

Curriculum Vitae

Scott Viteri

October 14, 2024

Address: 1215 Emerson St, Palo Alto, CA 94301

Phone: (480) 266-0242

Email: sviteri@stanford.edu

Google Scholar: Scott Viteri - Google Scholar

Profile

I am interested in training AIs to be human-like both in terms of the concepts they contain, as well as the values they hold. I think the key to this lies in training the language model to produce its own memory during training and to influence its future training data. In this vein, I have been optimizing pre-trained models to produce Chain-of-Thought reasoning which is *informative* to itself and to humans. In the future, I am interested in using a similar blend of unsupervised learning and reinforcement learning towards active learning, in the hopes that this sets the preconditions to raise the AI like a child, and have that result process in something child-like.

During my PhD, my focus has evolved from formal verification and programming languages to AI alignment. Prior to this, I majored in computer science and electrical engineering at MIT, contributing to AI and robotics research. After MIT, I explored interactive theorem proving at CMU with Simon Dedeo, publishing research on abduction in mathematics in the *Cognition* journal. My primary character trait is curiosity, and I really love math.

Education

- **2014-2018** Massachusetts Institute of Technology, BS Computer Science and Electrical Engineering
- **2020-Current** Stanford University, Doctoral Candidate

Employment

Stanford, CS PhD Candidate in AI Safety, Stanford CA

1/2022 – Current

- Developed a novel method to use Proximal Policy Optimization to train informative language models in ArXiv pre-print *Markovian Agents for Informative Language Modeling*, now *in submission at ICLR*.
- Created and taught Stanford course CS362: Research in AI Alignment.
- Won \$10K in Alignment Research Center contest for Eliciting Latent Knowledge (ELK) solution proposals.
- Presented talks at the Topos Institute and the Alignable Structures conference on ontology maps.
- Mentored for AGI safety fundamentals reading group and TA'd for Stanford's AI Alignment courses (STS 10SI, STS 20SI).

Stanford, CS PhD Student in Formal Verification, Stanford CA

9/2019 – 1/2022

- Created a flexible method for shifting between higher-order and first-order logic in program verification, with Oded Padon, using the Veribetrkv key-value store.
- Published to IJCAR on proof production in CVC5 SMT solver.
- Co-wrote neural decoder auto-completion framework using context-sensitive predictions.

Carnegie Mellon University, PL Researcher, Pittsburgh PA

10/2018 - 8/2019

- Analyzed proof trees of major theorems in Coq for automatic theorem proving patterns.
- Extended plugin to reify Coq AST, built a repository for analyzing ASTs from famous mathematical proofs.
- Published research on generative tree models for proof structure analysis in *Cognition*.

MIT, Artificial Intelligence Researcher, Cambridge, MA

1/2018 - 6/2018

- Worked on a Coq framework for knowledge representation and planning, treating sub-goaling as automatic theorem proving.
- Used coherent logic for theorem-proving “hints” in planning systems.

MIT, Robot Locomotion Group, Cambridge, MA

1/2017

- Used Vicon camera systems to analyze performance of state estimators for Atlas and Valkyrie robots.

Amazon, Software Engineering Intern, Seattle, WA

6/2016 – 8/2016

- Developed machine learning tools to predict user behavior on the Amazon website, reducing server requests with predictive modeling.
- Worked with Spark, Hadoop File Systems, and Elastic Map Reduce.

Roambotics, Software Development Intern, Phoenix, AZ

6/2015 - 8/2015

- Designed and implemented architecture for the company’s autonomous robot, integrating C and C++ libraries using Ant build tools.

Technical Skills

- **Computer Languages:** Python, Racket, Lean
- **Other Tools:** GNU/Linux (Ubuntu, Arch)
- **Hardware:** Comfortable with laser cutters, lathes, mills, 3D printers, and other shop tools.

Interests and Activities

Singing, guitar, Brazilian Jiu Jitsu, Zouk dancing, and math