Ryu Hirai

Nagoya University
Nagoya, Aichi
Japan
hirai.ryu.k6
@s.mail.nagoya-u.ac.jp

1 Research interests

Although task-oriented dialogue systems have improved, not all users can accomplish their tasks (Takanobu et al., 2020). The task success rate of the state-of-the-art model (Feng et al., 2023) on the MultiWOZ dataset (Budzianowski et al., 2018) is around 80%, indicating room for improvement. Even for dialogue systems built using large language models such as OpenAI's ChatGPT[1], the system performance is not always satisfactory (Hudeček and Dušek, 2023). One possible reason for this limited performance is that users fail to achieve their tasks because of limited knowledge about the system. Hence, I seek to offer a solution based on **tutorials**, which provide users with system knowledge, and **user adaptation**, which adapts the system's behavior to that of the user, thus enabling users to succeed in dialogues without changing their behavior.

1.1 Tutorials in task-oriented dialogue systems

To develop appropriate tutorials, I am currently conducting studies on estimating a user's task success ability. Among previous studies on user ability estimation, Ghazarian and Noorhosseini (2010) constructed an automatic skill classifier that uses mouse movements in desktop applications to adjust the interface or content provided by the system. Komatani et al. (2003) proposed a method for estimating user attributes, such as the skill level with respect to a system, and enabling the system to change its behavior accordingly. However, those studies focused on estimating user ability solely from user behavior. I believe that consideration of the characteristics of users' tasks would lead to better user ability estimation.

I proposed a method that estimates task success ability by applying item response theory (IRT) (Lord, 1980), which is commonly used in education for estimating examinee abilities, in slot-filling task-oriented dialogue systems (Hirai et al., 2023). Specifically, I first collect dialogues in which the system presents each user with a unique dialogue goal and the user must engage in a dialogue based on that goal. Next, by treating the correct filling of each designated slot as a problem, I apply IRT to estimate the item characteristics of slots. Finally, the user engages in the dialogue based on the given goal, and his/her task success ability is estimated by using the item characteristics of filled and unfilled slots. Through experiments on using the estimated task success ability to predict the probability of a correct answer for each slot, I found that the proposed method significantly outperformed baseline methods. In other words, the proposed method could accurately estimate a user's task success ability.

I now seek to improve the estimation accuracy by applying recent deep-learning-based IRT methods. Additionally, I aim to investigate methods for estimating task success ability more efficiently by not requiring the user to engage in a complete dialogue. I also want to create an interactive tutorial agent that poses a user with a certain dialogue goal and estimates the user's task success on the basis of how the goal is handled.

1.2 User adaptation by task-oriented dialogue systems

Tutorials can enable users to achieve tasks by changing their behavior. In an engineering sense, however, it is desirable to not require users to change their behavior, which makes user adaptation a viable option. That is, if a system can vary its behavior according to the user, then users will be able to accomplish tasks more easily.

I am particularly interested in exploring methods to adapt the system behavior according to a user's task success ability. For example, I want to develop a system that leverages a user's estimated task success ability to change the vocabulary level, adjust the amount of information included in an utterance, or adapt the parameters of recognition models such as those used in speech recognition and natural language understanding. For instance, Ohashi and Higashinaka (2022) proposed a method that uses reinforcement learning to generate adaptive utterances for users with a limited vocabulary. Such techniques using reinforcement learning could be applied in this research.

I previously participated in the Dialogue Robot Competition 2022 (DRC2022) (Minato et al., 2022). In that competition, participants developed systems for humanoid robots in a physical environment to act as counter

[1] https://openai.com/blog/chatgpt/

Proceedings of the 19th Annual Meeting of the Young Reseachers' Roundtable on Spoken Dialogue Systems, pages 64–65
September 11–12, 2023.

salespeople for travel agencies. I consider this setting ideal for developing user-adaptive task-oriented dialogue systems, because many types of users visit travel agencies, and salespeople must exhibit hospitality and adapt to users as much as possible to enable them to accomplish tasks in an efficient, satisfactory manner. In addition to the robot's dialogue content, I also want to implement multi-modal, user-dependent behaviors such as gestures and facial expressions.

