

Show Me What You Mean – Expressive Media for Online Communities

Toru Takahashi Yasuhiro Katagiri
ATR Media Information Science Research Labs.
2-2-2 Hikaridai, Keihanna Science City,
Kyoto, 619-0288 JAPAN
{toru, katagiri}@atr.co.jp

Christoph Bartneck
Technical University of Eindhoven
Faculty of Industrial Design
Den Dolech 2, 5600 MB Eindhoven
The Netherlands
christoph@bartneck.de

ABSTRACT

In this paper, we describe the TeMeA 2002 asynchronous online community, which uses embodied characters as expressive media to communicate messages. The functionality of the system and the challenges faced in designing it are discussed. Furthermore, we present the results of its first user evaluation.

Keywords

Online communities, Personified Media, Embodied Characters

INTRODUCTION

In face-to-face conversations, people use all of their natural modalities, such as speech, body language (gesture, pose, etc.) and facial expressions (gaze, emotion, nodding, etc.) to communicate with each other. Every conversation takes place in a shared context that may include the presence of other people and objects. The conversation is supported by the embodiment of all of its participants. This embodiment is still directly supported in videoconferences, but in Internet chat systems, only indirect representations of each participant, so called avatars [2], are available. These avatars help the participants to build a shared conversation context in a virtual chat environment. However, when people have a conversation with others through online communities, such as newsgroup and bulletin board systems, people are restricted to using textual information. Misinterpretations of these textual messages are common. The widespread application of emoticons [:-)] demonstrates that pure textual information lacks human embodiment and their communication modalities. Furthermore, it misses the conversational context, which might be compensated by including multimedia content in the messages. Although people are currently able to implicitly share context information by including links to web pages in their messages, they cannot include the web content explicitly inside the message itself.

One can distinguish two types of communications in online communities: asynchronous communication and synchronous communication. In the latter, the participants are present throughout the communication and react in real

time to messages. Videoconference systems and online chats are good examples of synchronous communication. In asynchronous communication, the participants are not present during the communication and several days may pass before a reaction to a message is posted. Newsgroups and bulletin board systems are instances of asynchronous communication. This study focuses on asynchronous communication because we regard it is still having considerable untapped potential for human-based new information society designs and the analysis of social conversation.

A synchronous communication poses stronger restrictions on communication modalities and awareness than synchronous communication. One way to overcome the restrictions is to employ avatar like embodied characters and let the characters express all non-textual information. The Media Equation [7] suggests that user will treat such characters as social actors and hence communicate with them as they would with other humans [12, 9]. The anthropomorphic appearance of the characters also helps the users to identify other participants and hence makes it easier to follow a discussion. Furthermore, the characters can help the users to understand the context of the conversation, including the involved personalities and their social relationships toward each other [10].

Based on this theoretical framework, we employed anthropomorphic characters in a prototypical asynchronous online community system called TeMeA 2002. These characters play the role of personal representations (avatars) and present the conversational contribution of its representative user. We call such characters personified media. The TeMeA 2002 system enables us to investigate the effect that expressive personified media have on the user's conversational behaviors. By analyzing logs of long-term online community activities, we hope to be able to find rules of social conversation from the viewpoint of usage and effects of expression with personified media. Such rules would be particularly helpful for autonomous character agents and enable them to act naturally and hence support them fulfilling their purpose, such as stimulating discussion between unacquainted users [4].

In this paper, we describe TeM eA 2002's functionality and its design challenges as well as report results of a usability test that preceded the upcoming long-term case study.

FUNCTIONALITIES OF TeM eA 2002

The basic functionality of TeM eA 2002 is based on a bulletin board system. Users can post their messages in an online community, and these messages become available to all other users.

The main improvement is the use of personalized media to present messages. Figure 1 shows the conversation process in TeM eA 2002. In the beginning, the users can choose to create a new topic or they can reply to an existing message.

