The Technical Specification and Architecture of a Virtual Support Partner

Sten Hanke[1,2], Emanuel Sandner[1], Andreas Stainer-Hochgatterer[1], Christiana Tsiourti[2] and Andreas Braun[3]

[1] AIT Austrian Institute of Technology GmbH, Vienna, Austria
[2] University of Geneva, Geneva, Switzerland
[3] Fraunhofer IGD, Darmstadt, Germany
sten.hanke@ait.ac.at,
Homepage: http://www.ait.ac.at

Abstract. Most elderly people prefer to live independent in their own homes for as long as possible. Needed support is delivered by someone else and/or via the use of technology. The current paper describes how so called conversational agents can be designed to provide a virtual support and help in daily life activities of the older adults. The paper describes the concept and the idea of an virtual support partner and the concrete realization of a virtual support partner in the EU funded Miraculous-Life project. It describes the deployment setup, the components as well as the architecture and gives some conclusion and lessons learned.

Keywords: avatar, system architectures, service architecture, virtual companion, embodied conversational agents

1 Introduction

Human like computer-animated characters, also known as Embodied Conversational Agents (ECAs) [1], have attracted a lot of attention over the past years in the field of artificial intelligence [2]. They are designed to simulate human face-to face conversation with their users and are typically represented as humans or animals, specifically lifelike and believable in the way they behave.

Cassell [1] defines ECAs as those virtual characters that have the same properties as humans in face-to-face conversation, including:

- The ability to recognize and respond to verbal and non-verbal input.
- The ability to generate verbal and non-verbal output such as mouth movements, eye movements, head movements, hand gestures, facial expressions, and body posture.
- The ability to deal with conversational functions such as turn taking, feedback, and repair mechanisms
- The ability to give signals that indicate the state of the conversation, as well as to contribute new propositions to the discourse.

2 Virtual support for quality of life

Empirical studies [3] reveal that ECAs can improve the natural interaction with elderly users in ambient intelligent environments. Specifically, older adults both with and without cognitive impairment, are capable of recognizing emotions in the facial expressions of an agent and follow instructions much better when interacting with an agent. Another study concluded that embodied agents allow the development of affinitive relationships with their human partners and can therefore help to fulfil the need of affiliation in ambient assisted living environments [4]. A number of recent systems based on ECAs aim to address the needs of older adults, to provide companionship and assist older adults in health related domains (i.e. physical exercise, medication adherence). Researchers have explored the design of ECAs that interact with users over multiple conversations, ranging from a handful of interactions to hundreds of interactions spanning months or years [5], [6] and [7]. Ring et al. [8] developed a conversational agent-based system to provide longitudinal social support to isolated older adults by means of empathic feedback. An exploratory short-term pilot study demonstrated significant reductions in loneliness of older adults based on self-reported affective state. Vardoulakis et al. investigated the use of an agent to provide social support and wellness counselling to older adults. Qualitative analysis of interactions with a remote-controlled agent (i.e., Wizard-of-Oz) installed in homes of older adults, identified multiple topics that users liked discussing and showed high acceptance ratings and a positive attitude towards the agent [9]. Other studies [10] and [11] explored relational agents, ECAs designed to form long-term social-emotional relationships with their users for health education and health behaviour change interventions. Results of a two month trial that investigated exercise promotion showed increased physical activity for participants using a virtual exercise coach compared to those using a conventional pedometer [12]. Nevertheless, this effect diminished when the coach was removed, suggesting that further research is needed to cause long-term behaviour change. In a similar study, Bickmore et al. developed a virtual laboratory to explore the longitudinal usage of a virtual exercise coach [13]. Older adult participants interacted with an agent from their home once a day for up to 120 days. Results showed that users who interacted with an agent that used variable dialogue exercised significantly more than those interacting with an ECA with non-variable dialogue. ECAs move beyond the paradigm of computer as a tool and allow for multimodal interaction reflecting natural human-to-human communication. By exhibiting a certain level of intelligence and autonomy as well as social skills, ECAs provide familiar and non-threatening interfaces, especially useful for building systems that are easy to use, engaging and gain the trust of older adults.

