



Junjie Nian 年骏杰

B.Eng. in Artificial Intelligence, Fudan University

EMAIL jjnian24@m.fudan.edu.cn

PHONE (+86) 19370555931

GITHUB github.com/JunjieNian

WEBSITE junjienian.com

PROFILES [Google Scholar](#) · [DBLP](#) · [Semantic Scholar](#) · [ACL Anthology](#) · [ORCID](#)

ABOUT

Undergraduate researcher at Fudan University majoring in Artificial Intelligence (ranked 1st in class, GPA 3.93/4.0). At FudanNLP ALEX Lab, advised by Prof. Yixin Cao, my research interests include reasoning models, AI agents, interpretability, and computational social science. My publications include work accepted to the EMNLP 2026 Main Conference and Findings of ACL 2026. Will be an exchange student at UC San Diego in Fall 2026.

EDUCATION

2024.9 – Present	Fudan University — B.Eng. in Artificial Intelligence School of Computer Science (Honors Track). GPA 3.93/4.0, ranked 1st in major.
2026.9 – 2027.1	UC San Diego — Exchange Student Computer Science & Engineering.
2021.9 – 2024.6	The High School Affiliated to Fudan University — Buqing Honors Class

RESEARCH EXPERIENCE

2025.10 – Present	FudanNLP ALEX Lab — Undergraduate Researcher Advised by Prof. Yixin Cao. Research on Large Reasoning Model capabilities, agent systems, and mechanistic interpretability. Publications include work accepted to the EMNLP 2026 Main Conference and Findings of ACL 2026.
2025.3 – 2025.6	Embodied Intelligence Research Project — Tengfei Innovation Program Advised by Prof. Shang Li. Built small-scale robots with asynchronous servo motor control and embedded development board integration.

PUBLICATIONS

Kang Chen*, **Junjie Nian***, Yixin Cao, Yugang Jiang. “Disagree to Explore, Agree to Commit: Routing-Guided Test-Time Scaling for Software Agents.” [\[arXiv\]](#) [\[Website\]](#) [\[Code\]](#)

Kang Chen, Minshen Yu, **Junjie Nian**, Yaoning Wang, Yixin Cao, Yugang Jiang. “Does the Same Token Mean the Same State? MoE Routing as Signal for Reasoning Control.” [\[arXiv\]](#) [\[Website\]](#)

Junjie Nian*, Kang Chen*, Ge Zhang, Yixin Cao, Yugang Jiang. “TraceGraph: Shared Decision Landscapes for Diagnosing and Improving Agent Trajectories.” [\[arXiv\]](#) [\[Code\]](#)

Kang Chen*, **Junjie Nian***, Yixin Cao, Yugang Jiang. “SliceGraph: Mapping Process Isomers in Multi-Run Chain-of-Thought Reasoning.” [\[arXiv\]](#) [\[Code\]](#)

Kang Chen*, Zhuoka Feng*, Sihan Zhao, Kai Xiong, **Junjie Nian**, Yaoning Wang, Changyi Xiao, Yixin Cao. “NEX: Neuron Explore-Exploit Scoring for Label-Free Chain-of-Thought Selection and Model Ranking.” [\[arXiv\]](#)

Kang Chen*, Fan Yu*, **Junjie Nian**, Sihan Zhao, Zhuoka Feng, Zijun Yao, Heng Wang, Minshen Yu, Yixin Cao. “Thinking Traps in Long Chain-of-Thought: A Measurable Study and Trap-Aware Adaptive Restart.” [Findings of ACL, ACL 2026](#). [\[arXiv\]](#) [\[ACL Anthology\]](#)

Zhuoka Feng*, Kang Chen*, Sihan Zhao†, Kai Xiong†, Yaoning Wang†, Minshen Yu, **Junjie Nian**, Changyi Xiao, Yixin Cao, Yugang Jiang. “ARM: Role-Conditioned Neuron Transplantation for Training-Free Generalist LLM Agent Merging.” [EMNLP 2026 Main Conference](#). [\[arXiv\]](#) [\[Website\]](#)

