

# Akanksha Sarkar

✉ akankshasarkar@gmail.com     Github     Website     LinkedIn akanksha-sarkar

## Education

<b>Cornell University</b> , Ithaca, New-York <i>Bachelor of Arts, Math (Cum Laude) and CS (Magna Cum Laude)</i> <i>Masters of Sciences, Computer Science</i>	<i>Aug 2021 – Aug 2027</i> GPA: 3.9
<b>Research interests:</b> AI4Science, Agentic workflows, Program Synthesis, NLP	

## Research Experience

<b>SunLab with Dr. Jennifer Sun</b>	<i>January 2024 - Present</i>
▶ <b>SciDUC-Scientific Data Utilization Challenge</b> Created SciDUC, a benchmark that tests LLM agents' ability to automate scientific workflows while quantifying their efficiency in data usage.	
▶ <b>FormulaCode</b> Worked on FormulaCode, a benchmark designed for evaluating agentic superoptimization on large codebases, with a focus on real-world performance optimization. Contributions included curating the dataset and processing for further task analysis.	
▶ <b>Novel Class Discovery</b> Proposed an evaluation framework to investigate what influences successful class discovery. Quantified effect of class coverage and number of known classes on discovery.	
<b>Networks Lab with Dr. Nate Foster</b>	

<b>Networks Lab with Dr. Nate Foster</b>	<i>June 2023 – Aug 2023</i>
○ Worked on Petr4, a formal semantics framework for P4 programming language. ○ Added an Intermediate representation (IR) language called Cimpl to the Petr4 compiler for conversion to C for program verification.	

## Scholarly Work

- A. Sarkar, B. Kim, J. Sun.** *How many classes for Novel Class Discovery?* Domain Generalization: Evolution, Breakthroughs and Future Horizons, CVPR workshop, April 2025
- Shortcut Learning and Distribution Shifts*, Ithaca, New York [Presentation] Milstein Undergraduate Research Symposium, May 2024.
- Implementing and Validating a Compiler from P4 to C*, Ithaca, New York. [Presentation] Bowers CIS Undergraduate Research Seminar, Summer 2023.

## Industry Experience

<b>Space System Design Studio with Dr. Curren Muhlberger</b>	<i>Jan 2023 – May 2023</i>
○ Worked on the Navigation and Modelling subteam in the Cislunar project focused on building CubeSat, a nanosatellite used to describe novel space technology.	
○ Designed and developed an object detection algorithm used to find the radius and center of nearby planetary objects seen by the CubeSat, optimized for arc detection. Used Hough Transformation and pixel-mapping from OpenCV to create the algorithm. Work went on to be patented by Ecliptic Enterprise. Responsible for testing and validation on the internal dataset.	
<b>Cornell Autoboat, Software Subteam Member</b>	

boat by calculating waypoints while considering obstacle geometry and boundary conditions.

#### Invsto, Software Developer Intern

June 2022 - Aug 2022

- Worked on the backend of the trading simulator SharkSigma, used to build and deploy strategies. Fetched real-time data by web-scraping and integrating external financial APIs using Python scripting. Used serialization with pickle and multi-threading to reduce data loading times by 20 %.

## Teaching Experience

---

Course Assistant for Math 1120 (Calculus II)	Jan 2022 - May 2023
Teaching Assistant for CS 3110 (Functional Programming with OCaml)	Aug 2023 - Dec 2023
Teaching Assistant for CS 4820 (Intro to analysis of Algorithms)	Jan 2025 - May 2025
Teaching Assistant for CS 6784 (Research Design for Machine Learning)	Aug 2025 - Dec 2025

## Other Involvements

---

Outreach Chair, Cornell Math Club	Sept 2024 - Present
Software Lead, Cornell Custom Silicon Systems	May 2023 - Dec 2023
Event Planner, Language Expansion Program	Oct 2021 - May 2023
Expand Your Horizon, Volunteer	Oct 2023
World Language Day, Volunteer	Oct 2022, 2023, 2024
Trinity College Grade 4 guitarist	March 2020

## Awards

---

- Computing Research Association, Outstanding Undergraduate Research Award nomination
- Cornell University Milstein Summer Research Award 2024 - \$7000
- Bowers CIS Undergraduate Research Experience Grant 2023 - \$5000
- 2023 HackNITR Hackathon, Excellence award for being in top 15 teams from over 300+ participating teams.

## Projects

---

CS 6784: Deep dive in LLMs, <i>Limited Memory Language Models</i>	Dec 2025
○ Building on existing work on limited memory language models which differentiate between specific and general knowledge by using an external database from the pre-training stage (unlike RAG which only uses external database during inference). ○ Extended the DSL used by the model and incorporated multi-hop reasoning. ○ Showed benefits in machine unlearning and more efficient use of parameters.	

CS 6840: Algorithmic Game Theory, <i>Sequential Hotelling Game</i>	Nov 2024
○ Worked on a research project that is a sequential variation of the traditional Hotelling game on a 1-d line. Found existence of optimal strategy and Nash for the $n$ -player setting when each player $p$ gets utility proportional to area of line closer to $p$ , under the rule that players are unable to change locations once chosen. ○ Defined the research problem by looking at Hotelling games more realistically where constant relocation of players is not feasible. Analyzed how relocation penalty affects optimal strategy.	

## References

---

Dr. Eva Tardos, Professor of CS at Cornell University. Email: et28@cornell.edu

Dr. Jennifer Sun, Assistant Professor of CS at Cornell University. Email: jjs533@cornell.edu