Gabriel Grand

grandg@mit.edu | gabegrand.com | Google Scholar | GitHub
MIT Schwarzman College of Computing, 51 Vassar Street (Building 45), Office 711-C, Cambridge, MA 02139

Education

Massachusetts Institute of Technology — Cambridge, MA
PhD in Electrical Engineering and Computer Science — *Sep 2021 – present*

- Co-advisors: Jacob Andreas and Josh Tenenbaum
- MIT Presidential Fellowship; NSF Graduate Research Fellowship

Harvard University — Cambridge, MA
B.A. in Computer Science and Mind, Brain, & Behavior; Summa Cum Laude — *Sep 2014 – Dec 2018*

- John Harvard Scholar (top 5% of class, GPA 3.98); Highest Honors in CS; Phi Beta Kappa.
- Hoopes Prize for Senior Thesis: *Learning Interpretable and Bias-Free Models for VQA*, advised by Sasha Rush.

Experience

Reverie Labs — Cambridge, MA
Senior Machine Learning Engineer — *Feb 2019 – June 2021*

- Founding team member at AI-driven drug discovery startup (acquired by Ginkgo Bioworks in 2024).
- ML engineering lead for multimillion-dollar cancer therapeutic research collaboration with Roche and Genentech.
- Developed and deployed models of molecular properties and protein-ligand interactions. Focus on graph convolutional neural networks (GNNs) and neural sequence models.
- Contributed to invention of novel cyclin-dependent kinase (CDK) inhibitor (US Patent No. 2022/0281863 A1).

Google Brain — Mountain View, CA
Associate Product Manager (APM) Intern, TensorFlow — *June – Aug 2018*

- Created a Git-like system in TensorFlow Extended (TFX) to facilitate collaborative development of ML pipelines.
- Coordinated with a team of 50+ engineers to integrate TFX features into 200+ Google products, including Ads, Cloud, Maps, Play, Shopping, and YouTube.

Software Engineering (SWE) Intern, Google OCR — *May – Aug 2017*

- Prototyped Google's first attention-based architecture for optical character recognition and trained it on a large multilingual document corpus; this model was subsequently productionized in Google Cloud OCR.

Publications

For the most up-to-date publication list, please see Google Scholar.

Preprints

★ Gabriel Grand*, Valerio Pepe*, Joshua B. Tenenbaum, Jacob Andreas. "Shoot First, Ask Questions Later? Building Rational Agents that Explore and Act Like People." arXiv:2510.20886, 2025. ***Top 1% of ICLR 2026 submissions by reviewer score***

Alexander K Lew, Tan Zhi-Xuan, Gabriel Grand, Vikash K Mansinghka. "Sequential Monte Carlo Steering of Large Language Models using Probabilistic Programs." arXiv:2306.03081, 2023.

★ Lionel Wong*, Gabriel Grand*, Alexander K. Lew, Noah D. Goodman, Vikash K. Mansinghka, Jacob Andreas, Joshua B. Tenenbaum. "From Word Models to World Models: Translating from Natural Language to the Probabilistic Language of Thought." arXiv:2306.12672, 2023.

Conference Proceedings

★ Gabriel Grand*, Joshua B. Tenenbaum, Vikash K. Mansinghka, Alexander K. Lew, Jacob Andreas. "Self-Steering Language Models." Conference on Language Modeling (COLM), 2025. ***Featured on Hugging Face Daily Papers***

João Loula, Benjamin LeBrun, Li Du, Ben Lipkin, Clemente Pasti, Gabriel Grand, Tianyu Liu, Yahya Emara, Marjorie Freedman, Jason Eisner, Ryan Cotterell, Vikash Mansinghka, Alexander K. Lew, Tim Vieira, Timothy J. O'Donnell. "Syntactic and Semantic Control of Large Language Models via Sequential Monte Carlo." International Conference on Learning Representations (ICLR), 2025. ***Oral Talk, ICLR 2025***

Kanishk Gandhi, Denise Lee, Gabriel Grand, Muxin Liu, Winson Cheng, Archit Sharma, Noah D. Goodman. "Stream of Search (SoS): Learning to Search in Language." Conference on Language Modeling (COLM), 2024. ***Oral Spotlight, COLM 2024***

Gabriel Grand*, Valerio Pepe, Jacob Andreas, Joshua B. Tenenbaum. "Loose LIPS Sink Ships: Asking Questions in Battleship with Language-Informed Program Sampling." Annual Meeting of the Cognitive Science Society (CogSci), 2024.

