

# PhD Thesis Supervisor's Report

**Thesis Title:** Low resource methods for dialogue systems applications  
**Thesis Author:** Mgr. Vojtěch Hudeček  
**Supervisor:** Mgr. Ondřej Dušek, Ph.D.

## Thesis Topic

The topic of Vojtěch Hudeček's PhD thesis are low-resource methods for task-oriented dialogue, which is an extremely important problem for practical applications of dialogue systems and chatbots serving as assistants connected to in-domain databases and APIs. In practice, data is always scarce, especially data with annotation, and being able to build a dialogue system with limited data makes the technology much more widely available. The thesis addresses certain problems within traditional modular dialogue systems (mostly focusing on the language understanding step), as well as the new trend of end-to-end neural dialogue models.

## Contents Summary

The main research contribution of the thesis are Chapters 4-7, each of them based on a conference paper previously published by the author, but all chapters also include more detailed results or extensions and previously unpublished experiments. The individual contributions are as follows:

- *Discovering dialogue slots – unsupervised language understanding (Chapter 4)*. Here, a method for extracting important attributes in task-oriented dialogue (slots) is introduced, which uses weak supervision from existing off-the-shelf natural language processing tools (named entity recognition and semantic parsing). The approach features steps that greatly increase its practical usefulness as opposed to any previous works on the topic: merging semantically related slots or training a standalone tagger independent of the weak supervision sources. The result thus allows to bootstrap natural language understanding from dialogue data with no additional annotation and produces a useful slot tagger, where only a simple handcrafted mapping to a database interface needs to be provided by the system developer, a task of a few minutes.

This work started in 2019 under supervision of Zhou Yu (then University of California Davis), then was gradually extended under my supervision, and the resulting chapter is mainly based on the final 2021 paper published at the top-tier ACL conference. Compared to the ACL paper, the thesis includes an extension comparing the weak supervision with fully unsupervised slot candidate generation via memory networks' text copying feature as well as large language model (LLM) prompting.

- *End-to-end dialogue modelling with latent variables (Chapter 5)*. This set of experiments processes task-oriented dialogue in an end-to-end fashion, with minimal data annotation, and includes discrete latent variables expressing the current dialogue state or system action, thus making the resulting system explainable. The first approach featured here is based on variational recurrent neural networks and trains from data with only database access annotation; it is thus able to track its own dialogue state. The evaluation on multiple datasets shows significant progress against other unsupervised approaches. The experiments are further extended with a different hierarchical model architecture based on variational autoencoders, aiming to produce different latent spaces at different levels of the hierarchy.

The work started in 2020 and the chapter is based on the final AACL 2022 conference paper. The hierarchical extension is based on co-supervision with Milica Gašić and Nurul Lubis from Heinrich-Heine University of Düsseldorf and is previously unpublished.

- *End-to-end sequence-based multi-domain modelling (Chapter 6)*. This chapter mostly focuses on data handling in multiple domains and domain transfer in dialogue training. These approaches aim at reducing the need for in-domain annotated data while reusing annotated data from other domains. Here, the author works with two-stage end-to-end neural decoders based on the GPT-2 pretrained language model. He shows that unifying input across multiple data sources helps, but domain differences in the data remain and non-trivial amounts of in-domain data are required for good performance. The research is based on two collaboration papers: in part on an EMNLP 2021 paper on the GPT-2-based model, where Vojtěch mostly focused on

data handling and training, but mainly on an LREC 2022 paper on data unification, where Vojtěch was the lead author. The experiments in the thesis are extended to include training on smaller portions of in-domain data. The DIASER dataset accompanying the LREC paper is still one of the largest resources available for training task-oriented systems to date.

- *LLMs for task-oriented dialogue (Chapter 7)*. The final experimental chapter tackles LLMs, the latest technology available in natural language processing, and their ability to conduct task-oriented dialogue, i.e., adhere to constraints of the dialogue topic and play the role of an assistant while correctly accessing and presenting database information. The author proposed a novel approach for using prompted LLMs in multiple steps to conduct the whole dialogue end-to-end, similar to previous end-to-end trained models. The chapter is based on a SIGDIAL 2023 paper, with some additional experimental details.

The experimental chapters are accompanied by three introductory Chapters 1-3, detailing the motivation and research questions of the thesis and including a detailed theoretical background, as well as an overview of related works in the area. The final concluding Chapter 8 summarizes progress on all of the research questions and discusses potential further works. Overall, the text is very well structured and written in well understandable English.

## Work Progress and Evaluation

Vojtěch's PhD path was relatively complex. He changed his PhD topic and supervisor to be able to work on dialogue systems specifically, and I am really grateful that I had the opportunity to supervise his research. He worked very diligently and dealt gracefully with various setbacks, always aiming to improve his experiments based on feedback. He was able to come up with inventive solutions to the problems he encountered and constantly showed lots of creativity and a deep interest in the subject. This also helped him to stay up-to-date on his research approaches in this rapidly changing field, including the latest switch to LLMs. I believe that the resulting thesis is excellent, basing on and extending four high-quality publications, including one paper published at the ACL main conference, i.e., the top venue in the field.

While working on his PhD, Vojtěch was also involved in other research efforts that ultimately did not make it into his thesis, especially works done during his two industry internships. He also took part in other important academic activities contributing to the community, such as teaching and conference organization. Most notably, he was one of the main organizers for the famous Young Researchers' Roundtable on Spoken Dialogue Systems for several years.

## Recommendation

In what clearly follows from the detailed account above, my unequivocal recommendation is that Vojtěch Hudeček's thesis be **approved** for a PhD.

Prague, 1 February 2024



Mgr. Ondřej Dušek, Ph.D.

Institute of Formal and Applied Linguistics  
Charles University, Faculty of Mathematics and Physics  
V Holešovičkách 747/2, 18000 Praha 8