required [17]. Indeed, in recent work involving social child-robot interactions, it has become clear that this process of adaptation works both ways (i.e. that the child adapts her behaviour to that of the robot), and that even a superficial adaptation on the part of the robot has profound effects, such as using an interactants name and acknowledging a prior interaction [9, 23]. Keeping with the generality of the example, adaptivity relies on the capacity to allow prior experience to influence current and future behaviour. Given our definition of memory, we can refine this: memory is required support the adaptation of the social robot to its interaction partner(s). In the simple case of acknowledging that a prior interaction took place, a boolean flag would suffice, but this is an encoding of some aspect of experience – very specific in this case.

Multimodal alignment. Taking this notion of adaptivity over extended periods one step further, we consider the more specific phenomenon of behavioural alignment, whereby an individual will modify their behaviour contingent on the behaviour of their interaction partner(s), and vice-versa [14]. This effect has been observed over multiple different interaction modalities, including linguistic [11], dialogue [25], non-verbal expressivity [32], and turn-taking (as a form of accomodation) [10]. It thus seems to be a fundamental feature of social interaction, and one that a long-term socially interacting robot will need to achieve. Consider for instance the role of robot proxemics (e.g. [29]): finding an appropriate distance/approach for one user would not necessarily apply to other users, necessitating a distinction, and thus requiring memory to retain this information for future encounters. Such information could for example be stored in a look-up table, where a distance (and/or set of movement parameters) could be stored for each of the individuals encountered. As with the first example, this solution would encode prior experience, but would also be restricted to the single context, being rather brittle to change of application.

In support of social cognition. In addition to the interaction itself, there are a number of supporting cognitive competencies that it is necessary to consider. For the goal of long-term sHRI, these aspects must be taken into account as they underlie the types of behaviour that go beyond the mere reactive[4]. For example, there is a need for the functionality afforded by episodic memory, i.e. the memory of a specific sequence of events that occured at a particular time [12]. Furthermore, the ability to learn, process and apply concepts (both grounded and abstract) is an important cognitive competence. The necessity for memory for each of these is clear. These could be conceivably implemented with a type of database: episodic memory has typically been encoded in this way using timestamps as a key for instance, though brittle representations (in terms of generality or task specificity) are a typical drawback.

3 The Pervasive Memory Perspective

The examples described above illustrate the requirement for memory in specific facets of sHRI, though we would contend that this could be extended to all facets. However, the definition of memory used though was intentionally general. We can now elaborate on it: memory is a mechanism (not a passive information repository) that encodes experience in a manner that facilitates its subsequent involvement in all cognitive processing [31], including those competencies required for social behaviour. For sHRI, Memory is thus Pervasive. In the examples above, plausible means of achieving this encoding of prior experience were suggested (boolean states, lookup tables, and databases). While somewhat contrived, these potential solutions are application or task specific, fulfilling the desired function but not facilitating generalisation to new types of behaviour. Three distinct domains can be distinguished for the application of memory to sHRI: (1) perception (such as face recognition, etc); (2) decision-making (such as using prior knowledge to decide on a topic of conversation); and (3) action (such as appropriate approaches, etc). However, sHRI implementations (such as in the examples considered above) typically only consider one or two of these.

Given that memory is intrinsically involved in each of these functions, what is desirable is an account of memory that can overcome this issue: to achieve this, inspiration was sought from a number of other disciplines.

Starting from both neuroscience [1], and philosophical [27] foundations, a memory-centred framework of cognitive processing has been formulated in the context of cognitive robotics [2]. The basic assumption of this framework is that memory consists of a single type of computational substrate, and that this remains consistent across application contexts. This substrate is assumed to be essentially distributed and associative, reflecting two central features of the human cognitive system [16]. Memory can therefore be considered independently of a given task context, with the set of mechanisms constituting memory instead acting as the starting point for the design of a system [7].

