Social Skills Training System in Social Settings using Multiple Animated Characters

Masahide Yuasa
Dept. of Applied Computer Sciences,
Shonan Institute of Technology,
Fujisawa, Japan
yuasa@sc.shonan-it.ac.jp

Abstract—In recent years, the development of information technology and the decrease in opportunities for direct communication among young adults have weakened "social skills," including the ability to verbally communicate with others, interpersonal skills, and stress responses. Social skills training systems using an animated character have been developed. However, the existing systems only involve one-on-one conversation; conflicts or pressure situations are not considered. Therefore, we propose a training system that reproduces various types of communication scenarios using multiple animated characters. Through role-playing and feedback, the user can experience various interactions among people and acquire additional social skills.

Keywords—social skills, character, animated agent, social scenes

I. Introduction

A. Background

Due to the diversification of family structures in recent years, declining birth rates, and increasing number of nuclear families, communications within families and with local communities have decreased. There are fewer opportunities for children and young adults to develop "social skills" through varied interactions with others. With widespread network information systems, we have a society in which direct communication and exchange of emotions between people have reduced. Consequently, children and young adults have not developed essential social skills, including conversational, interpersonal, and stress-handling skills. There are concerns that children and young adults will lose confidence after failing due to a lack of social skills and will eventually become completely helpless.

Social skills training (SST) is used to improve the social skills of people with developmental disorders. This training is also provided to people with poor communication skills. Reference and teaching materials have been created to support SST [1, 2, 3]. Although it is possible to acquire the basic skills through self-learning (e.g., textbooks or DVDs), practical exercises are required to master the skills cannot be learned by self-study.

A system that uses computer graphic (CG) characters on computers for mock drills has been developed for practical training and self-learning [2, 3]. The displayed CG character converses with the user. The user's tone, talking speed, and eye contact are analyzed, and feedback is provided to help the user learn social skills.

However, these systems do not include scenarios in which higher degrees of social interaction are required, such as speaking in situations where many people are present or confrontational situations where people express varied opinions (agreement vs. disagreement). The training systems are only for one-on-one conversation scenarios in which the user is asked how they would respond to a statement made by another person. Situations requiring advanced social skills and involving several people are not handled.

Therefore, in this study, we created communication scenarios and social situations involving several people by using multiple animated characters to act out their roles. We developed a support system for training social skills in such situations. The user can experience these social situations, learn the necessary skills through the conversational roleplay with multiple characters, and receive feedback. The novelty of this study lies in the fact that it focuses on the social skills involved in conversing with multiple individuals (characters), which considerably differs from those required when conversing with a single character (for example, the skills involved in a pressurizing situation or during a conversation involving several people.) In this study, we attempted to establish sets of social skills required while communicating with multiple individuals and realize a social training system using animated characters.

B. Proposed System

In social situations involving several people, the use of multiple characters presents situations that require a high level of social skills. These include starting and maintaining a conversation with multiple people and being assertive in confrontational situations. Such situations cannot be reproduced by textbooks and videos, or one-on-one conversations with a character. Therefore, the consideration of such a social situation is one major advantage of our proposal. Users will acquire confidence and assertiveness by attempting conversations with multiple people and will gradually gain confidence through step-by-step training. Fig. 1 shows an overview of our system. We created daily-conversation scenarios involving one user and multiple characters (interactive situations where multiple characters and a user converse), and confrontational situations for training the social skills of the participants.

The purpose of this study is to develop techniques that can help to improve these skills. Engineering technology such as line of sight input devices, image recognition, and voice detection, can be used to measure the skills of the user, monitor the behavior of the user, and verify the level of skill improvement.

Fig. 1. Overview of proposed system. We used a tool kit for the conversational agent [4].

C. Selecting Social Skills

As children and young adults need a wide range of social skills, the following skills are focused on by using "conversation situations involving multiple persons" which is the scope of the authors' work [5, 6]:

1. Skills to express opinions in confrontational and pressure situations involving several people, e.g., how to reject something in a pressure situation, or how to gracefully leave a conversation.

