Beyond Presentation - Employing Proactive Intelligent Agents as Social Catalysts

Madlen Wuttke*[1], Michael Heidt*[2]

* these authors contributed equally to this work

[1] Research Training Group crossWorlds, Faculty of Humanities, Chemnitz University of Technology, Thüringer Weg 5, 09126 Chemnitz, Germany

madlen.wuttke@phil.tu-chemnitz.de

[2] Research Training Group crossWorlds, Faculty of Informatics, Chemnitz University of Technology, Thüringer Weg 5, 09126 Chemnitz, Germany

michael.heidt@informatik.tu-chemnitz.de

Abstract. Despite long standing attention from research communities, the technology of intelligent agents still harbours a large amount of unrealised potential. In this text, we argue that agent technology can benefit from a shift in focus from presentation to possible functionalities. In doing this, our focus is on the provision of pro-activity: The ability of agents not to merely react but to predictively shape their environments. In order to illustrate our arguments, we present an instance of interactive technology, showing how pro-active intelligent agents can be employed in exhibition contexts.

1 Introduction

Scientific research regarding pedagogical agents has mainly been focused on analysing different forms of their depiction, rather than possible features. For example, numerous studies have analysed whether an agent should be designed as either male or female [13] or whether or not an agent should be displayed as realistic as possible, including facial animations [1]. In addition to this, Lusk et al. [18] tested if there was a positive effect on learning with an animated or a static agent and Baylor et al. [2] as well as Huang et al. [10] hypothesise a beneficial effect on learning as long as the agent is depicting one's own peer-group and ethnicity.

Regarding the features of an agent-system however, the focus is largely about establishing behavioural strategies. These focus on questions such as if agents are to be depicted as either polite or rude [25], whether the implementation of gestures and mimic behaviour changes the acquisition of learning material [6] or if social conver-

adfa, p. 1, 2011.

sations, not touching on the topic itself, help to create a positive learning environment [23]. What all these research projects have in common is the tendency to analyse passive features of an agent. But what appears to be missing from empirical discourse is research regarding active components of pedagogical agent designs such as active listening, observation of real-world surroundings and just-in-time information aggregation. Those active components would allow for an agent to analyse the environment and to react pro-actively to changes in it [26] as well as acting on behalf of the user.

In order to demonstrate the role pro-active agents are able to play with respect to design of interactive technologies, we discuss a series of design prototypes developed. These have been implemented in various degrees of fidelity, ranging from paper-prototypes to mid-fidelity digital artefacts. The devices conceived are targeted at the museum domain. Their goal is to strike up verbal interaction between previously unacquainted museum visitors. Embedded in the wider scope of a design ecology [7], the system comprises mobile components as well as a stationary wall mounted installation. The stationary setup is equipped with depth cameras used for monitoring of users. Museum visitors are provided with tablets which replace traditional printed museum documentation. On these tablets a personalised instance of an intelligent agent is presented. This agent acts in the capacity of a museum docent, providing both additional information as well as helpful incentives regarding the possibilities of the exhibition visited.

