Nicolas Wagner

Natural Language Generation and Dialogue Systems Group
University of Bamberg
96047 Bamberg
Bavaria, Germany

nicolas.wagner@uni-bamberg.de
www.uni-bamberg.de/en/ds/team/wagner/

1 Research interests

My research interests include **multi-user dialogue systems** with a focus on **user modelling** and the development of **moderation strategies**. Contemporary Spoken Dialogue Systems (SDSs) frequently lack the ability to interact with more than one user simultaneously. Moreover, I am interested in researching on the **Controllability of Language Generation** using **Large Language Models** (LLMs). Our hypothesis is that an integration of **explicit dialogue control signals** improves the Controllability and Reliability of generated sequences independently of the underlying LLM.

1.1 Multi-User Dialogue Systems

Although group interactions play an essential role in people's daily lives, research on multi-user dialogue systems is rather underrepresented. A reason for that could be the various challenges associated with this topic: Turn-taking strategies, addressee detection, and the lack of suitable training data, just to name a few. Researchers often face the issue that there are generally no guidelines or best practices for developing new multi-user SDSs, since most publications present systems which are strongly focusing on one specific task and are therefore not generalisable.

To gain insights into how users would like to be addressed by an SDS during a group conversation, we conducted a user study (Wagner et al., 2019). Here, we identified which system behaviours were perceived as less obtrusive and beneficial for the dialogue flow. As a next experiment, we evaluated moderation strategies for group chats (Wagner et al., 2022). The moderation strategies were intended to support groups in negotiating joint appointments and were rated as helpful.

Another topic I am interested in is human-robot interaction. We investigated recommendation strategies in a scenario with a household assistant robot (Kraus et al., 2022). In context of multi-user interaction, we examined how users perceived the usability of different dialogue strategies in a quiz game setup (Wagner et al., 2023). We intend to improve the developed strategies through the use of LLMs, which leads to the next section.

1.2 Controllable Language Generation

Task-oriented dialogue systems are designed to assist users in accomplishing specific tasks through natural language interactions. Traditional systems rely on a pipeline architecture with components for Natural Language Understanding, Dialogue Management, and Text Generation (Jokinen and McTear, 2009). Recently, LLMs are applied to substitute these components, as their generated outputs are much more flexible and appear more natural.

Since no task-specific data can be presented during pre-training, it is necessary to adapt models to a downstream task. Contemporary approaches include fine-tuning (Ouyang et al., 2022) and in-context learning (Brown et al., 2020). Although this equips models with certain capabilities to maintain context over conversations, it does not prevent the risk of incorrect responses or hallucinations. Further approaches like retrieval-augmented generation (Lewis et al., 2020) and automated generation of prompt templates (Sánchez Cuadrado et al., 2024) are supposed to provide additional knowledge and task-dependent prompt design.

However, none of these techniques consider the use of explicit control signals to control the dialogue flow, as they rely instead on implicit dialogue modelling within the neural net of transformer-based LLMs. To overcome these limitations, we propose equipping the system with a component for explicit dialogue control similar to the traditional pipeline architecture - the dialogue controller. For this, we have conducted a baseline experiment in which we showed that a dialogue controller improves the controllability of generated outputs (Wagner and Ultes, 2024). Specifically, the generated responses were more likely to correspond to human-annotated references. The proposed controller is designed to extract task-relevant data from a knowledge source and to provide a prompt instruction depending on the user input and intention. The system architecture is depicted in Figure 1. Furthermore, we plan to conduct experiments on constrained decoding as described in (Shin et al., 2021). This way, we expect to further enhance control over the generated output and provide users with more reliable responses.

The 20th Annual Meeting of the Young Researchers' Roundtable on Spoken Dialogue Systems, pages 70–72
September 16–17, 2024. ©2024 Association for Computational Linguistics

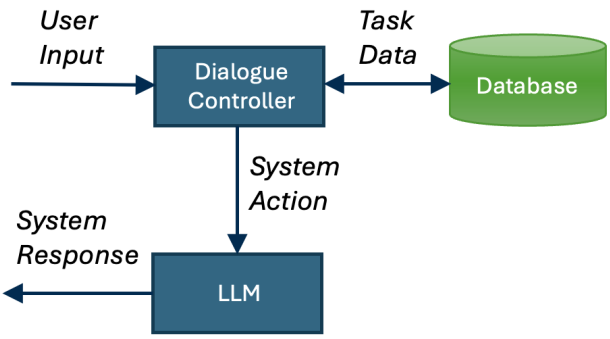

Figure 1: Depiction of the dialogue control architecture.

