Do I remember you? Memory and identity in multiple embodiments

Ruth Aylett, Michael Kriegel, Iain Wallace, Elena Segura, Johanna Mecurio, Stina Nylander, Patricia Vargas

Abstract — This paper investigates user perceptions of continuous identity as agents migrate between different embodiments. It reports an experiment seeking to establish whether migrating or not migrating the interaction memory of the agent would affect the user’s perception of consistent agent identity over different embodiments. The experiment involved a treasure hunt in which a virtual agent migrated from a screen to a mobile phone in order to accompany a user while they searched for clues. A total of 45 subjects took part in three different conditions with 15 subjects in each. The outcome showed that the presence of memory affected the competence users ascribed to the virtual agent but had no significant effect on a strong perception of consistent identity across multiple embodiments.

I. INTRODUCTION

It has long been established that embodied agents are seen by human users as social actors [7, 21]. Research in social agents [4, 7] examines the design issues for both robotic and virtual agents that are specifically intended for integration into human social environments.

A body allows an agent to display expressive behaviour and this in turn supports the user’s intentional stance [5] and their ability to use their theory of mind capabilities to interpret the agent’s behaviour in terms of internal states such as beliefs, plans and goals. The characteristics of a particular embodiment have been shown to affect the users’ perception and interpretation of, as well as behaviour towards the agent [8, 11, 16]. In turn, user perceptions and behaviour are significant design factors in developing social agents.

Embodiment is of course a functional as well as a social issue. While the higher physical presence of robot embodiments has been widely researched [8, 11, 19, 20, 24], still unsolved engineering problems of power sources, mobility and localization limit the ability of robots to accompany humans as they move from one social and physical environment to another - for example from home to work. The work reported here was carried out in the EU FP7 LIREC project [25] as part of a programme investigating long-term companionship, and as a result the disadvantages as well as the advantages of robot embodiment had to be confronted. A large-screen graphical agent cannot move at all in the real world but can use networks to move between screens and can appear as life-size, with behaviours that are not limited by sensors/actuators as for a robot. A mobile phone-based agent can be carried with a user as a personal companion into a variety of social environments, but lacks independent mobility, life-size and certain interaction modalities. With the aim of tailoring embodiment appropriately, the LIREC project carried out work on migration [14], the ability to transfer the internal state of a robotic embodiment into different robot embodiments as well as into graphical or handheld embodiments, and vice versa.

A key question for work on migration is: what migrates? We define this as transfer of the identity of the agent, by which we mean those features that persist and make it unique and recognizable from the user’s perspective. Which features should persist have been the subject of research. In spite of the importance of embodiment to interaction, an early experiment [10] in which a robot apparently migrated to a desk lamp, using an identical voice as as cue, showed that users can accept a transfer of identity even when appearance is vastly different. Conversely, a more recent experiment [1] showed that users could distinguish two identities appearing sequentially in one robot embodiment if the behaviour of the embodiment varied in a systematic way indicative of changed personality.

Cueing migration as a process is also important. Koay et al. [12] found that a majority of their participants did not perceive a common identity when an agent migrated. Participants found the appearance and sound of the two embodiments too dissimilar. However displaying a progress bar during migration supported users in understanding that the agent “moved” between embodiments.

Timing is also an issue. Migration between virtual (graphical) and physical (robot) embodiments of a Pleo dinosaur agent with child participants [16] showed that a long gap with both embodiments inactive led most to perceive one character, though some worried it might have ‘died’, while an overlap in migration with both versions active simultaneously made most of them perceive two characters. Thus to maintain the experience of one character migrating between two embodiments, only one embodiment should be active at a time. This work also maintained consistency of the experiences of the agent across embodiments. A hungry robot Pleo would still be hungry when it migrated to a virtual embodiment. If fed in this form, it would no longer be hungry when it migrated back to its robot form. Common needs and goals also bridged the virtual and physical embodiments: the agent wanted to be fed and petted in both. However hunger in the physical embodiment was expressed by sniffing and biting toward the ground (looking for food) while the virtual embodiment sat down and cried when it was hungry. These results suggest that if users can perceive continuous and identical goals, identical behaviour is not required.
What had not been tested until the work reported here was whether asserting continuous memory of interactions across embodiments acts as a migration cue. Autobiographical memory has been identified by several authors as a basis for a long-term relationship between a user and a virtual agent. It allows agents to remember significant past experiences and reconstruct their life stories from these experiences [9].

The Sarah agent is equipped with both a long-term and a short-term human-like autobiographical memory [15]: the former storing interaction episodes over periods of weeks or months, and the latter over hours. The logistics of testing long-term memory effects on interaction in the workplace are formidable, though this has been managed over the internet [3], where it significantly increased the persuasive impact on the user in the context of behavioural change for better health. However in geographically co-located interaction, even short-term memory effects do not appear to have so far been investigated. Thus we investigated whether reference back to the content of an interaction in an earlier embodiment would increase the believability of the agent and improve the user’s perception of its continuity across different embodiments.