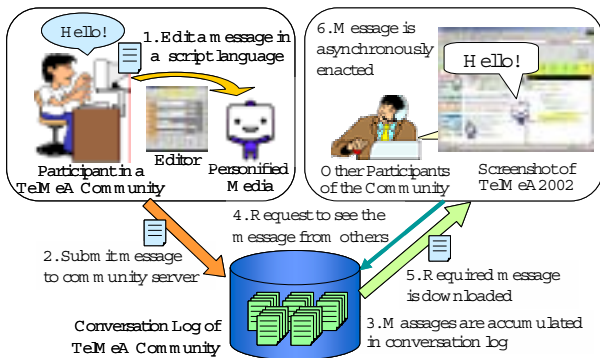


Figure 1. Conversation process in TeM eA 2002.

In both cases, the TeM eA 2002 Editor opens to allow the users to design their messages (Figure 2). The Editor provides five types of behaviors for the user's personalized medium; speech, affective expression, interpersonal attitude, document reference, and comments on document. The behavior components can be used multiple times and arranged in any sequence. This enables the users to create even complex messages



Figure 2. TeM eA 2002 Editor

The user may, for example, type in the text to be spoken ("Hello, bartneck") and choose a performative verb that describes the intention of the utterance ("greet"). To take advantage of the full potential of non-verbal cues, we defined a fine-grained set of 35 performative verbs. The verb "agree," for example, is further sub-instantiated to represent the entire range of agreeing from smiling to nodding and thumbing up. Next, the user may want to direct the attention of the audience to a certain website. Therefore, the user selects the web component and enters a URL. This web page will become the new background of the TeM eA 2002 stage. Afterwards, the user enters text again ("This is my favorite website") and chooses an affective expression for the personalized medium ("happiness"). The user could have chosen from a range of 48 other affective expressions or 13 types of interpersonal attitudes. Finally, the user submits the message. The user's message is transformed into a script language format [3] and sent to the community server, where it becomes available to all other users. A another user may select this message, for which only this script is transferred and executed on the client's computer. The posted messages are archived in a conversation log that presents the basis for further analyses of the communication, such as summarization of social conversations.

CHALLENGES

While building the TeM eA 2002 prototype, we encountered several problems that challenged the expressiveness and believability of the online community and its personalized media.

Personalization

Each participant in the TeM eA 2002 community is represented by a character, i.e., his/her personalized medium. To quickly identify the various participants in a conversation, a unique embodiment for each personalized medium is necessary. Therefore, a variety of eight personalized media is available in the TeM eA 2002 system, where these media vary in shape and color. We would like to expand this set to allow further personalization, but the high number of expressions required by each personalized medium puts a heavy load on the development resources. Consequently, we first focus on an analysis of the usage of the various personalized media before extending the grade of personalization.

Communication features

The expressions of the personalized media should cover all four features of human communication: facts, relationship, appeal and self-revelation [8]. The facts feature contains the content of the message, and the relationship feature contains the sender's opinion of the receiver and the sender-receiver relationship. The appeal feature contains the information on what the sender wants the receiver to do (intention), and the self-revelation feature contains information on the state of the sender, in particular his or her emotional state. The



Figure 3. Examples of the expressivity of the TeM ea 2002 characters. From left to right: greeting, happiness and complaint.

relationship, appeal and self-revelation features are not communicated through what is said but through how it is said.

The sender encodes all four features into his or her message and the receiver interprets the four features of the perceived message. Successful communication requires that the sent features of a message be similar to the interpreted features. A mismatch between the sent and interpreted features of the message can explain many failures of communication.

The TeM ea 2002 system enables the user to communicate facts through the spoken content of the messages. The relationship of the users toward each other can be expressed through the relative spatial distance and position of their personified media. The user might, for example, stand right next to a befriended user. The spatial distance might be a good indication of social distance. The appeal feature might be expressed through the various performative verbs, such as asks, agrees and declares. The self-revelation feature is communicated through the emotional expressions of the personified media, such as happiness, sadness and anger.