* Equal contribution † Equal second-author contribution

HONORS & AWARDS

Duckbill Special Scholarship, Fudan University (equivalent to First-Class Outstanding Student Scholarship)	2025
Third Prize, China Undergraduate Mathematical Contest in Modelling, Shanghai	2025
Third Prize, The Chinese Mathematics Competitions for College Students, Shanghai	2025
Freshman Scholarship, Fudan University (New Engineering Discipline)	2025
First Prize, Tengfei College Entrance Exam Freshman Scholarship, Fudan University	2024

SELECTED PROJECTS

Disagree to Explore, Agree to Commit: Routing-Guided Test-Time Scaling for Software Agents

Uses native sparse-MoE routing traces to steer exploration and commitment within an agent attempt, then select among four independently developed final patches without an external judge or selection-time tests.

[\[Details\]](#) [\[Paper\]](#) [\[Code\]](#)

TraceGraph: Shared Decision Landscapes for Diagnosing and Improving Agent Trajectories

Graph-based framework that converts multi-model agent rollouts into a shared decision landscape with productive cores, trap regions, and repair events. Uses this process view to diagnose benchmark behavior and support trap-aware recovery strategies for improving agent trajectories.

[\[Details\]](#) [\[Paper\]](#) [\[Code\]](#)

SliceGraph: Mapping Process Isomers in Multi-Run Chain-of-Thought Reasoning

Post-hoc graph view of multi-run Chain-of-Thought reasoning built from similarity between reasoning slices. Shows that runs reaching the same final answer often travel through distinct process families and shared intermediate reasoning states, revealing process geometry missed by final-answer aggregation.

[\[Details\]](#) [\[Paper\]](#) [\[Code\]](#)

From Fluency to Accountability: AI-Like Peer Reviews at ICLR 2026

Claim-lifecycle analysis of 53,463 public reviews and 416,862 claims, tracing whether review concerns receive author responses, reviewer follow-up, and meta-review uptake without treating AI-likeness as proof of AI use.

[\[Details\]](#) [\[Paper\]](#) [\[Code\]](#)

CoT Folding: Structural Analysis of LLM Reasoning Trajectories

Framework treating Chain-of-Thought as analogous to protein folding. Measures pairwise similarity between token segments using neuron activation patterns to reveal compact cores, long-range return connections, and drifting branches. Proposes Native Fold Score (NFS), a fully unsupervised, parameter-free quality metric for CoT reasoning.

[\[Details\]](#) [\[Paper\]](#) [\[Code\]](#) [\[Demo\]](#)

Reasoning Across Languages: Activation Agreement and Language Heuristics

Tests label-free trajectory selection across 2,250 mathematical reasoning traces in 18 languages. Activation-based ConsensusMin reaches 25.6% accuracy versus 20.8% for the fixed-first baseline, while language identity and entropy heuristics remain unreliable correctness proxies.

[\[Details\]](#) [\[Paper\]](#) [\[Code\]](#)

Entropy Is Not Frequency: Auditing the 80/20 Structure of Reasoning Tokens

Audits the minority high-entropy pattern across 51 caches, 539,466 trajectories, and 1.116 billion token positions. Separates position- and type-level estimates, compares five uncertainty metrics, and shows that frequency is only a weak entropy proxy.

[\[Details\]](#) [\[Paper\]](#) [\[Code\]](#)

SKILLS

Programming: Python, C/C++

AI / Research: PyTorch, NumPy/SciPy, neuron-level interpretability analysis, data analysis & visualization

Engineering: React, FastAPI, Three.js, Docker, LaTeX

Math & Physics: National Physics Olympiad Second Prize (x2); full marks in all core math & physics courses

ENGLISH PROFICIENCY

TOEFL 111 CET-6 619 CET-4 636