II. THE EXPERIMENT

The experiment we discuss was carried out within the LIREC showcase Spirit of the Building, involving a workplace companion called Sarah with three embodiments, seen in Figure 1. The first was a robot companion acting as a Team Buddy for a research group in a specific lab (S_R); the second a large-screen graphical companion able to interact with visitors entering the building (S_G); and the third, a mobile phone-based companion (S_M) able to direct visitors round the building using their own smart phone.

All Sarah embodiments ran a three-level architecture forming a common framework for different platforms [14]. Level 1 was wholly platform-dependent while Level 2 contained platform-specific competencies organised by a generic competency manager CMION [14]. Thus the S_G Level two supported menu-based interaction using a Kinect while that for S_M included competencies for robust indoor navigation around the building. Level three, common across all platforms, was provided by the FAtiMA affective agent architecture [6].

Migration was effected by transferring the FAtiMA instantiated parameter set so that the receiving embodiment was configured exactly as the sending one had been, together with a subset of the overall memory. This ensured that the goals, drives and affective state of the agent were transferred: if the agent was happy before migration it would still be happy after migration.

Experiment Design

The discussion above indicates that a variety of cues are significant in establishing a perception of a common identity over multiple embodiments. In order to isolate the impact of a common memory, careful attention was paid to keeping other cues unchanged.

The differential impact of robot versus graphical embodiment is well-established as already noted. For this reason, we chose the two graphical embodiments, S_G and S_M for this experiment. S_G and S_M use the same graphical model and animations. Both had text-to-speech output supplied by the same unit-selection based TTS system giving a natural female voice with a noticeable Irish accent. Neither had speech input: interaction with S_M used the smart phone touch-screen menus and text entry, and for S_G on-screen options were selected with hand gestures recognised with a Microsoft Kinect as seen in Fig. 1.

While the agent only ever appeared on one device at a time and migration only occurred when the user was in front of the screen with both devices visible, the term migration was not used at all in the experimental protocol to avoid biasing the user’s expectations. In other experiments we had found that a rising sound at the sender followed by a falling sound at the receiver helped establish the idea of a single agent in transit. In this experiment, no sound was used. The conversational style was chatty, friendly and informal across both platforms and the same people were responsible for editing the scripts for both, to ensure a consistent language style. Participants were asked specifically about differences between the different embodiments, and none said they had noticed differences in the behaviour.

The task the participants were set was a treasure hunt around one of the university buildings. Clues were provided by S_G embodied in a large TV screen, and participants had to hunt out the answers to these clues by navigating round the building, following directions, after migration, of S_M embodied in a mobile phone. Four clues were involved, so participants were exposed to a total of eight migrations, and considerable dialogue with the agent.

Testing perceptions of memory and identity

The change between the different experimental conditions was in the agent’s memory. Participants (N=45; M=28; F=17) were split into three groups, with 15 participants in each:
Group A – Memory was migrated every time, except for the final migration.

Group B – Memory was not migrated, except for the final migration where the agent remembered events from the prior embodiment.

Group C – Memory was migrated on alternating pairs of migrations, so the first migration to and from the mobile device the memory is preserved, the second it is not, and the third it is (but the earlier memories from the first embodiments are not restored).

Participant age ranged between 20 and 65 (mean 30, S.D 10.9). Most were computer science professionals or students with a mean of 35.6 hours a week of work-related computer use; 17 had some previous exposure to computer-based characters, of whom 7 cited Siri.

Migrating memory can only possibly have an impact if it changes what happens in interaction. Two approaches were taken. The explicit approach overtly signals the transfer or not of memory between embodiments. Here the agent would pass on information to the user in one embodiment, and then ask for the information again in the no-memory conditions. Thus at the start of each clue-search, $S_1$ passed on the clue. In the memory condition, $S_M$ would repeat the clue and make a suggestion about where to look. In the no-memory condition, $S_M$ would ask the user to supply the clue they had just heard and then make the suggestion. Additionally, $S_M$ made comments about the building and the task as it carried it out, e.g. “there’s an entrance down this end of the building, not many people know that”. $S_M$ would repeat the comment in the no-memory condition if the user returned there, but otherwise would comment that they’d been there before.

An implicit mechanism was used to evaluate if the user’s attitude changed during the course of the experiment as they were exposed to more migrations. Careful questions were inserted into in the agent’s dialogue to establish indirectly if the user believed they were interacting with the same agent as the previous embodiment, or a different agent. These questions were asked after the first migration, before the very last migration, and at the end. This was to establish an initial baseline for all participants after one migration (no training), after the majority of migrations (initial training) and at the end, where for cases A and B the memory condition is changed (further training).

