Socially Cooperative Behavior for Artificial Companions for Elderly and Cognitively Impaired People

Ramin Yaghoubzadeh, Hendrik Buschmeier, and Stefan Kopp

Social Cognitive Systems Group – CITEC and Faculty of Technology,
Bielefeld University, Bielefeld, Germany
{ryaghoubzadeh, hbuschme, skopp}@uni-bielefeld.de

Abstract. We present work towards a spoken-dialogue embodied companion system that behaves in a socially cooperative fashion, i.e., being attentive to communicative and social cues emitted by the user, inferring internal cognitive, affective and interactional states, and reciprocating with appropriately adapted behavior. It is argued that this capability is vital for fostering cooperative, effortless, and enjoyable interactions that maximize both efficacy and long-term acceptance with elderly or cognitively impaired users.

Keywords: artificial companion, embodied conversational agent, adaptive spoken dialogue, assistive systems, elderly users, users with cognitive impairments

1 Introduction

A growing number of elderly and cognitively impaired people is dependant on assistance for everyday activities. Some of these activities can be supported technologically to enable a prolonged self-determined lifestyle. While many technical services like calendar management, online video communication, or home automation, can potentially provide this valuable support, the interaction continues to pose challenges for many of the respective user groups. For example, elderly people often lack experience in using graphical user interfaces and may face problems learning them due to attitudinal (lack of interest or confidence), perceptual (decline of vision or hearing), cognitive (decrease in memory and attention), or physical barriers (motor impairments) [7,5]. People with cognitive impairments are often illiterate, and text-based interfaces pose almost insurmountable obstacles to them. They, as well as people with age-induced impairments, are quickly overburdened by the amount of functions or the complexity of interfaces that are not especially designed with these requirements in mind. Generally, spoken interaction is reported as a preferred interaction modality by older adults with little prior experience in new technologies [11]. Moreover, with regard to the practical use of assistive systems, there is still little work regarding concrete factors for actual long-term acceptance and motivation to operate those systems daily in one's own home environment, i.e., those that are not limited to (repeated) isolated sessions or controlled environments. It has been shown that variability in output increased user engagement in daily interactions with a support system over the course of one month [1]. In everyday settings and surroundings, designing assistive systems as *companions* can be a linchpin of successful interaction and assistance.

2 What should a companion be like?

Companion systems are envisioned not to be mere tools, but to cater to individual requirements and be aware of the current state of the user [2]; this includes using sensor technology to infer internal states such as their mood. Companions should evoke in the user a notion of increasingly seeing their preferences and idiosyncrasies reflected during assisted tasks, and a trust in the system solving tasks on their behalf, as opposed to purely following instructions. Additionally, research has shown that people expect the fulfillment of perceived social contingencies that arise in such assistive situations; the simplest cases for spoken interaction being a name that the system reacts to, and its capability to react to expressions of gratitude [12]. This is true even when no proper embodiment is present in the setup – an assistive system itself apparently has a potential for social affordance that goes beyond that of other items, although its exact strength and nature is not yet well-researched.

We argue that companions ought to go further to assume an enduring social presence. Unfolding social relationships in human dyads are not only characterized by the refinement of the mutual awareness of the individual requirements, motivations, and limitations, but also, centrally and from the very start, by behavioral synchronizations and adaptations that extend to the level of movement, posture, or facial cues [8]. These can even be employed strategically by human experts to foster the sense of rapport, which in turn facilitates all further interaction and can help to reduce social distance [10]. Crucially, to engage in these reciprocal coordination mechanisms, a system need not only process social signals, but be able to emit them as well. Thus, we employ a virtual assistant with a human-like appearance, which enables us to leverage anthropomorphic cues and foster natural multimodal communication [9].

