CS 59200: Can Machines Think? Reasoning with Large Language Models

Fall 2024 Instructor: Abulhair Saparov

Time and Location Time: TBD. Location: TBD

Overview:
Over the last few years, natural language processing (NLP) and artificial intelligence (AI), has gone through a paradigm shift, moving away from end-to-end training of task-specific models and moving towards large general-purpose foundational models. These models are not trained to perform a specific task, but rather allow for a general capability for dealing with language in very flexible ways. The key ingredients for these models are size (these models have an extremely large number of parameters) and massive datasets for training. The aim of this seminar is to study how these models work, the opportunities and risks they present and possible directions to avoid these risks.

Structure:
This seminar course will be a complementary mixture of lectures, paper discussions, and tool demonstrations. Lectures will provide an introduction to important concepts, while the paper discussions will provide an in-depth look at the state-of-the-art application of those concepts. Each paper will have a designated facilitator responsible for leading the discussion. To ensure a lively discussion, students will be responsible for reading assigned papers at sufficient level to summarize the research problem, the proposed solution, the relationship to existing work, and the evaluation of any claimed contributions. Tool demonstrations will similarly have a designated facilitator that is responsible for developing a demonstration of a tool that highlights both its strengths and weaknesses.

Evaluation:
Students will develop (in concert with the instructor) and work on an open-ended project over course of the semester. Ideally this will be an application of program reasoning to their personal interests. Students are welcome to pair up for more ambitious projects. The project will include a short final report and presentation during the last week of class.

Final grades will be assigned following
- Class Project   40%
- Facilitating Discussion   30%
- Participation   30%

Prerequisites:
As an advanced topic course, we assume that students already have a basic understanding of machine learning and NLP. The students must have taken either CS 373 (undergraduate ML), CS577 (graduate NLP), CS 578 (graduate ML) or an equivalent course. Project experience and good programming skills are a must, as the course project is an important part of this class.

Topics:*
- Background: Large language models
    - Symbolic language models
    - Neural language models
    - Large language models

- Background: Reasoning
    - Formal deductive reasoning
    - High-level intro to logic

- Reasoning with LLMs
    - Zero-/few-shot prompting
    - Chain-of-thought prompting
    - Prompt learning approaches
    - Neuro symbolic reasoning

- Learning Paradigms
    - Masked Language models
    - PPO and RL-based methods
    - Learning from Human Feedback

* This list of topics is tentative, and will be adjusted based on the interests of seminar participants.

Policy on Use of Generative AI
Students are permitted to use generative AI tools such as ChatGPT. However, we urge students to be wary of the output of such models on some tasks. These tools can be very effective for tasks such as paraphrasing or correcting grammar, but they do produce errors on other tasks, such as analysis of research papers or scientific scrutiny of an experimental setup.