3 The Miraculous-Life project

Modelling and designing a support system for elderly people raises several research questions including: the interaction between the user and the system, the computation and selection of verbal and non-verbal ways and their synchronization and representation resulting to emotional understanding. Important is also how to activate unambiguously and unobtrusively the appropriate support services, providing care and safety at home, without intruding into the users daily lifestyle. The objective is to simulate the communication of the elder with a human partner. Current solutions based on ECAs, however, focus only on a subset of communicative functions, modalities, and generation capacities. In the ICT based elderly care sector existing ECAs based initiatives are lacking mainly the personal, empathic and elderly behaviour understanding. They are not focused on providing a virtual character model able to engage and motivate the elder in a similar way a human carer does. Specifically, in the area of supporting the elderly in carrying out their daily activities such systems are not widespread. There are, however, new potentials for realizing such a novel real-time empathic elder care system considering also the proliferation of inexpensive portable devices, with high powered processors and sensors, such as cameras and depth sensing technology (i.e. Kinect sensor). However, for the development of such a system a multitude of technological areas is needed such as: multimodal sensing and processing; immersive human-computer interfaces including animated characters capable of capturing users attention and engaging users in active tasks; environmental context extraction and event interpretation; rich descriptions of human activities and emotional state and generation of behaviour models; ICT based services for elder home daily activities support and safety, social networks involving formal and informal carers facilitating elder daily care support. By employing a synergetic approach combining state-of the art research in areas mentioned above the Miraculous-Life project is contributing in advancing elder care support by modelling and developing a system playing the role of a Virtual Support Partner (VSP), that by analogy to a real life human partner, attends to the home daily activity and safety needs of the elder, while the elder goes about his normal daily life, and provides implicit support. The VSP is characterized by emotionally-enriched human computer interaction (recognition of users emotional state and expression of appropriate empathetic response) combined with behaviour analysis and understanding of the elder over the ageing process. Specifically, the VSP fuses together facial expressions, voice intonations, gestures and other contextual information of the users environment and provides intelligent responses and delivery of services through an Avatar-based interface exhibiting empathic response through facial emotions and associated voice intonations. The VSP is also simulating behaviour and dialogue patterns as established between the elder and a human partner over the life time and will play the role of an advisor and assistant motivating and stimulating the elder to remain longer active at home in carrying out daily life activities.

The validation of the system will be realized in two well selected use cases in two different countries. Up to 50 elderly people and their caregivers will use the system over a six month period. The system will be delivered to the user in form of a stand-alone consumer product, operating on a scalable distributed network of interconnected PCs, tablets and Kinect devices, focusing on minimum essential personalized elders daily activities care support at home. The system will provide benefits on a physical, psychological and social level enabling and motivating the elderly to remain longer active in carrying out their daily life at home, thus prolonging their independence and improving their wellbeing.

4 The Miraculous-Life system components

In this chapter we describe the technologies (hardware and software) employed in the Miraculous-Life project and provide a motivation for their use and an explanation on how they were used. Much care was taken to use technologies that are easy to use, reusable, and where possible are free of charge. As **workstation** we refer to the computer on which the Miraculous-Life core system runs on, that is the dialogue manager as well as the sensor acquisition and processing modules. All interactions and communications, from the sensor input to the tablet output are processed on the workstation. The specifications were chosen by incorporating different aspects like:

- User acceptance (unobtrusive, small, silent)
- Current and possible future system requirements

As operating system for the workstation Windows 8 64Bit was chosen, as this meets the requirements for addressing the Microsoft Kinect V2 sensor as well as the Microsoft speech recognition engine.

As main interaction device, hand-held and/or wall mounted **tablets** are used. On these tablets the Virtual Support Partner (VSP) is displayed next to detailed information from the provided services. The end-users will interact with the Miraculous-Life system via speech and direct user input (touch) if needed. The tablets which are used for the pre-trials are running Android 4.4.

In order for the system to interact with the user, it needs to get information from the user to determine the his or hers current status and to process an appropriate output. Therefore, different types of sensors are being used that collect data in an unobtrusive way. Special care has been taken in order to avoid too much intrusion and/or interference with the end-users daily life activities and their privacy. An important aspect is that the older adults feel not observed or monitored by the system, which would decrease the acceptance of the system and therefore undermines the benefit of using it. Sensors that are used include the Microsoft Kinect V2, tablet webcam, contact- and presence sensors. Contact sensors will be used to check if windows were and are open or closed, while presence sensors are used determine the persons sleeping habits. The contact- and presence sensors are connected through the HOMER platform [14] to the Miraculous-Life System. The gathered data are used for low level safety services and behaviour analysis of the users.

