Hugh Zhang

hughzhang@seas.harvard.edu

Last updated: December 2023 hughbzhang.com

Education

Harvard University, PhD Candidate in Computer Science Advisor: David Parkes	2021 - present
Stanford University, BA with Honors in Economics Advisor: Gabriel Carroll	2016 - 2020
Work Experience	
Google DeepMind, Student Researcher with Steven Zheng and Swaroop Mishra	2023 -
Facebook Al Research, Visiting Research Engineer with Noam Brown	2020 - 2021
Google Brain, Research Intern with Daniel Duckworth and Arvind Neelakantan	2019 - 2020
Stanford NLP Group, Research Intern with Percy Liang and Tatsunori Hashimoto	2018 - 2019
Asana, Software Engineer	2015 - 2016

Publications

Conferences and Journals (* denotes equal contribution or alphabetical ordering)

- ▶ **Hugh Zhang**, "No-regret Learning Dynamics for Sequential Correlated Equilibria." The 22nd International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2023. (Extended Abstract)
- Anton Bakhtin*, Noam Brown*, Emily Dinan*, Gabriele Farina*, Colin Flaherty*, Daniel Fried*, Andrew Goff*, Jonathan Gray*, Hengyuan Hu*, Athul Paul Jacob*, Mojtaba Komeili*, Karthik Konath*, Adam Lerer*, Mike Lewis*, Alexander H. Miller*, Sasha Mitts*, Adithya Renduchintala*, Stephen Roller*, Dirk Rowe*, Weiyan Shi*, Joe Spisak*, Alexander Wei*, David Wu*, **Hugh Zhang***, Markus Zijlstra.* "Human-level play in the game of *Diplomacy* by combining language models with strategic reasoning." Science 2022.
- Zhang, Hugh, Adam Lerer, and Noam Brown "Equilibrium Finding in Matrix Games Via Greedy Regret Minimization." 2022 AAAI Conference on Artificial Intelligence.
- Hashimoto, Tatsunori*, Hugh Zhang*, and Percy Liang. "Unifying Human and Statistical Evaluation for Natural Language Generation." 2019 Conference of the North American Chapter of the Association for Computational Linguistics (Oral).

Workshops

- **Zhang, Hugh** and David C. Parkes. "Chain-of-Thought Reasoning is a Policy Improvement Operator." 2023 Workshop on Instruction Tuning and Instruction Following at NeurIPS 2023.
- Zhang, Hugh*, Daniel Duckworth*, Daphne Ippolito, and Arvind Neelakantan. "Trading Off Diversity and Quality in Natural Language Generation." 2021 Workshop on Human Evaluation of Natural Language Processing Systems at the Conference of the European Chapter of the Association for Computational Linguistics.
- Zhang, Yuchen, Yi Sun, Jacob Steinhardt, Aditya Ganesh, **Hugh Zhang**, Philip Hwang, Florian Tramer, Percy Liang "Generating Transferable Adversarial Examples via Smooth Max Ensembling." 2017 NeurIPS Adversarial Attacks and Defenses Competition Track.

Preprints (* denotes equal contribution)

- ▶ **Zhang, Hugh*,** Luca D'Amico Wong* and David C. Parkes. "Easy as ABCs: Unifying Boltzmann Q-Learning and Counterfactual Regret Minimization." 2023.
- **Zhang, Hugh.** "A Simple Adaptive Procedure Converging to Forgiving Correlated Equilibria." 2020. Senior Thesis in Economics

Other Technical Writing

- ▶ **Hugh Zhang**, "OpenAl: Please Open Source Your Language Model", The Gradient, 2019.
- Hugh Zhang, "Speech Recognition Systems Are Now Vulnerable To Adversarial Attacks", The Gradient, 2018.

