ISAAC LIAO

iliao@andrew.cmu.edu https://iliao2345.github.io

EDUCATION

Carnegie Mellon University *Aug 2024 - May 2030 (Expected)*
PhD in Machine Learning.   Advisor: Albert Gu   GPA: 3.67/4.0

Massachusetts Institute of Technology *Sep 2023 - May 2024*
Master of Engineering in EECS.   Advisor: Max Tegmark   GPA: 5.0/5.0

Massachusetts Institute of Technology *Sep 2019 - Jun 2023*
Bachelor of Science, Double major in CS and Physics   GPA: 5.0/5.0

RESEARCH EXPERIENCE

Albert Gu Group *Aug 2024 - Present*
Graduate Researcher (PhD). Advising: Albert Gu

- Research on ARC-AGI Without Pretraining: training small models on target puzzles with inference-time-only learning; achieved 20% solve rate on ARC-AGI-1 with a 76k-parameter model.

Jane Street *May 2025 - Aug 2025*
Machine Learning Research Intern
- Built generative models to infer unseen activity in partially observed markets.
- Developed ensembling techniques for neural networks trained on market data.

Tegmark AI Safety Group *Sep 2023 - May 2024*
Graduate Researcher (MEng). Advising: Max Tegmark
- Discovered circular representations of days of the week in Mistral 7B and GPT-2-small.
- Simplified recurrent neural networks into standard forms using symmetry transformations.

PUBLICATIONS

ARC-AGI Without Pretraining Blog post, also submitted to NeurIPS 2025.
Isaac Liao, Albert Gu.
- Solving IQ-test-like visual puzzles by *extreme generalization*, training only on the target puzzle.
- 20% solve rate on ARC-AGI-1 despite training only at inference time, with a tiny 76K param model.

Not All Language Model Features Are One-Dimensionally Linear. arXiv 2024.
Josh Engels, Eric J. Michaud, Isaac Liao, Wes Gurnee, and Max Tegmark.
- Discovering circular representations of days of the week in Mistral 7B and GPT-2-small.

Opening the AI Black Box: Program Synthesis via Mechanistic Interpretability. arXiv 2024.
Eric J. Michaud, Isaac Liao, Vedang Lad, Ziming Liu, Anish Mudide, et al.
- Reducing RNN weights into interpretable python code through a series of simplifying steps.

Learning to Optimize Quasi-Newton Methods. TMLR 2023.
Isaac Liao, Rumen Dangovski, Jakob Nicolaus Foerster, and Marin Soljačić.
- Online learning a preconditioner for gradient descent; theoretical guarantees on preconditioner behavior.

Streamlining Physics Problem Generation to Support Physics Teachers in Using Gen. AI.
Shams El-Adawy, Isaac Liao, Vedang Lad, Mohamed Abdelhafez, et al. The Physics Teacher 2024.
- Techniques for chain-of-thought prompting an LLM to generate physics problems suitable for teaching.

Generating Interpretable Networks Using Hypernetworks. arXiv 2023.
Isaac Liao, Ziming Liu, and Max Tegmark.
- Designing a graph neural network to generate interpretable good weights for another neural network.

INVITED TALKS

ARC-AGI Without Pretraining NVIDIA LLM Reasoning Team, *Sept 2025*
Generating Interpretable Networks Using Hypernetworks Stanford Enigma Project, *May 2025*
ARC-AGI Without Pretraining Gwangju Institute of Science and Technology, *Apr 2025*

AWARDS AND HONORS

ARC Prize 2025: *3rd place Paper Award, $5000 prize.* *Dec 2025*
International Physics Olympiad: *Silver Medal. 2nd in Canada.* *July 2019*
International Physics Olympiad: *Honorable Mention. 5th in Canada.* *July 2018*
Citadel Securities PhD Summit Poster Competition: *3rd place, $5000 prize.* *March 2025*
MIT Battlecode swarm intelligence competition: *1st place, $8000 prize.* *Jan 2022*