Jiewen Hu

Phone: 412-251-8834 | Email: jiewenh@cs.cmu.edu | Personal Website: https://jiewenhu.com/

RESEARCH OBJECTIVE

My research focuses on advancing the logical reasoning abilities of language models, with a particular emphasis on mathematical reasoning and interpretability. I am passionate about developing models that can bridge the gap between human understanding and machine reasoning, enabling AI systems to assist in rigorous tasks such as theorem proving, complex planning, and decision-making. My interests span neural theorem proving, generalization ability in reasoning tasks, and mechanistic interpretability.

EDUCATION

Carnegie Mellon University, *Pittsburgh, PA*

- Master of Machine Learning   Sep 2024 - May 2025
- Bachelor of Science in Computer Science, GPA: 3.83/4.00   Sep 2020 - May 2024
  Additional Major: Mathematical Science   Minor: Statistics

Relevant Coursework: Advanced Natural Language Processing, Deep Reinforcement Learning, Probabilistic Graphical Models, Machine Learning with Large Datasets, Parallel Computer Architecture and Programming, Real Analysis (honors), Abstract Algebra (honors), Algorithm Design and Analysis, Operating Systems

RESEARCH EXPERIENCE

miniCTX Benchmark*, Carnegie Mellon University, Pittsburgh, PA*

Mentored by *Prof. Sean Welleck*   Feb 2024 – Present

- Developed the miniCTX benchmark to evaluate large language models in formal mathematics, focusing on real-world proof generation using context information.
- Fine-tuned deepseek-coder-1.3B model using file-tuning data, which incorporates context information alongside traditional state-tactic pairs, outperforming larger models like Llemma-7B and GPT-4o.
- Developed a Python wrapper for Lean REPL, simplifying interactions with Lean and enhancing usability.
- Published the first-author paper "miniCTX: Neural Theorem Proving with (Long-)Contexts" (https://cmu-l3.github.io/minictx/).
- Planning to extend the benchmark to areas beyond math, such as program verification, and to evaluate premise selection methods.

Large Language Model Specialized in Mathematics*, Shanghai Jiao Tong University, Shanghai, China*

Mentored by *Prof. Pengfei Liu*   May 2023 – Oct 2023

- Developed and refined step-by-step math problem prompts for training data based on error cases, enhancing the reasoning capabilities and output readability of Llama2, resulting in a 10% performance increase on the GSM8K dataset.
- Explored methods to integrate supplementary datasets with the primary dataset for supervised fine-tuning to address specific shortcomings, such as arithmetic calculations in main mathematical tasks.
- Created and executed a variety of supervised fine-tuning experiments to understand responses from large language models.
- Contributed to the development of Abel model (https://gair-nlp.github.io/abel/) and "Reformatted Alignment" paper (https://gair-nlp.github.io/ReAlign/).

Computational Evolutionary Dynamics Research, *Carnegie Mellon University, Pittsburgh, PA*

Mentored by Prof. Oana Carja and Yang Ping Kuo, PhD   Dec 2022 - Present

- Investigated the effects of spatial structures, as characterized by graph properties, on evolutionary dynamics.
- Explored the relationship between phylogenetic tree structures and graph properties, focusing on tree balance metrics.
- Utilized statistical tools to discern patterns of clonal evolutionary dynamics within complex graphs.
- Developed simulation programs to validate mathematical models.
- Conducted research into applications of these principles in network theory and biology.

Multitasking Facial Behavior Analysis Toolkit: Open Face 3.0, *Carnegie Mellon University, Pittsburgh, PA*

Mentored by *Prof. Paul Pu Liang* and *Prof. Louis-Philippe Morency* Sep 2023 – Present

- Developed a real-time multitasking framework for Open Face 3.0, supporting comprehensive facial behavior analysis including landmark detection, action unit detection, facial expression recognition, and gaze estimation.
- Enhanced toolkit accessibility by developing Python bindings that support both Open Face 2.0 and Open Face 3.0, enabling more flexible integration into computer vision applications.
- Implemented generative models for data augmentation, specifically targeting non-frontal face images, to enhance facial landmark detection and emotion recognition accuracy.

SKILLS

- **Computer:** Python, Lean, C/C++, PyTorch, R, SQL, Standard ML, LaTeX
- **Language:** Chinese (native), English (fluent)

PUBLICATIONS

- **Jiewen Hu,** Thomas Zhu, and Sean Welleck. miniCTX: Neural Theorem Proving with (Long-) Contexts. *In The 4th Workshop on Mathematical Reasoning and AI at NeurIPS'24*.
- Run-Ze Fan, Xuefeng Li, Haoyang Zou, Junlong Li, Shwai He, Ethan Chern, **Jiewen Hu,** and Pengfei Liu. 2024. Reformatted Alignment. *In Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 574–597, Miami, Florida, USA. Association for Computational Linguistics.

PRESENTATIONS

- miniCTX: Neural Theorem Proving with (Long-) Contexts Sep 2024
  Conference on Artificial Intelligence and Theorem Proving 2024, *Aussois, France*

- Rearrangements of Functions
  Mathematics Undergraduate Research Symposium, *Carnegie Mellon University* Sep 2022

HONORS

The William Lowell Putnam Mathematical Competition, *Top 10%* Dec 2020

Dean's List High Honors (Carnegie Mellon University) Fall 2020 – Spring 2024