2. Skills to speak up and express opinions at the correct time in conversations involving several people.

D. Social Skills Training Process

The general training process for acquiring social skills involves (I) teaching, (II) modeling, (III) roleplay, (IV) feedback, and (V) generalization [7]. The proposed system supports the modelling, roleplay, and feedback sections of the process by using multiple agents. In this system, conversations with several people are reproduced using agents, and the user is forced to select a type of behavior during the conversation. After the user says what he/she thinks is appropriate, the system gives feedback on his/her selection.

II. SOCIAL SKILLS AND DRILLS IN RESEARCH

A. Skill to Reject Something in a Pressure Situation

As described earlier, we focused on the skill to express an opinion in confrontational and pressure situations involving several people. First, we consider the case of "rejecting something in a pressure situation." For example, we deal with a typical case such as "rejecting a request made by a boss or a friend," or "turning down an invitation to attend an event (drinking party or home party)." For this scenario, we applied the knowledge gained from [7] and classified the coping skills in conflict situations as avoidant-type, harmony-type, individual-type, and integrated-type, as explained below.

(1) Avoidant type: this behavior is selected by many people, but it is difficult to obtain the understanding of the other person using this behavior, e.g., "A new game is launching today,.."

(2) Harmony-type: this person is not assertive and accepts the other person's opinion easily, even when he/she does not agree with the opinions or thoughts of the other person, e.g., "alright."

(3) Individual type: this person persuades, requests, or orders the other person to accept his/her opinion, e.g., "I can't come as I have to start early tomorrow."

(4) Integrated type: the opinions are asserted to each person and a compromise is proposed such that the opinions of both parties are respected. This is an advanced technique and ideal behavior however, it is hard to select this behavior on the first occasion, e.g., "Sorry, I have to start early tomorrow. I will definitely join you next time."

"Turning down an invitation to a drinking party from the boss" is used as an example. The user explains that he (you) are a new employee. The agent in the middle screen is the boss, and the agents to the left and right are new employees who have joined the company during the same period. In the conversation, the boss invites the user to a drinking party,

and the agents to the left and right also urge the user to participate. After that, four options (what the user should say) are displayed on the middle of the screen. The user utters one of the four statements. Fig. 2 indicates four options. After that, the characters react to this through gestures and dialogues, and in the end, a feedback screen describing the user's impression and advice is displayed (Fig. 3.) The user can experience the same conversation scenario again and again and can learn different types as different feedback screens are displayed for each statement (type).

B. Skill to Gracefully Leave the Conversation

In this situation, "the user must leave the conversation with same-age friends as he has a meeting scheduled with his superior." The user has to learn the need to skillfully leave the conversation without making the other participants uncomfortable, to ensure that future communications will not be affected. To skillfully leave the conversation, it is necessary to be assertive while being considerate towards the other participants. The dialogues in the chapter "Ending a Conversation" in [8] are examples. We define several models, as shown in Fig. 4. They are the assertive model, empathy model, forceful model, implicit model, and thoughtfulness model. The authors consider "non-verbal information" as an important part of the skills used in such a conversation. The "implicit model" was added as one of the options in which "non-verbal information" is used. The model should urge the user to learn by practicing "non-verbal information" in some form or other.
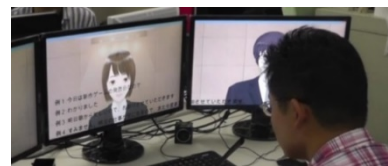

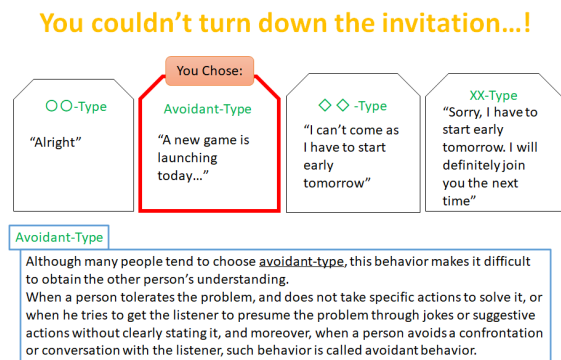Fig. 2.   Four options on the screen.

