International Workshop on Techniques Toward Companion Technologies (T2CT), Edinburgh, UK, August 2013.

An Always-On Companion for Isolated Older Adults

Candace Sidner[1], Timothy Bickmore[2], Charles Rich[1], Barbara Barry[2], Lazlo Ring[2], Morteza Behrooz[1], and Mohammad Shayganfar[1]

[1] Worcester Polytechnic Institute, Worcester, MA, USA
[2] Northeastern University, Boston, MA, USA
sidner@wpi.edu bickmore@ccs.neu.edu

Abstract. We summarize the status of an ongoing project to develop and evaluate a companion for isolated older adults. Four key scientific issues in the project are: embodiment, interaction paradigm, engagement and relationship. The system architecture is extensible and handles real-time behaviors. The system supports multiple activities, including discussing the weather, playing cards, telling stories, exercise coaching and video conferencing.

1 Introduction

The Always-On project[3] is a four-year effort, currently in its third year, supported by the U.S. National Science Foundation at Worcester Polytechnic Institute and Northeastern University. The goal of the project is to create a relational agent that will provide social support to reduce the isolation of healthy, but isolated older adults. The agent is "always on," which is to say that it is continuously available and aware (using a camera and infrared motion sensor) when the user is in its presence and can initiate interaction with the user, rather than, for example requiring the user login to begin interaction. The agent will help reduce the user's isolation not just by always being around but also by specific activities that connect the user with friends, family and the local community. Our goal is for the agent to be a natural, human-like presence that "resides" in the user's apartment for an extended period of time. Beginning in the fall of 2013, we will be placing our agents with about a dozen users for a month-long evaluation study.

Our project focuses on four key scientific issues:

- the embodiment of the agent,
- the interaction paradigm,
- the engagement between the user and the agent, and
- the nature of the social relationship between the user and the agent.

Embodiment. We are experimenting with two forms of agent embodiment. Our main study will employ the virtual agent Karen, shown in Fig. 1, that comes from the work of Bickmore et al. [1]. Karen is a human-like agent animated from

[3] `http://www.cs.wpi.edu/~rich/always`

Fig. 1. Virtual agent interface — "Karen"

a cartoon-shaded 3D model. She is shown in Fig. 1 playing a social game of cards with user. Notice that user input is via a touch-screen menu. Also, the speech bubble does not appear in the actual interface, which uses text-to-speech generation.

We are also planning an exploratory study substituting the Reeti[4] robot, shown in Fig. 2, for Karen, but otherwise keeping the rest of the system (i.e., the menus, text-to-speech and other screen graphics) as much the same as possible. One big difference we expect is that the effect of face tracking with the robotic agent will be much stronger than with Karen. On the other hand, because Reeti is not as human-like as Karen, it is possible that it will not be as well accepted overall as Karen.

By comparing people's responses to these two embodiments in situ, we hope to learn more about the nature of people's behavior with virtual versus robotic agents when they interact for an extended period of time. Due to the longevity of the interaction as well as the richness of activities (some purely conversational and some task-based), we hope that that users in both conditions will have extensive interactions, find those interactions satisfying, and develop a strong working relationship with their companion, in whichever embodiment. Part of our month long study will be to assess these experiences and determine if differences do occur.

Interaction Paradigm. The main interaction paradigm in our system is conversation, and in particular, dialog. The agent makes its contributions to the dialog using speech, and the user chooses his/her contribution from a menu of

[4] http://www.reeti.fr

utterances provided on the touch screen. Dialogs evolve around various activities and can extend for quite a long time (up to five or ten minutes) if the user chooses to continue the conversation. Some of the activities between involve additional on-screen graphics, such as the card game shown in Fig. 1, or a Week-At-A-Glance™ style planning calendar. When playing cards together, the user is allowed to directly manipulate the cards on-screen. However, we have otherwise eschewed other traditional GUI methods using icons, pull-down lists, etc., in favor of using speech and menu dialog interaction whenever possible. (Another exception, like direct manipulation of cards on-screen, is a virtual keyboard to allow typing in of proper names of people and places.) Our motivation for this design choice is to reinforce the relationship between the user and the agent.

We do support one activity in which the user is allowed to speak (rather than just choose from a menu). In our story acquisition activity, the user tells a personal story, which is audio and video recorded by the agent for later sharing with the user's family and friends via Vimeo™. During the recording, although the agent does not understand what is being spoken, it encourages the user by providing appropriate nodding back-channel behavior tied to pauses in the audio signal [9]. When the recording is completed, the agent has a menu-based dialog with the user regarding the content and disposition of the recording.

