

Yuhan Chi

Undergraduate, School of Mathematical Sciences, Fudan University

yhchi25@m.fudan.edu.cn | chi-shan0707.github.io | github.com/Chi-Shan0707 | Online CV

PROFILE

First-year undergraduate at Fudan University, majoring in Mathematical Sciences and pursuing a double degree in Information and Computing Science and Artificial Intelligence. My current research direction is math-first AI evaluation and trustworthy sequential decision-making: when model signals are reliable, when they are protocol- or domain-dependent, and how optimization, control, and reinforcement learning can give sharper language for trust and action under uncertainty.

EDUCATION

Fudan University

Shanghai, China

B.S. Candidate, Information and Computing Science & Artificial Intelligence, School of Mathematical Sciences Sep. 2025 – Present

- GPA: **3.96/4.00** in the most recent semester; ranked **10th** in cohort, top **5%**.
- Coursework and self-study interests: mathematical analysis, linear algebra, probability, optimization, algorithms, reinforcement learning, and LLM evaluation.

RESEARCH INTERESTS

- **Mathematical interpretability and AI trust:** grounding claims about model reliability in inspectable mathematical and statistical evidence.
- **Reasoning verification and evaluation:** studying when trace-level signals are diagnostic, misleading, protocol-dependent, or domain-specific.
- **Optimization, control, and sequential decision-making:** using mathematical tools to reason about uncertainty, feedback, constraints, and interaction with environments.

SELECTED RESEARCH AND OPEN-SOURCE PROJECTS

token-verification-mirage

[Paper](#) | [GitHub](#) | [AI4Math Workshop](#)

- Solo-author project with full pipeline ownership: dataset selection, model deployment, trace generation, evaluation design, analysis, figures, related-work organization, and writing.
- Poster at the **ICML 2026 Workshop on AI for Math (AI4Math)**.
- Audits how evaluation protocol choices such as pooling, in-sample scoring, and direction-agnostic AUROC can change apparent token-level verification performance.
- Finds that shallow token statistics can be useful diagnostics, but should not be treated as stable standalone verifiers without within-problem evaluation, fixed-direction baselines, and permutation-null calibration.

code-not-text

[GitHub](#) | [Interactive Demo](#)

- Solo measurement study of cheap CoT-surface and token-trajectory features on **DeepSeek-R1-0528-Qwen3-8B** across **31,040** math, science, and coding runs.
- Reports strong math signal (AoA **0.958**, AUROC@100% **0.982**), partial science signal (AoA **0.799**), and weak coding transfer (AoA **0.434**).
- Tests coding-side robustness through an 83-scalar feature sweep, grouped ablations, a CoT-only judge, nonlinear MLPs, SSL pre-training, semantic-knot annotation, and token-level de-knotting.
- Frames the result as **measurement non-invariance**: the same feature family tracks convergence-like behavior in math but not executable correctness in coding.

TinyLoRA-GRPO-Coder

[GitHub](#)

- Solo project adapting a small-parameter training idea toward verifiable competitive-programming code generation on Qwen2.5-Coder-3B.
- Uses compile-and-run rewards rather than static heuristics; built as a full research-system exercise covering data processing, training, multi-GPU setup, reward design, evaluation, and validation.

microgpt.cpp

[GitHub](#)

- Minimal GPT implementation from first principles in C++, built to understand model internals without relying on

high-level deep-learning frameworks.

ACADEMIC SERVICE

- Reviewer, [ICML 2026 Workshop on AI for Math \(AI4Math\)](#), 2026.

TECHNICAL SKILLS

Programming	Python, C++, C, Java, Node.js, Lean 4
ML / AI	PyTorch, ML experimentation, LLM evaluation, reinforcement-learning basics
Tools	Git, GitHub, Linux, LaTeX, Markdown, VS Code
Other	Mathematical proof writing, technical writing, self-directed research, competitive programming

COMMUNITY AND ADDITIONAL INFORMATION

- Maintains [github-unflag-playbook-cn](#), a Chinese playbook documenting GitHub account flagging, recovery cases, and appeal workflows.
- Contributor to [FDUGuideBook/nav-site](#), a Fudan community navigation project.
- Public work and notes are maintained at [chi-shan0707.github.io](#) and [github.com/Chi-Shan0707](#).