Shohei Tanaka

Nara Institute of Science and Technology (NAIST)
Takayama-cho 8916-5
Ikoma-shi, Nara-ken
Japan

tanaka.shohei.tj7@is.naist.jp
https://shohei-ta-ds7.github.io/

1 Research interests

My current research interest is **thoughtful dialogue agents** that respond with **thoughtful actions** to **ambiguous user requests**. The agents assume to be used in tasks that users sometimes can not clearly verbalize their requests such as sightseeing navigation or hotel reception. I also interested in how dialogue systems use external knowledge such as **event causality**. I have been studying these topics to develop a dialogue agent that can support users in diverse situations as well as thoughtful human concierges.

1.1 Ambiguous user requests and thoughtful system actions

Existing task-oriented dialogue systems assume that user intentions are clarified and uttered in an explicit manner; however, users often do not know what they want to request. User requests in such cases are ambiguous. Taylor (1968) categorizes user states in information search into four levels according to their clarity, as shown in Table 1. Most of the existing task-oriented dialogue systems (Madotto et al., 2018) convert explicit user requests (Q3) into machine readable expressions (Q4). Future dialogue systems need to take appropriate actions even in situations such as Q1 and Q2, where the users are not able to clearly verbalize their requests (Yoshino et al., 2017).

Figure 1 shows an example dialogue between a user and a developed dialogue agent. The user utterance "It is hot today!" is not verbalized as a request for a specific function. The dialogue agent responds with a thoughtful action, "Shall I search for a cafe around here?" and searches for a cafe. I collected a corpus and developed a model that classifies ambiguous user requests into corresponding system actions (Tanaka et al., 2021). The corpus collected in our study assumes cases where the user requests are ambiguous, such as Q1 and Q2 in Table 1.

However, the dialogue agent decides action candidates based on only one-turn user request. The agent can not decide one action with multi-turn dialogue when multiple actions could be regarded as thoughtful. For the user request in Figure 1, not only "cafe" but also "rest area" is thoughtful as system action. As future work, I will construct a agent that can deal with such cases by collecting a multi-turn corpus and updating the system architecture.

Level	Definition
Q1	The actual, but unexpressed request
Q2	The conscious, within-brain description of the request
Q3	The formal statement of the request
Q4	The request as presented to the dialogue agent

Table 1: Levels of ambiguity in requests (queries) (Taylor, 1968)

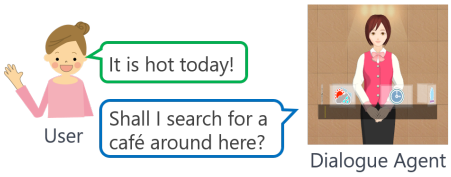

Figure 1: Example of thoughtful dialogue

1.2 Conversational response re-ranking based on event causality

While a variety of dialogue models such as the neural conversational model (NCM) (Vinyals and Le, 2015) have been researched widely, such dialogue models often generate simple and dull responses due to the limitation of their ability to take dialogue context into account. It is very difficult for these models to generate coherent responses to a dialogue history.

I tackled this problem with a new architecture by incorporating event causality relations between response candidates and a dialogue history (Tanaka et al., 2019). Typical event causality relations are cause-effect relations between two events, such as "be stressed out" precedes "relieve stress." The proposed method could select coherent and diverse responses to some extent.

As future work, I will apply event causality to task-oriented dialogue systems to estimate user intentions more accurately.

2 Spoken dialogue system (SDS) research

In this section, I will briefly present my opinions and point of views of current and future SDSs research for the following questions.

2.1 Where do you think the field of dialogue research will be in 5 to 10 years?

Tasks in more diverse and real situations become the research trends. Current task-oriented dialogue systems assume specific situations such as movie recommendation or sightseeing navigation. I think practical dialogue systems as partners for people need to work well in more diverse situations. Conventional end2end training does not work well because the task range is too diverse. The system have to imitate the deduction or estimation process of human beings with commonsense knowledge.

2.2 What are the most important things for users of SDSs

I think that users regards accuracy of SDS responses to user requests as most important point. Although neural network based systems can generate responses in diverse situations, their responses sometimes does not make sense. The users will stop use the systems in such cases. Then, SDS developers have to ensure accuracy even if the developed systems could not deal with diverse situations.

2.3 Is there a difference between SDS research in academia and industry?

One of the most different point between academia and industry is the purpose for developing SDSs. I have been studying SDSs for publishing papers which bring the related areas a profit. The system need to be for not making money within few years, but changing the world which SDSs are used in after ten years. If I develop SDSs in a company, the developed system bring specific users a profit. The development has also important value for me because I can immediately understand that I changed the world even if the area is spedific.

3 Suggested topics for discussion

In this section, I suggest three topics for discussion in the discussion panels during the event.

- Corpus collection: What is the best method to collect a corpus for difficult tasks such as suggesting thoughtful actions to ambiguous user requests. Simple WoZ method will fail because the tasks are too difficult for general people.

- How human beings estimate intentions of a person who they are talking to. What kind of knowledge is used in the process?

- Will dialogue systems be able to work as well as human concierges in sightseeing navigation or hotel reception within ten years? What problems should be solved to establish the goal?

References

Andrea Madotto, Chien-Sheng Wu, and Pascale Fung. 2018. Mem2Seq: Effectively Incorporating Knowledge Bases into End-to-End Task-Oriented Dialog Systems. In *ACL*. pages 1468–1478.

Shohei Tanaka, Koichiro Yoshino, Katsuhito Sudoh, and Satoshi Nakamura. 2019. Conversational response re-ranking based on event causality and role factored tensor event embedding. In *NLP for ConvAI*. Florence, Italy, pages 51–59.

Shohei Tanaka, Koichiro Yoshino, Katsuhito Sudoh, and Satoshi Nakamura. 2021. ARTA: Collection and Classification of Ambiguous Requests and Thoughtful Actions. In *SIGDIAL*. Singapore.

Robert S. Taylor. 1968. Question-negotiation and information seeking in libraries. *College & Research Libraries* 29(3):178–194.

Oriol Vinyals and Quoc V. Le. 2015. A Neural Conversational Model. In *ICML*.

Koichiro Yoshino, Yu Suzuki, and Satoshi Nakamura. 2017. Information navigation system with discovering user interests. In *SIGDIAL*. Saarbrücken, Germany, pages 356–359.

Biographical sketch

Shohei Tanaka earned B.E. from Nagoya Institute of Technology at 2018, M.E. from Nara Institute of Science and Technology at 2020. He has been a Ph.D. candidate at Nara Institute of Science and Technology since 2020. He has been engaged in research on natural language processing, especially on dialogue systems and event causality. He received a best paper award of the first workshop on NLP for conversational AI. He is a member of the Association for Computational Linguistics.