ILANA SHAPIRO

CONTACT

Email: ilshapiro@ucsd.edu
LinkedIn: linkedin.com/in/ilana-shapiro-157447170
Website: https://ilanashapiro.github.io
GitHub: github.com/ilanashapiro

RESEARCH INTERESTS

My interests lie in the integration of programming languages, generative AI, and human-computer interaction. I aim to develop human-interpretable structural constraints on sequence models for the controllable generation of well-formed sequence data. My goal is to create an interactive, human-machine co-creation process enabling end-users to produce structurally sound outputs.

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

Synthesizing Composite Hierarchical Structure from Music Corpora [preprint][code][slides] *2025*
- To appear in the 34th International Joint Conference on Artificial Intelligence 2025. Introduced a unified, hierarchical meta-representation of music structure as a *k*-partite DAG. Combined stochastic and formal logic techniques to frame and solve the dually NP-hard combinatorial optimization problem of music structure summarization.

Lexically Constrained Decoding of Transformers [paper][code][slides] *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*
- 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.

DNA to Music (MIDI) Translation [paper][code] *2019*

- Created original Python-based model translating DNA to MIDI. Analyzed result for harmonic sequences to classify species into taxonomic classes. Ranked species by musicality.

RESEARCH ARTIFACTS

* *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." To appear in *Proceedings of the 34th International Joint Conference on Artifical Intelligence (IJCAI '25),* Montreal, Canada, 2025. [paper9]

Ilana Shapiro, Shubham Saha, Diya Lakhani, Shree Venkatesh, and Runqiu Xu. "Grid Beam Search for Constrained GPT-2 Decoding" Unpublished manuscript, 2025. [paper8]

Ilana Shapiro, Michael Peng, and Andrew Lara. "The Impact of GitHub Copilot on Test-First Development." Unpublished manuscript, 2024. [paper7]

Cole Kurashige,* Savitha Ravi,* and **Ilana Shapiro**.* "pgen-rs: LLM-Aided Efficient and User-Friendly Genomic Data Wrangling." Unpublished manuscript, 2024. [paper6]

Kyle Thompson, **Ilana Shapiro**, Ani Canumalla. "ProCon: Continuous Enumeration for Just-In-Time Bottom-Up Synthesis." Unpublished manuscript, 2024. [paper5]

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. [paper4]

Kaitlin Pet, **Ilana Shapiro**, and Christopher Raphael. "Virtual Ensemble Assembly: Musicality in Separation." In *Web Audio Conference (WAC'22*), Cannes, France, 2022. [paper3]

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*. [paper2]

Ilana Shapiro. "Converting DNA to Music: Sonifying Structure, Splicing, and Translation." Unpublished manuscript, 2019. [paper1]

INDUSTRY EXPERIENCE

Research Intern, Microsoft (RiSE Group) *Summer 2025*

- Studying formal document processing at the intersection of formal logic (SMT), reinforcement learning, and LLMs. Supervised by Dr. Nikolaj Bjorner and Dr. Janardhan Kulkarni.

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

"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=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 • Java • Objective-C • SQL • GraphQL
Tools/Frameworks: LaTeX • Git • Functional and Object-Oriented Programming
Domain Knowledge: Domain Specific Languages • Parsing/Compilers • Stochastic/Combinatorial Optimization • User Studies • Human-Centered Design • Constrained Generative Models

OUTREACH AND MENTORSHIP

Presenter, Harmony Hacks @ CSU San Marcos *Spring 2025*
- NSF-funded event to broaden participation of women in computing. Inclusive, beginner-friendly events for high school girls combine computer programming and music in problem solving. I co-hosted a Q&A to inspire participants 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.