



The Hong Kong University of Science and Technology

Department of Mathematics

PhD THESIS EXAMINATION

Towards Better Reasoning in Large Language Models

By

Mr. Xin XU

ABSTRACT

Reasoning is a fundamental aspect of human intelligence. Yet evaluating whether large language models (LLMs) truly reason, and training them to reason better, remain fundamental open questions. This thesis advances both fronts in two parts, developing rigorous evaluation methodology alongside effective training methods for reasoning. In Part I, we introduce challenging benchmarks for undergraduate-level mathematics and physics. For mathematics, we design a dynamic benchmark that generates multiple randomized variants of each problem and measures whether LLMs solve them consistently, exposing striking inconsistency where models produce correct solutions in one case while failing on simple number perturbations. For physics, we construct a large-scale bilingual benchmark that extends reasoning evaluation beyond mathematics, revealing that reasoning ability acquired in mathematics transfers only weakly to physics, even for math-specialized models. In Part II, we advance reinforcement learning with verifiable rewards (RLVR) by improving both training efficiency and data utility. We identify two key bottlenecks in current practice: the prohibitive cost of generating long reasoning rollouts during training, and the diminishing training signal as prompts become too easy. We propose two complementary techniques: a policy initialization method that leverages the surprising finding that training on thinking-free data improves slow-thinking performance, yielding a stronger and more efficient starting point for RLVR; and a prompt composition strategy that combines easy problems into harder composite ones, sustaining the training signal throughout learning. Together with extensions on training dynamics, entropy stabilization, inference efficiency, and multimodal investigation, these methods form a cohesive framework for reasoning-oriented RLVR. By pairing systematic measurement of reasoning capabilities with practical methods that overcome key training bottlenecks, this thesis contributes towards better reasoning in LLMs.

Date : 16 July 2026, Thursday

Time : 4:00 pm

Venue : Rm 4472 (Lifts 25/26), 4/F Academic Building, HKUST

Thesis Examination Committee:

Chairman : Prof. Qi YING, ENVR/HKUST

Thesis Supervisor : Prof. Can YANG, MATH/HKUST

Member : Prof. Yang XIANG, MATH/HKUST

Member : Prof. Dong XIA, MATH/HKUST

Member : Prof. Jun ZHANG, ECE/HKUST

**External Examiner : Prof. Yixiang FANG, School of Data Science and School of Artificial Intelligence/ The Chinese University of Hong Kong, Shenzhen (via online mode)
(Open to all faculty and students)**

The student's thesis is now being displayed on the reception counter in the General Administration Office (Room 3461).