

Xuecheng Liu

Email: xuechen2@andrew.cmu.edu | [Google Scholar](#) | [Website](#)

EDUCATION

Carnegie Mellon University

January 2026 - December 2027

M.S. in Electrical and Computer Engineering (GPA: 4.00/4.00)

Zhejiang University & University of Illinois Urbana-Champaign

Dual Degree Program

September 2021 - June 2025

B.E. in Computer Engineering (GPA: 3.92/4.00), Zhejiang University

B.S. in Computer Engineering (GPA: 3.80/4.00), University of Illinois Urbana-Champaign

Dean's List every semester

PUBLICATIONS

- **Xuecheng Liu***, Daman Arora*, Gokul Swamy, and Andrea Zanette. "Message Passing Enables Efficient Reasoning." *COLM 2026, Oral Spotlight (24/856 accepted papers)*. [arXiv:2607.01077]
- Alvin Zhang*, **Xuecheng Liu***, Zixuan Wang*, Fahim Tajwar, Daman Arora, Ruslan Salakhutdinov, Daniel Khashabi, Yuda Song, and Andrea Zanette. "Harness Learning Enables Generalizable Test-Time Adaptation." *COLM 2026 Workshop on Lifelong Agents (LLA), Oral Presentation*. [arXiv:2609.35738]

* Equal contribution.

RESEARCH EXPERIENCE

Message Passing Parallel Reasoning in Large Language Models

September 2025 - April 2026

Advisor: Andrea Zanette, Carnegie Mellon University

- Designed Message Passing Language Models (MPLMs), an inference-time parallel reasoning framework with persistent LLM threads and explicit point-to-point communication; implemented the execution harness on top of vLLM.
- Designed SFT and evaluation pipelines for Sudoku and 3-SAT; fine-tuned Qwen3-0.6B on 1,000 puzzles to achieve 72% accuracy on 25×25 Sudoku, compared with 20% for GPT-5 Pro prompted without tools.
- Extended MPLMs to training-free long-context reasoning on LongBench-v2, achieving a 2.2× end-to-end inference speedup over Recursive Language Models (RLMs) at comparable accuracy using Qwen3.6-35B-A3B.

Harness Learning for LLM Agents

July 2026 - Present

Advisor: Andrea Zanette, Carnegie Mellon University

- Designed and implemented the harness evolution pipeline for multi-hop question answering, including the retrieval environment, harness execution, proposer training, and iterative evaluation.
- Trained a Qwen3-4B proposer with Markov GRPO over 10 successive harness-revision steps, using the current harness and execution feedback at each step with per-step rewards. Improved HotpotQA exact match by 19.9%.
- Demonstrated out-of-distribution generalization to MuSiQue (+7.8%) and 2WikiMultihopQA (+25.5%) exact-match gain with no additional training.

Training Language Models to Fail Gracefully

August 2026 - Present

Advisor: Andrea Zanette, Carnegie Mellon University

- Designed and implemented the evaluation framework for failure behavior and fallback in LLMs, measuring success-error trade-offs and reliability frontiers across mathematical reasoning and sequential decision-making tasks.
- Evaluated generalization and failure behavior across MATH500, AIME, and AbstentionBench, including comparisons with RLCR, RLOO/GRPO, and other policy-gradient baselines.
- Analyzed model abstention and uncertainty behavior under different caution levels, and out-of-distribution transfer.

PROJECT EXPERIENCE

Upgrade Testing for Distributed Systems

- Built an AST-based pipeline to convert mini-cluster tests into distributed-system upgrade tests, identifying six potential upgrade failures in Apache Cassandra.

A Linux-like Basic Operating System for x86 Architecture Processors

- Implemented a 32-bit operating system in C and x86 assembly, including virtual memory, process scheduling, system calls, interrupt handling, synchronization, and I/O.

SKILLS

Programming Languages: Python, C, C++, x86 assembly, Java, MATLAB

Frameworks & Tools: PyTorch, CUDA, vLLM, Transformers, verl, Linux, Slurm, Weights & Biases

PHOTOGRAPHY (Portfolio: xuechengliu.com)

- 2025 Sony World Photography Awards - Open Competition, Architecture Winner
- 2025 Le Prix de la Photographie de Paris (PX3) - Gold Winner
- 2025 Weibo Landscape Photographer of the Year
- 2024 NASA Astronomy Picture of the Day (APOD)