3 The importance of socially cooperative behavior

In previous work, which has been conducted in the BMBF-funded projects 'VASA' and 'VERSTANDEN', we found that elderly and cognitively impaired users reacted positively to a virtual assistant and were readily able and willing to engage in spoken-language interaction with the assistant [14] (Fig. 1, left). A short-term social effect of the agent was evident through users' spontaneous social smiles, honoring perceived social contingencies that were not explicitly elicited (such as spontaneous apologies uttered by users), as well as cursory social comments which they interwove into the task-related dialogue (appointment scheduling), albeit not hindering it. Anecdotally, long-term social effects were reflected in participants referring back to their last 'hands-on' interactions with the agent, their eagerness for future interactions, and their reports of exchanges about the agent and their presenting of keepsakes with the agent's image.

Regarding the efficacy of spoken interaction in a task-related domain, we found in initial Wizard-of-Oz (WOz) studies that special care must be taken in information presentation and confirmation strategies to avoid the propagation of, and inadvertent commitment to, wrong information resulting from mis-communication [14]. We not only analyzed interactions with older adults, but extended this to people with congenital cognitive impairments (with an IQ lower than about 70). For both groups, and most pronouncedly for the latter, information presentation and confirmation strategies were least error-prone when delivered in the smallest possible chunks (for the control group, packaging thrice the information per unit did not degrade accuracy). Additionally, their usability ratings were equivalent between verbose and terse presentations, respectively.

These results informed the development of an incremental dialogue management framework and an autonomous spoken dialogue system in the same domain, which has been evaluated with initial participants (Fig. 1, right), who did not rate the system less favorably than the WOz prototype [15]. The autonomous system successfully induced error-awareness in the user groups, matching the performance of the previous system in this respect. For most participants the system performed adequately; yet a few subjects uttered verbose elaborations inside the task, which had the strongest detrimental effect on the efficacy of spoken interaction. The agent's capabilities to emit natural and socially acceptable signals for dialogue flow control to mitigate this effect are the subject of ongoing research.

Fig. 1. Interaction studies. **Left:** WOz setup, participant with cognitive impairments (anonymized). Staff was present for reasons of safety. **Right:** Autonomous system, older participant (anonymized).

4 The KOMPASS project

The aforementioned work, has led to the current project 'KOMPASS' ("Sozial kooperative virtuelle Assistenten als Tagesbegleiter für Menschen mit Unterstützungsbedarf")[1] with partners from computer science, linguistics and psychology. Our special focus in this project, which started in April 2015, is to ensure assistance and acceptance by enabling what we call 'socially cooperative behavior' in an artificial companion. By this we mean the approach of consequently treating communication and human–companion interaction as an instance of social cooperation and collaboration [6], in which the system is highly sensitive and responsive to the states of the interaction and the user. The companion is to be able to recognize and interpret multimodal cues on short and longer timescales, the former to ensure robustness of communication and be aware of affective reactions to communicated content, and the latter to assess mood and the progression

[1] For an organizational overview, please refer to the project page: https://purl.org/net/kompass

of the interaction experience. More specifically, short-term assessment will enable the system to react rapidly in an appropriate way, e.g., to adapt or elaborate its presentations, initiate repairs and confirmations, and reciprocate following social affordances. Long-term assessment will steer strategy selection; both can serve as supervisory signals for long-term preference learning.

In addition to the dialogue management infrastructure mentioned above, the project will draw on previous work on modeling a user's mental state during listening (e.g., does she perceive and understand what is said?) as revealed by verbal/vocal and non-verbal communicative feedback. The agent maintains and updates a probabilistic representation – an 'attributed listener state' [4] – which provides a basis for adapting and (re-)planning its communicative behaviour based on its estimate of the user's levels of perception, understanding, etc. Among other things, the agent can almost instantly vary the information density of ongoing utterances using an incremental natural language generator [3] or elicit feedback from its user if it is uncertain about her mental state of listening [4].

