Wenjun Li

wjli.2020@phdcs.smu.edu.sg • +65 8941 2807 • homepage

EDUCATION

Singapore Management University, Singapore   **Aug 2020 - Mar 2025**
PhD in Computer Science, Supervisor: Pradeep Varakantham
- Presidential Doctoral Fellowship for 2023 and 2024

University of Southern California, USA   **Aug 2018 - May 2020**
MSc in Electrical Engineering, GPA 3.6/4.0

China Jiliang University, China   **Sep 2014 - Jun 2018**
BEng in Opto-electronics Information Science and Engineering, GPA 3.7/4.0
- Excellent Graduate; 1st Class Scholarship for Academic Excellence

Work Experience

Huawei Noah's Ark Lab – Singapore   **Aug 2024 - present**

AI Researcher (Mar 2025 – present)
- **Deep Research Project**: Contributed to the full lifecycle of a deep research project, including methodology proposal, pipeline design and optimization, model training, evaluation, and deployment.

Intern Researcher (Aug 2024 – Dec 2024)
- **Adaptive Tool Use in LLMs**: Introduced a novel algorithm, MeCo, for adaptive tool use in LLMs. MeCo captures emergent representations of high-level cognitive phenomena that quantify the LLM's meta-cognitive scores, guiding decisions on external tool usage.

Geely – China   **Jun 2019 – Aug 2019**

Intern Algorithm Engineer
- Applied the Kanade-Lucas-Tomasi (KLT) algorithm for moving object detection (MOD) within the Active Parking Assist (APA) framework. Designed an algorithm to filter noisy flows, mitigated the impact of moving vehicles, and employed clustering methods to identify and extract information about moving objects.

Research Experience

Large Language Models   **Jan 2024 - present**

LLM × RL Training
- **Critical Analysis and Replication of R1-Zero Training (Github repo with more than 1k stars)**: investigated the impact of base models and RL on LLM reasoning, identifying pretraining biases and optimization inefficiencies in GRPO. Proposed Dr. GRPO to improve token efficiency and developed a minimalist R1-Zero recipe achieving 43.3% accuracy on AIME 2024 with a 7B model, setting a new state-of-the-art.

LLM Agent
- **Adaptive Tool Use in LLMs**: Introduced a novel algorithm, MeCo, for adaptive tool use in LLMs. MeCo captures emergent representations of high-level cognitive phenomena that quantify the LLM's meta-cognitive scores, guiding decisions on external tool usage.

LLM Fine-tuning and Data Quality Control
- **Enhanced Fine-tuning Strategies**: Extensively explored the impact of fine-tuning on the planning capabilities of LLMs. Developed the Clustering-based Maximum Diversity Sampling (CMDS) algorithm to select diverse and representative data in the embedding space, significantly improving the sample efficiency for fine-tuning.

Reinforcement Learning: Training Generalizable Agents and Humans   **Jun 2021 – May 2024**

Training Humans
- **Developed 3D Emergency Response Training Game:** Conceptualized and created an immersive 3D emergency response training game using Unity. Designed scenarios that challenge players to acquire and apply emergency response knowledge, including the use of diverse medical equipment for patient rescue.
- **Designed Adaptive Human Training Algorithm:** Developed an adaptive training algorithm based on the similarity and transferability of training tasks, enhancing training efficiency and overall outcomes.

* Equal contribution

Training Agents
- **Enhanced Environment Generation with Novelty Quantification:** Employed the GMM model to dynamically compute state-action space coverage and quantify the novelty score of environments. Developed a novelty-driven framework to improve existing environment generation algorithms.
- **Integrated Meta-Reinforcement Learning Techniques:** First, encode student policy with validation performance on environments selected by Quality Diversity (QD). Then, integrated meta-reinforcement learning with environment generation to train a versatile teacher agent capable of generalizing across student agents with diverse learning properties. The student policy is encoded with Quality Diversity.
- **Developed Novel Distance Metric:** Introduced a distance metric based on the occupancy measure of agent trajectories to quantify environment distance. Utilized this metric to measure the diversity of training environments, achieving state-of-the-art generalization performance.
- **Proposed a Novel Reward Signal:** Designed a reward mechanism based on marginal benefits to guide teacher agent in generating "right" training environments, resulting in improved computational efficiency and leading performance among RL-based environment generation algorithms.

Selected Publications

Large Language Models
- Liu*, Chen*, **Li Wenjun***, et al. Understanding R1-Zero-Like Training: A Critical Perspective. *[COLM 2025]*
- **Li Wenjun**, et al. Adaptive Tool Use in Large Language Models with Meta-Cognition Trigger. *[ACL 2025]*
- **Li Wenjun**, Chen Changyu, Varakantham P. Unlocking Planning Capabilities of Large Language Models with Clustering-Based Maximum Diversity Sampling. *[NAACL-findings 2025]*

Reinforcement Learning
- **Li Wenjun***, Sidney Tio*, Varakantham P. RL as a Surrogate: Teacher Algorithms for Human Learning. [preprint]
- Teoh Jayden*, **Li Wenjun***, Varakantham P. Improving Environment Novelty Quantification for Effective Unsupervised Environment Design. ***[NeurIPS 2024 Oral]***
- Li D, **Li Wenjun**, Varakantham P. Marginal Benefit Driven RL Teacher for Unsupervised Environment Design. *[AAAI 2025 Oral]*
- **Li Wenjun**, Varakantham P. Unsupervised Training Sequence Design: Efficient and Generalizable Agent Training. *[AAAI 2024]*
- **Li Wenjun**, Li D, Varakantham P. Generalization through Diversity: Improving Unsupervised Environment Design. *[IJCAI 2023]*

Additional

- Languages: English, Mandarin, German
- Hobbies: football, badminton, reading

* Equal contribution