Internship project, master 2, 2024

Title: Interactive Teaching of System 2 Reasoning for Large Language Models

Supervision: Cédric Colas, Pierre-Yves Oudeyer
Host team: Flowers team, Inria Bordeaux
Duration: 6 months around march – august 2024

How to apply: contact pierre-yves.oudeyer@inria.fr and ccolas@mit.edu with CV and letter of motivation, and using the **[application]** tag in the email object.

Keywords: language models, interactive teaching, prompt engineering, automatic prompting, retrieval-augmented LLMs, reasoning programs, program synthesis

Scientific context: The Flowers Lab at Inria Bordeaux studies open-ended learning in machines and humans, with a focus on curiosity-driven autotelic exploration (= learning by setting one's own goals) and on the use of language, and large language models, as cognitive tools (see Colas et al., 2023 paper for the directions we are exploring, see also Colas et al., 2020). The current project aims to address some of the challenges one needs to address to scale up these approaches.

TL;DR: The objective of this internship is to study different (adaptive and automated) prompting methods and interactive processes to improve the reasoning capabilities of large language models.

Project:
Pretrained large language models (LLMs) often perform well in arbitrary tasks defined in text at test time (e.g. Brown et al, 2020) but still fail in more challenging tasks involving multi-step thinking such as mathematical reasoning, decision-making, social reasoning, or commonsense (e.g. Wei et al, 2022). Borrowing from Kahneman's *dual process model*, we could see LLMs as implementing a form of System 1 (fast, intuitive thinking processes) but struggling with System 2 (slow, deliberate, rational thinking processes).

This project aims at studying possible ways a non-expert user could naturally teach forms of system 2 reasoning to LLMs. This will involve reviewing the existing literature on multi-step reasoning approaches (Wei et al., 2022, Nye et al., 2022, Wang et al., 2022, Dohan et al., 202, Yao et al., 2023), cognitive architectures for LLMs (Summers et al., 2023), reasoning in code (Maadan et al., 2022) and the links between reasoning and code, and prompt engineering approaches, identifying their limitations and suggesting new approaches.

One particular relevant approach is to frame reasoning as the generation of programs to structure and decompose reasoning tasks into smaller steps. This internship could look into new ways of leveraging real-time natural feedback from non-expert users to guide the generation of these reasoning programs. As LLMs learn to generate them, they gain the ability to think for themselves in more systematic ways.

Requirements

We are looking for motivated Master 2 (MSc) students with a background in machine learning, in particular in deep NLP (experience with LLMs and prompt engineering will be a plus), as well as interest in cognitive sciences. Additionally the student should know Python and have good oral and written communication skills.

References

- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, *33*, 1877-1901.
- Colas, C., Karch, T., Moulin-Frier, C., & Oudeyer, P. Y. (2022). Language and culture internalization for human-like autotelic AI. Nature Machine Intelligence, 4(12), 1068-1076.
- Colas, C., Karch, T., Lair, N., Dussoux, J. M., Moulin-Frier, C., Dominey, P., & Oudeyer, P. Y. (2020). Language as a cognitive tool to imagine goals in curiosity driven exploration. Advances in Neural Information Processing Systems, 33, 3761-3774.
- Dohan, D., Xu, W., Lewkowycz, A., Austin, J., Bieber, D., Lopes, R. G., ... & Sutton, C. (2022). Language model cascades. *arXiv preprint arXiv:2207.10342*.
- Sumers, T., Yao, S., Narasimhan, K., & Griffiths, T. L. (2023). Cognitive architectures for language agents. *arXiv preprint arXiv:2309.02427*.
- Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, et al. Chain-of-thought prompting elicits reasoning in large language models. Advances in Neural Information Processing Systems, 35:24824–24837, 2022
- Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc Le, Ed Chi, Sharan Narang, Aakanksha Chowdhery, and Denny Zhou. Self-consistency improves chain of thought reasoning in language models. arXiv preprint arXiv:2203.11171, 2022.
- Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Thomas L Griffiths, Yuan Cao, and Karthik Narasimhan. Tree of thoughts: Deliberate problem solving with large language models. arXiv preprint arXiv:2305.10601, 2023.
- Maxwell Nye, Anders Johan Andreassen, Guy Gur-Ari, Henryk Michalewski, Jacob Austin, David Bieber, David Dohan, Aitor Lewkowycz, Maarten Bosma, David Luan, et al. Show your work: Scratchpads for intermediate computation with language models. arXiv preprint arXiv:2112.00114, 2021.
- Madaan, A., Zhou, S., Alon, U., Yang, Y., & Neubig, G. (2022). Language models of code are few-shot commonsense learners. arXiv preprint arXiv:2210.07128.