Prelegent: dr Maciej Besta, ETH Zurich

Temat: Reasoning Language Models: A Blueprint

Abstrakt:

Reasoning language models (RLMs), also known as Large Reasoning Models (LRMs), such as OpenAI's o1 and o3, DeepSeek-R1, and Alibaba's QwQ, have redefined AI's problem-solving capabilities by extending LLMs with advanced reasoning mechanisms. Yet, their high costs, proprietary nature, and complex architectures - uniquely combining reinforcement learning (RL), search heuristics, and LLMs - present accessibility and scalability challenges. To address these, we propose a comprehensive blueprint that organizes RLM components into a modular framework, based on a survey and analysis of all RLM works. This blueprint incorporates diverse reasoning structures (chains, trees, graphs, and nested forms), reasoning strategies (e.g., Monte Carlo Tree Search, Beam Search), RL concepts (policy, value models and others), supervision schemes (Outcome-Based and Process-Based Supervision), and other related concepts (e.g., Test-Time Compute, Retrieval-Augmented Generation, agent tools). We also provide detailed mathematical formulations and algorithmic specifications to simplify RLM implementation. By showing how schemes like LLaMA-Berry, QwQ, Journey Learning, and Graph of Thoughts fit as special cases, we demonstrate the blueprint's versatility and unifying potential. To illustrate its utility, we introduce x1, a modular implementation for rapid RLM prototyping and experimentation. Using x1 and a literature review, we provide key insights, such as multi-phase training for policy and value models, and the importance of familiar training distributions. Finally, we discuss scalable RLM cloud deployments and we outline how RLMs can integrate with a broader LLM ecosystem. Our work demystifies RLM construction, democratizes advanced reasoning capabilities, and fosters innovation, aiming to mitigate the gap between "rich AI" and "poor AI" by lowering barriers to RLM design and experimentation. Finally, our overview will dive deeper into the reasoning "in the real world" by presenting a recent design for effective and low-cost AI assistants (Knowledge Graph of Thoughts).