2 Spoken dialogue system (SDS) research

In my opinion, LLMs will increase the performance and popularity of SDSs and thus have an impact on research. However, I believe spoken language has its greatest potential in domains where information cannot be conveyed more efficiently or conveniently by other means. This includes voice assistants, smart speakers, health sector applications, or smart home environments. Personalisation and context-awareness will also play a vital role in the future, which may enable SDSs to become more and more useful in everyday life. Moreover, the research on multi-user dialogue systems needs to be intensified and common design rules established.

For the young research community, I would welcome the idea of actively participating in the development and application of ethical guidelines for the use of artificial intelligence, as this is where I expect major differences between the aims of industry and academia.

3 Suggested topics for discussion

My suggestions centre on the field of multi-user dialogue systems, and controllable natural language generation using Large Language Models.

- **Multi-User Dialogue Systems**: What are the best practices for development and evaluation? How to deal with the challenges in user modelling and dialogue state tracking? What turn-taking and moderation strategies should be used?
- **Natural Language Generation**: How to design explicit control signals to improve the controllability of language generation? Can constrained decoding enhance the response quality?
- **Actionable Evaluation Metrics**: Which metrics are applicable to measure the perceived naturalness of SDSs? How can they be used for policy optimisation? Can they also be applied for multi-user SDS?

References

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems* 33:1877–1901.

Kristina Jokinen and Michael McTear. 2009. *Spoken dialogue systems*. Morgan & Claypool Publishers.

Matthias Kraus, Nicolas Wagner, Wolfgang Minker, Ankita Agrawal, Artur Schmidt, Pranav Krishna Prasad, and Wolfgang Ertel. 2022. Kurt: A household assistance robot capable of proactive dialogue. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction (HRI 2022)*. ACM/IEEE.

Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, et al. 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems* 33:9459–9474.

Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. 2022. Training language models to follow instructions with human feedback. *Advances in neural information processing systems* 35:27730–27744.

Jesús Sánchez Cuadrado, Sara Pérez-Soler, Esther Guerra, and Juan De Lara. 2024. Automating the development of task-oriented llm-based chatbots. In *Proceedings of the 6th ACM Conference on Conversational User Interfaces*. Association for Computing Machinery, New York, NY, USA, CUI '24.

Richard Shin, Christopher Lin, Sam Thomson, Charles Chen, Subhro Roy, Emmanouil Antonios Platanios, Adam Pauls, Dan Klein, Jason Eisner, and Benjamin Van Durme. 2021. Constrained language models yield few-shot semantic parsers. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, pages 7699–7715.

Nicolas Wagner, Matthias Kraus, Niklas Lindemann, and Wolfgang Minker. 2023. Comparing multi-user interaction strategies in human-robot teamwork. In *In Proceedings of the 13th International Workshop on Spoken Dialogue Systems Technology (IWSDS '23)*. Springer.

Nicolas Wagner, Matthias Kraus, Niklas Rach, and Wolfgang Minker. 2019. How to address humans: System barge-in in multi-user hri. In *Proceedings of the*

10th International Workshop on Spoken Dialog Systems Technology (IWSDS 2019).

Nicolas Wagner, Matthias Kraus, Tibor Tonn, and Wolfgang Minker. 2022. Comparing moderation strategies in group chats with multi-user chatbots. In *In Proceedings of the fourth ACM Conference on Conversational User Interfaces (CUI '22)*. ACM.

Nicolas Wagner and Stefan Ultes. 2024. On the controllability of large language models for dialogue interaction. In *To be published at SIGDIAL 2024*.

Biographical sketch

Nicolas Wagner received his B.Sc. and M.Sc. degrees in communications and computer engineering with focus on human-machine interaction at Ulm University, Germany, in 2018. From 2018 to 2024, he was a research assistant at the Dialogue Systems Group of Ulm University, Germany. Since 2019, he is also enrolled as a Joint-PhD candidate at the University of Granada, Spain. He is currently pursuing his PhD at both institutions. He joined the Natural Language Generation and Dialogue Systems Group at the University of Bamberg in 2024. His research interests include the development of multi-user dialogue systems and controlled language generation.