

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, that is, the extent to which competence acquired from finite experience extends beyond the distribution that produced it. Existing theory characterises generalisation within a fixed distribution, but for regimes absent from training no comparable account exists, and in its absence the requirements on model architecture and the objectives of training remain underdetermined. Three questions follow, namely how this boundary should be characterised over a distribution of tasks, how it can be extended, and how it can be measured on held-out environments rather than asserted. My position is that what a policy achieves in environments outside the training distribution is determined by the form of the model it has learned of the environment dynamics, and not by the reward it was trained to maximise. I therefore study model-based formulations in which the learned dynamics are represented explicitly and revised during interaction, expressed as executable programs, corrected from the discrepancy between predicted and observed transitions, and assessed through the policy's own rollouts rather than through an externally supplied reward. Generative code world models and self-organising multi-agent systems are the environments in which these questions are currently examined.

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 Engineer 2011–2015	Jilin University

EMPLOYMENT

Research Scientist 2026–Present	Huawei Technologies R&D UK, London - Generative code world models: agents explore open-ended environments and encode the rules they discover as executable, self-improving code. 100% RHAE on the ARC-AGI-3 community leaderboard. - Research and system design for OneManCompany (OMC) , an open-source agentic AI platform for self-organising multi-agent systems. https://github.com/1mancompany/OneManCompany (Github 400+ star)
---	--

Member of Technical Team 2025–2026	Inephany Ltd., London - Multi-agent reinforcement learning Infra for deep learning models training, including large language models and vision transformers.
Algorithm Expert 2024–2025	Apsara Lab (now Tongyi Lab), Alibaba Cloud - Research on LLM post-training (Reinforcement learning based methods): efficiency, alignment and agentic tool use, related work published at NeurIPS and AAAI. - Delivering Bespoke LLM agent systems with dynamic reasoning and tool use for enterprise operations-optimisation problems. - Led a feature team, supervised research interns and junior engineers.
Senior Algorithm Engineer 2021–2024	Apsara Lab (now Tongyi Lab), Alibaba Cloud - Research on multi-agent reinforcement learning, related works published in IEEE TKDE and IJMLC. - Progressive transfer learning under distribution shift, published in IEEE TIP. - Shipped five or more large-scale commercial AI projects and proof-of-concept demos, contributing to over RMB 25M in cumulative GAAP-recognised revenue.
Research Intern 2018–2021	DAMO Academy, Alibaba Group - Large-scale multi-agent reinforcement learning and deep representation learning. Granted 7+ national invention patents.

SELECTED PUBLICATIONS

The * symbol denotes equal contribution.

Peer-Reviewed Papers

- C4** 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” *ISWC*, 2026.
- 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 Supply-Demand Modeling” *International Joint Conference on Artificial Intelligence (IJCAI)*, 2020.

Journal

- 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.
- J1** Zhengxu Yu, Zhongming Jin, Long Wei, Jianqiang Huang, Deng Cai, Xiaofei He, Xian-Sheng Hua. “Progressive Transfer Learning” *IEEE Transactions on Image Processing (TIP)*, 2022.

Technical Reports

- 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](#).