Fig. 2. Robotic interface — "Reeti"

Engagement. Our system continuously maintains a model of the state of engagement[10] between the user and the agent. For example, when the agent senses nearby motion (via infrared) followed by the appearance of a face in its vision system, it decides that the user is initiating engagement. Disengagement can come about at the natural conclusion of the conversation or when the user leaves for an unexpected reason, e.g., to answer a ringing door bell. Because our agent cannot understand sounds in the environment, it may not know why the user has disengaged, but it does have simple strategies for dealing with unexpected interruptions. Generally, the agent does not initiate disengagement, although it may attempt to hurry the conclusion of a session if some event in the user's calendar is about to start.

Relationship. Since the user and agent have conversations over an extended period of time, it is natural to consider that they have some kind of social relationship [2, 5]. To reason about this relationship, we have implemented a planning system [4] that decides which activities are appropriate to suggest to the user each time they interact (in what we call a *session*). This planning system uses a relationship model based on the *closeness* between the agent and user. Their closeness increases as they do activities together. Closeness decreases when the user and agent do not interact for a period of time, such as a few days.

Each available activity has a *required* closeness in order to be undertaken. Only those activities whose required closeness is less than or equal to the current closeness between the user and agent will be suggested for the current session. Activities that, although suggested, do not actually occur (due to user choice or other reasons) are reported to the relationship planning system for planning the next session.

2 User Studies

To better understand what users want to talk about, the effect of being always on, and to evaluate some candidate activities, the Northeastern team has conducted a series of preliminary studies. These studies include a Wizard of Oz (WOZ) study in which several users interacted with a virtual agent controlled remotely by an experimenter, and an early prototype of a fully autonomous virtual agent. Both studies involved having the agent in the homes of older adults for a week. In the WOZ study [11] we discovered that more than half of the participants wanted to talk about the weather, family and personal stories (telling stories to the agent). Additionally, half of all the conversations involved discussing the participant's future plans and activities.

The preliminary study with a fully autonomous agent [8] used a motion sensor, but no camera, to detect the user's presence. The agent could thus initiate interactions. Users could converse about the weather and exercise, and could hear humorous anecdotes told by the agent. In these studies, users indicated high levels of acceptance of the agent. Participants did not find the agent intrusive, and almost all participants had short conversations with the agent every day.

In our upcoming field studies, in addition to the motion sensor, the computer's built-in web camera will be used for face detection and tracking. However, no data from the camera will be permanently stored. Other uses of the web camera include an activity called story acquisition, in which the agent invites the user to record a personal story (which can then be shared with friends and family via Vimeo™), and Skype™ (in which live video and audio data is streamed, but not recorded). Other than for Skype™, our system does not use audio input (i.e., no speech recognition).

The participant consent protocol for our studies is being approved by the Institutional Review Boards at both Worcester Polytechnic Institute and Northeastern University. During the consent process, potential study participants are apprised of all aspects of video and audio use. Potential participants are also screened for emotional and psychological well-being.

3 System Architecture

Fig. 3 shows the high-level architecture of our system, which addresses two main challenges. The first challenge is modularity and extensibility with respect to activities. To address this challenge, we developed a plugin approach, which

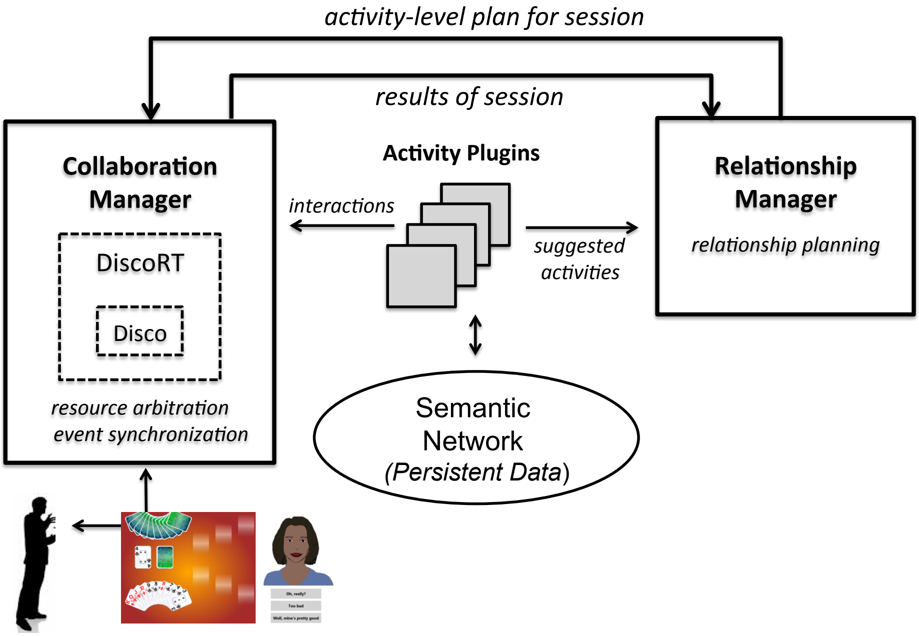

Fig. 3. Always-On system architecture

has allowed different members of the research team to separately develop new activities for the agent.