Expressive Repertoire

Humans have a wide repertoire of conversational and emotional expression, ranging from subtle frowns to ecstatic dances of joy. Personified media need to cover the entire scale of expression to become believable entities. Unfortunately, many of the current implementations of personified media exaggerate their emotional expressions or do not have enough variations in their expressions and are therefore perceived as comic characters. The TeM ea 2002 system employs 35 performative verbs (explains, agrees, complains, etc.), 48 affective expressions (likes, sadly, worries, etc.), and 13 interpersonal attitudes (yes, I know, forgotten, etc.). This variety should enable to the users to find a suitable expression for almost any situation. Figure 3 shows some examples.

In addition, the TeM ea 2002 system has certain conversational expressions to direct attention, such as pointing to objects and the distances of personified media from each other and objects. The relatively higher importance of the conversational expressions, such as glance and nods, over emotional expressions [1] might not be the same in the TeM ea 2002 system since turn taking is regulated automatically.

Communication modalities

Personified media should express their emotions consistently through all modalities available to them to ensure high believability [6]. It would not be convincing if the personified medium showed a sad face but talked with a neutral voice. Systematical manipulation of emotions in speech remains difficult, and unfortunately the speech synthesizer used for TeM ea 2002 is not able to perform this task. Therefore, we are planning to make some rough manual adjustments in the pitch and speed of the synthesized speech to acquire a minimum level of consistency.

Logging and analysis

The TeM ea 2002 community enables its participants to use rich personified media for their messages. All messages are encoded in an XML-based script language [11] and stored in log files on the TeM ea2002 server. The highly structured nature of this scripting language is optimized comprehensive analyses of the messages, including their content, performative verbs, affective words and animations. Ideally, such analysis will enable us to gain a better understanding of social communication.

EVALUATION

A qualitative usability test of the TeM ea 2002 system was performed at the Technical University of Eindhoven, Holland, in December 2002. The goal of the test was to identify major usability problems and suggest design solutions. Five participants were given the representative task to show their favorite website to another user. The other user was the second experimenter, located in Kyoto, Japan. The experimenter played the conversation partner of the participant. He took a passive role and thus only reacted to the messages of the participant. A videoconference link connected the experiment room in Eindhoven with the experimenter in Kyoto and enabled him to observe the progress from a distance and gain insight into the activities of the user. Only when he observed that the participant appeared to be stuck would he take the initiative and send a new message. Two cameras filmed the participants and their screen activities. The participants used the "Thinking-out-loud" method [3] to allow the experimenter to gain insight into their goals and activities. The experimenter in Eindhoven also observed the participants and made notes during the experiment. Afterwards he reviewed the videotapes to cross check the initial notes. Several usability problems could be identified and classified into general graphical user interface (GUI) problems, technical problems, and communication problems.

The GUI problems included problems with missing or unclear labels, wrong visualization of buttons, and redundant interface elements. Most of these problems were easily resolved by a redesign of the respective elements. The technical problems of the system consisted of excessively long response times, instability of the servers,

and scripting problems in the client software. If the user, for example, wanted to compose a new message he or she would click on the compose button, which would bring up a composer window. The loading of all elements of this window took several seconds, and the window was only operational if the loading was complete. Many participants clicked on elements before the completion of the loading process and hence caused a scripting error that in some cases disabled the entire interface. As the result, the participant would have to go back to the login screen and start over. An ongoing effort is being put into the technical improvement of the system, and we hope to have solved most problems before the upcoming case study.

The most interesting but also most difficult to solve problems are the communication problems. Several participants had problems understanding that TeMeA2002 is an asynchronous communication system (bulletin board) and not a synchronous system (chat). They tried to use the system as they would use chat systems, which resulted in several process problems. The replies to their messages, for example, appeared too late. We believe that the participants might have been misled by the constant presence of the personified media. Since the personified medium of the other user was visible all the time, the participants assumed that the other user him self or herself was online all the time and hence that they could chat with the other user. The constant presence also had the effect of making the participant believe that they could literally show a certain webpage to the other user by showing it to his or her personified medium. They put the other user's personified medium on top of a page and scrolled it up and down to show it to the other user. The participants assumed that the interface would be a shared space and that the other user could see exactly what they themselves saw on the screen. Another problem was the expectations of the participants toward the conversational abilities of the personified media. Due to the anthropomorphic gestalt of the personified media and their ability to synthesize speech, they expected the personified media to also be able to recognize speech. The participants started to talk back to the personified media after they finished their utterances.

