

# ILANA SHAPIRO

## CONTACT

---

Email: [ilshapiro@ucsd.edu](mailto:ilshapiro@ucsd.edu)

Website: <https://ilanashapiro.github.io>

LinkedIn: [linkedin.com/in/ilana-shapiro-157447170](https://linkedin.com/in/ilana-shapiro-157447170)

GitHub: [github.com/ilanashapiro](https://github.com/ilanashapiro)

## RESEARCH INTERESTS

---

My interests lie in the integration of programming languages, automated reasoning, generative AI, and human-computer interaction. I aim to develop usable structural constraints on sequence models for the controllable generation of well-formed sequence data, and to apply formal reasoning methods to improve the automated reasoning tools that underlie these systems.

## EDUCATION

---

**University of California, San Diego**, La Jolla, CA

2023-present

Ph.D. Computer Science

Programming Systems Group, advised by Professor Sorin Lerner

GPA: 3.96/4.0

**Pomona College**, Claremont, CA

2018-2022

B.A. Computer Science/Music (Flute) double major, minor in Mathematics

GPA: 4.0/4.0, Summa Cum Laude, Distinction in Senior Exercises

- **Computer Science Thesis:** "MusAssist: A Domain Specific Language for Music Notation"
  - Advised at Harvey Mudd College by Professor Ben Wiedermann
- **Music Thesis and Recital:** "Mieczysław Weinberg: Music Transcending Tragedy"
  - Advised by Professors Alfred Cramer, Joti Rockwell, and Eric Lindholm

## RESEARCH PROJECTS

---

**Parallelizing Z3: Adaptive Cubing via Online Sampling of CDCL Conflict Traces** [[code](#)] [[slides](#)] 2025

- At Microsoft Research, I worked with Nikolaj Bjorner on a novel parallelization algorithm for the SMT solver Z3. We developed an online cube-and-conquer approach: cubes are dynamically sampled during solving from CDCL conflict-variable heuristics and distributed to worker threads based on similarity.

**Training LLMs for Verified Programming** [[code](#)] 2025

2025

- At Microsoft Research, I helped train a 32B LLM specialized in program verification. I prepared and augmented Lean datasets for SFT and RL training, built a custom Dockerized Lean verification server, and evaluated model checkpoints.

**Synthesizing Composite Hierarchical Structure from Music Corpora** [[paper](#)] [[code](#)] [[slides](#)] [[talk](#)] 2025

- Published at 34th International Joint Conference on Artificial Intelligence 2025. Combined stochastic and SMT techniques to frame and solve the nested NP-hard combinatorial optimization problem of music structure summarization as an extension of the Generalized Median Graph problem.

**Lexically Constrained Decoding of Transformers** [[paper](#)] [[code](#)] [[slides](#)] 2025

2025

- Adapted the constrained decoding algorithm Grid Beam Search (GBS) to impose lexical constraints on GPT2, and fine-tuned GPT2 on a corpus of Chekhov's stories. GBS + fine-tuned GPT2 subjectively outperformed GBS + GPT2 alone.

**The Impact of GitHub Copilot on Test-First Development** [[paper](#)] [[code](#)] 2024

2024

- Conducted between-subjects pilot study to determine impact of Copilot on Test-First Development. Thematic analysis revealed that while Copilot enhanced coding speed, it resulted in superficial problem comprehension and decreased scope of the test suites.

**pGen-rs: LLM-Aided, Efficient, User-Friendly Genomic Data Wrangling** [[paper](#)] [[code](#)] [[slides](#)] 2024

- Developed pgen-rs, a tool enabling end-users to write genomic data wrangling requirements in natural language and execute with Rust-based high-performance genomic data processor.

**ProCon: Continuous Enumeration for Just-In-Time Bottom-Up Synthesis** [[paper](#)] [[code](#)] 2024

- Introduced continuous, rule-based enumeration for just-in-time bottom-up search in SyGuS problems, where programs are enumerated in order of continuous, nonrounded weights as determined by a probabilistic weighting function.

