

ZHANGIR AZERBAYEV

(+1)2039017566 ◊ zhangir.azerbayev@yale.edu

EDUCATION

Yale University

B.S Mathematics (Expected May 2023)

August 2019-Present

Stanford Online High School

Part-time student. Courses: linear algebra, abstract algebra.

August 2018-May 2019

PAPERS

Preprints:

Zhangir Azerbayev, Bartosz Piotrowski, Hailey Schoelkopf, Edward W. Ayers, Dragomir Radev, Jeremy Avigad. ProofNet: Autoformalizing and Formally Proving Undergraduate-Level Mathematics. *arXiv:2302.12433*

Zhangir Azerbayev, Ansong Ni, Hailey Schoelkopf, Dragomir Radev. Explicit Knowledge Transfer for Weakly-supervised Code Generation. *arXiv:2211.16740*

Marlene Berke, **Zhangir Azerbayev**, Mario Belledone, Zenna Tavares, Julian Jara-Ettinger. MetaCOG: Learning a metacognition to recover what objects are actually there. *Preprint*. <http://www.zenna.org/publications/metaCOG2022.pdf>

Publications:

Zhangir Azerbayev, Bartosz Piotrowski, Jeremy Avigad. ProofNet: A Benchmark for Formally Proving and Formalizing Undergraduate-level Mathematics. *MATH-AI workshop at NeurIPS 2022 contributed talk*.

Ansong Ni, **Zhangir Azerbayev**, Mutethia Mutuma, Troy Feng, Yusen Zhang, Tao Yu, Ahmed Hassan Awadallah, and Dragomir Radev. SummerTime: Text Summarization Toolkit for Non-experts. *EMNLP 2021 System Demonstration*.

OPEN SOURCE

proofGPT models

- Open-source pre-trained language models for mathematical tasks.
- Trained for 8 billion tokens on the [proof-pile](#), a curated dataset of mathematical text.
- Huggingface: [1.3B weights.](#), [6.7B weights.](#)

Mathlib Semantic Search

- Semantic search engine for Lean mathlib built with the OpenAI API.
- Github: [zhangir-azerbayev/mathlib-semantic-search](#). [Web demo](#).

Lean Chat

- A VS code extension for translating natural language theorem statements into Lean.
- Visual Studio Marketplace: [Lean Chat](#)
- Uses OpenAI's Codex API as a backend.

SummerTime: Toolkit for Automatic Text Summarization

- GitHub: Yale-LILY/SummerTime, 216 stars.
- One of four people responsible for the concept and early-stage design decisions.
- Changed 13,000 lines of code.

particle_filtering

- Github: zhangir-azerbayev/particle_filtering
- Raw numpy implementation of particle filtering for sequential Bayesian inference.

Lean Mathematical Library

- GitHub: leanprover-community/mathlib
- Contributed formalizations of [alternating maps](#) and [exterior algebras](#).

WRITING

The Future of Interactive Theorem Proving?

- Guest post on [Kevin Buzzard's Xena Project blog](#).
- Reached top 10 on Hacker News.

AWARDS AND HONORS

Wu Tsai Institute for Neuroscience MIND Fellow

Summer 2021

Research fellowship awarded for work in the Computational Social Cognition Lab

GROUPS

Yale Rugby Football Club

2019-2022

Player (2019-2021), Social Chair (2021), Treasurer (2022)

15 hour-a-week commitment including trainings, team lifts, and games.

LABS

Visiting Researcher, Hoskinson Center for Formal Mathematics at Carnegie Mellon University

June - August 2022

Supervised by Prof. Jeremy Avigad

Researcher, Yale Language, Information & Learning lab

December 2020-Present

Supervised by Prof. Dragomir Radev

Researcher, Yale Computational Social Cognition Lab

December 2020-September 2021

Supervised by Prof. Julian Jara-Ettinger.