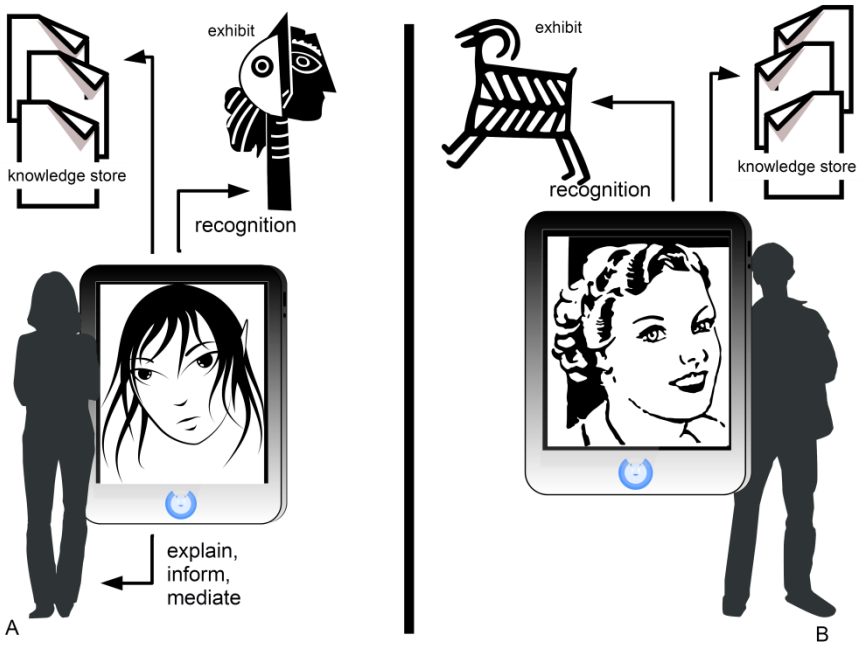

Fig. 1. Exploration phase: Agents on mobile devices

The system's main functionality is localised at the wall mounted installation. It serves in analogy to information plaques, displaying personalized multimedia content. Designed to accommodate two visitors at the same time, its screen setup realizes a split-screen configuration. When users approach the display, respective individual agents migrate from the tablet into the stationary screen space, taking up position at the left and right periphery of the screen.

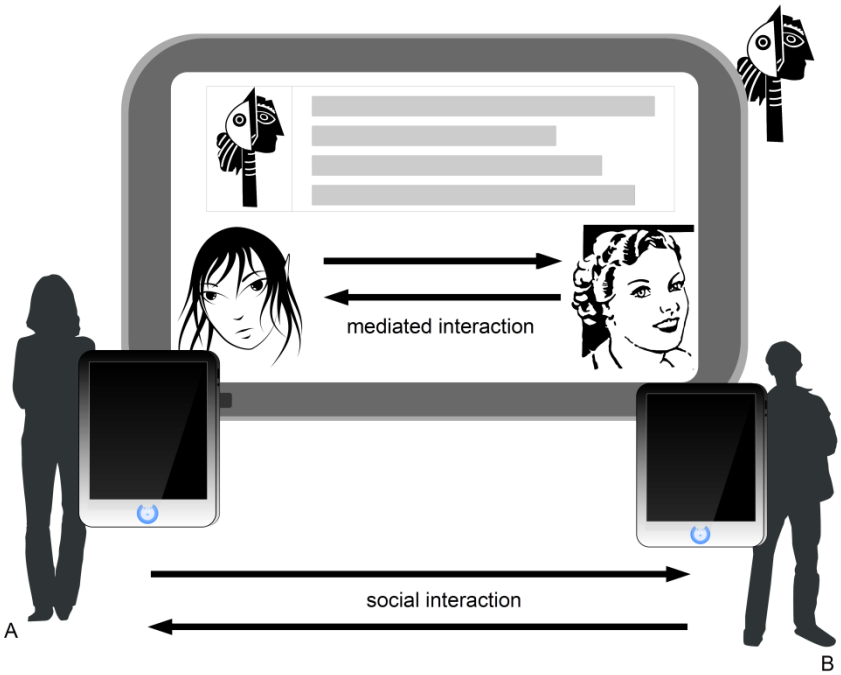

Fig. 2. Wall-mounted display: Agents as facilitators of social interaction.

Whenever two visitors use the station concurrently, this marks the critical part of system operation. User monitoring is employed in order to assess if users are orientated towards each other communicatively or not. Should the system infer communicative interest, an attempt is made to connect both visitors by supplying a communicative incentive. This is provided as follows: Individual agents situated within screen space leave their position at the periphery of the screen and meet at the lower centre. Here, they engage in pseudo-social interaction with one another. Hereby an element of surprise is provided, acting as a helpful catalyst for interaction in exhibition spaces [14]. The intended effect is for the agents' owners to react to the surprising behaviour of their "virtual pets", ideally by engaging in direct discourse with one another. The situation is constructed in analogy to phenomena such as dog-owners striking up conversations posterior to a meeting of the animals they were walking.

A crucial part of system operation lies in judging if respective users are communicatively inclined during the critical phase. To this end, we intend to employ Hidden Markov Models trained with manually annotated data sets. Among the markers to be analysed are eye-movement behaviour, body posture as well as complementing proxemics features [19].

2 Museum Scenario

The depicted scenario consists of a two level system. One is installed on a device which is handed out to the visitors at the beginning of their tour while the other, the main program, is located on a server which is administrating the informational database and which is able to initiate crossovers between the interests of individual visitors. In addition, the exhibits have explanatory screens which present additional information like the basic description, usage, relevance up to videos of seeing the object in context of, for example, everyday life or whatever purpose.

