



Seminar Hora Informaticae

Institute of Computer Science, Prague

Tuesday, January 23, 2024, 14.00 – 15.30 (2 - 3:30 PM) CET

Meeting room 318, Address: Pod Vodárenskou věží 2, Prague 8

ZOOM Meeting ID: 954 7823 4977 , Passcode: 712564

ZOOM: <https://cesnet.zoom.us/j/95478234977?pwd=dXoyekFHbDJ0MkNrTjVVS3F2STZqUT09>



Ondřej Dušek (Institute of Formal and Applied Linguistics, Charles University Prague):

Getting Structure in Dialogue with Large Language Models

The current state of the art in text generation are large language models (LLMs) pretrained on vast amounts of text and finetuned to produce solutions given instructions. LLMs represent significant progress, allowing the user to request outputs for various tasks by stating a query in natural language and being able to follow examples provided by the user (in-context learning/prompting), without the need for further training (finetuning) on task-specific data. However, they retain some of the problems of the previous generation of language models, in particular their opacity and lack of controllability. This talk will show experiments on using LLMs with prompting only for multiple tasks: data-to-text generation, task-oriented dialogues, and dialogue evaluation. All these tasks operate with structure (structured data input, structured outputs, structured dialogue), which is not what these LLMs were specifically pretrained for. I show that LLMs are usable for these tasks, but also point out their limitations and potential areas of improvement.

References:

1. Are Large Language Models All You Need for Task-Oriented Dialogue?

<https://aclanthology.org/2023.sigdial-1.21>

2. Mind the Labels: Describing Relations in Knowledge Graphs With Pretrained Models.

<https://aclanthology.org/2023.eacl-main.176>

3. Three Ways of Using Large Language Models to Evaluate Chat.

<https://aclanthology.org/2023.dstc-1.14>

Ondřej Dušek (<https://ufal.mff.cuni.cz/ondrej-dusek>) is an Assistant Professor at Charles University in Prague, focusing on natural language generation and human-computer dialogue. His recent research focuses on generative language models, mostly applied to the data-to-text and dialogue response generation tasks. He is specifically interested in semantic accuracy and grounding in language generation, as well as ways of evaluating generation accuracy. After obtaining his PhD in Prague, he spent 2 years as a postdoc at Heriot-Watt University in Edinburgh in 2016-2018, where he also co-advised the university team in the Amazon Alexa Prize chatbot competition. He is currently the principal investigator of an ERC Starting Grant titled Next-Generation Natural Language Generation, which aims to adapt neural models in order to produce fluent, accurate and explainable NLG systems.

HORA INFORMATICAЕ (meaning: TIME FOR INFORMATICS) is a broad-spectrum scientific seminar devoted to all core areas of computer science and its interdisciplinary interfaces with other sciences and applied domains. Original contributions addressing classical and emerging topics are welcome. Founded by Jiří Wiedermann, the seminar is running since 1994 at the Institute of Computer Science of the Czech Academy of Sciences in Prague.

<https://www.cs.cas.cz/horainf>