The second main challenge is the need to operate at multiple time scales[6] from hard real-time (milliseconds) to long-term (days and weeks). At the real-time end of this spectrum, our solution is an extension, called DiscoRT (Disco for Real-Time), to our collaborative dialogue system Disco [7]. DiscoRT implements an arbitration-based parallel schema architecture that handles such phenomena as barge-in (the agent immediately stops speaking in the middle of an utterance when the user touches a menu item on the screen) and time-outs (the agent repeats an utterance when the user doesn't respond after some time). DiscoRT also coordinates the inputs from the agent's sensors and helps manage Disco's dialog focus. At the other end of the spectrum, the relationship manager handles the per-sesson planning, as discussed above.

The system is implemented in a combination of Java and .NET on Windows™.

4 Activities for User and Agent

One take-away of our preliminary studies is that users would like to have more activities to do with the agent. For example, they wanted to be able to tell the agent about their friends and family.

Different activities can serve different goals, either improving the user's well-being by reducing social isolation (which is the system's main goal), or just doing something the user wants to do. Activities such as talking about the weather or playing a social game of cards are meant as ice breakers, to give the user and agent a way to develop relational closeness. Other activities are more instrumental, such as the agent's self-introduction dialogs, in which it interactively explains its capabilities to the user, or the user's enrollment dialogs, in which he/she provides information about family members, such as their relationship, geographic location, birthday, etc. Once the agent knows about the user's family and friends, it can support the user's desire to talk about them, albeit in a very limited fashion. To provide some light-hearted enjoyment, the agent can tell short humorous anecdotes to the user, an activity that users reported liking in preliminary studies.

Several activities directly address the project goal of connecting the user to other people. We have implemented what we call the "Skype™ buddy," an activity in which the agent arranges video calls for the user with family and friends. The activity to coach outdoor walking, an area well explored by Bickmore et al. in earlier work [1], can result in more social connection by getting the user out of his/her apartment and into the community.

Finally, the story acquisition activity addresses both a project goal and a user desire. Seniors tell stories about their lives in part to make sense of their full life experience [3]. Our system supports sharing these stories, thereby also mitigating social isolation.

Developing an adequate set of activities has been our biggest challenge in the Always-On project, because of the extensive programming needed and complexity of the heterogenous real-time operating environment. We expect that not all users will use every activity. After our main study, we will understand better when and how often each activity is used, and the effects of their their use on the participants' sense of isolation and their development of a working alliance with the agent.

Acknowledgments. This work is supported in part by the National Science Foundation under award IIS-1012083. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

1. Bickmore, T., Caruso, L., Clough-Gorr, K., Heeren, T.: "It's just like you talk to a friend"– Relational agents for older adults. Interacting with Computers **17**(6) (2005) 711–735
2. Bickmore, T., Schulman, D.: Empirical validation of an accomodation theory-based model of user-agent relationship. In: Proc. Int. Conf. on Intelligent Virtual Agents, Santa Cruz, CA (2012)
3. Butler, R.: The life review: An interpretation of reminiscence in the aged. Psychiatry **26** (1963) 65–70
4. Coon, W., Rich, C., Sidner, C.: Activity planning for long-term relationships. In: Proc. Int. Conf. on Intelligent Virtual Agents, Edinburgh, UK (2013)

5. Kidd, C., Breazeal, C.: A robotic weight loss coach. In: Proc. 22nd National Conference on Artificial Intelligence, Vancouver, Canada (2007)
6. Rich, C., Sidner, C., Nooraei, B., Coon, W.: Operating in a hierarchy of time scales for an always-on relational agent. In: Workshop on Real-Time Conversations with Virtual Agents, Santa Cruz, CA (September 2012)
7. Rich, C., Sidner, C.L.: Using collaborative discourse theory to partially automate dialogue tree authoring. In: Proc. Int. Conf. on Intelligent Virtual Agents, Santa Cruz, CA (September 2012)
8. Ring, L., Barry, B., Totzke, K., Bickmore, T.: Addressing loneliness and isolation in older adults. In: Int. Conf. on Affective Computing and Intelligent Interaction, Geneva, Switzerland (2013)
9. Scherer, S., Marsella, S., Stratou, C., Xu, Y., Morbini, F., Egan, A., Rizzo, A., Morency, L.P.: Perception markup language: Towards standardized representation of nonverbal behaviors. In: Proc. Int. Conf. on Intelligent Virtual Agents, Santa Cruz, CA (2012)
10. Sidner, C.L., Lee, C., Kidd, C., Lesh, N., Rich, C.: Explorations in engagement for humans and robots. Artificial Intelligence **166**(1-2) (2005) 104–164
11. Vardoulakis, L., Ring, L., Barry, B., Sidner, C., Bickmore, T.: Designing relational agents as long term social companions for older adults. In: Proc. Int. Conf. on Intelligent Virtual Agents, Santa Cruz, CA (2012)