★ Gabriel Grand*, Lionel Wong, Matthew Bowers, Theo X. Olausson, Muxin Liu, Joshua B. Tenenbaum, Jacob Andreas. "LILO: Learning Interpretable Libraries by Compressing and Documenting Code." International Conference on Learning Representations (ICLR), 2024.

Benjamin Lipkin, Lionel Wong, Gabriel Grand, Joshua B. Tenenbaum. "Evaluating Statistical Language Models as Pragmatic Reasoners." Annual Meeting of the Cognitive Science Society (CogSci), 2023.

Cedegao Zhang, Lionel Wong, Gabriel Grand, Joshua B. Tenenbaum. "Grounded Physical Language Understanding with Probabilistic Programs and Simulated Worlds." Annual Meeting of the Cognitive Science Society (CogSci), 2023.

Matthew Bowers, Theo X. Olausson, Lionel Wong, Gabriel Grand, Joshua B. Tenenbaum, Kevin Ellis, Armando Solar-Lezama. "Top-Down Synthesis for Library Learning." Proceedings of the ACM on Programming Languages (POPL), 2023.

Catherine Wong*, William P. McCarthy*, Gabriel Grand*, Yoni Friedman, Joshua B. Tenenbaum, Jacob Andreas, Robert D. Hawkins, and Judith E. Fan. "Identifying concept libraries from language about object structure." Annual Meeting of the Cognitive Science Society (CogSci), 2022.

Workshops

Dev Patel, Gabrielle Gervacio, Diekola Raimi, Kevin Zhu, Ryan Lagasse, Gabriel Grand, Ashwinee Panda, Maheep Chaudhary. "Alignment-Constrained Dynamic Pruning for LLMs: Identifying and Preserving Alignment-Critical Circuits." Socially Responsible and Trustworthy Foundation Models Workshop, Conference on Neural Information Processing Systems (ResponsibleFM @ NeurIPS 2025) & Deployable AI Workshop, Association for the Advancement of Artificial Intelligence (DAI @ AAAI 2026), 2025. ***Oral Talk, DAI at AAAI 2026***

Joongho Kim, Xirui Huang, Zarreen Reza, Gabriel Grand, Kevin Zhu, Ryan Lagasse. "Chopping Trees: Semantic Similarity Based Dynamic Pruning for Tree-of-Thought Reasoning." Efficient Reasoning Workshop, Conference on Neural Information Processing Systems (ER @ NeurIPS), 2025.

Gabriel Grand*, Valerio Pepe, Jacob Andreas, Joshua B. Tenenbaum. "A Llama Sunk My Battleship! Asking Rational Questions with LLMs via Bayesian Inference." System-2 Reasoning at Scale Workshop, Conference on Neural Information Processing Systems (Sys2 @ NeurIPS), 2024.

Walid Ahmad, Elana Simon, Seyone Chithrananda, Gabriel Grand, and Bharath Ramsundar. "ChemBERTa-2: Towards Chemical Foundation Models." Machine Learning for Molecules Workshop, European Laboratory for Learning and Intelligent Systems (ML4Molecules @ ELLIS), 2021.

Seyone Chithrananda, Gabriel Grand, and Bharath Ramsundar. "ChemBERTa: Large-Scale Self-Supervised Pretraining for Molecular Property Prediction." Machine Learning for Molecules Workshop, Conference on Neural Information Processing Systems (ML4Molecules @ NeurIPS), 2020.

Gabriel Grand and Yonatan Belinkov. "Adversarial Regularization for Visual Question Answering: Strengths, Shortcomings, and Side Effects." 2nd Workshop on Shortcomings in Vision and Language, North American Chapter of the Association for Computational Linguistics (SiVL @ NAACL), 2019. ***Best Paper Award***

Gabriel Grand, Aron Szanto, Yoon Kim, and Alexander Rush. "On the Flip Side: Identifying Counterexamples in Visual Question Answering." Deep Learning Day, ACM SIGKDD Conference on Knowledge Discovery and Data Mining (Deep Learning Day @ KDD), 2018.

