

## EDUCATION

---

- **University of Illinois Urbana-Champaign** Urbana-Champaign, IL, USA  
*B.S. of Mathematics and Computer Science; GPA 3.91/4.0 Jan. 2024 - Dec. 2025*
  - **Relevant Course:** Natural Language Processing. Machine Learning System. Machine Learning. Algorithms and Models of Computation. Numerical Analysis.
- **Beijing Jiaotong University** Beijing, China  
*B.Eng. of Computer Science and Technology; GPA 89.3/100 (3.74/4.0). Sept. 2021. - Dec. 2023 (Transferred)*
  - **Honors:** Ranked 5th in my major and obtained First-Class Scholarship for Academic Excellence (Top 3%).

## PUBLICATIONS & PRESENTATIONS

---

1. **Gao, Zijun**, Xu, Z., Ye, X. & Zhou, B. CORE: Concept-Oriented Reinforcement for Bridging the Definition–Application Gap in Mathematical Reasoning. The International Conference on Learning Representations (**ICLR 2026**), under review
2. Zhuang\*, Y., Ren\*, J., Ye\*, X., He, X., **Gao, Zijun**, et al. SimWorld: A World Simulator for Scaling Photorealistic Multi-Agent Interactions. The IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR 2025**), Demo Track

## RESEARCH HIGHLIGHTS

---

- **MAGEN: Training AI Agents to Collaborate** *May. 2025 - Present.* *Mentor: Manling Li*
  - Designed **MAGEN-Gym**, an action-centric environment suite featuring a **unified structure** that supports both sequential and simultaneous interactions across variable agent counts and modalities (LLM/VLM). Developed a reverse-playing mechanism leveraging jitter step curriculum and L-path geometric constraints to synthesize **ground-truth trajectories** with progressive difficulty. Implemented **Dual Sokoban** and **Overcooked** scenarios to enforce **necessary cooperation** in both IID and OOD settings.
  - Formulated the task as a **Dec-POMDP** with role-sharing policies, optimized via **Multi-Agent PPO** with shared parameters. Established a **two-stage SFT-then-RL pipeline** where agents generate structured thought traces (<think>) to internalize a **Multi-Agent World Model**. Designed a hierarchical reward system combining fine-grained local signals and sparse team rewards to resolve credit assignment via **Turn-Level GAE**.
  - Outperformed strong baselines in collaborative success rates across Dual Sokoban and Overcooked in both **IID** and **OOD** settings. Validated that the SFT-then-RL pipeline significantly accelerates convergence compared to training from scratch. Demonstrated via ablation studies that the explicit thinking mechanism shifts the policy from superficial memorization to causal reasoning, enabling robust generalization across **LLM and VLM** architectures.
  - Targeting submission to **ICML 2026**
- **CORE: Concept-Oriented Reinforcement for Bridging the Definition–Application Gap in Math Reasoning** *Feb. 2025 – Dec. 2025. Paper Page* *Mentor: Ben Zhou*
  - Observed that mainstream LLMs can describe mathematical concepts but often fail to apply them in reasoning. Identified this issue as the **Definition–Application Gap**. Hypothesized that robust reasoning requires learning like humans: moving from definitions to applications instead of relying on rote memorization. Proposed **CORE** to bridge this gap.
  - Developed **CORE**, a reinforcement learning framework grounded in rigorous mathematical definitions. It transforms static textbooks into dynamic supervision signals via **quiz-style probes**. This mechanism forces active concept retrieval and reshapes internal representations. Introduced **Concept Injection** to automatically localize and reinforce persistently failing concepts for targeted training.
  - Achieved significant performance gains (up to **+14.8%**) across 10+ stylistically diverse mathematical benchmarks on both base and instruction-tuned models. Verified through **perturbation tests** that CORE significantly improves robustness against irrelevant distractors. Analysis confirms gains stem from genuine concept application rather than superficial heuristic matching.
  - Submitted to **ICLR 2026** (under review)

- **SimWorld: A World Simulator for Scaling Photorealistic Multi-Agent Interactions**  
*Jul. 2024 – Jan. 2025. **Project Page*** *Mentor: Zhiting Hu, Lianhui Qin*
  - *SimWorld* is a photorealistic world simulator used to study multi-agent interactions in large, city-like environments.
  - Contributed to the simulation framework by helping develop a diffusion-based human motion module and implementing Python APIs that call LLMs to control and customize city generation. Also created 3D assets and scenes in Blender.
  - The system has been accepted as a demo at **CVPR 2025** and is used as a platform for testing interactive agents.

## AWARDS

---

- Dean's List, UIUC Urbana-Champaign, IL, USA, 2024
- National 3rd Prize, National College Mathematics Competition Beijing, China, 2023
- Provincial Award, Entrepreneurship and Innovation Competition for College Students Beijing, China, 2023
- The First Prize Scholarship, Beijing Jiaotong University (3%) Beijing, China, 2022
- Outstanding Student, Beijing Jiaotong University Beijing, China, 2021-2023

## SERVICE

---

- Volunteer Proctor, ICPC Online Programming Contest
- Core Member, Youth League Committee, Beijing Jiaotong University