Following the extensive examples of Lieberman and Selker [16] regarding an agent's depiction and usage as a helpful tool inside a virtual environment, the primary directionality of the aspired social catalyst would be that of an 'advisor' instead of being an 'assistant'.

Although in later steps it might be necessary to not keep this explicit distinction, we employ it here for the sake of conceptual clarity. Once stepping into the museum the aforementioned tablets are handed out together with the admittance ticket. The tablet would ideally be a small one in the range of current 7" display sizes in order to be able to keep it in one hand or to easily store it in a pocket.

The screen would be populated by applications like an in-door positioning system, providing for an accurate 'you-are-here'-button at all times. Additionally, there would be different routes presented, available by pushing a button on the side of the screen.

This would allow to find:

- the nearest exit and other points of interest (coffee spots, sanitary installations, phone spots, souvenirs etc.)
- an information officer, a real human 'agent' to talk to and help in case of any problems with the device or the exhibition
- a personalised route through the exhibition, perhaps even planned ahead from home

In addition to the device being able to plot routes, the system would be represented by an embodied agent being able to react and offer conversational topics regarding the museum and presented objects. A conversational database in the background would continually track the user's interaction with the agent and compare it to other visitor's inquiries. Due to this, the administrating program can check for similar requests to the database and proximity of visitors based on their location inside the museum. It could

provide access to personal information about the visitors via their social network connections.

Based on those two cornerstones of information, the system would engage visitors in a conversation by pre-structuring conversations towards similar interests. Once two devices and their associated visitors converge to a distance which would allow for a regular volume speaking voice, the agents initiate their social catalyst routine. This would happen in three steps.

1. The agents individually acknowledge each other and their respective individual or group to the person using the device. This happens by virtue of a visible turn of the agent towards the other one.
2. A user then has the option to either confirm the upcoming interaction or deny it, resulting in a courteous discontinuation of the initiated process. As soon as one party denies, the respective agents would suggest politely continuing the tour at another point of the exhibition – further away from the other group.
3. If the interaction is confirmed by both parties, the agents visually leave the tablet space and appear on the screen in front of the exhibition.

At first the agents start to interact with each other, which enables them to get the respective parties up to speed about their individual pathways through the exhibition. Afterwards, the significance of the current exhibit would be explained in the context of the whole exhibition and conversational pointers engage the humans in front of the screen to interact with each other.

If the catalyst worked and the visitors continue their journey through the museum together, the agents continue to provide conversational incentives by engaging the humans and each other to keep the discussion going.

If it did not work as intended, then the agents return to their previous state as an informational advisor about the exhibit.

3 Research Directions

The aforementioned scenario and requirements provide numerous research opportunities. A conversational agent is already very well researched as well as regarding agents working in groups as companions [12]. But due to the necessity of implementing new ways of interconnected databases and agent's behaviour, the need for an interdisciplinary approach is obvious. Social sciences for analyzing and categorizing human conversational behaviour and information sciences for implementing the software infrastructure for the agent's behaviour and the administrating instance governing the database.

The goal has to be the implementation of a pro-active conversational agent which is capable of gathering various forms of input. As mentioned before, numerical statistics are able to elicit certain behaviour, like in the case of proximity to another agent system. Environmental information like shambles in the vicinity or auditory superimposi-

tions which would hinder a conversation, can be used to either get away from such incidents or avoid plotting through such areas beforehand. These behaviours would be in accordance with the postulations of Lieberman and Selker [17] as they urge to enable computer systems to be able to grasp the context of a situation.

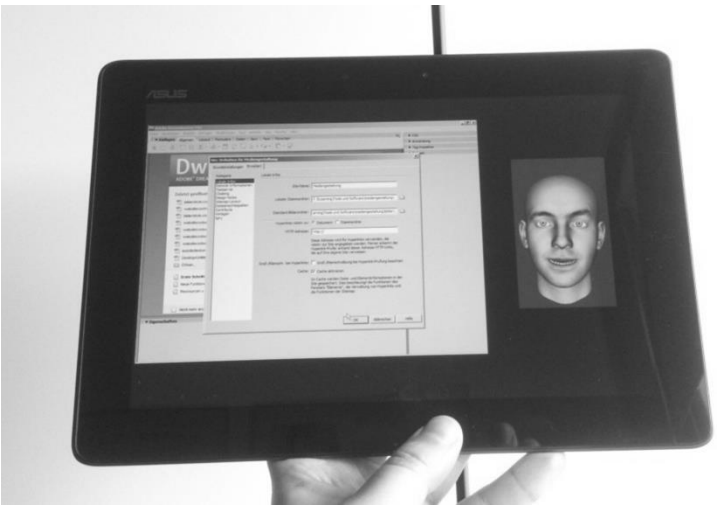

Fig. 3. Tablet with webcam and microphone and pedagogical agent capable of pro-actively reacting to environmental disturbances while transferring knowledge