Journals

Anna A. Ivanova, Aalok Sathe, Benjamin Lipkin, Unnathi Kumar, Setayesh Radkani, Thomas H. Clark, Carina Kauf, Jennifer Hu, R.T. Pramod, Gabriel Grand, Vivian Paulun, Maria Ryskina, Ekin Akyürek, Ethan Wilcox, Nafisa Rashid, Leshem Choshen, Roger Levy, Evelina Fedorenko, Joshua Tenenbaum, Jacob Andreas. "Elements of World Knowledge (EWOK): A cognition-inspired framework for evaluating basic world knowledge in language models." Transactions of the Association for Computational Linguistics (TACL), 2024.

★ Gabriel Grand, Idan Blank, Francisco Pereira, and Evelina Fedorenko. "'Semantic projection' recovers rich human knowledge of multiple object features from word embeddings." Nature Human Behaviour, 2022.

Theses

Gabriel Grand. "Discovering Abstractions from Language via Neurosymbolic Program Synthesis." MIT Master's Thesis, advised by Jacob Andreas and Joshua Tenenbaum, 2023.

Gabriel Grand. "Learning Interpretable and Bias-Free Models for Visual Question Answering." Harvard Undergraduate Thesis, advised by Alexander Rush and presented to the Department of Computer Science, 2018. ***Hoopes Prize***

Invited Talks & Presentations

2025 "Towards Meta-Cognitive Reasoning Models." Visions of Language Modeling Workshop, COLM, Montreal, Canada, October 10, 2025. [talk]

2025 "Self-Steering Language Models." Intel-NSF Research Collaboration on Scalable Machine Programming (ScaMP), virtual, November 13, 2025. [talk]

2025 "Self-Steering Language Models." COLM, Montreal, Canada, October 7, 2025. [poster]

2025 "Self-Steering Language Models." VerifAI: AI Verification in the Wild Workshop, ICLR, Singapore, SG, April 27, 2025. [recorded talk]

2025 "Self-Steering Language Models." New England NLP Symposium (NENLP), New Haven, CT, April 11, 2025. [poster]

2025 "Self-Steering Language Models." GenLM Consortium Meeting, virtual, March 5, 2025. [talk]

2024 "From Word Models to World Models." Philosophy of Deep Learning Group (hosts: Raphaël Millière, David Chalmers), virtual, November 6, 2024. [talk]

2024 "LILO: Learning Interpretable Libraries by Compressing and Documenting Code." MIT Research Qualifying Examination (RQE) (Committee: Michael Carbin and Armando Solar-Lezama), MIT, August 7, 2024. [talk]

2024 "Asking Questions in Battleship with Language-Informed Program Sampling." CogSci, Rotterdam, Netherlands, July 24, 2024. [poster]

2024 "LILO: Learning Interpretable Libraries by Compressing and Documenting Code." ICLR, Vienna, Austria, April 17, 2024. [recorded talk]

2024 "Asking Questions in Battleship with Language-Informed Program Sampling." MIT Lincoln Laboratory - Department of the Air Force (DAF) - MIT AI Accelerator (AIA) Meeting, virtual, March 7, 2024. [talk]

2024 "LILO: Learning Interpretable Libraries by Compressing and Documenting Code." Harvard Center for Mathematical Sciences and Applications (CMSA) Seminar, Cambridge, MA, March 6, 2024. [talk]

2024 "Asking Questions in Battleship with Language-Informed Program Sampling." CoCoSci Lab Meeting, MIT, March 6, 2024. [talk]

2023 "LILO: Learning Interpretable Libraries by Compressing and Documenting Code." Intrinsically Motivated Open-ended Learning (IMOL) Workshop (NeurIPS), New Orleans, LA, December 16, 2023. [poster]

2023 "LILO: Learning Interpretable Libraries by Compressing and Documenting Code." NSF Neurosym Site Visit, MIT, November 6, 2023. [talk]

2023 "LILO: Learning Interpretable Libraries by Compressing and Documenting Code." Guest lecture, Harvard CS252R, Program Synthesis (host: Nada Amin), October 19, 2023. [talk]

2023 "From Word Models to World Models." Gershman Lab, Harvard, September 8, 2023. [talk]

2023 "From Word Models to World Models." LLMs Meet CogSci Workshop, CogSci 2023, Sydney, Australia, July 26, 2023. [recorded talk]

2023 "From Word Models to World Models." Kanwisher Lab, MIT, July 13, 2023. [talk]

2023 "From Word Models to World Models." Society for Philosophy and Psychology (SPP), Pittsburgh, PA, June 26, 2023. [talk]

2023 "LILO: Learning Interpretable Libraries by Compressing and Documenting Code." LINGO Lab Meeting, MIT, June 8, 2023. [talk]

