Zachary **Robertson**

Graduate Student

zroberts@stanford.edu | zrobertson466920.github.io/ | zrobertson466920 | zrobertson466920

"Play is the highest form of research."

Education

Stanford
Stanford, U.S.

Ph.D. in Computer Science — *Sept. 2022 -*

- School of Engineering Fellowship
- EDGE Fellowship
- NSF Honorary Mention

University of Illinois at Urbana-Champaign
Champaign, U.S.

MS in Computer Science — *Aug. 2020 - May. 2022*

- Thesis on Performance Metric Elicitation
- GEM fellowship
- Wing Kai Cheng Fellowship

University of Chicago
Chicago, U.S.

B.S. in Computational and Applied Math — *Oct. 2016 - Present*

- Graduated with Honors
- Jackie Robinson Scholarship

LoneStar Community College
Spring, TX, U.S.

Associate Degree in Science — *Oct. 2013 - May. 2016*

- President's List
- Home schooled with dual enrollment at local community college

Skills

Programming	Python, LaTeX, C++, Haskell
Machine Learning	PyTorch,TensorFlow
Math	Analysis, Linear Algebra, Statistical Learning Theory, Optimization, Statistics

Experience

STAIR Lab at Stanford (advised by Sanmi Koyejo)
Stanford, U.S.

Research Assistant — *Sept. 2022 - Present*

- Developing scalable oversight mechanisms and aligning AI systems with human preferences
- Contributing to the lab's work on AI superalignment, supported by an OpenAI award

Google
Mountain View, U.S.

Student Researcher — *Apr. 2022 - Aug. 2022*

- Designing tractable surrogates for welfare maximization
- Applications to the training of predicted click-through-rates models used in Ads

Lam Research
Fremont, U.S.

Research Intern — *June 2020 - March 2021*

- Lam is a major player in the design, manufacture, marketing, and service of semiconductor processing equipment used in the fabrication of integrated circuits
- Applied knowledge of reinforcement learning and machine learning to optimize the production flow of semi-conducting wafers
- Used Python, Open AI Gym, and Tensorflow

Robot Intelligence through Perception Lab at TTIC
Chicago, U.S.

Research Assistant — *Jun. 2019 - June 2020*

- Implemented a model to do sparse-depth completion on a robot using PyTorch, Docker, and SLURM.
- Trained a simulated UR5 arm to pick up blocks from a table with natural language using PyTorch, OpenGym, and Mujoco.
- Engaged in theoretical study of reinforcement learning using regret based approach for Honor's Thesis

Exo-Planets Group at UChicago *Chicago, U.S.*

Research Assistant *June 2017 - Oct. 2018*

- Worked on increasing calibration precision using machine learning (EM and pattern matching) to help align known spectra to those reported on the spectrograph pixel output
- Published a paper describing more efficient calibration software for spectrographs

Allen and Company LLC. *New York, U.S.*

Analyst *June 2018 - Aug. 2018*

- Allen and Co. is a boutique investment banking firm that has advised on deals such as Facebook's acquisition of WhatsApp (2014) and hosts the annual Sun Valley Conference in Idaho
- Brought domain knowledge to help evaluate quantum computing startups and cryptocurrencies, automated formulaic internal reporting, and analyzed investor sentiment

Algorithmic Dynamics Lab *Stockholm, Sweden*

Software Contributor *June 2018 - Aug. 2018*

- Developed an algorithm to approximately calculate algorithmic complexity for array data structures using Python.
- Completed remote course-work on algorithmic information dynamics hosted by the Santa Fe Institute and taught by Hector Zenil

Honors & Awards

2022	**Awardee**, Superalignment Fast Grant (OpenAI)	*Stanford*
2022	**Awardee**, Accelerating Foundation Models Research (Microsoft)	*Stanford*
2022	**Recipient**, School of Engineering Fellowship	*Stanford*
2022	**Recipient**, EDGE Fellowship	*Stanford*
2021	**Honorary Mention**, NSF GRFP Fellowship	*UIUC*
2021	**Recipient**, Wing Kai Cheng Fellowship	*UIUC*
2020	**Recipient**, GEM Fellowship	*UChicago*
2018	**Recipient**, Dean's List	*UChicago*
2017	**Winner**, Illinois Blockchain Hackathon	*Chicago/Urbana*
2016	**Scholar**, Jackie Robinson Foundation	*New York*

Extracurricular Activity

Alignment Newsletter *CHAI*

Contributor *Sept. 2019 - Present*

- Writes summaries and opinions on advances in machine learning for the Alignment Newsletter
- Over 1k subscribers

Inter-House Council *UChicago*

Chair of Philanthropy Committee *Oct. 2016 - June 2018*

- Planned and implemented an update to the Saturday night meal pilot program for students with financial need.
- Used a portion of the saved money to help feed local people in need.

Chicago Maroon *UChicago*

Data Analysis & Writer *Oct. 2016 - 2018*

- Conducted student polls and analyzed the results

Publications and Talks

1. Zachary Robertson and Sanmi Koyejo. Layer-wise alignment is conserved in deep neural networks. *ICML Workshop on Localized Learning (LLW)*, 2023

2. Boxiang Lyu, Zhe Feng, Zachary Robertson, and Sanmi Koyejo. Pairwise ranking losses of click-through rates prediction for welfare maximization in ad auctions. *ICML*, 2023

3. Zachary Robertson, Hantao Zhang, and Sanmi Koyejo. Cooperative inverse decision theory for uncertain preferences. *AISTATS*, 2023

4. Zachary Robertson, Hantao Zhang, and Sanmi Koyejo. Probabilistic performance metric elicitation. *1st Workshop on Human and Machine Decisions (WHMD 2021) at NeurIPS 2021*, 2022

5. Zachary Robertson and Matthew Walter. Concurrent training improves the performance of behavioral cloning from observation. *arXiv preprint arXiv:2008.01205*, 2020

6. Julian Stürmer, Andreas Seifahrt, Zachary Robertson, Christian Schwab, and Jacob L Bean. Echelle++, a fast generic spectrum simulator. *Publications of the Astronomical Society of the Pacific*, 131(996):024502, 2018