The computational implementation we employ to instantiate these principles is based on a parallel distributed processing (PDP) model, e.g. [20], which makes use of a localist knowledge representation scheme [24]. Termed the Distributed Associative and Interactive Memory system (DAIM), this implementation encodes the experience of the agent in a series of associative links which are incrementally acquired and updated. Using the resulting associative network

[4] In this sense, we take reactive include the kind of behavioural alignment described above since the adaptation of the robot is in response to the behaviour of the human interaction partner.

as a substrate for activation dynamics, encoded information can be '*recalled*' and '*processed*' through the reactivation of associations [7]. Being a distributed sub-symbolic representation scheme, in common with other connectionist approaches, one natural side-effect is the presence of a degree of generalisation in both the encoding and retrieval processes. Details of the computational implementation may be found in [5] and [4].

We have applied DAIM to the issue of multi-modal alignment in social interaction. With experimentation currently under way, the working hypothesis is that if a robot aligns its behaviour with its human interaction partner, the human will tend to remain engaged in the interaction for longer than if the robot does not display such contingent behaviour. Set in the context of a touchscreen-based sorting game with no imposed interaction structure [6], the robot can modify various aspects of its behaviour (modalities), such as movement speed, time taken between sorting attempts, and performance accuracy [4]. Alignment here would involve the characterisation of the human's behaviour, and the replication of this by the robot across each of the modalities. DAIM can successfully learn such a characterisation [4]. This task on its own could be solved using alternative means (see the lookup table example above). However, the same system has been employed to provide an account of concept acquisition through experience (thus also demonstrating the capacity for generalisation) [5], and augmented processing in dialogue management [3], which would not be possible for a task-specific solution. Returning to the three domains described above, these cases show that DAIM has been applied to perception, decision-making, and action, though not simultaneously. In sum, while not yet providing a complete account, DAIM provides a promising task-independent implementation of the Pervasive Memory perspective.

4 Looking through the Past into the Future

That memory is important for (s)HRI is not per se a novel claim in this paper. Indeed, among others, a model of embodied sensorimotor histories applied to human behaviour understanding explicitly encodes prior experience to facilitate future prediction [21]. What we do however claim is not just that memory is pervasive across all aspects of social behaviour and cognition, but that it should thus be *central* to any account of sHRI. The type of computational model we have developed shares a number of similarities with other implementations, from a range of disciplines. For example, from developmental robotics, the Epigenetic Robotics Architecture is similarly based on the PDP principles, and is used to model aspects of infant and child language acquisition as facilitated by embodiment [22]. For the Pervasive Memory perspective, it is the principle of memory as the substrate of cognition that is key.

This emphasis on principles is in common with research into cognitive architectures, both natural and synthetic, which seek to provide models of general cognition [19]. Not restricted to sHRI, they emphasise the general, rather than specific task functionality, and typically take inspiration from humans. It has, for example, been persuasively argued that it is necessary for a robot to take into account the sub-optimalities of human behaviour in interaction, with a robot cognitive architecture based on principles of human cognition facilitating this process [30].

The widening of scope entailed by cognitive architectures would seem to be a useful tool for sHRI, since it forces the system designers to maintain a view on the broader issues involved in social interaction, thus encouraging a coherent implementation approach. It is this that we are attempting to achieve through the application of the Pervasive Memory perspective.

On a final note, to a certain extent, the particular design decisions we have taken in the implementation of DAIM and its subsequent application to various problems should be seen as independent from the idea of Pervasive Memory. As a model, DAIM requires further refinement to account for other cognitive and social phenomena; in this regard, the Pervasive Memory perspective provides a consistent guideline rather than a template for a computational system. For example, using DAIM to implement an episodic memory system is in principle plausible [7], but in practice requires implementational detail.

5 Conclusion

In this brief position paper, we have advanced the idea that memory is necessarily and fundamentally involved in all aspects (though we have described only a few) of sHRI: it is pervasive. We suggest that for the design of sHRI systems, there are a number of implications: centrally, that a common memory should pervade all stages of processing, from perception to action. In so doing, we have drawn on evidence, models and examples from a range of disciplines to illustrate the broad base of the approach. Nevertheless, the utility of such a perspective will ultimately only find validation in its successful application to sHRI. Whilst work to this end is as yet in its infancy, we hope to have provided here a persuasive argument in support of the Pervasive Memory perspective. In the simplest terms, long-term sHRI requires the notion of contingent adaptivity, and such adaptivity requires memory. We suggest that, in this sense, the future development of social robots capable of long-term interaction really does lie in the past.

