Ann Arbor, MI
leeay@umich.edu

Ayoung Lee

Website:
aylee2008.github.io

My research focuses on **evaluation and post-training of LLMs on nuanced, non-verifiable tasks**: I have designed a structure that makes decision-making in dilemma situations verifiable [2], and introduced rubrics to assess and improve culture-aware moral decision-making [5] and coding agents [6]. I am also interested in **multi-agent and human-AI collaboration**: I aim to build high-fidelity simulations and interaction strategies that enable LLM agents to work effectively and safely in teams.

Education

University of Michigan **Aug 2024 - Apr 2029 (expected)**
Ph.D. in Computer Science and Engineering (Advisor: Prof. Lu Wang)

Seoul National University **Mar 2020 - Feb 2024**
B.S. in Computer Science and Engineering
GPA: 4.18/4.3, Summa cum laude, rank: 1st in the department

Seoul Science High School **Mar 2017 - Feb 2020**
The school for gifted students in mathematics and science, teaching college-level math and science
GPA: 4.18/4.3

Skills

Hands-on Experience	Rubrics as rewards, RLVR, SFT, LLM-as-a-judge, Self-improvement, Synthetic data generation, Moral reasoning, Multi-agent simulation
Tools & Frameworks	vLLM, LLaMA-Factory, Tinker, OpenRouter, Slurm

Publications

[6] Rubrics as Test Cases: Test-Grounded Execution-Free Verification of Coding Agents
Ayoung Lee, Justus Mattern
Under review at ARR Oct 2026

[5] MET: Theory-Grounded and Culture-Aware Multilingual Moral Reasoning
Ayoung Lee, Ryan Sungmo Kwon, Yunxiang Zhang, Yuxuan Liu, Peter Railton, Lu Wang
COLM 2026 [paper]

[4] LiveOIBench: Can Large Language Models Outperform Human Contestants in Informatics Olympiads?
Kaijian Zou, Aaron Xiong, Yunxiang Zhang, Xinliang Frederick Zhang, Yueqi Ren, Jirong Yang, Ayoung Lee, Shitanshu Bhushan, Lu Wang
ICML 2026 [paper]

[3] Logit Arithmetic Elicits Long Reasoning Capabilities Without Training
Yunxiang Zhang, Muhammad Khalifa, Lechen Zhang, Xin Liu, Ayoung Lee, Xinliang Frederick Zhang, Farima Fatahi Bayat, Lu Wang
Findings of ACL 2026 [paper]

[2] CLASH: Evaluating Language Models on Judging High-Stakes Dilemmas from Multiple Perspectives
Ayoung Lee, Ryan Sungmo Kwon, Peter Railton, Lu Wang
ICLR 2026 [paper]

[1] On Consistency Training for Language-Based Image Editing Interface
Youngwon Lee, Ayoung Lee*, Yeonjoon Jung, Seung-won Hwang (* denotes equal contribution)*
IJCNLP-AACL 2023, Second Workshop on Natural Language Interfaces (Oral) [paper]

Honors and Awards (Selected)

Best Presentation Award, NLP @ Michigan Day **Mar 2025**
Excellent Bachelor's Thesis Presentation Award **Feb 2024**
Outstanding Tutor Award **Aug 2021**
Presidential Science Scholarship **Mar 2020 - Feb 2024**

Work Experience

Proximal **May 2026 - Aug 2026**
Research Intern (Host: Justus Mattern) *San Francisco, United States*

NAVER Cloud HealthCare AI **Mar 2024 - Jun 2024**
Research Intern (Host: Heeyoung Kwak) *Seongnam, South Korea*

Language and Data Intelligence Lab **Mar 2023 - Feb 2024**
Undergraduate Research Intern (Advisor: Prof. Seung-won Hwang) *Seoul, South Korea*

DeepMetrics **Jun 2022 - Jan 2023**
Machine Learning Engineering Intern (Advisor: Prof. Hyun Oh Song) *Seoul, South Korea*

Samsung Device Solutions Memory **Jul 2021 - Aug 2021**
Software Intern *Hwaseong, South Korea*

Talks

Injecting Math Reasoning Abilities in Language Models **Jun 2024**
@ Deepest, Seoul National University

Study Abroad Presentation for Students **May 2024**
@ Seoul Science High School

Teaching

(EECS 497) Human-Centered Software Design and Development **Jan 2026 - Apr 2026**
Graduate Student Instructor *University of Michigan*

(M1522.000700) Logic Design **Sep 2021 - Dec 2021**
Teaching Assistant *Seoul National University*

(034.005) Foundation of Physics 1 **Mar 2021 - Aug 2021**
Peer Tutor *Seoul National University*

(M1522.000700) Logic Design **Mar 2021 - Aug 2021**
Tutor *Seoul National University*