To make users tacitly understand the asynchronous nature of TeMeA2002, we are planning to change the interface concept to the metaphor of a theatre. The various personified media would only be visible on the stage and perform their acts on it. This metaphor appears to better suit the concept of TeMeA2002. A second usability test will be necessary to confirm this assumption.

CONCLUSIONS

We presented a first prototype of an asynchronous online community that enables its participants to communicate by using personified media in the form of embodied screen characters. Participants can enrich their messages with a wide range of animations and expressions. A first usability

test was performed and resulted in several redesign suggestions. Currently, we are working on the implementation of these suggestions to prepare the system for a long-term case study. Some challenges remain, such as the consistent role of the characters. Characters are commonly used for assistants, such as the Microsoft Office Assistants [5], and as avatars to represent the user in a virtual environment. Personified media in TeMeA2002 are avatars of their respective human users. In addition, the personified medium of each individual user also fulfill the role of an assistant. The media help the user with problems and gives suggestions on how to overcome them. This dual role caused some confusion in the usability test, and we intend to make the special role of the users' own personified media clearer through their continuous presence. All other characters should only be present on the TeMeA2002 stage. This is a major conceptual change in the system, and a second usability test will be necessary to assess its success.

ACKNOWLEDGMENTS

We thank to Tsutomu Kanegae and Hideaki Fujii for their technical support and Keiko Nakao for her design of TeMeA2002. This research was partly supported by the Telecommunications Advancement Organization of Japan.

REFERENCES

1. Cassell, J. & Thorisson, K. R., The Power of a Nod and a Glance: Envelope vs. Emotional Feedback in Animated Conversational Agents. *Journal of Applied Artificial Intelligence*, 13 (3), 519-538, 1999.
2. Damer, B., *Avatars: Exploring and Building Virtual Worlds on the Internet*. Berkeley: Peachpit Press, 1997.
3. Dumais, J.S. & Redish, J.C., *A Practical Guide to Usability Testing*. Portland: Intellect Books, 1993.
4. Izbister, K., Nakanishi, H., Ishida, T., & Nass, C., Helper Agent: Designing an Assistant for Human-Human Interaction in a Virtual Meeting Space. Paper presented at CHI2000, The Hague, 2000.
5. Microsoft Corporation, *Microsoft Agent Documentation*, 1998. Available at <http://msdn.microsoft.com/workshop/imedia/agent/alldocs.zip>
6. Nass, C. & Gong, L., Is Maximization or Consistency the More Social? The Case of Synthesized Voices and Faces. Paper presented at the CHI2000, The Hague, 2000.
7. Nass, C. & Reeves, B., *The Media Equation*. Cambridge: SLI Publications, Cambridge University Press, 1996.
8. Schulz, F. v. T., *miteinander Reden - Störungen und Klärungen*. Reinbeck bei Hamburg: Rowolth Taschenbuch Verlag GmbH, 1981.
9. Takahashi, T., Takeuchi, Y., & Katagiri, Y., Change in Human Behaviors Based on Affiliation Needs – Toward the Design of Social Guide Agent System -, *Proc. of KES2000*, Vol.1, pp. 64-67, 2000.
10. Takahashi, T. and Takeda, H., TeMeA: An Asynchronous Community System with Avatar-like Agents, *Proc. INTERACT2001*, pp. 480-487, 2001.
11. Takahashi, T. and Takeda, H., Proposal of a Script Language for Embodied Conversational Agents as Asynchronous Conversational Media, *Proc. AAMAS2002*, pp. 1387-1388, 2002.
12. Takeuchi, Y. & Katagiri, Y., Social Character Design for Animated Agents, In *Proc. ROMAN '99*, pp. 53-58, 1999.