Other pro-active components are facial interpretations by camera systems, finger-tip temperature through sensors on the surface of the tablet, gait information, body posture as well as gaze and eye tracking. These person centred information can be used to indicate a user's current emotional, vigilance and inquisitiveness state. The cognition of emotional states via facial action recognition, as shown by Kapoor, Qi and Picard [11], provide a reliable prognosis of human reactions to certain events. While Breazeal [5] developed a humanoid robot's emotional model which is able to register affective intents based on a user's voice.

Regarding the interaction between two users, the system should be able to 'read' users reactions to the initiation process. Even for humans this is not an easy task since non-verbal cues are often polysemantic. To adequately register the emotion, context is once more of relevance to the process. As stated by Olsson and Ochsner [21] the prior experiences with the persons become relevant which ideally have been tracked by the agent system along the way up to the point of becoming acquainted with the other visitor.

Regarding the depiction of the agent, empirical research postulates an agent to be able to act socially intelligent. Meaning it knows about cultural peculiarities and possesses the ability to detect and act on them. It has to be perceived as being polite [15, 22, 25] which also extends to the choice of clothing and grooming. The user should be offered a choice of agent representations since learning from a representative of one's

own peer group seems to be beneficial [12, 13, 20]. If the agent is equipped with a voice, then the choice of words and tone of the voice should be polite as well but also it must not be identifiable as a text-to-speech software. Although immense progress has been visible over the last decade, it still is not comparable to a human voice, which might even be more important than appearances [23, 24].

Facial animations of the agent seem to be an issue as well. Static agents still transfer information but some studies [1, 4, 18] postulate a positive effect on motivation, retention and transference of learning material. Gestures and body postures of an agent however apparently do not have a positive effect. Once added to an already facially acting agent, the gestures either showed no [6] or even hindering effects [3].

4 Experiential Structure

Design efforts are guided by a three-partite construal [9] of the museum experience:

Prior to their visit, users become aware of the museum and the possibilities contained within it. During this phase, users utilise information offered in order to decide which institution to visit.

During their visit, users interact with installations, and with each other. After the visit, users possibly relive experiences made and reflect on knowledge gained.

Following this structure, interactive artefacts as well are grouped into three interaction ecologies [8]:

- A web and app-based ecology, allowing for information about the museum to be gathered.
- A spatially structured ecology within the museum, allowing for incentives to be generated in the context of interactive installations and mobile devices.
- A web and app-based ecology, allowing for additional information on exhibits to be obtained and for furthering of social contacts made.

Likewise, distinct content presentation strategies are adopted in order to address different requirements during the phases. I.e. consumption of time-based media potentially creates problems within a museum setting, running the danger of distracting visitors from the experientially rich environment around them.

However, watching videos or listening to historical recordings can be a useful activity during a train ride antecedent to the actual visit. They refer to experiences already made while prolonging the possibility to exist within the historical space encountered.

Agents provide an experiential tie between all three phases. They can be gently introduced in the first phase, provide helpful incentives during the second and act as gentle reminders in the last one.

5 Conclusion

We have argued for a shift in focus from presentation to functionality within interactive agent research. Numerous scenarios exist where proactively behaving agents could be beneficial. We detailed one such scenario within the domain of interactive installations in museums.

The discussion points to a broader issue. Refocusing agent research onto the level of functionality forces us to reopen the design space. Many of the tacit assumptions present within existing discourse surrounding pedagogical agents have to be re-examined. This will provide both for new possibilities while creating new challenges for the agent research community.