This approach was implemented through a recall test, inserted where interaction provided a natural opportunity for the agent ask the user a question. The questions related to information given by the agent in a previous embodiment. For example, in the initial interaction with the user, the agent asks about favourite colours, saying it needs to check that the user can interact properly. During this it reveals its own preference is for purple.

After migration to the mobile device, the first thing the agent does is to ask the user to set the background to the agent’s favourite colour. This answer in itself tells us nothing – the user may choose purple because they believe all such agents like purple. However it is followed by a free-text question on why they chose the colour, which provides the necessary information with which to assess the user’s perception. We typically received replies such as “because you told me it was your favourite”, which implied they perceived a single migrating agent. Further such recall questions referred to the origin of the agent’s accent, which was distinctively Irish, and the total distance it had covered that day.

Apart from using the behaviour of the agent to gather data, a standard questionnaire-based approach was also taken, with pre- and post-questionnaires and finally a structured interview. No questionnaire item directly asked how many agents the user perceived so as not to influence their opinion: they might not have even considered the possibility.

The post-interview first tries to elicit responses indicating participant views indirectly (do they say “the agents” or “the robot moved to the phone” etc.). The emphasis in the interview was on getting the subjects to express what happened in the experiment in their own words, to tease out their opinion of the agent(s) and perception of the agent identity. The questions started as fairly general, and progressed into more specific questions about agent memory, finally explicitly asking if one or two agents were perceived.

The design of the two questionnaires and the choice of questions primarily came from those used in related experiments in the same project. The pre-questionnaire provides an indication of the person’s previous experience and familiarity with virtual agents, computers more generally, and some indication of their personality according to five-factor theory. The post-interview contains a similar evaluation, but of their opinion of the agent they interacted with (summaries of both in the Appendix).

III. RESULTS

A detailed analysis of the results was not needed in order to assess the answer to our main research question. In all three conditions, every single participant gave strong evidence of perceiving a consistent identity across the embodiments. Thus the differences in short-term memory capabilities made no difference to the perception of identity. This can be seen most clearly in the explicit question asked at the end of the structured interview: How many characters were there in the treasure-hunt. *Every* participant answered ‘One’ to this question.

A slightly – but not very much – more nuanced response can be seen in the responses to the recall questions asked after the first migration, before the very last migration, and at the end, as shown below in Table 1. This table records whether the responses to each of the three recall questions during interaction indicated the participant thought there was one continuing agent (Yes) or several discontinuous agents (No). A result of (Unknown) means it was not possible to tell from their answers what they believed.

From this we can see that at no point did anyone definitely believe there to be multiple agents. At the end, the 100% result mirrors that revealed in the interview. The most interesting result is the ‘After 7 migrations’ case.
The structured interview also revealed a greater range of views about the human-likeness or not of the agent than the questionnaire data or the recall questions had suggested. Participants were asked: ‘How do you think the information transfer between the devices worked’? Twenty-two of the forty-five responses focused on the technology or the information, for example ‘A common database’; ‘Transferred a clue’; ‘Some technology that makes the devices communicate’; ‘Some network communication’. This is not a surprise given the high number of computer specialists in our sample and that the question itself was posed technologically. However four of these also referenced the agent as ‘she’ when asked if they had any further comments on this topic.

On the opposite pole, nineteen of the responses focused on the agent, for example: ‘The same agent moves’; ‘She disappears from the screen and appears on the phone’; ‘Felt like the same agent moving to the TV’; ‘Agent appeared and disappeared’; ‘Same person going from one device to the other’; ‘She’s transferring over the internet like Dropbox’. All respondents that ascribed a gender described the agent as ‘she’, and this must have been information derived from the female voice given that the appearance is robot-like and minus any specific gender cues (no name was given to the agent by experimenters during the experiments).

Interestingly, two responses expressed some doubt about the continuity of the agent: ‘Either two agents and a central system or they are the same’; ‘The same agent moves, though having to ask for clues suggests there are two, but I assume it is the same as it looks and sounds the same’. A further point about replies to this question was that some participants offered a technological explanation for the failures of memory that participants noticed and others a human-like explanation: ‘Some information transferred, could have been more – some got lost’; versus ‘forgetful, tells it twice and asks again; she’s testing I’m listening, people do that too’.

This division in perspective between technological tool and social agent is consistent with results for appearance preferences found in [25], where some participants preferred fewer humanoid features in a robot (for example no ‘face’) and other preferred more humanoid features.

