

Zayne Sprague

zaynesprague@gmail.com • zrs2020@nyu.edu • +1 (918) 521 3229

www.zaynesprague.com • <https://github.com/Zayne-sprague>

EDUCATION

PH.D. IN COMPUTER SCIENCE

Thesis advisor: Greg Durrett

New York University

Expected Degree 2028

B.S. AND M.S. IN COMPUTER SCIENCE

GPA: 3.6 / 4.0 | Masters GPA: 3.5 / 4.0

University of Texas in Austin

2017-2023

PUBLICATIONS

Zayne Sprague, Jack Lu, Manya Wadhwa, Sedrick Keh, Mengye Ren, Greg Durrett. 2025, SkillFactory: Self-Distillation For Learning Cognitive Behaviors. Proceedings of the Fourteenth International Conference on Learning Representations (ICLR 2026)

Etash Guha, Ryan Marten, Sedrick Keh, Negin Raoof, Georgios Smyrnis, Hritik Bansal, Marianna Nezhurina, Jean Mercat, Trung Vu, **Zayne Sprague**, Ashima Suvarna, Benjamin Feuer, et al. 2025, OpenThoughts: Data Recipes for Reasoning Models. Proceedings of the Fourteenth International Conference on Learning Representations (ICLR 2026)

Liyan Tang, Grace Kim, Xinyu Zhao, Thom Lake, Wenxuan Ding, Fangcong Yin, Prasann Singhal, Manya Wadhwa, Zeyu Leo Liu, **Zayne Sprague**, Ramya Namuduri, Bodun Hu, Juan Diego Rodriguez, Puyuan Peng, Greg Durrett. 2025, ChartMuseum: Testing Visual Reasoning Capabilities of Large Vision-Language Models. NeurIPS Datasets and Benchmarks Track 2025

Manya Wadhwa, **Zayne Sprague**, Chaitanya Malaviya, Philippe Laban, Junyi Jessy Li, Greg Durrett. 2025, EvalAgent: Discovering Implicit Evaluation Criteria from the Web. Proceedings of COLM 2025

Zayne Sprague, Fangcong Yin, Juan Diego Rodriguez, Dongwei Jiang, Manya Wadhwa, Prasann Singhal, Xinyu Zhao, Xi Ye, Kyle Mahowald, Greg Durrett. 2025, To CoT or not to CoT? Chain-of-thought helps mainly on math and symbolic reasoning. Proceedings of the Thirteenth International Conference on Learning Representations (ICLR 2025)

Zayne Sprague, Xi Ye, Kaj Bostrom, Swarat Chaudhuri, and Greg Durrett. 2023, MuSR: Testing the Limits of Chain-of-thought with Multistep Soft Reasoning. Proceedings of the Twelfth International Conference on Learning Representations (ICLR 2024) As a **Spotlight Presentation**

Rohan Chandra*, **Zayne Sprague***, Jarrett Holtz, Joydeep Biswas. 2023, SOCIALGYM 2.0: Simulator for Multi-Agent Social Robot Navigation in Shared Human Spaces. Proceedings of the AAAI Conference on Artificial Intelligence 2024

Zayne Sprague, Kaj Bostrom, Swarat Chaudhuri, and Greg Durrett. 2023, Deductive Additivity for Planning of Natural Language Proofs. *Natural Language Reasoning and Structured Explanations Workshop at ACL*

Zayne Sprague, Kaj Bostrom, Swarat Chaudhuri, and Greg Durrett. 2022, Natural Language Deduction with Incomplete Information. *Proceedings of the Conference on Empirical Methods for Natural Language Processing (EMNLP)*. (Acceptance Rate 26%)

Kaj Bostrom, **Zayne Sprague**, Swarat Chaudhuri, and Greg Durrett. 2022. Natural Language Deduction through Search over Statement Compositions. *Findings of the Conference on Empirical Methods for Natural Language Processing (EMNLP)*. (Acceptance Rate 37%)

RESEARCH EXPERIENCE

Text Analysis, Understanding, and Reasoning (TAUR) Lab	New York City, New York
<i>Advised by Dr. Greg Durrett</i>	2021-Present

- Released multiple models post-trained via SFT and advanced RL methods such as GRPO and PPO to enable advanced reasoning on challenging benchmarks like GPQA and AIME through self-distillation of cold-start data.
- Published an analysis of reasoning methods such as Chain-of-Thought on state-of-the-art LLMs empirically showing the limitations of current reasoning methods.
- Wrote a paper and released a new multi-step commonsense reasoning dataset over long narratives (MuSR) along with a creation algorithm to automatically generate datasets as LLMs improve reasoning abilities.
- Published multiple papers on natural language reasoning, introducing LLM-based planning for deduction, validated abductive reasoning methods, and a step-by-step deductive model achieving state-of-the-art results.

Autonomous Mobile Robotics (AMRL) Lab	Austin, Texas
<i>Advised by Dr. Joydeep Biswas</i>	2022-2023

- Wrote a paper and built a novel simulation called SocialGym using Docker, ROS, UT AMRL stacks, Stable Baselines and Petting Zoo for social navigation with multi-agent Reinforcement Learning
- Researched Deep Reinforcement Learning with Game Design Mechanisms for solving Multi-Agent Social Navigation in constrained indoor environments.

Developmental Intelligence Lab (DIL)	Austin, Texas
<i>Advised by Dr. Chen Yu</i>	2021-2022

- Implemented and Researched topics from computer vision, including image inpainting & pose estimation.
- Build a GPU-accelerated data pipeline in Docker for processing motion and image data with ML algorithms.

RELEVANT INDUSTRY EXPERIENCE

Copilot Search

Austin, Texas

AI Consultant (founding member)

2025-Present

- Designed and deployed advanced LLM-based systems, including multi-agent workflows, structured reasoning pipelines, and internal prototyping frameworks (of which cut down manual overhead by 66%).
- Conducted applied research and technical evaluations of state-of-the-art LLM methods to inform product strategy, architecture decisions, and model selection.
- Advised engineering and leadership teams on integrating AI agents and LLM-driven tooling to improve developer productivity and accelerate feature development.

Senior Fullstack Engineer (founding member)

2015-2022

- Responsible for implementing the Machine Learning stack, priority front-end funnels, and mid-stack APIs
- Managed multiple teams on high-level projects, led onboarding of new hires, and maintained documentation of systems and code-release cycles used throughout the company.
- Wrote code spanning many areas, including AWS services and deployments, IOS and Android stacks, Python libraries and Flask APIs, React, data collection and storage with Elastic Search and PostgreSQL.

CONFERENCE AND WORKSHOP ORGANIZATION

- Reviewer, International Conference on Learning Representations (ICLR): 2026
- Reviewer, International Conference on Learning Representations (ICLR): 2025
- Reviewer, International Conference on Learning Representations (ICLR): 2024
- Reviewer, AAAI Conference on Artificial Intelligence (AAAI): 2024
- Reviewer, Conference on Empirical Methods in Natural Language Processing (EMNLP): 2023
- Reviewer for *The First Natural Language Reasoning and Structured Explanations Workshop at ACL* 2023
- Reviewer, Conference on Empirical Methods in Natural Language Processing (EMNLP): 2022
- Reviewer for EMNLP (Empirical Methods for Natural Language Processing) 2022

PROFICIENCIES

- Python
- React / Javascript / HTML / css