Kenta Yamamoto

SANKEN, Osaka University
Ibaraki, Osaka
Japan

kentayamamoto@sanken.osaka-u.ac.jp

1 Research interests

The author is interested in building dialogue systems with **character** and **user adaptaton**. The goal is to create a dialogue system capable of establishing deeper relationships with users. To build a trustful relationship with users, it is important for the system to express its character. The author particularly aims to convey the system's character through multimodal behavior.

Although users currently try to speak clearly to avoid speech recognition errors when interacting with SDSs, it is necessary to develop SDSs that allow users to converse naturally, as if they were speaking with a human. The author focused on user adaptation by considering user personality and proposed a system that adjusts its manner of speaking according to the user's personality. In addition, the author is interested not only in adjusting the system's speaking style to match the user but also in making **the system's listening style** more conducive to natural conversation.

1.1 Character expression for SDSs

The character expression of robots leads to increasing user engagement in dialogue. In this work, "personality" is used as a psychological dimension for classifying users, and "character" notes the impression that the robot gives to the user. Generally, the character (personality) of a system is set based on enumerated personas or the Big Five personality traits. However, the characteristics of a personality that are easier to convey to users will differ depending on the modality. Therefore, the author focused on features in spoken dialogue such as backchannels, fillers, and switching pause length, and constructed a character expression model for a spoken dialogue system (Yamamoto et al., 2022). For example, an extroverted system is programmed to give frequent backchannels, while an emotionally unstable system is controlled to use more fillers. The results of dialogue experiments showed that the system was able to give the impression of executing its role more appropriately in the dialogue by expressing a character according to the task.

1.2 User adaptation based on user personality

The goal with user adaptation of SDSs is to make the systems generate behaviors appropriate to the user, which leads to increasing user satisfaction in the dialogue. There are several approaches when it comes to achieving dialogue suitable for the user such as selecting topics of interest to the user, synchronizing with the user's behavior, and predicting the user's internal state to facilitate dialogue. The author is interested in a system's manner of speaking that makes the user feel comfortable talking. In other words, it should evoke an impression of "getting along well" or "feeling at ease."

The author has previously demonstrated that having an SDS express a character that matches the user's personality can enhance user satisfaction with the dialogue (Yamamoto et al., 2023). In this earlier work, an "extroverted system" and "introverted system" were constructed using the techniques explained in Section 1.1, and the author analyzed the tendency of users to prefer interacting with each system based on the user's personality.

1.3 Control of system behaviors as a listener

There are have been many studies on the operation of spoken dialogue robots, but for users to have a pleasant dialogue, it is also necessary for the robot to behave appropriately as a listener. In the past, the author has been involved in the development of a listening dialogue system Inoue et al. (2020) that ensures users can speak comfortably by appropriately utilizing responses such as backchannels, evaluative feedback, and questions. On the other hand, humans perform various actions and reactions while listening, such as nodding and giving backchannel responses. By appropriately modeling such behaviors, the author aims to adequately express the sense that the agent is actively listening during spoken dialogues.

2 Spoken dialogue system (SDS) research

The author discusses the important topics for future studies on SDSs.

2.1 User adaptation in first-time interactions

Just as it is implausible for humans to like every single person they meet, there is no dialogue system that can be

The 20th Annual Meeting of the Young Researchers' Roundtable on Spoken Dialogue Systems, pages 76–77

liked by all people. Therefore, a dialogue system needs to understand the characteristics of the conversation partner from the dialogue and adapt accordingly. However, such user adaptation is typically assumed for dialogue systems used by the same user over a long period.

Conversely, systems designed for first-time interactions, such as with store clerks, should be designed to speak in a manner suitable for the role (satisfying many users). User adaptation based on persona information, such as individual user preferences, makes it difficult to handle first-time interactions.

However, the author believes that it is possible to achieve user-appropriate dialogues even in first-time interactions. This can be done by recognizing the user's personality at the beginning of the conversation and then adjusting the speech style to match it during the conversation. Although human personalities vary, it is believed that preparing several personality groups for specific situations can handle these variations. Another advantage of this approach is that, unlike when using personas, the system does not explicitly communicate the recognized result of the user's personality to the user. Therefore, even if the recognition result is incorrect, the conversation itself can avoid collapsing.

2.2 Understanding human relationships

Two types of human relationships are discussed here: first, the relationship between the user in front of the system and the system itself, and second, the relationships between users. When implementing user adaptation in SDSs, it is necessary to model the relationship between the user and the system. In other words, it is essential to constantly monitor how much trust the user and the system have built. This is because, in scenarios where the user interacts with the system over an extended period, the manner and content of the conversation will change. This change cannot simply be measured by the length of the interaction time because it switches in accordance to changes in the relationship with the user. Therefore, it is crucial to model the relationship between the user and the system and conduct the dialogue accordingly.

Currently, it is sufficient to continue the dialogue by considering only the relationship with the person in front of the system in a one-on-one interaction. Indeed, most of the datasets collected for learning purposes assume one-on-one dialogues. However, when SDSs or robots are used in society, it becomes necessary to conduct dialogues that consider multi-person interaction scenarios and relationships with people not present in the conversation.

In such cases, it is necessary to make utterances that consider the relationships between users. However, there is a lack of data and methods for constructing dialogue systems that handle multi-person interactions or are utilized by multiple users. For example, in multi-person dialogues, predicting the next speaker can vary depending on the relationships between users. It is thus necessary to collect multi-party dialogue datasets in various everyday situations to build models that capture the relationships between users.

3 Suggested topics for discussion

The author suggests three topics for discussion in the discussion panel during the event.

- Is it necessary for spoken dialogue systems to possess human-like dialogue? Are there more appropriate methods for dialogue with dialogue systems?
- How can insights gained from theoretical studies of dialogue, such as conversation analysis, be incorporated into our studies?
- Does a SDS's agreement fulfill the user's need for approval? Can it serve as a substitute for human friends?

Acknowledgement

This work was supported by Grant-in-Aids for JSPS (23K20005, 24K20839).

References

Koji Inoue, Divesh Lala, Kenta Yamamoto, Shizuka Nakamura, Katsuya Takanashi, and Tatsuya Kawahara. 2020. An attentive listening system with android ERICA: Comparison of autonomous and WOZ interactions. In *SIGDIAL*. pages 118–127.

Kenta Yamamoto, Koji Inoue, and Tatsuya Kawahara. 2022. Character expression for spoken dialogue systems with semi-supervised learning using variational auto-encoder (79):101469–101469.

Kenta Yamamoto, Koji Inoue, and Tatsuya Kawahara. 2023. Character adaptation of spoken dialogue systems based on user personalities. In *IWSDS*.

Biographical sketch

Kenta Yamamoto received his Ph.D. in 2023 from the Graduate School of Informatics in Kyoto University, Kyoto, Japan. He was a JSPS Research Fellow (DC1) from 2020 to 2023. Currently, he is an Assistant Professor of SANKEN (The Institute of Scientific and Industrial Research), Osaka University. His research interests include spoken dialogue systems (SDSs) and character expression for SDSs.