Lynnix Zou

Email: qzou24@wisc.edu
Mobile: (608) 621-8517
Webpage: lynnixz.github.io

Education Background

University of Wisconsin-Madison — Madison, WI
- *Major: Computer Engineering, Computer Science; GPA: 3.98/4.0* — *Sep. 2023 – May. 2026*
  By 04/2024

Sichuan University — Chengdu, China
- *Major: Materials Science and Engineering; GPA: 3.69/4.0* — *Sep. 2021 – Jun. 2023*
  By 02/2023

Research

- **Group-Testing Model Soups for Safety Auditing of Fine-Tuned LLMs**
  Leader; Advisor: Prof. Kangwook Lee — *Nov. 2025 – Present*
  - Analyzed how fine-tuning organizes models in weight space and how model soups (weight averages of fine-tuned models) change the weight-space structure.
  - Designed a group-testing style safety check: form model soups over subsets of candidate models, subtract a clean anchor model to cancel shared structure, and probe the anchored soups to find misaligned models with fewer safety evaluations.

- **Benchmarking LLMs in Embodied AI with SLAM**
  Leader; Advisor: Prof. Kangwook Lee — *July 2025 – Present*
  - Constructed a MiniGrid-based room environment and designed agent trajectories that ensure enough loop closures for SLAM tasks.
  - Converted each trajectory into step-wise textual descriptions of egocentric views, and injected sensor noise, control noise, and log blackout.
  - Compared several LLMs with a Rao–Blackwellized particle filter and found that LLM accuracy drops much more under realistic noise, indicating weak belief maintenance.

- **Jump Representation for Analyzing and Improving LLM Reasoning**
  Key contributor; Advisor: Prof. Kangwook Lee — *March 2025 – June 2025*
  - Represented chain-of-thought as a tree-jump trace (structure + action), labeling steps as calc/verify/backtrack to quantify exploration, exploitation, overthinking, and forgetting.
  - Built a two-stage ReJump-Extractor to convert free-form CoTs into structured graphs.
  - Showed that models with similar accuracy can exhibit distinct reasoning styles on the same benchmarks.
  - **Paper:** ReJump: A Tree-Jump Representation for Analyzing and Improving LLM Reasoning. (under review at ICLR 2026)

- **Length Generalization for Transformers in Arithmetic Tasks**
  Leader; Advisor: Prof. Grigoris Chrysos — *August 2024 – September 2025*
  - Proposed Aligned Blankspace Augmentation (ABA) to zero-pad numbers and insert synchronized blanks across operands and results, so that corresponding digits always align.
  - Extended ABA to six arithmetic tasks where vanilla transformers do not length-generalize and outperformed previous methods on most of the tasks.
  - **Paper:** Data Augmentations for Arithmetic Length Generalization in Small Transformers (under review at ICLR 2026); NeurIPS 2025 workshop version.

Teaching & Mentoring

- **Bayview Foundation, Inc.** — Madison, WI
  Volunteer Tutor — Teen Program — *February 2025 – June 2025*
  - Help high-school students with math homework and strengthen problem-solving strategies.
  - Support after-school programming and supervise youth activities.

- **University of Wisconsin–Madison** Madison, WI
  Peer Mentor *January 2025 – Present*
  - Worked one-on-one with students to develop problem-solving skills.
  - Tutored linear algebra, machine learning algorithms, and PyTorch.