References

1. Agada, R., Yan, J.: Research to Improve Communication by Animated Pedagogical Agents. J. Gener. Inf. Technol. 3, 1, (2012).
2. Baylor, A.L. et al.: Interface Agents As Social Models: The Impact of Appearance on Females' Attitude Toward Engineering. CHI '06 Extended Abstracts on Human Factors in Computing Systems. pp. 526–531 ACM, New York, NY, USA (2006).
3. Baylor, A.L., Kim, S.: Designing nonverbal communication for pedagogical agents: When less is more. Comput. Hum. Behav. 25, 2, 450–457 (2009).
4. Baylor, A.L., Ryu, J.: The Effects of Image and Animation in Enhancing Pedagogical Agent Persona. J. Educ. Comput. Res. 28, 4, 373–394 (2003).
5. Breazeal, C.: Emotion and sociable humanoid robots. Int. J. Hum.-Comput. Stud. 59, 1–2, 119–155 (2003).
6. Craig, S.D. et al.: Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features and redundancy. J. Educ. Psychol. 94, 2, 428–434 (2002).
7. Heath, C. et al.: Crafting participation: designing ecologies, configuring experience. Vis. Commun. 1, 1, 9–33 (2002).
8. Heidt, M. et al.: Diverse Ecologies – Interdisciplinary Development for Cultural Education. In: Kotzé, P. et al. (eds.) Human-Computer Interaction – INTERACT 2013. pp. 539–546 Springer Berlin Heidelberg (2013).
9. Heidt, M.: Examining Interdisciplinary Prototyping in the Context of Cultural Communication. In: Marcus, A. (ed.) Design, User Experience, and Usability. Health, Learning, Playing, Cultural, and Cross-Cultural User Experience. pp. 54–61 Springer Berlin Heidelberg (2013).
10. Huang, H.-H. et al.: Toward a multi-culture adaptive virtual tour guide agent with a modular approach. AI Soc. 24, 3, 225–235 (2009).
11. Kapoor, A. et al.: Fully automatic upper facial action recognition. IEEE International Workshop on Analysis and Modeling of Faces and Gestures, 2003. AMFG 2003. pp. 195–202 (2003).
12. Kim, Y. et al.: Pedagogical agents as learning companions: the impact of agent emotion and gender. J. Comput. Assist. Learn. 23, 3, 220–234 (2007).

13. Kim, Y., Wei, Q.: The impact of learner attributes and learner choice in an agent-based environment. Comput. Educ. 56, 2, 505–514 (2011).
14. Vom Lehn, D. et al.: Exhibiting Interaction: Conduct and Collaboration in Museums and Galleries. Symb. Interact. 24, 2, 189–216 (2001).
15. Lepper, M.R. et al.: Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. Comput. Cogn. Tools. 1993, 75–105 (1993).
16. Lieberman, H., Selker, T.: Agents for the user interface. Handb. Agent Technol. 1–21 (2003).
17. Lieberman, H., Selker, T.: Out of context: Computer systems that adapt to, and learn from, context. IBM Syst. J. 39, 3.4, 617–632 (2000).
18. Lusk, M.M., Atkinson, R.K.: Animated pedagogical agents: does their degree of embodiment impact learning from static or animated worked examples? Appl. Cogn. Psychol. 21, 6, 747–764 (2007).
19. Mead, R. et al.: Recognition of spatial dynamics for predicting social interaction. Proceedings of the 6th international conference on Human-robot interaction. pp. 201–202 ACM, New York, NY, USA (2011).
20. Moreno, R., Flowerday, T.: Students' choice of animated pedagogical agents in science learning: A test of the similarity-attraction hypothesis on gender and ethnicity. Contemp. Educ. Psychol. 31, 2, 186–207 (2006).
21. Olsson, A., Ochsner, K.N.: The role of social cognition in emotion. Trends Cogn. Sci. 12, 2, 65–71 (2008).
22. Reeves, B., Nass, C.: The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. Cambridge University Press, New York, NY, USA (1998).
23. Veletsianos, G.: How do learners respond to pedagogical agents that deliver social-oriented non-task messages? Impact on student learning, perceptions, and experiences. Comput. Hum. Behav. 28, 1, 275–283 (2012).
24. Veletsianos, G.: The impact and implications of virtual character expressiveness on learning and agent–learner interactions. J. Comput. Assist. Learn. 25, 4, 345–357 (2009).
25. Wang, N. et al.: The politeness effect: Pedagogical agents and learning outcomes. Int. J. Hum.-Comput. Stud. 66, 2, 98–112 (2008).
26. Wuttke, M.: Pro-Active Pedagogical Agents. Int. Summerworkshop Comput. Sci. 2013. 17, 59 (2013).