

NAICHENG YU

Computer Science & Engineering Department, University of California, San Diego, USA
Email: n7yu@ucsd.edu

EDUCATION

University of California, San Diego

Master of Science in Computer Science and Engineering

La Jolla, CA, United States

Sep 2025 – Mar 2027 (Expected)

University of California, Berkeley

Visiting Student

Berkeley, CA, United States

Jan 2024 – May 2024

- Courses: Introduction to Artificial Intelligence, Data Structures, Efficient Algorithm and Intractable Problems

Zhejiang University

Bachelor of Engineering in Software Engineering

Hangzhou, Zhejiang, China

Sep 2021 – Jul 2025

- Selected awards: Zhejiang University Scholarship - First Prize, Zhejiang Provincial Government Scholarship, Gold Award of the 8th Zhejiang College Students “Internet Plus” Innovation and Entrepreneurship Competition

RESEARCH EXPERIENCE

University of California, San Diego (CSE Department)

Research Assistant to Professor Jingbo Shang

San Diego, U.S.

Jun 2026 – Present

Meta-Delta: Evolving Knowledge through Global Review for LLM-Driven Optimization (NeurIPS 2026)

Workshop on Meta-Agents, co-first author)

- **Led the implementation and experimental evaluation of Meta-Delta**, an LLM-driven optimization framework that jointly evolves candidate solutions and structured search knowledge through evidence-grounded Delta Cards and Global Review.
- **Built and ran large-scale evaluations** across 11 mathematical, scientific, and systems optimization tasks and both 118-task SCILAWS-REAL and SCILAWS-PARALLEL settings, achieving best or tied-best results on 10 of 11 tasks and leading aggregate performance on key SCILAWS metrics.

Can LLMs Discover Scientific Laws in Real and Parallel Settings? (ICLR 2027 submission)

- Implemented and ran the **search-and-selection analysis**, conducting inference-time search experiments across **118 scientific-law tasks** to compare self-selection with oracle best-of-N; also served as one of five expert annotators for human validation of the benchmark’s LLM judges.

University of Illinois Urbana-Champaign (Electrical & Computer Engineering Department)

Research Assistant to Professor Huan Zhang

Remote

Sep 2025 – Present

Your Benchmark Is an Empirical Measure Over Difficulty (NeurIPS 2026, co-first author)

- Formalized LLM benchmarks as empirical measures over task difficulty, deriving Wasserstein and max-gap diagnostics for distribution mismatch, rank instability, and missed local improvements.
- Built an IRT-based benchmark analysis pipeline over 25 benchmarks across 7 domains with Bayesian difficulty estimation and robustness diagnostics. Designed BENCHFILL, an adaptive LLM-based gap-filling algorithm that generates and calibrates new items at targeted difficulty levels.

How do Visual Attributes Influence Web Agents? A Comprehensive Evaluation of User Interface Design Factors (Findings of ACL 2026, co-first author)

- Developed WebAgentsLook, an automated evaluation pipeline that generates 200+ CSS-controlled HTML variants (color/typography/layout/position/clarity) to find how visual design alone biases multimodal web agents’ decisions while keeping content and task semantics fixed.
- Conducted large-scale empirical studies on real-world e-commerce/travel scenarios and SOTA agents, identifying sensitivities such as strong color-saliency effects and attention misalignment under positional shifts.

MonitorBench: A Comprehensive Benchmark for Chain-of-Thought Monitorability in Large Language Models (COLM 2026)

- Designed the **agent-task component** of MonitorBench by adapting ShadeArena environments into paired malicious and benign dual-objective scenarios, enabling evaluation of long-horizon agent monitorability while controlling for safety-alignment confounds.

University of California, San Diego (CSE Department)

Research Assistant to Professor Jishen Zhao

San Diego, U.S.

Jan 2026 – Mar 2026

Multi-Agent Memory from a Computer Architecture Perspective: Visions and Challenges Ahead (Architecture 2.0 2026 Workshop, second author)

- Articulated multi-agent memory consistency as a core open problem, separating it into update-time visibility/ordering and read-time conflict resolution for semantically heterogeneous artifacts.
- Formulated an architecture-inspired memory hierarchy for agent systems (I/O, cache, and memory layers) and analyzed two underexplored protocol problems: inter-agent cache sharing and structured memory access control.

TEACHING & MENTORING EXPERIENCE

Teaching Assistant

San Diego, U.S.

Principles in Computer Architecture (CSE 240A), University of California, San Diego

Sep 2026 – Dec 2026

- Graded and designed course assignments, held office hours, and answered student questions on Piazza, providing guidance on computer architecture concepts and coursework.

Teaching Assistant

Zhejiang, China

Fundamentals of Software Engineering, Zhejiang University

Sep 2023 – Dec 2023

- Assisted the instructor in grading assignments, leading lab sessions, and providing project guidance. Explained core software engineering principles during practice sessions and helped students debug and refine their code.
- Conducted group discussions and offered personalized feedback to improve student engagement and teamwork.

ADDITIONAL INFORMATION

- **Language Proficiency:** Mandarin Chinese (Native), English (Fluent – TOEFL 108)
- **Programming Language and Tools:** C, C++, Java, Python, SQL, PyTorch, Pandas, NumPy
- **Development Frameworks:** Spring Boot (Back-end), Vue.js (front-end), Node.js (server-side)