Title: Instructing Language Models to Do Reasoning Wisely

Abstract:
Reasoning in natural language is an amazing human ability. The NLP community has been collecting numerous data to train and test language models on various reasoning tasks such as mathematical reasoning, commonsense reasoning, abductive reasoning, and counterfactual thinking. The training was conventionally tuning the model parameters on input-output pairs, like a math word problem and an answer. Large language models, hitting the world with their emergent abilities of doing tasks in a chat mode, have changed the methodology of teaching machines to do reasoning. In this talk, I will start from the key techs behind the large language models, introducing why "instructing" models to do reasoning in a chat model, instead of "training", is quite effective and becomes a fashion. I'll then present a few studies *very briefly* where psychological knowledge and/or human learning skills have inspired algorithm design to instruct the large language models to do reasoning wisely. The studies vary from solving complex math word problems in English, answering questions in a different language, explaining a verification on statements, to answering questions under counterfactual presuppositions. These works are accepted to top NLP or AI venues in 2023-2024. One received the outstanding paper award in EMNLP 2023.

Bio:
Meng Jiang is an Associate Professor in the Department of Computer Science and Engineering at the University of Notre Dame. He received B.E. and PhD from Tsinghua University. He was a visiting scholar at CMU and a postdoc at UIUC. He is interested in data mining, machine learning, and natural language processing. His data science research focuses on graph and text data for applications such as reasoning and planning, question answering, online education, mental healthcare, and recommender systems. He received the CAREER Award from the National Science Foundation. He has delivered 14 conference tutorials and organized seven workshops. He is a Senior Member of ACM and IEEE.