

Large-scale auto-formalization of mathematical theories: why, how, and why now?

vendredi 12 juillet 2024 12:15 (20 minutes)

A machine readable and verifiable account of a large portion of human mathematics would change the way mathematicians can work, learn and collaborate. While impressive progress has been made in the mathematical standard libraries of proof assistants like Lean, Isabelle and Coq, the proportion of mathematical results formalized in such systems remains tiny overall. In the talk, I will argue that it is only with automated machine learning systems and hybrid human-machine systems that we will be able to clear this backlog, and that making large corpora of natural language mathematics available in proof assistants will be critical for progress in automated theorem proving.

An automated system to formalize entire theories requires several components: large language models specialized for tasks like statement formalization, proof sketch translation or proof refactorings, smaller language models for tree-search based proof search [1] as well as data-efficient reinforcement learning loops for continual learning. Drawing inspiration from the field of code generation [2], I argue that contemporary models are potent enough to make a significant contribution to formalization efforts when moving from zero-shot inference on artificial benchmarks to hierarchical and iterative methods on challenging real-world use cases (cf. [3],[4]).

[1] Gloeckle, F., Roziere, B., Hayat, A., & Synnaeve, G. (2023). Temperature-scaled large language models for Lean proofstep prediction. In The 3rd Workshop on Mathematical Reasoning and AI at NeurIPS'23.

[2] Roziere, B., Gehring, J., Gloeckle, F., Sootla, S., Gat, I., Tan, X. E., ... & Synnaeve, G. (2023). Code llama: Open foundation models for code. arXiv preprint arXiv:2308.12950.

[3] Yang, J., Jimenez, C. E., Wettig, A., Lieret, K., Yao, S., Narasimhan, K., & Press, O. (2024). SWE-agent: Agent Computer Interfaces Enable Software Engineering Language Models.

[4] Jiang, A. Q., Welleck, S., Zhou, J. P., Li, W., Liu, J., Jamnik, M., ... & Lample, G. (2022). Draft, sketch, and prove: Guiding formal theorem provers with informal proofs. arXiv preprint arXiv:2210.12283.

Auteur principal: GLOECKLE, Fabian (Ecole des Ponts ParisTech)

Orateur: GLOECKLE, Fabian (Ecole des Ponts ParisTech)

Classification de Session: Contributed Talks

Classification de thématique: Contributed talks