Jianhao RUAN

MetaGPT, *Research Engineer*

aurorra1123.github.io | Email | +1 (213) 269–8016

Shenzhen, China & Los Angeles, US

Education

- **The Hong Kong University of Science and Technology (Guangzhou)** *Sep 2023 – Jun 2027 (Expected)*
  Bachelor of Science in Data Science Guangzhou, China
  Selected Coursework: Introduction to Artificial Intelligence (A), Deep Learning (A, Top Grade)

- **University of California, Los Angeles (UCLA)** *Jan 2026 – Jun 2026*
  Visiting Student / Visiting Researcher Los Angeles, CA, US

Selected Publications

- **Exposing Weaknesses of Large Reasoning Models through Graph Algorithm Problems** *June 2025 – Sept 2025*
  Qifan Zhang, **Jianhao Ruan***, Aochuan Chen, Kang Zeng, Nuo Chen, Jing Tang, Jia Li* *2026 ICLR*
  - Proposed a **graph-algorithm-based benchmark** for evaluating large reasoning models (LRMs), covering tasks such as shortest path, connectivity, cycle detection, and combinatorial graph traversal, with controlled difficulty and reasoning depth.
  - Designed **multi-step reasoning probes** that decouple surface-level pattern matching from true algorithmic reasoning, enabling fine-grained diagnosis of reasoning robustness.
  - Conducted large-scale evaluations across multiple LRMs, revealing **systematic failure modes** including error accumulation, brittle intermediate state tracking, and sensitivity to graph size and topology.

- **AOrchestra: Automating Sub-Agent Creation for Agentic Orchestration** *2025 – 2026*
  ***Jianhao Ruan**, Zhihao Xu, Yiran Peng, Fashen Ren, Zhaoyang Yu, et al.* *arXiv preprint* **HF: #2 Paper of the Day**
  - Proposed a unified, framework-agnostic **sub-agent abstraction** that models any agent as a tuple $\langle \text{MODEL}, \text{TASK}, \text{TOOLS}, \text{CONTEXT} \rangle$, enabling on-demand composition of specialized executors.
  - Introduced **AOrchestra**, an orchestration system where a central orchestrator concretizes the tuple at each step by curating task-relevant context, selecting tools/models, and delegating execution via **automatic sub-agent creation**.
  - Demonstrated **plug-and-play** compatibility with diverse agent executors, reducing human role engineering while improving adaptability for long-horizon, multi-turn tasks.
  - Achieved a **16.28% relative improvement** over the strongest baseline (paired with **Gemini-3-Flash**) across **GAIA, SWE-Bench, and Terminal-Bench**, while enabling controllable **performance–cost** trade-offs approaching Pareto efficiency.

- **AutoWebWorld: Synthesizing Infinite Verifiable Web Environments via Finite State Machines** *2025 – 2026*
  *Yifan Wu, Yiran Peng, Yiyu Chen, **Jianhao Ruan**, et al.* *Under Review*
  - Developed a **verifiable web environment synthesis** pipeline by modeling websites as **Finite State Machines (FSMs)** with explicit state/action/transition rules.
  - Generated **11,663 verified trajectories** across **29** environments at **$0.04 per trajectory**, improving downstream Web GUI agent performance on benchmarks such as **WebVoyager**.

- **AutoEnv: Automated Environments for Measuring Cross-Environment Agent Learning** *2025*
  *Jiayi Zhang, Yiran Peng, Fanqi Kong, Cheng Yang, Yifan Wu, . . . , **Jianhao Ruan**, et al.* *arXiv preprint*
  - Built an automated benchmark framework for evaluating **cross-environment agent learning** via diverse generated environments.
  - Released **AutoEnv-36** with **36 environments** and **358 validated levels**, enabling systematic comparison of learning methods and generalization.

Research Internship

- **DeepWisdom AI, MetaGPT** *Sept 2025 – Present*
  Role: Agent Research Engineer Intern
  - Developed and evaluated **GUI-based agents** for web navigation, reproducing **FARA** results on **WebVoyager** and **Online-Mind2Web**, with a focus on visual grounding and action space design.
  - Conducted systematic analysis of the **VisualWebArena** benchmark, including environment abstraction, observation structure, and action space constraints for scalable GUI-agent evaluation.
  - Led a core research project on **dynamic sub-agent creation and orchestration**, enabling a main agent to instantiate task-specialized sub-agents with customized instructions and action spaces.
  - Evaluated the proposed framework across multiple agentic benchmarks including **GAIA** (**Avg@1: 80.17**), **SWE-Bench**, and **TerminalBench**, demonstrating improved cross-environment generalization.
  - Reproduced and compared strong agent baselines such as **OpenHands**, **ReAct**, and **Mini-SWE-Agent**, performing controlled analyses on planning, tool usage, and execution robustness.
  - Collaborated closely with senior researchers on **agent framework abstraction and system-level design**, contributing to experimental pipelines and research discussions aligned with ongoing **ICLR 2026** submissions.

Skills

- **Agent Frameworks & Paradigms:** ReAct, ReCode, OpenHands, Mini-SWE-Agent, Web Agents, Claude Code, CodeX
- **Benchmarks & Environments:** GAIA, SWE-Bench, TerminalBench, VisualWebArena, WebVoyager, Online-Mind2Web
- **Machine Learning & LLMs:** PyTorch, LLM Inference (vLLM), Prompting, Tool-augmented LLMs
- **Systems & Tooling:** Docker, Bash, Git, Linux, LaTeX, SSH, Tmux
- **Programming Languages:** Python, C++

Honors and Awards

- **National Finalist (Top 2.23%)** *Jul 2025*
  National Undergraduate Electronics Design Contest
- **Gold Medal & Nomination Prize for Best New Basic Part** *Oct 2024*
  International Genetically Engineered Machine (iGEM) Competition
- **Freshman Scholarship** *Sep 2023*
  The Hong Kong University of Science and Technology (Guangzhou)
- **Outstanding Graduate Award (Top 10%)** *Jun 2023*
  Shenzhen Senior High School

Leadership & Service

- **Senate Member** *Jul 2025 - Present*
  The Senate of HKUST(GZ)
- **Community Leader** *Oct 2023 - Jun 2024*
  Residential College of HKUST(GZ)
- **Vice President** *Sep 2021 - Jun 2022*
  Student Union, Shenzhen Senior High School (Group)

Additional Information

Languages: English (Fluent), Chinese (Native)
Interests & Activities: Reading, Traveling, Photography, Music; Member of the Varsity Basketball Team, Student Press Squad, and Photography Society.