Any kind of microphone, whether it is a built-in, USB or analogous connected, is used for audio data capturing. Miraculous-Life is using this data to enable a speech driven interaction modality for the user. Speech recognition software detects input commands from the user for the system, while an emotion recognition module analysis the speech input and derives the users emotional state based on certain features within that audio sample. Microsoft's speech recognition engine, which is also used with the Kinect sensor, will be used for the speech recognition part.

The emotion recognition from audio data as well as the avatar rendering is performed on an external server, which is hosted by one of our project partners Zoobe.

Modularity has been realized by using an OSGi framework in Java. Java based applications are platform independent, as they are executed by the Java Runtime Environment (JRE), which can be installed on various operating systems. The usage of an OSGi framework provides remote maintenance and individual adaptability of the system. The components, coming in the form of bundles for deployment, can be remotely installed, started, stopped, updated and uninstalled during runtime, without requiring to shutdown or reboot the system. Thus, the framework is flexible in terms of expanding its functionality and updating single modules during runtime. The interactions and dependencies between bundles are handled by the framework itself. It manages searching and binding of required services, which are exposed functionalities within the OSGi environment, even when the service is activated at later time. Fine grained configuration options allow detailed access to functionalities in each OSGi bundle.

Apache Karaf is used as the OSGi framework, which is a lightweight container onto which various components and applications can be deployed. It sits on top of the OSGi container and adds additional functionalities to the OSGi environment via add-on features. The software maintenance and management is handled by Apache Maven. Based on the concept of a project object model (POM), Maven can manage the compilation of source code, the projects building process, reporting and documentation from a central configuration file. The maven-eclipse plug-in facilitates creation of Eclipse projects to import source code packages in Eclipse easily.

For centralized data management within the trial sites at the care organisations, Microsoft's SQL Server 2008 is used. This database stores project relevant data for service usage as well as for user context and behaviour analysis.

The VSP user interface is be provided as an HTML5 web page, that does not limit the usage of the system to proprietary rendering platforms but opens the usability to various devices and browser implementing the HTML5 standard. The initial connection to the web page is established using a HTTP connection, while the dynamic update of the user interface is managed by an WebSocket connection, JSON encoded messages and extensive use of JavaScript.

5 The integration of the Miraculous-Life System

The Miraculous-Life system integration setup is optimized for different requirements of the installation site (residential houses, private living flats etc.). The system can be easily setup and deployed for a various number of clients.

The system uses a central server (Microsoft Windows Server 2008) running in the network of the residential house. It is important to note that this server runs locally at the trial site and provides the OSGi- and web-services, while also hosting the database (also referred to as knowledge base). This setup assures that all private data are stored in a local network environment, so no user data is transferred to a remote (cloud) location. By running the OSGi- and web-services on one local server, they can be used by many clients at the same time. Hence, this server is the place where services can be updated for all clients at once and the performance of the system can be optimized in a central location.

Additionally, an Asterisk Server is used which provides the phone call and SMS functionality. This telephone server can call ordinary phone numbers as well as internal client numbers. The clients are connected to the Asterisk Server using the SIP protocol. The whole calling functionality on the client side is already integrated into the HTML5 interface (the user interface), and is therefore not a native functionality of the tablets operating system. Depending on the performance and outcome of multi-user tests, the Miraculous-Life system will either run with one Asterisk Server accessible by all installations or with locally installed Asterisk Server instances.

Different client flats are connected to the locally available Windows Server. In each of the flats, a small PC (i.e. the workstation) is running, to which the sensors installed in the flats are connected.

The user is accessing the services and the avatar via a tablet. Since the avatar is rendered as an HTML5 interfaces, the choice and the operating system of the tablets is very flexible. In the current deployment we use tablets running Android version 4.4.

In the current setup, the only server outside the local environment and network is the server provided by the partner Zoobe, which performs the avatar generation and rendering as well as the emotion recognition based on speech.