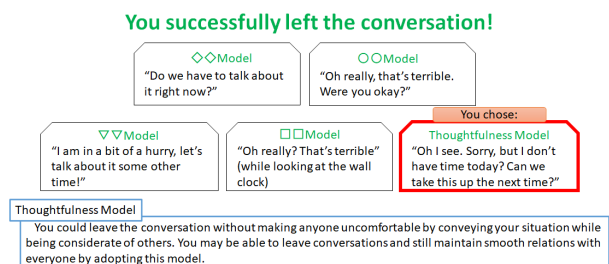Fig. 3.   Content of the feedback presented

Fig. 4. Feedback screen after selecting "thoughtfulness model." The models from top-left to bottom-right are "assertive model," "empathy model," "forceful model," "implicit model,", and "thoughtfulness model."
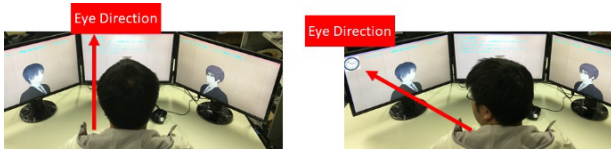
Fig. 5. Location of the option of "implicit model" and eye direction of the user.

The option of the "implicit model" is not displayed in the center of the screen, but beside the clock at the top on the left-hand side. This prompts the user to converse while looking at the clock beside the option rather than looking directly at the character. Fig. 5 shows the location of the option and the expected eye direction of the user. The aim is to enable the user to learn that they can leave a conversation by encouraging their conversation partner to "take a hint" by focusing on something else other than the partner ("I would like to leave now"). However, the partner might not always understand the clue. Therefore, the user is encouraged to learn the "thoughtfulness model" to ensure that the user can leave the conversation promptly.

In addition, as the system offers the opportunity to learn from mistakes, the "empathy model" is also added [7], which shows that the user is very considerate towards the partner and he/she is involved in the conversation. The "forceful model" is a way in which the user can almost always leave the conversation, but the user's actions are forced on the partner without any consideration for the partner's feelings. The "empathy model" and "forceful model" might be used frequently in daily lives, but they are not skillful ways of leaving a conversation. They are added to the system to help the user to understand that.

In this situation, the three characters start conversing as if they were having a chat till now and turn towards the user to start a new topic. A total of five options are then displayed on the three screens, and the user selects an option and repeats it. The agents give their reactions, and the feedback screen is displayed.

C. Skill to Express Opinion in Conversation

Skill training systems in the previous study asked the user for a response and the system gave feedback on their response. This is one-on-one communication. This method is simple for users entering a conversation, because the user is given the opportunity to start talking to the conversation partner. However, in a multi-party conversation, opportunities to start talking are normally not provided.

For example, when there are three people in conversation and the speaker asks others for their ideas, there is not exact opportunity of turn-taking, no person is automatically selected as next speaker. A person must take his/her turn assertively to start speaking. In one-on-one communications, the opportunity to take turns is provided automatically when one speaker finishes.

When there is a lack of opportunity for taking turns, people need to start to speak on their own initiative, however, it is hard for some people who do not have enough experience or communication skills, to forcefully take their turn and start to speak [9]. Therefore, we created a scenario where the system did not provide the opportunity for taking

turns in the conversation, and we manufactured the following situations:

(1) Identified question, e.g. the chairperson says, "what is everyone's opinion about this garbage problem?"

(2) Not identified question, e.g. the chairperson says, "any questions or comments on this topic?"

(3) No opportunity is provided, and the user must interrupt the conversation, e.g. the chairperson says, "This topic is finished. The next agenda item is …"

In the case of situation (1), it is easy to start talking because the chairperson suggests questions and makes time available for participants to talk. On the other hand, in the case of situation (3), it is difficult to start talking because the user needs to interrupt the conversation of several people.

III. PRELIMINARY EXPERIMENTS

Firstly, we conducted a preliminary experiment for assessing the skill required to gracefully leave a conversation, which has been discussed earlier in Section II-B.

