

# ATTICUS GEIGER

Stanford University  
atticusg@stanford.edu

## EDUCATION

---

### Ph.D. Linguistics

Stanford University, Stanford, CA

September 2019 -

### M.S. Computer Science

Stanford University, Stanford, CA

September 2016 - June 2019

### B.S. Symbolic Systems

Stanford University, Stanford, CA

September 2015- June 2019

## CAREER HISTORY

---

### PhD Research, Stanford University

September 2019 -

- Advised by Christopher Potts and Thomas Icard
- Primary author publications at NeurIPS, EMNLP, and Black Box EMNLP
- Collaborated with and mentored several master's students and undergraduates

### Honors Thesis Research, Stanford University

June 2018 - September 2018

- Acquired grant for self led natural language inference research project advised by Chris Potts, Thomas Icard, and Lauri Karttunen
- Constructed artificial natural language inference dataset using logic models
- Designed task specific neural model with standout performance on the generated datasets

### Symbolic Systems Research Intern, Stanford University

June 2017 - September 2017

- Worked with Lauri Karttunen and Ignacio Cases to create a Natural Language Inference dataset focused on implicatives
- Implemented neural network models for natural language inference in Tensor Flow

### Software Engineer Intern, Alaska Satellite Facility

June 2016 - September 2016

- Created a AWS cloud processing infrastructure and website interface
- Provided a service that automatically processes incoming satellite data
- Accomplished significant, largely self led software implementation

## SKILLS

---

**Programming:** Python, C++, C, TensorFlow, PyTorch, AWS

**Social Skills:** I love collaborating and find great joy in mentoring!

## AWARDS AND HONORS

---

**Firestone Medal** Award for B.S. honors thesis titled *Can Natural Language Inference Models Perform Natural Logic Reasoning?* and advised by Chris Potts and Thomas Icard.

## GRANTS

---

PI, Fair Adversarial Tasks for Natural Language Understanding. Facebook Research: Robust Deep Learning for Natural Language Processing. Co-PI Chris Potts. 20192020

1. Zhengxuan Wu\*, **Atticus Geiger**\*, Josh Rozner, Elisa Kreiss, Hanson Lu, Thomas Icard, Christopher Potts, and Noah D. Goodman. Causal distillation for language models. arXiv:2112.02505, 2021
2. **Atticus Geiger**\*, Zhengxuan Wu\*, Hanson Lu\*, Josh Rozner, Elisa Kreiss, Thomas Icard, Noah D. Goodman, and Christopher Potts. Inducing causal structure for interpretable neural networks. arXiv:2112.00826, 2021
3. **Atticus Geiger**\*, Hanson Lu\*, Thomas Icard, and Christopher Potts. Causal abstractions of neural networks. In *Advances in Neural Information Processing Systems*, 2021
4. Christopher Potts, Zhengxuan Wu, **Atticus Geiger**, and Douwe Kiela. DynaSent: A dynamic benchmark for sentiment analysis. In *Proceedings of the Association for Computational Linguistics*, 2021
5. Douwe Kiela, Max Bartolo, Yixin Nie, Divyansh Kaushik, **Atticus Geiger**, Zhengxuan Wu, Bertie Vidgen, Grusha Prasad, Amanpreet Singh, Pratik Ringshia, Zhiyi Ma, Tristan Thrush, Sebastian Riedel, Zeerak Waseem, Pontus Stenetorp, Robin Jia, Mohit Bansal, Christopher Potts, and Adina Williams. Dynabench: Rethinking benchmarking in NLP. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4110–4124, Online, June 2021. Association for Computational Linguistics
6. **Atticus Geiger**, Kyle Richardson, and Christopher Potts. Neural natural language inference models partially embed theories of lexical entailment and negation. In *Proceedings of the Third BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP*, pages 163–173, Online, November 2020. Association for Computational Linguistics
7. **Atticus Geiger**, Alexandra Carstensen, Michael C. Frank, and Christopher Potts. Relational reasoning and generalization using non-symbolic neural networks. Ms., Stanford University, 2020
8. **Atticus Geiger**, Ignacio Cases, Lauri Karttunen, and Christopher Potts. Posing fair generalization tasks for natural language inference. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 4475–4485, Stroudsburg, PA, November 2019. Association for Computational Linguistics
9. **Atticus Geiger**. Can natural language inference models perform natural logic reasoning? B.s. thesis, Stanford University, 2019
10. Ignacio Cases, Clemens Rosenbaum, Matthew Riemer, **Atticus Geiger**, Tim Klinger, Alex Tamkin, Olivia Li, Sandhini Agarwal, Joshua D. Greene, Dan Jurafsky, Christopher Potts, and Lauri Karttunen. Recursive routing networks: Learning to compose modules for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Stroudsburg, PA, June 2019. Association for Computational Linguistics
11. **Atticus Geiger**, Ignacio Cases, Lauri Karttunen, and Christopher Potts. Stress-testing neural models of natural language inference with multiply-quantified sentences. Ms., Stanford University. arXiv 1810.13033, 2018
12. Steve A. Arko, Rose Hogenson, **Atticus Geiger**, Jake Herrmann, Brian Buechler, and Kirk Hogenson. Sentinel-1 Archive and Processing in the Cloud using the Hybrid Pluggable Processing Pipeline (HyP3) at the ASF DAAC. In *AGU Fall Meeting Abstracts*, volume 2016, pages G43A–1040, December 2016
13. Kirk. Hogenson, Steve A. Arko, Brian Buechler, Rose Hogenson, Jake Herrmann, and **Atticus Geiger**. Hybrid Pluggable Processing Pipeline (HyP3): A cloud-based infrastructure for generic processing of SAR data. In *AGU Fall Meeting Abstracts*, volume 2016, pages IN21B–1740, December 2016

## INVITED TALKS

---

1. 2022. Causal Abstraction and Computational Explanations in Artificial Intelligence. Seminar Series in Stanford Psychology for cognitive and neuroscience areas. January 21.
2. 2021. Causal Abstractions and Interchange Intervention Training. Allen AI Aristo Group. September 23.
3. 2021. Causal Abstraction for Neural Network Analysis. McDonnell plenary. July 8.
4. 2020. Modular Representation in Neural Natural Language Inference Models. Allen AI Aristo Group. June 19.

## TEACHING

---

1. 2021 (Fall). Probabilistic Pragmatics. Guest Lecture for LINGUIST 245 *Psycholinguistics* at Stanford. (Graduate and Undergraduate)
2. 2020 (Spring). Causal Abstraction with Applications to Computational Implementation and Neural Network Analysis. Guest Lecture for PHIL 359 *Logic Seminar* at Stanford. (Graduate)
3. 2020 (Spring). Natural Logic. Guest Lecture for LINGUIST 130B *Lexical Semantics* at Stanford. (Undergraduate)
4. 2019 (Spring). Evaluating NLU Models with Harder Generalization Tasks. Guest Lecture for CS 224U *Natural Language Understanding* at Stanford. (Graduate)
5. Teaching Assistant Positions: CS 224U *Natural Language Understanding* at Stanford 2019 (Spring), LINGUIST 130B *Lexical Semantics* at Stanford 2020 (Spring) , and LINGUIST 245 *Psycholinguistics* at Stanford 2021 (Fall).