ACKNOWLEDGEMENTS

This work is funded by the EU FP7 ALIZ-E project (grant 248116): www.aliz-e.org.

REFERENCES

[1] Moshe Bar, Elissa Aminoff, Malia Mason, and Mark Fenske, 'The Units Of Thought', *Hippocampus*, **17**, 420–428, (2007).

[2] Paul Baxter and Will Browne, 'Memory as the substrate of cognition: a developmental cognitive robotics perspective', in *10th International Conference on Epigenetic Robotics*, pp. 19–26, (2010).

[3] Paul Baxter, Heriberto Cuayahuitl, Rachel Wood, Ivana Kruijff-Korbayova, and Tony Belpaeme, 'Towards Augmenting Dialogue Strategy Management with Multimodal Sub-Symbolic Context', in *KI 2012*, pp. 49–53, Saarbruecken, Germany, (2012).

[4] Paul Baxter, Joachim de Greeff, and Tony Belpaeme, 'Cognitive architecture for humanrobot interaction: Towards behavioural alignment', *Biologically Inspired Cognitive Architectures*, **6**, 30–39, (2013).

[5] Paul Baxter, Joachim De Greeff, Rachel Wood, and Tony Belpaeme, 'Modelling Concept Prototype Competencies using a Developmental Memory Model', *Paladyn Journal of Behavioral Robotics*, **3**(4), 200–208, (2012).

[6] Paul Baxter, Rachel Wood, and Tony Belpaeme, 'A Touchscreen-Based Sandtray to Facilitate, Mediate and Contextualise Human-Robot Social Interaction', in *HRI'12*, pp. 105–106, Boston, MA, U.S.A., (2012).

[7] Paul Baxter, Rachel Wood, Anthony Morse, and Tony Belpaeme, 'Memory-Centred Architectures: Perspectives on Human-level Cognitive Competencies', in *AAAI Fall 2011 Symposium on Advances in Cognitive Systems*, ed., Pat Langley, pp. 26–33, Arlington, Virginia, U.S.A., (2011). AAAI Press.

[8] Tony Belpaeme, Paul Baxter, Joachim De Greeff, James Kennedy, Rosemarijn Looije, Mark Neerincx, Ilaria Baroni, and Mattia Coti, 'Child-Robot Interaction: Perspectives and Challenges', in *International Conference on Social Robotics*, pp. 452–459, Bristol, U.K., (2013). Springer.

[9] Tony Belpaeme, Paul Baxter, Robin Read, Rachel Wood, Heriberto Cuayahuitl, Bernd Kiefer, Stefania Racioppa, Ivana Kruiff-Korbayova, Georgios Athanasopoulos, Valentin Enescu, Rosemarijn Looije, Mark Neerincx, Yiannis Demiris, Raquel Ros-Espinoza, Aryel Beck, Lola Canamero, Antione Hiolle, Matthew Lewis, Ilaria Baroni, Marco Nalin, Piero Cosi, Giulio Paci, Fabio Tesser, Giacomo Sommavilla, and Remi Humbert, 'Multimodal Child-Robot Interaction: Building Social Bonds', *Journal of Human-Robot Interaction*, **1**(2), 33–53, (2012).

[10] Štefan Benuš, Agustín Gravano, and Julia Hirschberg, 'Pragmatic aspects of temporal accommodation in turn-taking', *Journal of Pragmatics*, **43**(12), 3001–3027, (September 2011).

[11] Holly P. Branigan, Martin J. Pickering, Jamie Pearson, and Janet F. McLean, 'Linguistic alignment between people and computers', *Journal of Pragmatics*, **42**(9), 2355–2368, (2010).

[12] Nicola S Clayton and James Russell, 'Looking for episodic memory in animals and young children: prospects for a new minimalism.', *Neuropsychologia*, **47**(11), 2330–40, (September 2009).

[13] Kerstin Dautenhahn, 'Socially intelligent robots: dimensions of human-robot interaction.', *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **362**, 679–704, (April 2007).

