

Socialite in derSpittelberg: Incorporating animated conversation into a Web-based community-building tool

Brigitte Krenn and Barbara Neumayr

¹ Austrian Research Institute for Artificial Intelligence, Vienna, Austria
brigitte@oefai.at

² sysis interactive simulations ag
barbara.neumayr@sysis.at

Abstract. In this paper, we describe an application where embodied conversational characters are integrated into an existing application which functions as a community building tool accessible via the Web. We discuss a number of design criteria which arise on the one hand from the task to simulate animated human-like conversation (which is watched by the user), and which arise on the other hand from technical restrictions imposed by the Web and the computing facilities of an average non-expert computer user.

1 Introduction

The use of the Internet has dramatically increased and nowadays we cannot imagine the Internet without its Web-communities. sysis has developed a platform (sysis NetLife) which serves as the basis for a special kind of Web-community, which is avatar-driven and focussed on bringing users together under a certain metaphor of community. Each user creates her own avatar as her virtual representation, and sends it off to the community, which might be presented as a cruise (Flirtboat), a holiday making resort on an island (Manata Island) or a flat shared by students (derSpittelberg). The platform is a client-sever application implemented in Java (JDK, JRun). The tools work on Windows NT and Linux. The Web interface is optimised for Microsoft Internet Explorer. Flash is applied for illustrating and animating the avatars. All information concerning the community including the characteristics and the “life” of the avatars, the history of their encounters with other avatars and the user is stored in a mysql database.

On the first log in, the user creates her avatar representative by answering an online questionnaire including questions about gender, age, personality, preferences and attitudes of the avatar, and selects the graphical representation of her avatar from a number of pre-designed characters. The thus created avatar has its autonomous existence in the application and seeks to satisfy its needs and pursue application-specific goals. When the user logs in and visits her avatar, she is presented with information about the current situation of and encounters made by the avatar in the community. This information is generated by the system based on the database

entries. Questionnaire, graphical design of the avatars and presentation of information to the user are made to measure for the target group of the specific application.

The users communicate with the system via key input. In part they select from system generated actions which are related to inviting other avatars to meetings with one's own avatar or with accepting invitations for the own avatar. In part free text input is possible which however is not analysed by the system, but directly delivered to the other user(s). Free text input is possible via a text window which allows a few lines of free text to accompany the invitation actions and is displayed when the user of the invited avatar logs in for the next time. In addition, an e-mail facility is available which allows users to send e-mail on behalf of their own avatars to the mailbox of other avatars in the application, and users can directly meet via chat.

2 derSpittelberg and Socialite

derSpittelberg has been developed in cooperation with an Austrian newspaper and was targeted at a young urban audience, with an affinity for flat sharing and student life. With the reference to the Viennese area Spittelberg, a particular socio-cultural context is set. See Figure 1 for an example of the user interface. In the upper middle frame the avatar and his surroundings are presented. In the text window below, the user is given some suggestions what she could do, such as improve the avatar profile, alter the presentation of the avatar's room, check the calendar and access presentations of previous encounters with other agents in the application (see left frame). In addition, the user gets information on how her avatar is doing, such as how many days the avatar is already in the application ("Tag 74 als Student", day 74 as a student), how many new people and how many friends he has met ("23 neue Leute kennen gelernt", 23 new people got to know; "5 Freunde getroffen", 5 friends met). In the frame above the representation of the avatar we find the name of the avatar (Walter), his current score in the application (1259) and his mood (super).

In addition to the text-based monologue version as shown in Figure 1, in Socialite the user is also presented with animated dialogues where she can listen and watch her avatar gossiping with another avatar about a third one (Figure 2).¹ The text window below the animated scene is only an add-on where the user can reread the dialogue which is spoken by the avatars. The screenshot shown is taken from a demonstrator for an international audience therefore the text displayed is an English translation of the German spoken dialogue. In the online version, the German text is displayed. Like in derSpittelberg, the user communicates with the system via key input.

For the Socialite application, the sysis NetLife platform has been adapted to the pipeline architecture developed in the NECA project for the generation of animated scenes of face-to-face communication between two avatars, see [2]. In particular, the text generation in sysis NetLife has been enhanced with multimodal information, including directives for speech, facial expressions and gestures. The dialogue text possibly including SSML markup and information on the emotion with which a

¹ This work is part of the NECA project, <http://www.oefai.at/NECA/>.

certain dialogue part should be expressed is fed into a speech synthesizer [7]. The timing information produced by the synthesis component is then used as a basis for fine-tuning the temporal integration of verbal and nonverbal expression down to a level of granularity where an eyebrow raise is aligned with a phoneme. All information available during processing is represented with an XML-compliant rich representation language (RRL) developed in NECA, [6] and for the full specification of the RRL see <http://www.ai.univie.ac.at/NECA/RRL/index.html>. Via the RRL we have access to the information produced at each stage of processing. This is an important precondition for the platform to be used as a test bed for animated character design, as we are able to influence/change the text to be generated, manually fine-tune speech parameters, change the voice quality and align particular gestures or facial expressions with selected passages of the spoken dialogue.

