

1 Research interests

My research interests can be broadly categorized into the areas of **natural language understanding** and **multi-modal dialogue systems**.

1.1 Global context for natural language understanding

My main topic of interest is **global context for natural language understanding**. As a shallow analogy to human conversation, the understanding of the user's utterances by a dialogue system can be improved by supplying it with additional context constituting the greater „whole” of the conversation, such as dialogue history.

In my ongoing master's thesis, I describe various sources of global context and analyze the effect appending some of their configurations to user utterances has on a natural language understanding module's performance. I consider intra-systemic sources where the extra information is already present in the system, for example, dialogue history or previously identified dialogue acts and slot values, and – inspired by such approaches as Xu et al. (2021) – extra-systemic ones, where the information is taken from the outside, e.g., WordNet definitions of words present in a given utterance. As the main part of the thesis, I conduct a series of experiments with the use of a T5-based natural language understanding module and the MultiWOZ dataset. By comparing the performance of several versions of the module fine-tuned with samples enriched with global context in various ways against a baseline, I show that overall the use of global context translates into better performance of the module.

1.2 Multi-modal dialogue systems

Conversations with dialogue systems are often not easy for the users: as various evaluation studies show (Adamopoulou and Moussiades, 2020), they tend to give up mid-conversation, get frustrated easily, or end up rating the whole ordeal poorly. My second topic of interest is the theory and practice relating to **multi-modal dialogue systems**, which are one of the potential solutions to such problems. By engaging the user through multiple modalities, it is easier to draw his attention and keep him

focused on the task at hand, allowing the conversation to continue in cases where it might have ended prematurely.

Recently, along with a team of other students from my university, I finished the development of AMUseBot, a task-oriented dialogue system envisioned as a cooking assistant, previously presented at the 4th Polish Conference on Artificial Intelligence in Łódź in the form of a poster and a publication (Christop et al., 2023) describing the work-in-progress stage of the project. AMUseBot communicates with users simultaneously through text, voice, and a graphical display, putting the principles of multi-modal communication into practice. It also employs a mix of rule-based and machine learning-based modules, enabling a controlled „main scenario” dialogue progress while simultaneously being able to understand and reply to more open-ended user utterances in a robust manner.

2 Spoken dialogue system (SDS) research

- **Where do you think the field of dialogue research will be in 5 to 10 years?** I think that the field of dialogue research is bound to rise in importance by a large margin. Besides personal assistants becoming vastly more capable and, ergo, more ubiquitous over the next few years, I suspect that a relatively new avenue for dialogue system research will open up – embedded SDSs. By that, I mean systems integrated with websites, electronic appliances, buildings, etc., acting as an interface layer of sorts, enabling querying dedicated databases containing information about a company, product, and the like, in natural language.
- **What do you think this generation of young researchers could accomplish in that time?** I think it will be a perfect time to be a young researcher. Even now, there is a vast array of opportunities relating to both implementing recent research findings in practice and also pushing the theoretical side of the field forward. The current climate of natural language processing research might be a bubble, at least to some extent, but it need not devalue the accomplishments of the researchers in the near future. It is a unique opportunity that should not be missed.

- **What kind of questions need to be investigated to get the field to that point?** I believe that the field can go far just with the current momentum. Nevertheless, I would like to see more attention being brought to ethical and social issues of SDSs deployment so that the field does not only go forward but also in the right direction. As the capabilities of such systems increase, so too do their influence on society and the responsibility of their creators.
- **What are the most important things for users of SDSs** In my opinion, one of the most important things for users of SDSs is the feeling that the system inhabits the same world as they do. I feel like there is still work to be done in regard to creating robust systems that not only talk in a natural, engaging way but also do not include false, nonsensical, or ambiguous information „not from this world” in their utterances without resorting to rule-based architectures.
- **Is there a difference between SDS research in academia and industry?** In the broader context of natural language processing research, currently, a growing divide between academia and industry can be observed in terms of available resources, with the industry leaving academia behind. Considering the current paradigm revolving around large language models, some predict academia to become relegated to a „secondary” role, being limited to, e.g., evaluation of models developed by private companies. I am very interested in alternative, more optimistic perspectives regarding this issue or ideas relating to preventing the pessimistic one.
- **Will SDSs be more widely used in the future? How? In what scenarios?** With the somewhat recent advances (and, more importantly, public recognition) of generative AI, I think that in the future, SDSs will be used way more widely, albeit we can expect to mainly see more conversational systems focused only on holding a natural-sounding conversation with a user without a task component involved. Nevertheless, I think that in the long run, this will also bring attention to research aimed at supplementing task-oriented SDSs with controlled (grounded) generative modules.

3 Suggested topics for discussion

- The importance of additional sources of knowledge in SDSs: global context, ontologies, external knowledge bases, grounding, etc. Are such solutions necessary in the long run? Or is further pure compute scaling sustainable?

- Is developing a robust test for measuring high-level natural language understanding that will not fall prey to the AI effect like the Turing test possible? What could it look like?
- Long-term perspectives for transformer-based models in the dialogue domain. Can we expect future breakthroughs without changing the current paradigm?
- Is there potential in integrating SDSs with embodied AI agents? What are the implications of such a synthesis?

References

- Eleni Adamopoulou and Lefteris Moussiades. 2020. Chatbots: History, technology, and applications. *Machine Learning with Applications* 2:100006. <https://doi.org/10.1016/j.mlwa.2020.100006>.
- Iwona Christop, Kacper Dudzic, and Mikołaj Krzymiński. 2023. Amusebot: Towards making the most out of a task-oriented dialogue system. *Progress in Polish Artificial Intelligence Research* 4, forthcoming.
- Ruo Chen Xu, Yuwei Fang, Chenguang Zhu, and Michael Zeng. 2021. Does knowledge help general nlu? an empirical study. *CoRR* abs/2109.00563. <https://arxiv.org/abs/2109.00563>.

Biographical sketch



Kacper Dudzic is a student at the Adam Mickiewicz University in Poznań (Poland), in his last year of a master’s degree in Artificial Intelligence. As a part of his master’s thesis, he is investigating the role of global context for user utterance disambiguation in dialogue systems. He also holds a master’s degree in Japanese Linguistics from the same university with a diploma thesis focused on linguistic issues of Japanese-Polish machine translation. His academic interests include natural language understanding in the context of dialogue systems, large language model research, and various other topics at the intersection of computer science and linguistics. Currently, he is employed at the Center for Artificial Intelligence at his home university. Before that, he worked as a language modeling engineer at VoiceLab.AI¹, the creator of the first Polish ChatGPT-like AI assistant, TRURL. He wants to start a Ph.D. in natural language processing next year.

¹<https://voicelab.ai/>