IV. CONCLUSION

We interpret this experiment as showing that of all the cues relating to migration, appearance overwhelmingly dominates. Thus rather than minimising appearance differences, it would probably make more sense to maximise them in looking for effects from other migration cues. Using SR – the robot embodiment - and SM with a completely different appearance would make a good follow-up

<table>
<thead>
<tr>
<th>Question</th>
<th>Mean agent competence responses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group A</td>
</tr>
<tr>
<td>Intelligent- Unintelligent</td>
<td>2.7</td>
</tr>
<tr>
<td>Competent - Incompetent</td>
<td>2.1</td>
</tr>
<tr>
<td>Expert-Inexpert</td>
<td>2.5</td>
</tr>
<tr>
<td>Bright-Stupid</td>
<td>2.2</td>
</tr>
<tr>
<td>Trained -Untrained</td>
<td>1.6</td>
</tr>
<tr>
<td>TOTAL</td>
<td>11.1</td>
</tr>
</tbody>
</table>
experiment. We know from [10] that migration between very different embodiments but the same voice supports perception of a single entity, so using voices with two different accents (Irish, Scottish) would also be sensible.

It is also possible that the computer literacy of the participant sample affected results, since the concept of transferring information from one place to another electronically was very familiar to them. Nevertheless the discussion above shows that a variable degree of anthropomorphism was taking place, with some participants taking a more instrumental view and some a more human-like view of the agent.

Though multiple-embodiments are not found as such in the natural world, migration seemed an intuitive concept to participants such that with a small set of cues, consistency of identity over embodiments seems to be the default assumption. Our work therefore suggests that incorporating migration into social robotics is a useful idea. One should note here that distinguishing ‘soul’ from ‘body’ is a long-standing and widespread cultural construct: examples include the concepts of ‘possession’ and ‘life after death’ in Western European and African religions, and of reincarnation in the Indian subcontinent. More currently, users of social computing applications are now familiar with virtual/digital representations of self and with the idea that a virtual representation (avatar) need not look like the human embodiment of its owner.

The impact of including memory in an agent is inherently difficult to measure where the agent is not dedicated to a recall task but carries out more generic social interaction. We are tolerant of degrees of short-term memory error in our human interaction partners. Arguably it is only when memory failure becomes an obstacle to smooth interaction, as in the repetitions of dementia sufferers and their inability to follow a dialogue structure, that we really feel it challenges identity.

More important in human relationships is shared long-term memory of past interactions and common experiences, a joint autobiographical memory of a specific relationship. This cannot however be either established or tested over a short interaction session such as a treasure hunt, but only over long interaction periods of weeks, months or years. Such experiments are difficult to set up since it requires an agent that can successfully integrate itself into the social life of a user in the first place. This is certainly non-trivial for robot embodiments. The internet-based support agents for the elderly of [3] is a rare example of an environment in which such studies could be located.

APPENDIX - QUESTIONNAIRES

A. Pre-questionnaire questions

- Age, gender, profession/field of study
- Hours/week using computer for work/study
- Hours/week using computer for pleasure/recreation
- Do you program computers? If so, how many hours?
- Experience of interactive characters: web, SIRI etc?

- TIP: Ten-factor personality test questions

Agree/disagree with statements about AI/robots scale 1-5:
- 12 questions (not analysed in this paper)

B. Post-questionnaire questions

10 assertions about agent personality attributes, agree 1-5:
- Extravert, critical, dependable, anxious, open, quiet, warm, careless, calm, conventional

Statements about agent(s)/experience, agree 1-5:
- Was/were capable at tasks
- Had lots of knowledge about task
- …was/were reliable
- I was confident about the agent(s)’ ability
- I trusted the agent(s) to do tasks
- I trusted the actions of the agent(s)
- I could rely on the agent(s)’ skills and abilities
- The agent(s): was/were annoying
- ..was/were disruptive
- I was willing to perform tasks requested by agent(s)
- I thought it important to react to agent(s)’ requests
- The mobile app was useful in finding answers
- App navigation instructions were clear
- App navigation instructions were useful
- It was easy to complete the treasure hunt
- The treasure hunt was fun
- Instructions at the big screen were helpful
- Dialogue at the big screen was clear
- Interacting with agent(s) at the screen was easy

I enjoyed using the agent(s)

(12 questions about the Treasure hunt not used in this paper)

I perceived the agent as (scale 1-5)

- Unselfish/selfish
- Intelligent/Unintelligent
- Competent/Incompetent
- Honest-Dishonest
- Expert-NonExpert
- Bright-Stupid
- Trained-Untrained
- Sympathetic-Unsympathetic
- Trustworthy-Untrustworthy
- Warm-Cold
- Approachable-Unapproachable

ACKNOWLEDGMENT

This work was partially supported by the European Commission (EC) and was 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


[22] Rossen, B; Johnsen, K; Deladisma, A; Lind, S. & Lok, B (2008) IVA '08 Proceedings of the 8th international conference on Intelligent Virtual Agents Pages 237-244