2023 "Grounded Physical Language Understanding with Probabilistic Programs and Simulated Worlds." CoCoSci Lab Meeting, MIT, February 8, 2023. [talk]

2022 "Translating From Natural Language to the Language of Thought." Harvard LangCog (Language and Cognition) Seminar (hosts: Hayley Ross, Jesse Snedeker, Kathryn Davidson), Cambridge, MA, October 20, 2022. [talk]

2022 "Identifying Concept Libraries from Language about Object Structure." CogSci, Toronto, Canada, July 30, 2022. [poster]

2022 "From Natural Language to the Many Languages of Thought." McDonnell Foundation Plenary Meeting, Venice, Italy, July 6, 2022. [talk]

2020 "Serving Molecular Prediction Models at Scale." Django Boston Meetup, Boston, MA, January 23, 2020. [recorded talk]

Press

Enabling small language models to solve complex reasoning tasks. *MIT News*, December 12, 2025.

Making AI-generated code more accurate in any language. *MIT News*, April 18, 2025.

Natural language boosts LLM performance in coding, planning, and robotics. *MIT News*, May 1, 2024.

What words can convey. *MIT McGovern Institute*, April 26, 2022.

Reverie Labs raises $25 million for its AI drug discovery platform. *VentureBeat*, February 23, 2021.

A Virtual Drug Discovery Company—With No Lab Of Its Own—Partners With Roche To Fight Cancers In The Brain. *Forbes*, August 25, 2020.

A better way to train machine learning models: Teaching AI to overcome human bias. *Harvard SEAS*, July 30, 2019.

Professors to Examine Proposed Cognitive Science Concentration. *The Harvard Crimson*, February 28, 2018.

Teaching, Mentorship, & Service

Teaching

- Teaching Assistant, MIT 6.8611 (Natural Language Processing), Fall 2025.
- Final Project Supervisor, MIT 9.660 (Computational Cognitive Science), Fall 2023.

Mentorship

- MIT Summer Program for Undergraduate Research (MSRP) Mentor, Summer 2023, Summer 2024, Summer 2025.
- MIT EECS Graduate Application Assistance Program (GAAP), Fall 2023.
- Algoverse AI Research, Principal Investigator / Mentor (Summer – Fall 2025).

Undergraduate Research Supervision

- 2022-2025: Valerio Pepe (Harvard College → OpenAI, Member of Technical Staff)
- 2023-2024: Muxin (Maxine) Liu (Harvey Mudd College → University of Pennsylvania, PhD Student)

Reviewing

- *Conferences:* International Conference on Learning Representations (ICLR), 2025, 2026; Conference on Language Modeling (COLM), 2024, 2025; Cognitive Science Society Annual Meeting (CogSci), 2024, 2025.
- *Workshops:* NeurIPS CogInterp Workshop, 2025; NeurIPS System-2 Reasoning Workshop, 2024.
- *Journals:* Computational Linguistics Journal Special Issue on Language Learning, Representation, and Processing in Humans and Machines, 2024; Open Mind: Discoveries in Cognitive Science, 2023.

Awards & Fellowships

National Science Foundation Graduate Research Fellowship	2021 – 2026
MIT Presidential Fellowship	2021 – 2022
Harvard Thomas T. Hoopes Prize	2019
Phi Beta Kappa	2019
Summa Cum Laude	2019
John Harvard Scholar (top 5% of class)	2016
Harvard College Scholar (top 10% of class)	2015

Skills & Interests

Programming: Python expert (top 1% of Python engineers globally on Algora.io). Comfortable with ML frameworks including PyTorch, HuggingFace, TensorFlow (contributor), and Sklearn. Experienced in functional programming (OCaml, Haskell, Scheme) and scientific computing (MATLAB, R, C++). Proficient in modern web frameworks (Django, React, Ruby) and databases (PostgreSQL, DynamoDB, Amazon RDS).

Infrastructure: AWS, Google Cloud, Azure, Docker, Kubernetes, SLURM, Prefect, MCP.

Language Proficiency: English (native), Spanish (fluent), Mandarin (basic).

Hobbies: Backpacking, skiing, sailing, rowing, music, comedy. I play jazz guitar and perform with several groups, including the MIT CMS Jazz Combos and a Cambridge-based jazz/funk group called *The Institute Policy*. For several years, I've served as Program Chair and co-host of SIGTBD, MIT CSAIL's annual humor conference.