香港浸會大學
HONG KONG BAPTIST UNIVERSITY

DEPARTMENT OF
COMPUTER SCIENCE
計算機科學系

TMLR Young Scientist SEMINAR

2025 SERIES

Trustworthy Machine Learning and Reasoning Group

Dr. Xuandong Zhao

Postdoctoral Researcher,
Computer Science,
UC Berkeley.

Date: 18 June 2025 (Wednesday)

Time: 10:00 – 11:00 (HKT)

Meeting: https://hkbu.zoom.us/j/6603117755

Learning to Reason without External Rewards

ABSTRACT

Recent studies have demonstrated that training large language models (LLMs) for complex reasoning via Reinforcement Learning with Verifiable Rewards (RLVR) is effective. However, such a paradigm is limited by reliance on costly and domain-specific supervision. This talk presents methods for enhancing Large Language Model (LLM) reasoning without external supervision. The first part introduces self-certainty, a scalable metric for Best-of-N selection that leverages a model's own probability distribution to estimate response quality, outperforming traditional approaches on both closed- and open-ended tasks. The second part introduces Reinforcement Learning from Internal Feedback (RLIF), where models are trained using intrinsic confidence (via self-certainty) as the sole reward. We showcase *Intuitor*, an RLIF algorithm that promotes structured reasoning and generalization without external rewards. The talk concludes with a survey of concurrent work and open challenges in unsupervised LLM reasoning and learning.

BIOGRAPHY

Dr. Xuandong Zhao is a postdoctoral researcher at the University of California, Berkeley, working with Prof. Dawn Song. He is affiliated with the Responsible Decentralized Intelligence (RDI) and Berkeley Artificial Intelligence Research (BAIR) labs. Xuandong earned his Ph.D. in Computer Science from UC Santa Barbara, where he was advised by Profs. Yu-Xiang Wang and Lei Li. Prior to that, he completed his Bachelor's degree in Computer Science at Zhejiang University. Xuandong's research interests focus on Machine Learning, Natural Language Processing, and AI Safety, with a particular emphasis on responsible and reliable generative AI.

ENQUIRY

Email: bhanml@comp.hkbu.edu.hk