Zhengxu Yu

Huawei Technologies R&D UK
London, United Kingdom

yuzxfred@gmail.com
https://zhengxuyu.github.io/
https://dblp.org/pid/246/3155

RESEARCH OVERVIEW

My research concerns the generalisation limits of learning systems: how far competence acquired from finite experience extends beyond the distribution that produced it. Existing accounts of distribution shift buy their guarantees with assumptions that cannot be checked at deployment, and the methods built on them have proved hard to separate from plain empirical risk minimisation. My interest is in the case they do not reach, where the new regime is not a perturbation of the training one and no amount of interpolation within what was seen will produce the right behaviour.
What is missing there is exploration that can leave the region the training data describes rather than search inside it, and I take that to depend on how the agent represents its environment. Dynamics learned as executable programs, written from the agent's own interaction and revised wherever prediction and observation disagree, form a hypothesis that can be rejected outright and replaced in a way a distributed representation does not allow. Whether this is what governs behaviour in an unfamiliar regime is a question about held-out environments and has to be settled by running there: currently on ARC-AGI-3, and in self-organising multi-agent populations.

EDUCATION

Ph.D. Computer Science 2017–2021	Zhejiang University Advisors: Prof. Deng Cai and Prof. Xiaofei He Thesis: Stable Learning for Non-I.I.D Data
M.Sc. Computer Science 2015–2016	University of Surrey Advisor: Prof. H. Lilian Tang Thesis: CNN-based Mycobacterium Cells Segmentation for Time-lapse Images
B.Eng. Electrical Engineering 2011–2015	Jilin University

EMPLOYMENT

Research Scientist
Project Leader (PL)
Jan 2026–Present

Huawei Technologies R&D UK, London
- Led development of a generative code world model in which agents explore open-ended environments and encode the rules they discover as executable, self-improving code, reaching 100% Relative Human Action Efficiency (RHAE) on the ARC-AGI-3 community leaderboard.
- Project Leader for two feature teams, code world models (10 researchers and engineers) and OneManCompany (8), owning technical direction and delivery for each.
- Led research and system design for **OneManCompany**, an open-source agentic AI platform for self-organising multi-agent systems; 400+ GitHub stars, with the design published as arXiv:2604.22446.
https://github.com/1mancompany/OneManCompany

Member of Technical Team
Aug 2025–Jan 2026

Inephany Ltd., London
- As the company's first hire, built the multi-agent reinforcement learning training infrastructure for large language models and vision transformers, supporting distributed training runs on clusters of 500+ GPUs.

Algorithm Expert
2024–2025
Senior Algorithm Engineer
2021–2024

Apsara Lab (now Tongyi Lab), Alibaba Cloud
- Developed RL-based post-training methods for LLM reasoning efficiency, alignment and agentic tool use, cutting reasoning-token usage by 50% on Qwen-7B at no accuracy loss (NeurIPS 2025, AAAI 2026).
- Developed multi-agent reinforcement learning for city-scale traffic signal control, reducing mean travel time by 2.8% against the CoLight state of the art on real-city road networks of up to 196 intersections, and deployed it in production urban traffic signal control (IJCAI, IEEE TKDE, IJMLC).
- Developed progressive transfer learning under distribution shift, raising Rank-1 accuracy by up to 4.5 points over the MGN backbone across four person re-identification benchmarks (MSMT17, Market-1501, CUHK03, DukeMTMC-reID) (IEEE TIP).
- Delivered three bespoke LLM agent systems with dynamic reasoning and tool use for decision support and information retrieval, deployed for the International Olympic Committee and several municipal government bureaus in China.
- Shipped those three systems and two further large-scale enterprise solutions as tech lead and core algorithm owner, contributing to over RMB 25M in cumulative GAAP-recognised revenue.
- Led feature teams over four years, latterly 13 engineers and a 40-person data-annotation group, and supervised 8 research interns and junior engineers; promoted to Algorithm Expert in 2024.

