A Socially-Aware Memory for Companion Agents

Mei Yii Lim[1], Ruth Aylett[1], Wan Ching Ho[2], Sibylle Enz[3], and Patricia Vargas[1]

[1] School of Mathematical and Computer Sciences,
Heriot Watt University,
Edinburgh, EH14 4AS, Scotland
{myl, ruth}@macs.hw.ac.uk, p.a.vargas@hw.ac.uk
[2] Department of Computing Science,
University of Hertsfordshire, England, UK
w.c.ho@herts.ac.uk
[3] Otto-Friedrich-Universitaet Bamberg,
Kapuzinerstrasse 16,
D-96045 Bamberg, Germany
sibylle.enz@uni-bamberg.de

Abstract. Memory is a vital capability for intelligent social Companions. In this paper, we introduce a simple memory model that allows a Companion to maintain a long-term relationship with the user by remembering past experiences in order to personalise interaction. Additionally, we implemented a situational forgetting mechanism that gives the Companion the ability to protect the user's privacy by not disclosing sensitive data. Two test scenarios are used to demonstrate these abilities in our Companions.

1 Introduction

Memories are part of what makes up our personality, shapes our reactions to life situations and often influences our mood [1]. Besides giving us identity, memory is fundamental to intelligence. Learning and memory are the basis of our knowledge and abilities that allow us to consider the past, place us in the present and help us to predict the future. Additionally, memories help us to determine who and what are important to us. They allow us to behave in socially appropriate ways and hence maintain a long term relationship with our interaction partners. Since memory is such an important aspect of human intelligence and social life, we argue that in order to create intelligent artificial social Companions, memory is crucial. A 'human-like' memory in these Companions will help them to comprehend their world, focus their attention on important information relevant to the current interaction situation and to make predictions about it. These Companions will be able to act consistently and hence exhibit a 'personality', a reflection of 'self' that is important in social communication [2].

In this paper, we discuss the important aspects of memory for intelligent artificial social Companions. Section 2 summarises aspects of the background in human memory research on remembering and forgetting. Section 3 discusses issues related to the memory of social Companions. Section 4 describes our initial memory prototype complete with example scenarios of personalisation and social awareness. Section 5 presents some conclusions and describes future work.

2 Remembering and Forgetting

It is yet unclear and controversial among scientists, how exactly memory works, but the following review introduces some of the ideas that are more widely agreed upon. Three fundamental stages in the formation of memory are encoding, storage and retrieval which take place on sensory, short-term and long-term levels [3]. Information from STM is stored in LTM through repeated exposure and generalisation (reconstruction).

Bartlett's work [4] emphasized the reconstructive view of LTM showing that memories are often reconstructed based upon world knowledge and schemata. He rejects the notion that memory representations consist of accurate traces that are consistent over long durations. Alba and Hasher [5] proposed a prototypical schema theory of memory, assuming the operation of four central encoding processes: "selection, abstraction, interpretation and integration". In addition, a fifth process, reconstruction, occurs when an individual attempts to reproduce a memory episode.

On the other hand, forgetting is useful to improve efficiency, scalability and adaptability of cognitive systems operating in dynamic task environments. Forgetting is also important to prevent stale information from interfering with fresh information and can be explained by decay theory, displacement, reconstruction process, interference and repression.

When it comes to recalling information from memory, contextual cues are crucial. Tulving and Psotska [6] have shown that the absence of a valid cue for recall causes forgetting (cued recall) and if contextual information is missing, memory recall fails. Bouton and colleagues [7] suggest that retrieval is most effective when a match exists between encoding and retrieval conditions. They add that a mismatch might occur with the passage of time due to the fluctuation of internal and external contextual cues, hence, reducing the likelihood of the target material being retrieved.

3 Artificial Social Companions and Memory

In recent years, artificial Companions have become increasingly popular especially as a form of entertainment and to assist the elderly in maintaining an acceptable standard of life. These Companions can include digital pets, such as the popular Tamagotchi, or robots, such as the PARO[4] and the Sony AIBO[5]. In this paper, the term "companion" refers to an artificial agent who is frequently

[4] http://www.parorobots.com/
[5] http://support.sony-europe.com/aibo/

in the company of the user. Up to date, the social, psychological and cognitive foundations and consequences of such technological artifacts entering our daily lives over an extended period are not well understood. Many scientists and philosophers [8, 9] discuss the potential ethical danger of these artifacts. In the field of robotics, a new discipline named Roboethics [10] has evolved as a result.