1) Experimental method: Users experienced the system in the flow described earlier (in II-B). In our scenario, three characters start a chat with the user; however, he/she must leave it soon to join another meeting soon. The interview was conducted twice, once with all the options displayed, and the second time with all the options and the feedback having been seen by the participants.

2) Results: 11 male students from our university participated. Almost of all participants selected the "empathy model" before experiencing the system but selected the "thoughtfulness model" after experiencing it.

As a result, we thought that the learning objective was achieved. However, to make it easy to use the system for the user, it is possible that the design of the learning system was made too simple. Thus, there is the possibility that users have learnt that the "thoughtfulness model" is correct, and selected it without understanding why it should be selected. It is also possible that even though the users can answer the questions on the spot, they will not retain the learning; thus, further study is required.

Secondly, we conducted a preliminary experiment to assess the skill required to express one's opinion during a conversation, which has been discussed earlier in Section II-C.

1) Experimental method: We considered two scenarios described as follows. First, "In a meeting including three characters and you (the user), you notice a mistake in a document, which you have to point out in the meeting". Second, "In a meeting including three characters and you (the user), the other participants (three characters) are talking and determining the assignment of new jobs to members of the sales section, and you have to convince them to act againt their judgement because the sales section is very busy and no one but you knows this." These situations are examples of case (3) outlined in

Section II-C. Moreover, we prepared a time bar indicating that the user must interject within a limited time. When the time lefts, the valid (indicated) time is shrunken, and the user must talk to in the time.

2) Results: 8 out of 12 users commented that the bar prompted them to start speaking. Out of 30 tests by 15 users, 21 tests were answered with the comment that the system advice was useful for their learning, and they could recognize when and how to start to speak.

We thought that the certain objective was achieved. We will continue to conduct similar experiments and create realistic situations. We will confirm that our proposed system helps people to learn how to start to speak, how to interrupt in a conversation involving several people, and so contribute to the refinement of their social skills.

IV. CONCLUSION

Our work involves building support technologies for improving social skills. We have reported on the concept of our social skill training system that uses multiple characters and described which social skills were used in our system. We also explained how to assist the learning process for users in our training system. By providing opportunities to develop socially using multiple characters, it is possible to contribute towards forming a society where the problems associated with a lack of social skills in children and young adults is addressed. We expect that such skills will enable them to become independent and able to express their opinions with confidence.

REFERENCES

[1] O. Golan, S. Baron-Cohen, "Systemizing empathy: teaching adults with Asperger syndrome or high-functioning autism to recognize complex emotions using interactive multimedia," Development and Psychopathology, vol. 18, no. 2, pp. 591–617, 2006, DOI: 10.1017/s0954579406060305.

[2] H. Tanaka, S. Sakti, G. Neubig, et al. "Automating Social Skills Training," vol. 114, no. 305, ET2014-61, pp. 1-6 (in Japanese).

[3] M. E. Hoque, M. Courgeon, B. Mutlu, J-C. Martin, R. W. Picard, "MACH: My Automated Conversation Coach," In the 15th International Conference on Ubiquitous Computing (Ubicomp), September 2013.

[4] MMDAgent, a toolkit for building voice interaction systems. http://www.mmdagent.jp/

[5] M. Yuasa, "Can violation of conversational behavior maintain a sense of unity in informal situations? - A study on perception of conversational behavior using interactive robots/agents" -, CogSci, 2018.

[6] M. Yuasa, "Can Animated Agents Help Us Create Better Conversational Moods? - An Experiment on the Nature of Optimal Conversations" -, HCII2014, 8th International Conference, UAHCI 2014, Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, Proceedings, Part IV, pp. 634-640, in Universal Access in Human-Computer Interaction. Design for All and Accessibility Practice, 2014.

[7] M. Aikawa, "Psychology of Social Skills," 2000 (in Japanese).

[8] N. Katou, Communication Training Manual for the Adult Autism Spectrum, Seiwa-Shoten, 2017 (in Japanese).

[9] R. Hale-Evans, Mind Performance Hacks, O'Reilly Media, Inc. 2006.