

QINGYAO LI

Email: ly890306 AT sjtu.edu.cn

Address: 800 Dongchuan Road, Shanghai, China, 200240

EDUCATION

Shanghai Jiao Tong University

2022.9-present

Ph.D. candidate in Computer Science and Technology

Zhiyuan Honorary Doctoral Program

Advisor: Weinan Zhang, Yong Yu

GPA: 3.91/4.00

Nanyang Technology University

2025.9 - 2026.3

Ph.D. Exchange in College of Computing and Data Science

Advisor: Bo An

Xi'an Jiao Tong University

2018.9 - 2022.8

B.S in Automation (Qian Xuesen Class)

GPA: 4.05/4.3 (Rank: 1/25)

Xi'an Jiao Tong University Young Talent Class

2016.9 - 2018.8

Preparatory studies

Suzhou High School (2016-2017) and Xi'an Jiao Tong University (2017-2018)

RESEARCH INTERESTS

My research focuses on **LLM reasoning and code intelligence**. I develop **search- and learning-based algorithms** (e.g., MCTS, RL) that integrate structured test-time reasoning with fine-grained execution feedback. By connecting planning, verification, and learning, my work aims to guide exploration and enhance the robustness of LLM-based code agents on complex programming tasks.

Previously, I also worked on intelligent education problems, particularly **learning path recommendation using reinforcement learning**, which shaped my interest in sequential decision making and optimization—skills that now directly inform my research on LLM reasoning and code generation.

PUBLICATIONS / PREPRINTS

AdverMCTS: Combating Pseudo-Correctness in Code Generation via Adversarial Monte Carlo Tree Search

Qingyao Li, Weiwen Liu, Weinan Zhang, Yong Yu, Bo An

ICML 2026 [[Paper](#)]

We propose AdverMCTS, an adversarial Monte Carlo Tree Search framework for robust code generation. It couples a Solver agent that searches for code candidates with an Attacker agent that actively discovers targeted corner test cases, forming a dynamic and progressively hostile verification environment that reduces pseudo-correctness and improves generalization beyond public tests.

ATGen: Adversarial Reinforcement Learning for Test Case Generation

Qingyao Li, Xinyi Dai, Weiwen Liu, Xiangyang Li, Yasheng Wang, Ruiming Tang, Yong Yu, Weinan Zhang

ICLR 2026 [[Paper](#)]

We introduce ATGEN, a novel framework that trains the test case generator using adversarial reinforcement learning. By creating a dynamic curriculum of challenging bugs, ATGEN learns to produce highly effective test cases that significantly outperform prior methods reliant on static datasets. Our experiments validate ATGEN’s superiority and demonstrate its practical utility as a powerful filter for selecting correct code and as a high quality reward source for training more capable code generation models.

Learning Structure and Knowledge Aware Representation with Large Language Models for Concept Recommendation

Qingyao Li, Wei Xia, Kounianhua Du, Qiji Zhang, Weinan Zhang, Ruiming Tang, Yong Yu

AAMAS 2026 Extended Abstract [Paper]

This work introduces human knowledge systems into concept recommendation by leveraging the world knowledge of large language models (LLMs) in the form of textual representations. To address the anisotropy issue in these representations, we further propose a graph-based text adapter. Experimental results demonstrate that the proposed method outperforms previous ID-Text, Text-only, and ID-only recommendation methods.

RethinkMCTS: Refining Erroneous Thoughts in Monte Carlo Tree Search For Code Generation

Qingyao Li, Wei Xia, Kounianhua Du, Xinyi Dai, Ruiming Tang, Yasheng Wang, Yong Yu, Weinan Zhang

EMNLP 2025 [Paper]

We present RethinkMCTS, a novel approach for enhancing code generation in LLMs using Monte Carlo Tree Search (MCTS). It conducts thought-level searches prior to code generation, leveraging fine-grained execution feedback to refine errors and guide reasoning along correct paths. Empirically, it improves pass@1 from 70.12 to 89.02 on GPT-3.5-turbo and from 87.20 to 94.51 on GPT-4o-mini.