In the design of an artificial social Companion memory, important ethical issues have to be addressed alongside the aim of providing the companion with a memory that can maintain long term relationships. Significant questions would be 'what' the companion should remember, 'how' these data are processed and to 'whom' the information should be disclosed. One of the important aspects of memory modelling for an ethical Companion is forgetting [11], particularly motivated forgetting or repression [12] which allows the Companion to repress ('forget') sensitive information under specific circumstances to prevent harmful consequences to its relationship with the user. To maintain long-term relationships, it is important for the Companion to be adaptive so that its interaction with the user can be personalised.

4 Initial Prototype

4.1 The Architecture

Based on the reviews and discussions in the previous sections, we have developed an initial memory prototype for intelligent artificial social Companions. The prototype is built on top of FAtiMA [13], an emotional model for virtual agents that incorporates reactive and deliberative appraisal components responsible for agents' decision making. We take advantage of this functionality to help the agent determine the appropriateness of its behaviour. Additionally, FAtiMA includes an autobiographic memory [14] where previous events are stored as well as a knowledge base that stores information related to the interaction environment. This provides a basis for our Companion memory design, testing and development. However, FAtiMA memory does not distinguish between short-term and long-term storage and the existing memory is comparable to LTM. Additionally, FAtiMA stores all actions and events without any forgetting.

In order to create a more human-like memory, FAtiMA memory was restructured. Since we adopt an iterative prototyping approach, our initial memory model is very simple as shown in Figure 1. The autobiographic memory (AM) and knowledge base (KB) are considered as LTM and a STM is added. The KB in the LTM and the STM are comparable to the recently enhanced ACT-R components: the LTM-DM (declarative memory) and the LTM buffer (working memory) respectively [15]. However, the ACT-R theory does not include motivated forgetting mechanism and the model has not yet been used in human-agent interaction scenarios. Our STM acts as a buffer to actions and events before they are transferred to the AM and consists of a working memory that holds information related to current goal processing. This information includes properties of objects and people in the interaction and the Companion's internal representations of its evaluation of potential events that might take place. When a goal either succeeds or fails, the companion will update its knowledge base (LTM) with the new evaluation information from its working memory. The flow of information between these memory components can be seen in the dotted box component of Figure 1.

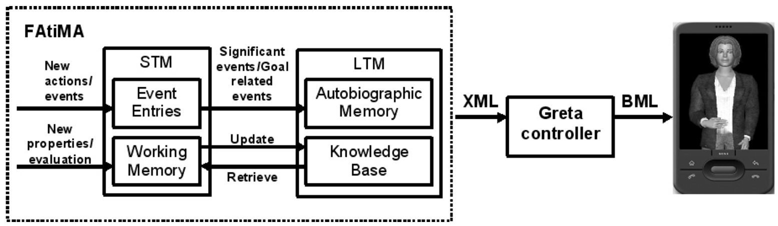

Fig. 1. Linking FAtiMA to Greta

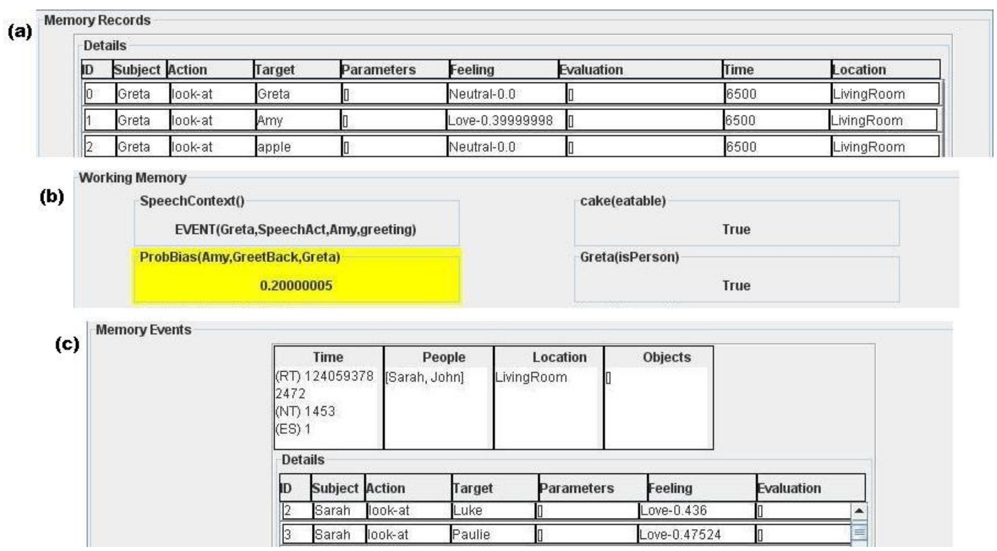

Fig. 2. (a)The Structure of Memory Records in Short Term Memory; (b)Working Memory and (c)Autobiographic Memory