6 Conclusion

The virtual support partner of the Miraculous-Life project provides a promising approach of a virtual conversational agent which can support elderly people in their familiar environment and provide them with services which enhance their quality of life. More importantly this approach evaluates how elderly people accept the conversation with an artificial intelligence and are willing to use services provided by the avatar. First evaluations showed promising results. Although a dialogue between an avatar and a human being cannot replace a dialogue between two human beings, this kind of solution might be more natural way to interact with artificial systems. This of course not only applies for the setting in elderly support but also in any interaction with ICT systems. The described system setup has been designed to provide a lightweight as well as adaptable system. Due to the system design it is possible to use the VSP on a single, stand alone device independent from the operation system (tablets etc.). The services and database components which might need regular updates and extensions are centralized on a server which is serving many clients. This provides the possibility to update as well to extend the system (by services, clients etc.) in an easy way. The technology chosen is state of the art and easy to use. The service deployment in OSGi supports the idea of an easy software life cycle management. Beside the open research question which need to be followed regarding the interaction of ECAs we think that a well designed architecture provides the possibility of a stable running system as well as the possibility of easy extensibility and update of the components.

Acknowledgement

The presented work has been funded by the European Commission under FP7-ICT-2013-10. For further information, please visit the Miraculous-Life homepage http://www.miraculous-life.eu.

References

1. Cassell, J.: More Than Just Another Pretty Face: Embodied Conversational Interface Agents. Commun. ACM, vol. 43, pp. 70-78, (2000)
2. Wilks, Y.: Close Engagements with Artificial Companions Key social, psychological, ethical and design issues. John Benjamins Pub. Co., Amsterdam/Philadelphia, (2010)
3. Ortiz, A., Carretero, P., Oyarzun, D., Yanguas, J.J., Buiza, C., Gonzalez, M.F. and Etxeberria,I.: Elderly Users in Ambient Intelligence: Does an Avatar Improve the Interaction?. Intelligence, vol. 4397, pp. 99-114, (2002)
4. Nijholt, A.: Disappearing computers, social actors and embodied agents. in Proceedings of the International Conference on Cyberworlds, pp. 128- 134, (2003)
5. Bickmore, T.W., Schulman, D. and Yin,L.: Maintaining Engagement in Long-term Interventions with Relational Agents. Appl. Artif. Intell., vol. 24(6), pp. 648-666, (2010)
6. Bickmore, T.W.: Relational Agents: Effecting Change through Human- Computer Relationships. Massachusetts Institute of Technology, (2003)
7. Pinto, H., Wilks, Y., Catizone, R. and Dingli, A.: The senior companion multiagent dialogue system. In Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems, vol. 3, pp. 1245-1248, 2008.
8. Ring, L., Barry, B., Totzke, K. and Bickmore, T.: Addressing Loneliness and Isolation in Older Adults. In Affective Computing and Intelligent Interaction, (2013)
9. Vardoulakis, L.P., Ring,L., Barry, B., Sidner, C.L. and Bickmore, T.: Designing relational agents as long term social companions for older adults. In Proceedings of the 12th international conference on Intelligent Virtual Agents, vol. 7502, pp. 289-302, (2012)
10. Bickmore, T.W., Gruber, A. and Picard, R.W.: Establishing the computerpatient working alliance in automated health behavior change interventions. Patient Educ. Couns., vol. 59(1), pp. 21-30, (2005)
11. Bickmore, T.W. and Picard, R.W.: Establishing and Maintaining Long- Term Human-Computer Relationships. ACM Trans. Comput. Hum. Interact., vol. 12, pp. 293-327, (2005)

12. Bickmore, T.W., Silliman,R.A., Nelson, K., Cheng,D.M., Winter, M., Henault, L. and Paasche-Orlow, M.K.: A randomized controlled trial of an automated exercise coach for older adults. J. Am. Geriatr. Soc., vol. 61(10), pp. 1676-83, (2013)
13. Bickmore, T.W. and Schulman, D.: A Virtual Laboratory for Studying Long-term Relationships between Humans and Virtual Agents. In Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems, pp. 297-304, (2009)
14. Fuxreiter, T., Mayer, C., Hanke, S., Gira, M., Sili, M., and Kropf, J.: A modular platform for event recognition in smart homes. In 12th IEEE International Conference on e-Health Networking Applications and Services (Healthcom), pp. 16, (2010)