Search, Reason or Recombine? Paradigms for Scaling Formal Proving

Fabian Glöckle

Meta / Ecole des Ponts

In the effort to scale test-time computation for language models on mathematical benchmarks, two prominent paradigms have emerged: large-scale search with reinforcement learning, exemplified by methods like AlphaProof, and long chain-of-thought reasoning with emergent self-verification, as seen in models like o1. For the future of reinforcement learning in formal theorem proving, this opens up a spectrum of hybrid methods. These range from line-level tree search with environment feedback to multi-turn iterative whole proof generation, with and without integrated informal reasoning, to hierarchical problem decompositions and recombination of partial proofs. I will explain these methods as inference methods and discuss the challenges faced when applying reinforcement learning to them.