The structure of STM memory records and the working memory are presented in Figure 2(a) and Figure 2(b) respectively. Instead of storing all actions and events taking place in the environment, only significant events (eg. events with emotional impact) and those related to goal processing (activation, success and failure) are transferred to the AM as shown in Figure 2(c). In addition, the AM stores general information about a particular episode of interaction such as time, location, participants and objects. For more information about the memory architecture, please refer to [16].

4.2 Testbed and Test Scenarios

In order to embed the memory into an interaction environment, the version of FAtiMA used has been connected to the Greta [17] agent. The components link is presented in Figure 1. Information from FAtiMA is sent to the Greta controller in XML format that is translated to the Greta understandable format, BML, so that actions can be performed and texts can be synthesised to speech accordingly. The user interacts with the system through a graphical user interface by choosing actions from a drop-down list.

Two scenarios have been devised to test our initial memory prototype, one relates to personalisation and the other addresses the user privacy issues [11]. In the first scenario, we aim to show how an agent, Greta, remembers its interactions and uses this information in later encounters with the user. The interaction involves the agent offering food to the user, Amy, and remembering her preference. Additionally, we included the notion of time which is currently represented as a property of the agent about the environment. Greta selects different actions based on the time of interaction - offering fruits in the morning and cake in the afternoon. The interaction takes place at different locations in Amy's house. Details of the scenario are presented below:

- Case 1: Location - Living Room; Time - Morning; Description - Greta will offer Amy fruits - either apple, banana or orange, one at a time and will stop offering when the user accepts one of the fruit
- Case 2: Location - Study Room; Time - Afternoon; Description - Greta will offer Amy cake
- Case 3: Location - Kitchen; Time - Morning; Description - Greta retrieves from its memory previous interactions and offers the fruit that Amy has accepted before

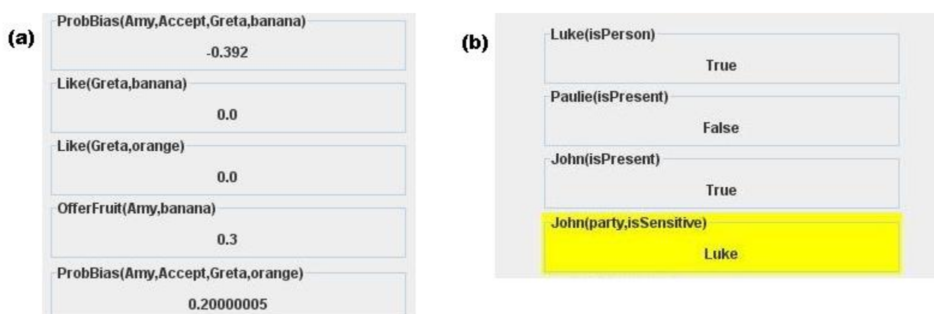

Fig. 3. (a)Greta's evaluation of probabilities user accepting certain fruit; (b)Tagging of sensitive information

After each interaction in this scenario, Greta updates its knowledge base with new information from current processing. So, after case 1, suppose Amy rejected banana and accepted orange, it will have the information (probability that Amy accept banana is -0.392 and orange is 0.20000005) as shown in Figure 3(a) in its memory. During case 3, Greta retrieves this information into its working memory and it is this information that helps Greta to decide on what to offer the user next, hence personalising its interaction with the user.

In the second scenario, we aim to show how an agent, Sarah (using the same graphical representation as in Figure 1) tags its memory about information sensitive to a particular user and how it handles queries related to the information when being asked by the respective user. The interaction takes place in an office setting and three users are involved - John, Luke and Pauline. Details of the scenario are as below:

- Case 1: Location - Reception; Agents - Sarah and John; Description - John tells Sarah about him being drunk at a party and told Sarah that she should keep it secret from Luke, a good friend of his wife. Additionally, he also told Sarah that he is getting a new job and his existing boss Pauline should not know about this
- Case 2: Location - Common Room; Agents - Sarah, Luke and Pauline; Description - Luke and Pauline ask Sarah about John regarding the party and the new job. Sarah will hide the information since each piece of the information is sensitive to one of them - the party to Luke, and the new job to Pauline
- Case 3: Location - Pauline's office; Agents - Sarah and Pauline; Description - Pauline ask Sarah again about the related topics. This time Sarah will expose the information about John being drunk to Pauline but not about him getting a new job

When John tells Sarah his personal information and to whom this information is sensitive, Sarah tags its memory entry as shown in Figure 3(b). Later, when it is being asked about this information, it will retrieve its memory of the entry and will know to whom the information should or should not be disclosed.