2 Spoken dialogue system (SDS) research

Multi-modal dialogue systems have the characteristic of being able to convey information that cannot be conveyed through text or speech alone. However, the research on multi-modal dialogue systems is not especially extensive when compared with research on text- or speech-based dialogue systems. Additionally, most of the current research on multi-modal dialogue systems focuses on systems that use images (Sun et al., 2022), whereas there is limited research on dialogue robots in physical environments.

I believe that the scarcity of large-scale, multi-modal dialogue datasets is one reason for the limited progress in the field. It is anticipated that virtual-, augmented-, or mixed-reality systems will be useful in constructing multi-modal dialogue datasets at a lower cost, thus enabling the development of large multi-modal datasets. Consequently, there will be an increase in research on multi-modal dialogue systems, including those involving robots.

3 Suggested topics for discussion

I would like to discuss the following topics:

- What information should be obtained from users when adapting task-oriented dialogue systems to them?
- Deep learning is commonly used in task-oriented dialogue systems but involves high costs for dataset construction. What methods are available to collect annotated, large-scale datasets efficiently?
- Can large language models be used for accurate annotation of task-oriented dialogue datasets?

References

Paweł Budzianowski, Tsung-Hsien Wen, Bo-Hsiang Tseng, Iñigo Casanueva, Stefan Ultes, Osman Ramadan, and Milica Gašić. 2018. MultiWOZ — a large-scale multi-domain Wizard-of-Oz dataset for task-oriented dialogue modelling. In *Proc. of EMNLP*. pages 5016–5026.

Yihao Feng, Shentao Yang, Shujian Zhang, Jianguo Zhang, Caiming Xiong, Mingyuan Zhou, and Huan Wang. 2023. Fantastic Rewards and How to Tame Them: A Case Study on Reward Learning for Task-oriented Dialogue Systems. *arXiv preprint arXiv:2302.10342* .

Arin Ghazarian and S Majid Noorhosseini. 2010. Automatic detection of users' skill levels using high-frequency user interface events. *User Modeling and User-Adapted Interaction* 20:109–146.

Ryu Hirai, Ao Guo, and Ryuichiro Higashinaka. 2023. Applying Item Response Theory to Task-oriented Dialogue Systems for Accurately Determining User's Task Success Ability. In *Proc. of SIGDIAL*.

Vojtěch Hudeček and Ondřej Dušek. 2023. Are LLMs All You Need for Task-Oriented Dialogue? *arXiv preprint arXiv:2304.06556* .

Kazunori Komatani, Shinichi Ueno, Tatsuya Kawahara, and Hiroshi G Okuno. 2003. User Modeling in Spoken Dialogue Systems for Flexible Guidance Generation. In *Proc. of Eurospeech*. pages 745–748.

Frederic M Lord. 1980. *Applications of Item Response Theory to Practical Testing Problems*. Routledge.

Takashi Minato, Ryuichiro Higashinaka, Kurima Sakai, Tomo Funayama, Hiromitsu Nishizaki, and Takayuki Nagai. 2022. Overview of Dialogue Robot Competition 2022. *arXiv preprint arXiv:2210.12863* .

Atsumoto Ohashi and Ryuichiro Higashinaka. 2022. Adaptive Natural Language Generation for Task-oriented Dialogue via Reinforcement Learning. In *Proc. of COLING*. pages 242–252.

Qingfeng Sun, Yujing Wang, Can Xu, Kai Zheng, Yaming Yang, Huang Hu, Fei Xu, Jessica Zhang, Xiubo Geng, and Daxin Jiang. 2022. Multimodal Dialogue Response Generation. In *Proc. of ACL*. pages 2854–2866.

Ryuichi Takanobu, Qi Zhu, Jinchao Li, Baolin Peng, Jianfeng Gao, and Minlie Huang. 2020. Is Your Goal-Oriented Dialog Model Performing Really Well? Empirical Analysis of System-wise Evaluation. In *Proc. of SIGDIAL*. pages 297–310.

Biographical sketch

Ryu Hirai is a master's student at the Graduate School of Informatics, Nagoya University. He is interested in making task-oriented dialogue systems user-friendly and participated in the Dialogue Robot Competition 2022. He is supervised by Prof. Ryuichiro Higashinaka.