[14] Hanne De Jaegher, Ezequiel Di Paolo, and Shaun Gallagher, 'Can social interaction constitute social cognition?', *Trends in cognitive sciences*, **14**(10), 441–7, (October 2010).

[15] Ezequiel Di Paolo and Hanne De Jaegher, 'The interactive brain hypothesis', *Frontiers in Human Neuroscience*, **6**, 1–16, (2012).

[16] J M Fuster, 'Network Memory', *Trends in neurosciences*, **20**(10), 451–9, (1997).

[17] Michael A. Goodrich and Alan C. Schultz, 'Human-Robot Interaction: A Survey', *Foundations and Trends in Human-Computer Interaction*, **1**(3), 203–275, (2007).

[18] C D Kidd and Cynthia Breazeal, 'Robots at Home: Understanding Long-Term Human-Robot Interaction', in *Intelligent Robots and Systems IROS*, pp. 22–26. IEEE, (2008).

[19] Pat Langley, John E. Laird, and Seth Rogers, 'Cognitive architectures: Research issues and challenges', *Cognitive Systems Research*, **10**(2), 141–160, (June 2009).

[20] James L McClelland and David E Rumelhart, 'An Interactive Activation Model of Context Effects in Letter Perception: Part 1, an account of basic findings', *Psychological Review*, **88**(5), 375–407, (1981).

[21] N. a. Mirza, C. L. Nehaniv, K. Dautenhahn, and R. te Boekhorst, 'Grounded Sensorimotor Interaction Histories in an Information Theoretic Metric Space for Robot Ontogeny', *Adaptive Behavior*, **15**(2), 167–187, (2007).

[22] Anthony F. Morse, Joachim De Greeff, Tony Belpaeme, and Angelo Cangelosi, 'Epigenetic Robotics Architecture (ERA)', *IEEE Transactions on Autonomous Mental Development*, **2**(4), 325–339, (2010).

[23] Marco Nalin, Ilaria Baroni, Ivana Kruijff-Korbayova, Lola Canamero, Matthew Lewis, Aryel Beck, Heriberto Cuayahuitl, and Alberto Sanna, 'Childrens adaptation in multi-session interaction', in *RoMAN 2012*. IEEE, (2012).

[24] M Page, 'Connectionist modelling in psychology: a localist manifesto', *The Behavioral and brain sciences*, **23**(4), 443–67; discussion 467–512, (2000).

[25] Martin J Pickering and Simon Garrod, 'Toward a mechanistic psychology of dialogue', *The Behavioral and brain sciences*, **27**(2), 169–90; discussion 190–226, (April 2004).

[26] Karola Pitsch, Hideaki Kuzuoka, Yuya Suzuki, Luise Sussenbach, Paul Luff, and Christian Heath, 'The first five seconds: Contingent stepwise entry into an interaction as a means to secure sustained engagement in HRI', *RO-MAN 2009*, 985–991, (2009).

[27] Alexander Riegler, 'Constructive memory', *Kybernetes*, **34**(1), 89–104, (2005).

[28] Ron Sun, 'Memory systems within a cognitive architecture', *New Ideas in Psychology*, **30**(2), 227–240, (August 2012).

[29] Elena Torta, Raymond H Cuijpers, and James F Juola, 'Design of a parametric model of personal space for robotic social navigation', *International Journal of Social Robotics*, **5**(3), 357–365, (2013).

[30] J Gregory Trafton, Laura M Hiatt, Anthony M Harrison, P Tamborello, Sangeet S Khemlani, and Alan C Schultz, 'ACT-R/E : An Embodied Cognitive Architecture for Human-Robot Interaction', *Journal of Human-Robot Interaction*, **2**(1), 30–54, (2013).

[31] Rachel Wood, Paul Baxter, and Tony Belpaeme, 'A Review of long-term memory in natural and synthetic systems', *Adaptive Behavior*, **20**(2), 81–103, (2012).

[32] Akiko Yamazaki, Keiichi Yamazaki, Matthew Burdelski, Yoshinori Kuno, and Mihoko Fukushima, 'Coordination of verbal and non-verbal actions in humanrobot interaction at museums and exhibitions', *Journal of Pragmatics*, **42**(9), 2398–2414, (2010).