Research Intern
2018–2021

DAMO Academy, Alibaba Group
- Developed large-scale multi-agent reinforcement learning and deep representation learning methods for the traffic engine of Alibaba City Brain; 7 granted national invention patents.

SELECTED PUBLICATIONS

The $*$ symbol denotes equal contribution.

Conference Papers

C3 Weihang Pan, Zhengxu Yu, Yong Wu, Xun Liang, Zhongming Jin, Qiang Fu, Penghui Shang, Binbin Lin, Xiaofei He, Jieping Ye. "FGD-Align: Pluralistic Alignment for Large Language Models via Fuzzy Group Decision-Making" *AAAI Conference on Artificial Intelligence (AAAI)*, 2026.

C2 Yuxiang Zhang*, Zhengxu Yu*, Weihang Pan, Zhongming Jin, Qiang Fu, Deng Cai, Binbin Lin, Jieping Ye. "TokenSqueeze: Performance-Preserving Compression for Reasoning LLMs" *Neural Information Processing Systems (NeurIPS)*, 2025.

C1 Zhengxu Yu*, Shuxian Liang*, Long Wei, Zhongming Jin, Jianqiang Huang, Deng Cai, Xiaofei He, Xian-Sheng Hua. "MaCAR: Urban Traffic Light Control via Active Multi-agent Communication and Action Rectification" *International Joint Conference on Artificial Intelligence (IJCAI)*, 2020.

Journal Articles

J4 Weihang Pan, Binbin Lin, Yafei Wang, Zhengxu Yu, Xinkui Zhao, Xiaofei He, Jieping Ye. "Cooperative Driving at Multiple Unsignalized Intersections in Fully Autonomous Driving Scenarios" *IEEE Transactions on Intelligent Transportation Systems (T-ITS)*, vol. 26, no. 12, 2025.

J3 Chao Xiang, Zhongming Jin, Zhengxu Yu, Xian-Sheng Hua, Yao Hu, Wei Qian, Kaili Zhu, Deng Cai, Xiaofei He. "Optimizing Traffic Efficiency via a Reinforcement Learning Approach Based on Time Allocation" *International Journal of Machine Learning and Cybernetics (IJMLC)*, vol. 14, no. 10, 2023.

J2 Xin Guo, Zhengxu Yu, Pengfei Wang, Zhongming Jin, Jianqiang Huang, Deng Cai, Xiaofei He, Xian-Sheng Hua. "Urban Traffic Light Control via Active Multi-Agent Communication and Supply-Demand Modeling" *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, vol. 35, no. 4, 2023. (Journal extension of C1.)

J1 Zhengxu Yu, Dong Shen, Zhongming Jin, Jianqiang Huang, Deng Cai, Xian-Sheng Hua. "Progressive Transfer Learning" *IEEE Transactions on Image Processing (TIP)*, vol. 31, 2022.

Technical Reports

P3 Zhengxu Yu, Ke Xu, Meng Fang, Haitham Ammar, Weilin Luo, Jun Wang. "The Value of Thinking: Learning When to Reason in Hybrid Decision Models" *arXiv preprint arXiv:2611.16961*, 2026. Paper.

P2 Weihang Pan, Zhengxu Yu, Yuxiang Zhang, Zhongming Jin, Binbin Lin, Xiaofei He, Jieping Ye. "ChainPrune: Evaluating and Reducing Redundancy in Long Chain-of-Thought Reasoning" *arXiv preprint arXiv:2608.21860*, 2026. Paper.

P1 Zhengxu Yu, Yu Fu, Zhiyuan He, Yuxuan Huang, Ka Yiu Lee, Meng Fang, Weilin Luo, Jun Wang. "OneManCompany: An Open-Source Operating System for Self-Organizing Multi-Agent Systems" *arXiv preprint arXiv:2604.22446*, 2026. Paper.