NL-Debugging: Exploiting Natural Language as an Intermediate Representation for Code Debugging

Weiming Zhang*, **Qingyao Li***, Xinyi Dai, Jizheng Chen, Kounianhua Du, Weiwen Liu, Yasheng Wang, Ruiming Tang, Yong Yu, Weinan Zhang

EMNLP 2025 [Paper]

We introduce NL-DEBUGGING, which leverages natural language as an intermediate representation to enhance code debugging efficiency and accuracy significantly. Experimental results show that NL-DEBUGGING outperforms traditional code-level debugging, especially in addressing deep algorithmic flaws. Our results suggest that using natural language sketches is the most effective format for code debugging.

CodePRM: Execution Feedback-enhanced Process Reward Model for Code Generation

Qingyao Li, Xinyi Dai, Xiangyang Li, Weinan Zhang, Yasheng Wang, Ruiming Tang, Yong Yu

ACL 2025 Findings [Paper]

We introduced CODEPRM, a novel framework that leverages execution feedback to build a process reward model for code generation. Our approach collects thought steps enriched with code pass rates, code snippets, and execution feedback, and trains a feedback-enhanced code PRM to score individual thought steps. At inference time, CODEPRM serves as a process verifier within a Generate-Verify-Refine (GVR) pipeline, enabling dynamic error correction during code search. Experimental results demonstrate that CodePRM significantly outperforms strong baselines. Further analysis identifies key factors for building code PRMs.

Privileged Knowledge State Distillation for Reinforcement Learning-based Educational Path Recommendation

Qingyao Li, Wei Xia, Liang Yin, Jiarui Jin, Yong Yu

KDD 2024 [Paper]

We address the challenge of dynamic environments in reinforcement learning (RL) for learning path recommendation by utilizing privileged feature distillation. The dynamic nature arises because each student has a unique knowledge state, making it difficult for RL models to achieve effective training. To tackle this, we propose incorporating the knowledge states as privileged features and utilizing a distillation approach to ensure the model acquires and retrain the ability to adapt. Experimental results show significant improvements in the recommendation performance of representative RL models.

Graph Enhanced Hierarchical Reinforcement Learning for Goal-oriented Learning Path Recommendation

Qingyao Li, Wei Xia, Li'ang Yin, Jian Shen, Renting Rui, Weinan Zhang, Xianyu Chen, Ruiming Tang, Yong Yu

CIKM 2023 [Paper]

We propose a Graph-Enhanced Hierarchical Reinforcement Learning (GEHRL) framework for learning path recommendation. In this framework, a high-level model assigns sub-goals to a low-level model, while the low-level model recommends specific concepts to students. To promote the completion of these sub-goals, the low-level model incorporates a test-based intrinsic reward mechanism. GEHRL achieves state-of-the-art (SOTA) performance on three student simulation benchmarks.

Adapting Large Language Models for Education: Foundational Capabilities, Potentials, and Challenges

Qingyao Li, Lingyue Fu, Weiming Zhang, Xianyu Chen, Jingwei Yu, Wei Xia, Weinan Zhang, Ruiming Tang, Yong Yu

Arxiv [Paper]

This work provides a comprehensive survey of the educational capabilities of large language models (LLMs). We conceptualize the problem-solving process of LLMs as the integration of five key abilities: mathematical reasoning, writing proficiency, programming skills, logical reasoning, and knowledge-based question answering. Additionally, we summarize the common methodologies employed to enhance these abilities, offering a reference framework to guide the future development of large-scale educational models.

RESEARCH INTERNSHIP

Algorithm Research Intern, Xiaohongshu

2026.3 - present

Mentor: Wenxiang Jiao, Yuan Lu

Student Researcher, Huawei Noah's Ark Lab

2024.7 - 2025.5

Mentor: Wei Xia, Xinyi Dai, Ruiming Tang

HONORS AND AWARDS

Huawei Excellent Intern

2025.5

Outstanding Graduate of Xi'an Jiao Tong University

2022.7

International Excellence Award in Baidu Big Data Competition

2020.9

National Second Prize in Mathematical Modeling

2019.10

Qiu Changrong Second-Class Scholarship

2019.9

SERVICES

Conference Reviewer: ICML 2026, KDD2024, CIKM2023, AAAI2023

PATENTS

Concept Recommendation System Integrating Human Knowledge Structures

2024

(Under Application)

A Wireless Charger for Pacemakers

2022.12

(Utility Model Patent)