

Changling Li (李常凌)

+41 079-613-4902 | changlingli.mpi@gmail.com | changlingli.com

[in](#) LinkedIn | [G](#) GitHub | [g](#) Google Scholar | [X](#)

“And now I see with eye serene, the very pulse of the machine.” – William Wordsworth

RESEARCH INTEREST

AI Safety · AI Governance · Multi-Agent Systems · Reinforcement Learning

EDUCATION

- **Max Planck Institute for Intelligent Systems** starting Sept 2026
PhD Student Tübingen, Germany
 - Supervisor: Maksym Andriushchenko and Rediet Abebe
- **ETH Zürich** Sept 2022 - June 2026
Master of Science in Computer Science Zurich, Switzerland
 - Grade: ~ 5.4/6
 - Major: Machine Intelligence; Minor: Theoretical Computer Science
- **Colby College** Sept 2018 - May 2022
Bachelor of Arts in Physics (Astrophysics) and Computer Science with Honors Maine, United States
 - GPA: 3.99/4 (Top 5%, Distinction in both majors)
- **Li Po Chun United World College** Sept 2016 - June 2018
International Baccalaureate Bilingual Diploma Hong Kong
 - Grade: 41/45

PUBLICATIONS

*=EQUAL CONTRIBUTION, C=CONFERENCE, J=JOURNAL, S=IN SUBMISSION

- [S.1] **Changling Li**, Terry Jingchen Zhang, Jie Zhang, Zhijing Jin, Sahar Abdelnabi, Maksym Andriushchenko. “**Decomposing and Measuring Evaluation Awareness.**” *arXiv preprint arXiv:2605.23055.*, 2026.
- [C.4] David Guzman Piedrahita, **Changling Li**, Dave Banerjee, Terry Zhang, Kevin Blin, Samuel Simko, Punya Pandey, Irene Strauss, Rada Mihalcea, Bernhard Schölkopf, Zhijing Jin. “**Position: Safe Models Do Not Guarantee Safe Societies.**” *Proceedings of the 43rd International Conference on Machine Learning. (ICML Spotlight)*, 2026.
- [C.3] Choi Younwoo*, **Changling Li***, Yongjin Yang, and Zhijing Jin. “**Agent-to-Agent Theory of Mind: Testing Interlocutor Awareness among Large Language Models.**” *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing (EMNLP Main)*, 2025.
- [C.2] **Changling Li**, Zhang-Wei Hong, Pulkit Agrawal, Divyansh Garg, and Joni Pajarinen. “**ROER: Regularized Optimal Experience Replay.**” *Reinforcement Learning Conference (RLC)*, 2024.
- [J.1] **Changling Li**, Ying Li. “**Scaling up Energy-Aware Multi-Agent Reinforcement Learning for Mission-Oriented Drone Networks with Individual Reward.**” *IEEE Internet of Things Journal (IoT-J)*, 2024.
- [C.1] Ying Li, **Changling Li**, Jiyao Chen, and Christine Roinou. “**Energy-Aware Multi-Agent Reinforcement Learning for Collaborative Execution in Mission-Oriented Drone Networks.**” In *2022 International Conference on Computer Communications and Networks (ICCCN)*, pp. 1–9. IEEE, 2022.

GRANTS, AWARDS & SCHOLARSHIPS

Research: FAR AI travel grant; Future of Life Institute travel grant; Cooperative AI sponsorship; Coefficient Giving sponsorship.

Undergrad Honors: Distinction in both majors; Magna Cum Laude (top 5%); Phi Beta Kappa; Sigma Pi Sigma; UWC Davis Scholar; Dean’s List F’18, S’19, F’19, F’21 (2020 and S’21 - cancelled due to COVID-19).

Colby College Institutional Grant: 45,065\$/year × 4 years undergraduate scholarship (90% scholarship, acceptance rate < 5%), Colby College, 2018 - 2022.



UWC Davis Scholarship: 20,000\$/year × 4 years undergraduate scholarship, Davis UWC Scholars, 2018 - 2022.

UWC National Committee Scholarship: HKD 278,000/year × 2 years high school scholarship (full scholarship, acceptance rate < 1%), United World College, 2016 - 2018.

ACADEMIC SERVICE

- **Trustworthy AI for Good Workshop at ICML 2026, Co-organizer**, 2026.
- **European Workshop on Reinforcement Learning, Reviewer**, 2026.
- **Reinforcement Learning Conference, Technical Reviewer**, 2025, 2026.

TEACHING EXPERIENCE

- **Teaching Assistant** *Sept 2019 - May 2022*
Department of Computer Science, Colby College 
 - Courses: CS 353 Interactive System; CS 251 Data Analysis and Visualization; CS 231 Data Structure and Algorithm; CS 152 Computational Thinking: Science; CS 151 Computational Thinking: Visual Media.
- **Teaching Assistant** *Sept 2019 - May 2021*
Department of Physics and Astronomy, Colby College 
 - Courses include: PH 241 Modern Physics I; PH 242 Modern Physics II.

COMMUNITY SERVICE

- **Zurich AI Safety**, Volunteer *Sept 2025 - Present*
- **WatervilleCreates!**, 3D Printing Instructor *Sept 2021 - May 2022*
- **The Bridge (LGBTQIA+) Club of Colby College**, Co-leader *Feb 2019 - Sept 2021*
- **Coral Monitoring in Collaboration with WWF and CoralWatch**, Co-leader *Sept 2016 - June 2018*

SKILLS & INTERESTS

- **Programming Languages:** Python, C++, Java.
- **Frameworks:** PyTorch, JAX, TensorFlow, Sklearn, Git, SQL, JSONiq.
- **Domains:** AI, Machine Learning, Reinforcement Learning, Multi-Agent Systems, AI Governance, Ethics.
- **Natural Languages:** English (Native), Mandarin (Native), Spanish (Intermediate), German (Elementary).
- **Interests:** Play Squash, Horror Film, Visual Arts, Scuba Diving, Effective Altruism.

REFERENCES

Upon request