Scaling Thought: Efficient Reasoning with LLM Agents for Complex Real-World Tasks

PI: Jiawei Zhou, Co-PI: Niranjan Balasuramanian

Project Overview

Abstract. Recent advances in Large Language Models (LLMs) have enabled impressive gains in complex reasoning, powering applications in math, coding, and real-world decision-making agents. Models like OpenAI's o1/o3 and DeepSeek R1 demonstrate the use of *reinforcement learning (RL)* to unlock advanced reasoning through slow, deliberative thinking—generating long chains of intermediate steps across thousands of tokens. This paradigm, known as test-time scaling, shows that increased inference effort can improve task performance.

However, this trend introduces significant cost. As reasoning length grows, so do the computational and energy demands, leading to rising financial and environmental impact. At scale, these inefficiencies threaten the deployability and sustainability of LLM-based systems. There is a critical need for models that can reason effectively with fewer steps and lower compute.

To address this, we propose to develop RL-driven methods that enable LLMs to perform advanced reasoning more efficiently. Our approach will design algorithms that shorten reasoning chains by compressing intermediate steps and guiding the model to retain only essential information. These techniques will adapt to the structure of different problems and reduce both token length and memory usage while preserving accuracy.

We will validate our methods in LLM-based agents for complex real-world tasks such as AppWorld, an interactive environment developed by our team, where agents must perform long-horizon reasoning for tasks involving planning, tool use, and adaptation. These real-world scenarios offer a strong benchmark for evaluating efficiency and generalization.

Intellectual Merit and Broader Impact. This project addresses a core challenge in LLM development: enabling advanced reasoning with reduced computational cost. We propose a reinforcement learning framework that jointly optimizes for correctness and efficiency, guided by rewards shaped by reasoning length, token usage, and intermediate step utility. The approach integrates reasoning pattern compression to distill long autoregressive traces into minimal, effective paths, and uses curriculum learning to gradually scale task complexity. By combining imitation learning with RL fine-tuning on open-source models like DeepSeek R1, we aim to develop adaptive, task-aware LLMs. These methods will be evaluated in AppWorld, a benchmark environment requiring tool use, planning, and long-horizon decision-making.

The research advances both the theoretical foundations and practical deployment of efficient LLM agents, with broad implications for sustainable AI. By reducing compute and energy demands, it addresses key challenges in scalable AI deployment. The resulting methods will support real-world applications in education, accessibility, and automation. Open-source outputs and integration into graduate education and student research will extend the project's impact while reinforcing Stony Brook's leadership in frontier AI research.

Alignment with Seed Grant Objectives. Given the rapid pace of this research area, seed funding is essential to generate preliminary results and establish feasibility. These early outcomes will position us to pursue competitive external funding and **help maintain Stony Brook University's leadership in cutting-edge AI research.**