Invited Talks

Chain-Of-Thought Reasoning is a Policy Improvement Operator

Scale AI - San Francisco, CA, USA - December 7, 2023

Unifying Human and Statistical Evaluation for Natural Language Generation

- NAACL Oral Minneapolis, MN, USA June 4, 2019
- Oxford Applied and Theoretical Machine Learning Group Oxford, UK June 20, 2019
- Apple Siri Lab Cambridge, UK June 25, 2019
- Naver Machine Learning Seoul, KR June 28, 2019
- Seoul National University Vision and Learning Lab Seoul, KR July 1, 2019

Teaching

Optimized Democracy (CS 238) Teaching Fellow

Spring 2023

▶ Teaching fellow for Ariel Proccacia's Optimized Democracy course (https://sites.google.com/view/optdemocracy23)

Summer Program in Applied Rationality (SPARC) Instructor

Summer 2020/2022

- High school summer program teaching real-life quantitative thinking skills (<u>https://sparc-camp.org/</u>).
- After spending two years as a student in 2014-2015, I returned as a junior counselor in 2017 and an instructor in 2020 and 2022, leading classes on game theory, unintended consequences, improv, and Go.

Code in Place Section Leader

Spring 2020

- A special one-time online offering of Stanford's introductory CS program open to anyone in the world.
- Course taught by Professor Chris Piech (https://compedu.stanford.edu/codeinplace/announcement/)
- As a section leader, responsible leading a weekly section and ensuring the success of my sectionees.

Computer Organization & Systems (CS 107) Course Assistant

Spring 2020

- Course taught by Professor Chris Gregg (http://web.stanford.edu/class/cs107/)
- As a course assistant, responsible for grading, holding weekly office hours, and running a weekly (virtual) lab

Natural Language Processing with Deep Learning (CS 224N) Course Assistant

Winter 2020

- Course taught by Professor Christopher Manning (http://web.stanford.edu/class/cs224n/)
- As a course assistant, responsible for grading / advising course projects and holding weekly office hours

Structured Liberal Education (SLE) Writing Tutor

Winter 2019

- SLE is a year-long course in the humanities for first years at Stanford similar to the Great Books programs at other universities. Students live together in the same dorm and read classics in philosophy, religion and literature spanning from ancient Greece to the present day.
- I participated in the program as a first year and returned in my third year as a writing tutor.

Programming Abstractions (CS 106B) Section Leader

Fall 2018

- Course taught by Professor Julie Zelenski (http://web.stanford.edu/class/cs106b/)
- As a section leader, responsible leading a weekly section, grading, and ensuring the success of my sectionees.

Programming Methodologies (CS 106A) Section Leader

Spring 2018

- Course taught by Professor Chris Piech (http://web.stanford.edu/class/cs106a/)
- As a section leader, responsible leading a weekly section, grading, and ensuring the success of my sectionees.

Other Activities

The Gradient, Co-founder and Editor

2018 - present

Digital magazine on the cutting edge of machine learning. Our articles have received 1M+ total views and are frequently shared and referenced by top AI researchers in industry and academia on Twitter, Reddit, and Hacker News.

Stanford Al Salon, Co-organizer and Moderator

2019 - 2020

Al Salon is a gathering of members of the Stanford Al Lab to discuss high-level topics in Al and machine learning, and their applications beyond the level of day to day research. We operate in the spirit of French enlightenment era salons by prohibiting the usage of electronics during discussions. I moderated discussions on topics spanning climate change, robotics, government policy and self driving cars.

Bridge Peer Counseling Center, Counselor and President

2017 - 2020

The Bridge is a student-run peer-counseling center at Stanford University that offers free anonymous counseling 24/7. We primarily cater to students, but will not turn away members from the general public. During my junior year, I served as the organization's president and was responsible for ensuring all shifts were filled despite a severe staffing shortage.

Parliamentary Debate

2017 - 2019

Member of the Stanford delegation to Worlds 2018 (WUDC) and Nationals 2018 (USUDC). While studying abroad at Oxford, I joined the Oxford Union and delivered the winning floor speech on a debate about US Foreign Policy.

High School Football

2011

Running back and outside linebacker for football team my freshman year.

Honors

Go (also known as weigi or baduk)

•	US Open, 10th place	2022
•	North American Master's Tournament, 6th place (tied).	2013
•	Youngest Team Member on the US Team at the World Mind Sports Games	2008
•	North American U12 Champion	2008

Other

•	Kempner Institute Graduate Research Fellowship	2023
•	National Science Foundation Graduate Research Fellowship	2022
•	John G. Sobieski Award for Creative Thinking (for my senior honors thesis in Economics)	2020
•	USA Computing Olympiad Finalist (top 24 high school students nationally)	2014