For our purposes of animation rendering, Flash has been identified as most appropriate, because of (i) the possibility to fine-tune the integration of animation and speech at various levels of granularity;² (ii) the moderate size of the animation information that needs to be sent to the client;³ (iii) the Flash player being free of charge and its easy installation at any standard home pc. Another advantage of Flash over other lean rendering approaches, e.g. Charamel (www.charamel.de) and the living actor technology by Cantoche (www.livingactor.com) is that the developers have full control over Flash, whereas the other approaches are proprietary. Full control over the animation is essential when one is interested in studying the integration of verbal and nonverbal behaviour at a fine level of granularity. Another disadvantage of living actor is that only one character is supported at a time, and thus dialogical scenes cannot be animated.

3 Conclusion

We have presented two integrated Web applications, derSpittelberg and Socialite. derSpittelberg is a commercial product and serves as a basis for bringing users together under a certain metaphor of community. Socialite is developed within a research project and aims at the development and integration of components for generating animated dialogue presented by embodied conversational characters. As regards the expression of conversational behaviour, our work relates to the BEAT system [1] and to the work on virtual humans [4]. However, in our current system, dialogical scenes are generated in one go, and are then played to the user. In other words, user input affects the generation of the whole scene/dialogue, but there is no direct interaction between the user and the evolving animated dialogue. As users watch the animated dialogues via the Internet, our approach has some communalities with Agneta and Frida [5]. A major difference, however, is that with the work on Socialite we aim at fully automatic generation of animated dialogue including speech synthesis and animation directives, whereas in Agneta and Frida the dialogues are

² This however requires expertise in Flash programming.

³ Currently initial loading time is impaired by the fact that the whole scene is fully generated in one go at the server side and then transferred to the client.

predefined and spoken by humans. As the animated scenes are integrated with sysis NetLife applications which offer a number of facilities to form and maintain Web communities, a large number of users⁴ can be accessed from which we can get feedback on Socialite's ability to generate appropriate communicative behaviour. A further advantage is that the collection of user feedback can be integrated naturally with the dramaturgy of the application-specific scenario.⁵

Acknowledgements

The Austrian Research Institute for Artificial Intelligence (ÖFAI) is supported by the Austrian Federal Ministry for Education, Science and Culture. This research is supported by the EC Project NECA IST-2000-28580.⁶

References

- [1]Cassell, J., Vilhjálmsón, H., Bickmore, T.(2001) "BEAT: the Behavior Expression Animation Toolkit." Proceedings of SIGGRAPH '01, pp. 477-486. August 12-17, Los Angeles, CA.
- [2]Krenn B., Grice M., Piwek P., Schröder M., Klesen M., Baumann S., Pirker H., van Deemter K., Gstrein E.: Generation of Multi-modal Dialogue for Net Environment, In Proceedings of KONVENS-02, 30 September - 2 October 2002, pp.91-98, Saarbrücken, Germany.
- [3]Krenn B., Neumayr B., Gstrein E. and Grice M. Life-Like Agents for the Internet: A Cross-Cultural Case Study. In Payr S., Trapp R. (eds). Agent Culture: Designing Human-Agent Interaction in a Multicultural World. Lawrence Erlbaum Associates, NJ, to appear.
- [4]Marsella S., Gratch J. and Rickel J., "Expressive Behaviors for Virtual Worlds," Life-like Characters Tools, Affective Functions and Applications, Prendinger H. and Ishizuka M. (Editors), Springer Cognitive Technologies Series, 2003
- [5]Persson P. AGNETA & FRIDA: A Narrative Experience of the Web? AAAI Fall Symposium on Narrative Intelligence, North Falmouth, Massachusetts, 1999.
- [6]Piwek P., Krenn B., Schröder M., Grice M., Baumann S., Pirker H.: RRL: A Rich Representation Language for the Description of Agent Behaviour in NECA, In Proceedings of the Workshop "Embodied conversational agents - let's specify and evaluate them!", held in conjunction with AAMAS-02, July 16 2002, Bologna, Italy.
- [7]Schröder M. and Trouvain J. (to appear). The German Text-to-Speech Synthesis System MARY: A Tool for Research, Development and Teaching. to appear in: International Journal on Speech Technology, Special Issue following the 4th ISCA Workshop on Speech Synthesis.

⁴ derSpittelberg was first launched in May 2002 and had 4.500 users within three months. Socialite was launched beginning of March 2003, and advertised to the former derSpittelberg community. After 30 days runtime the user group has reached 370 users.

⁵ See [3] for examples on collecting user data from NetLife applications.

⁶ Disclaimer: The information in this document is provided as is and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.



Fig. 1. Screen shot: derSpittelberg.



Fig. 2. Screen shot: Socialite (www.derspittelberg.at)