**MusAssist: A Domain Specific Language for Music** [[TENOR paper](#)] [[thesis2](#)] [[code](#)] [[demo](#)]

2022

- Published at 8th International Conference on Technologies for Music Notation and Representation 2023. Created DSL bridging the abstraction gap between music theoretical structures and notation. Wrote Haskell-based compiler to MusicXML.

**Mieczysław Weinberg: Music Transcending Tragedy** [\[thesis1\]](#)[\[recital\]](#)

2022

- Published in Scholarship@Claremont. Wrote extensive musicology thesis examining narrative and memory in Weinberg's *Kaddish Symphony*. Presented flute recital of my transcriptions of Weinberg's cello works.

**Markov Chain Music Generation** [\[JHM paper\]](#)[\[code\]](#)

2021

- Published in *Journal of Humanistic Mathematics*. Created a novel system of Markov chains using inverse transform sampling, enabling end-users to rapidly generate musical sketches.

**Virtual Ensemble Assembly: Musicality in Separation** [\[WAC paper\]](#)[\[code\]](#)

2020

- Published at Web Audio Conference 2022. Assisted on Prof. Christopher Raphael's research at Indiana University Bloomington exploring synchronizing audio tracks without click tracks.

---

## PUBLICATIONS

\* equal contribution

**Ilana Shapiro**, Ruanqianqian (Lisa) Huang, Zachary Novack, Cheng-i Wang, Hao-Wen Dong, Taylor Berg-Kirkpatrick, Shlomo Dubnov, and Sorin Lerner. "Synthesizing Composite Hierarchical Structure from Symbolic Music Corpora." In *Proceedings of the 34th International Joint Conference on Artificial Intelligence (IJCAI '25)*, Montreal, Canada, 2025.

[\[paper8\]](#)

**Ilana Shapiro**, Shubham Saha, Diya Lakhani, Shree Venkatesh, and Runqiu Xu. "Grid Beam Search for Constrained GPT-2 Decoding" Unpublished manuscript, 2025.

[\[paper7\]](#)

**Ilana Shapiro**, Michael Peng, and Andrew Lara. "The Impact of GitHub Copilot on Test-First Development." Unpublished manuscript, 2024.

[\[paper6\]](#)

Cole Kurashige,\* Savitha Ravi,\* and **Ilana Shapiro**.\* "pgen-rs: LLM-Aided Efficient and User-Friendly Genomic Data Wrangling." Unpublished manuscript, 2024.

[\[paper5\]](#)

Kyle Thompson, **Ilana Shapiro**, Ani Canumalla. "ProCon: Continuous Enumeration for Just-In-Time Bottom-Up Synthesis." Unpublished manuscript, 2024.

[\[paper4\]](#)

**Ilana Shapiro**. "MusAssist: A Domain Specific Language for Music Notation." *Proceedings of the International Conference on Technologies for Music Notation and Representation (TENOR'23)*, pp. 75-82, Northeastern University, Boston, MA, 2023.

[\[paper3\]](#)

Kaitlin Pet, **Ilana Shapiro**, and Christopher Raphael. "Virtual Ensemble Assembly: Musicality in Separation." In *Web Audio Conference (WAC'22)*, Cannes, France, 2022.

[\[paper2\]](#)

**Ilana Shapiro**. 2022. *MusAssist: A Domain Specific Language for Music Notation*. Bachelor's thesis. Pomona College.

[\[thesis2\]](#)

**Ilana Shapiro**. 2022. *Mieczysław Weinberg: Music Transcending Tragedy*. Bachelor's thesis. Pomona College.

[\[thesis1\]](#)

[\[recital\]](#)

**Ilana Shapiro** and Mark Huber. "Markov Chains for Computer Music Generation." In *Journal of Humanistic Mathematics, Volume 11 Issue 2 (July 2021)*, pp. 167-195.

[\[paper1\]](#)

---

## INDUSTRY EXPERIENCE

**Research Intern, Amazon Web Services (Automated Reasoning Group)** Summer 2026

- Incoming research intern in the AWS Automated Reasoning Group working on SAT/SMT tasks

**Research Intern, Microsoft (Research in Software Engineering/RiSE Group)** Summer 2025

- Researching SMT parallelization algorithms and natural language reasoning for verifiable code generation. Supervised by Nikolaj Bjorner.

