

Selim Jerad

selimjerad@gmail.com | [Linkedin](#) | [Google Scholar](#)

I am an incoming PhD student at TTIC, with a strong interest in understanding the theoretical underpinnings of language models.

EDUCATION

2026 - 2031 (expected)	TOYOTA TECHNOLOGICAL INSTITUTE AT CHICAGO <i>PhD student in Computer Science. Supervised by Prof. William Merrill.</i> My research focuses on establishing the abilities and limitations of language modeling architectures (transformers, hybrid models) through tools from formal language theory and circuit complexity , with an eye towards developing more expressive architectures .	Chicago, USA
2023 - 2025	ETH ZÜRICH <i>Master's degree in Computer Science</i> - Major: Machine Intelligence, Minor: Theoretical Computer Science - Grade average: 5.33 / 6 - Relevant coursework: Large Language Models, Optimization for Data Science, Deep Learning, Advanced Formal Language Theory, Probabilistic AI	Zürich, Switzerland
2020 - 2023	EPFL <i>Bachelor's degree in Communication Systems</i>	Lausanne, Switzerland

WORK EXPERIENCE

Winter 2025	RESEARCH INTERN AT OKINAWA INSTITUTE OF SCIENCE AND TECHNOLOGY Investigating the impact on transformer expressivity of various positional embeddings	Okinawa, Japan
Spring 2025	VISITING ACADEMIC AT NYU CENTER FOR DATA SCIENCE 6-month research thesis on the ability of transformers to recognize context-free languages	New York City, USA
Summer 2023	RESEARCH INTERN AT GEBEL GROUP (MARKET MAKING START-UP) Trained various models on forecasting cryptocurrency returns	Paris, France
Summer 2022	RESEARCH INTERN AT INSTADEEP (LEADER IN APPLIED ML) Trained DNNs to produce Antibiotic Resistant Gene annotations given a protein sequence	London, UK

PUBLICATIONS

[UNIQUE HARD ATTENTION: A TALE OF TWO SIDES](#)* (*Oral, ACL 2025*)

- Theoretically characterized the expressive power of unique-hard attention transformers

[CONTEXT-FREE RECOGNITION WITH TRANSFORMERS](#)* (*ICML 2026*)

- Analyzed how theoretically and empirically, dynamically-scaling transformers can process grammatical syntax

[DISENTANGLING THE EXPRESSIVITY OF RoPE](#)* (*COLM 2026*)

- Characterized the expressive power of rotary position embeddings

*first-author publication

TEACHING EXPERIENCE

TEACHING ASSISTANT AT ETHZ

- Stochastics and Machine Learning (Spring 2024), Algorithms Lab (Fall 2024)

TEACHING ASSISTANT AT EPFL

- Linear Algebra (Fall 2021), Discrete Mathematics 1 (Fall 2021 and 2022), Theory of Computation (Spring 2023)