The KOMPASS companion is projected to both process and employ effective subsets of all those social signals that make human face-to-face interactions effortless and enjoyable. Centrally, this excludes the companion actively suggesting a personal relationship by, e.g., telling (faux) back stories – this is due both to our focus on the interaction as well as principal ethical objections by the project partners, including a large health-care provider. Thus, the companion is envisioned to be a pleasant, cooperative – and when needed proactive – helper, referring back to past common episodes when required, yet being reserved with respect to interaction outside a specific set of tasks, and especially with respect to mutual self-disclosure [13].

Acknowledgements This research was partially supported by the German Federal Ministry of Education and Research (BMBF) in the project 'KOMPASS' (FKZ 16SV7271K) and by the Deutsche Forschungsgemeinschaft (DFG) in the Cluster of Excellence 'Cognitive Interaction Technology' (CITEC).

References

1. T. Bickmore, D. Schulman, and L. Yin. Maintaining engagement in long-term interventions with relational agents. *Applied Artificial Intelligence*, 24:648–666, 2010.
2. S. Biundo and A. Wendemuth. Von kognitiven technischen Systemen zu Companion-Systemen. *Künstliche Intelligenz*, 24:335–339, 2010.
3. H. Buschmeier, T. Baumann, B. Dosch, S. Kopp, and D. Schlangen. Combining incremental language generation and incremental speech synthesis for adaptive information presentation. In *Proceedings of the 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 295–303, Seoul, South Korea, 2012.
4. H. Buschmeier and S. Kopp. A dynamic minimal model of the listener for feedback-based dialogue coordination. In *SemDial 2014: Proceedings of the 18th Workshop on the Semantics and Pragmatics of Dialogue*, pages 17–25, Edinburgh, UK, 2014.
5. N. Charness and W. R. Boot. Aging and information technology use: Potential and barriers. *Current Directions in Psychological Science*, 18:253–258, 2009.
6. H. H. Clark. *Using Language*. Cambridge University Press, Cambridge, UK, 1996.

7. R. D. Ellis and J. C. Allaire. Modeling computer interest in older adults: The role of age, education, computer knowledge, and computer anxiety. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 41:345–355, 1999.
8. S. Kopp. Social resonance and embodied coordination in face-to-face conversation with artificial interlocutors. *Speech Communication*, 52:587–597, 2010.
9. N. C. Krämer. Social communicative effects of a virtual program guide. In *Proceedings of the 5th International Working Conference on Intelligent Virtual Agents*, pages 442–453. Kos, Greece, 2005.
10. J. L. Lakin and T. L. Chartrand. Using nonconscious behavioral mimicry to create affiliation and rapport. *Psychological Science*, 14:334–339, 2003.
11. P. Langdon, P. Biswas, T. Angélique, M. Gonzales, C. Jung, G. Heinrich, and L. Almeida. User Interaction & Application Requirements. Technical Report Deliverable D2.1, GUIDE Consortium, Darmstadt, Germany, 2011.
12. M. Meis. Nutzerzentrierte Entwicklung eines Erinnerungsassistenten. Abschlusssymposium Niedersächsischer Forschungsverbund Gestaltung altersgerechter Lebenswelten, 2013.
13. Y. Wilks. On being a Victorian Companion. In Y. Wilks, editor, *Close Engagements with Artificial Companions. Key social, psychological, ethical and design issues*, pages 121–128. John Benjamins, Amsterdam, The Netherlands, 2010.
14. R. Yaghoubzadeh, M. Kramer, K. Pitsch, and S. Kopp. Virtual agents as daily assistants for elderly or cognitively impaired people. Proceedings of the 13th International Conference on Intelligent Virtual Agents, pages 79–91, Edinburgh, UK, 2013.
15. R. Yaghoubzadeh, K. Pitsch, and S. Kopp. Adaptive grounding and dialogue management for autonomous conversational assistants for elderly users. In *Proceedings of the 15th International Conference on Intelligent Virtual Agents*, pages 28–38, Delft, The Netherlands, 2015.