**Freelance Software Engineer, Stainless** Feb. 2023-Oct. 2024

- Make open-source contributions to codebases such as Stoplight Prism, node-tree-sitter, Microsoft Pyright, NPM Trends, and json-schema-benchmark.

**Software Engineer, Meta**

Oct. 2022-Nov. 2022

- Improve type safety of Python and Hack code in engineering bootcamp. Impacted by the 13% company layoff as a new hire.

### 3x Software Engineering Intern, Facebook

Summers of 2019, 2020, 2021

- iOS/serverside fullstack intern on Facebook Events and Groups.

---

### Talks

“Parallelizing Z3: Adaptive Cubing via Online Sampling of CDCL Conflict Traces.” *Programming Systems Group, UC San Diego*, Nov. 2025.

“Synthesizing Composite Hierarchical Structure from Symbolic Music Corpora.” *The 34th International Joint Conference on Artificial Intelligence (IJCAI)*, Aug. 2025.

“Synthesizing Composite Hierarchical Structure from Symbolic Music Corpora.” *The 19th SoCal Programming Languages and Systems Workshop (SoCaL PLS)*, Feb. 2025.

“Deriving Structure from Music Corpora.” *Programming Systems Group, UC San Diego*, Apr. 2024.

“MusAssist: A Domain Specific Language for Music Notation.” *International Conference on Technologies for Music Notation and Representation*, May 2023.

---

### TEACHING EXPERIENCE

**Teaching Assistant, CSE 130: Programming Languages, UCSD (N=69)**

Fall 2025

**Teaching Assistant, CSE 130: Programming Languages, UCSD (N=126)**

Spring 2025

**Teaching Assistant, CS 133: Database Systems, Pomona College (N=20)**

Spring 2020

---

### ACADEMIC HONORS

**NSF Graduate Research Fellowship, Honorable Mention**

2025

- The NSF GRFP recognizes and supports outstanding graduate students who are pursuing full-time research-based master's and doctoral degrees in STEM fields.

**The Phi Beta Kappa Award**

2022

- Endowed by the Pomona Chapter of Phi Beta Kappa, awarded to one senior selected for high quality of scholarship and promise of future distinction.

**The Rena Gurley Archibald High Scholarship Prize**

2022

- Awarded to the member(s) of the Pomona College graduating class ranking highest in scholarship.

**Distinction in Senior Exercise (Computer Science)**

2022

- Exceptional work on the senior exercise is awarded based on review by the entire faculty of the Computer Science Department at Pomona College.

**Distinction in Senior Exercise (Music)**

2022

- Exceptional work on the senior exercise is awarded based on review by the entire faculty of the Music Department at Pomona College.

**The Katherine J. Hagedorn Prize**

2022

- Awarded annually to the Pomona College student(s) demonstrating exceptional loyalty and dedication to their music studies.

**Phi Beta Kappa Induction (Junior Year) - Pomona College Chapter**

2021

- 1 of 8 juniors awarded for "good moral character," distinguish in "breadth of culture," and "excellence of scholarship."

**The William F. Russell Prize**

2020

- Awarded annually to the Pomona College prospective music major(s) showing substantial accomplishment and significant promise in their study of music.

---

### SKILLS

**Programming Languages:** Python • Haskell • TypeScript • C++ • Java • Objective-C • SQL

**Tools/Frameworks:** LaTeX • Git • Functional and Object-Oriented Programming

**Domain Knowledge:** Programming Languages • Automated Reasoning • SAT/SMT Solvers • Stochastic/Combinatorial Optimization • Neurosymbolic Generation • User Studies

---

### OUTREACH AND MENTORSHIP

**Presenter, Harmony Hacks @ CSU San Marcos***Spring 2025*

- NSF-funded event to broaden participation of women in computing. I co-hosted a Q&A to inspire high school girls to pursue careers in CS.

**Mentorship Co-Chair, GradWIC UCSD***Fall 2024-Spring 2025*

- Manage the UCSD Graduate Women in Computing Mentorship Program. Pair 174 mentees with mentors, personally mentor 2 students, host inclusive group activities.