5 Conclusion and Future Work

This paper demonstrates a simple artificial social Companion memory model that allows personalisation of interaction and takes ethical issues into consideration. The design of a full working memory model is still on-going, therefore an initial prototype of the memory model has been implemented and two case scenarios were used to show how the above issues are addressed. Future work includes a more detailed specification of the Companion memory model that addresses issues such as general forgetting, generalisation and retrieval. This would entail modification to the existing memory organisation as it currently only allows storage based on events and does not employ any ontological or hierarchical structure.

Acknowledgements

This work was partially supported by the European Commission (EC) and is currently funded by the EU FP7 ICT-215554 project LIREC (Living with Robots and Interactive Companions). The authors are solely responsible for the content of this publication. It does not represent the opinion of the EC, and the EC is not responsible for any use that might be made of data appearing therein.

References

[1] Carver, J.M.: Emotional memory management: Positive control over your memory. Burn Survivors Throughout the World Inc. (2005) http://www.burnsurvivorsttw.org/articles/memory.html.

[2] Dautenhahn, K.: The art of designing socially intelligent agents – science, fiction and the human in the loop. Applied Artificial Intelligence **12**(7-8) (1998) 573–617

[3] Atkinson, R., Shiffrin, R.: Human memory: A proposed system and its control processes. The Psychology of learning and motivation: Advances in Research and Theory **2** (1968)

[4] Barlett, F.C.: Remembering: A Study in Experimental and Social Psychology. Cambridge University Press, Cambridge, Great Britain (1932)

[5] Alba, J.W., Hasher, L.: Is memory schematic? Psychological Bulletin **93** (1983) 203–231

[6] Tulving, E., Psotka, J.: Retroactive inhibition in free recall: inaccessibility of information available in the memory stores. Journal of Experimental Psychology **87** (1971) 116–124

[7] Bouton, M.E., Nelson, J.B., Rosas, J.M.: Stimulus generalization, context change, and forgetting. Psychological Bulletin **125** (2008) 171–186

[8] Sparrow, R.: Killer robots. Journal of Applied Science **24**(1) (2006)

[9] Walters, M.L., Otero, N.R., Koay, K.L., Syrdal, D.S., Dautenhahn, K.: He knows when you are sleeping - privacy and the personal robot. Technical report (2007)

[10] Veruggio, G.: The birth of roboethics. In: ICRA 2005, IEEE International Conference on Robotics and Automation Workshop on Robo-Ethics. (2005)

[11] Vargas, P.A., Ho, W.C., Mei Yii Lim, S.E., Fernaeus, Y., Aylett, R.: To forget or not to forget:towards a roboethical memory control. In: Killer Robots or Friendly Fridges: the Social Understanding of Artificial Intelligence, AISB'09, Edinburgh (2009) 18–23

[12] Freud, A.: The Ego and the Mechanisms of Defence. Hogarth Press and Institute of Psycho-Analysis, London (1937)

[13] Dias, J., Paiva, A.: Feeling and reasoning: A computational model for emotional agents. In: 12th Portuguese Conference on Artificial Intelligence (EPIA 2005), Portugal, Springer (2005) 127–140

[14] Ho, W.C., Dautenhahn, K., Nehaniv, C.L.: Computational memory architectures for autobiographic agents interacting in a complex virtual environment: A working model. Connection Science **20**(1) (2008) 21–65

[15] Schultheis, H., Lile, S., Barkowsky, T.: Extending act-r's memory capabilities. In: Proc. of EuroCogSci'07. (2007) 758–763

[16] Ho, W.C., Lim, M.Y., Vargas, P.A., Enz, S., Dautenhahn, K., Aylett, R.: An initial memory model for virtual and robot companions supporting migration and long-term interaction. In: ROMAN 2009, IEEE International Symposium on Robot and Human Interactive Communication. (to appear).

[17] de Rosis, F., Pelachaud, C., Poggi, I., Carofiglio, V., Carolis, N.D.: From Greta's mind to her face: Modeling the dynamics of affective states in a conversational embodied agent. Special Issue on Applications of Affective Computing in Human-Computer Interaction, The International Journal of Human-Computer Studies **59** (2003) 81–118