

Thomas Jiralerspong

Université de Montréal
Mila
Montreal, Canada

thomasjiralerspong@gmail.com
superkaiba.github.io
+1 (514) 625-9308

[Google Scholar](#)
[LinkedIn](#)
[GitHub](#)

Education

Université de Montréal

PhD - Computer Science

In progress

Supervisors: [Yoshua Bengio](#) & [Doina Precup](#)

Vanier Canada Graduate Scholarship (150 000\$)

FRQNT Scholarship (40 000\$) (Rank #1 among all applicants in category)

NSERC Canada Graduate Scholarship (17 500\$)

Hydro-Québec Excellence Scholarship (10 000\$)

Arbour Scholarship (7 500\$)

Massachusetts Institute of Technology

Brains, Minds, and Machines Summer Course

2024

McGill University

B.Sc., Honours Computer Science

2023

Supervisors: [Blake Richards](#) & [Doina Precup](#)

GPA: 4.00/4.00

Exchange semester at the **National University of Singapore**

J.W. McConnell Major Entrance Scholarship (9 000\$)

Refereed Conferences

Eric Elmoznino*, **Thomas Jiralerspong***, Yoshua Bengio, Guillaume Lajoie. “A Complexity-Based Theory of Compositionality.” In *Forty-Second International Conference on Machine Learning (ICML)*. 2025.

Jin Hwa Lee*, **Thomas Jiralerspong***, Lei Yu, Emily Cheng. “Geometric Signatures of Compositionality Across a Language Model’s Lifetime.” In *The 63rd Annual Meeting of the Association for Computational Linguistics (ACL)*. 2025.

Ezekiel Williams, Avery Ryoo*, **Thomas Jiralerspong***, Matt Perich, Guillaume Lajoie. “Expressivity of neural networks with random weights and learned biases.” In *The 13th International Conference on Learned Representations (ICLR)*. 2025.

Jean-Pierre Falet, Hae Beom Lee, Nikolay Malkin, Chen Sun, Dragos Secieru, **Thomas Jiralerspong**, Dinghuai Zhang, Guillaume Lajoie, Yoshua Bengio. “Delta-AI: Local Objectives for Amortized Inference in Sparse Graphical Models” In *Twelfth International Conference on Learning Representations (ICLR)*. 2024.

Chen Sun, Wannan Yang, **Thomas Jiralerspong**, Dane Malenfant, Benjamin Alsbury-Nealy, Yoshua Bengio, Blake Richards. “Contrastive Retrospection: honing in on critical steps for rapid learning and generalization in RL.” In *Thirty-seventh Annual Conference on Neural Information Processing Systems (NeurIPS)*. 2023.

Flemming Kondrup*, **Thomas Jiralerspong***, Elaine Lau, Nathan de Lara, Jacob Shkrob, My Duc Tran, Doina Precup, Sumana Basu. “Towards Safe Mechanical Ventilation Treatment Using Deep Offline Reinforcement Learning.” In *Thirty-seventh AAAI Conference on Artificial Intelligence (AAAI)*. 2023.

Marshall Wang, John Willes, **Thomas Jiralerspong**, Matin Moezzi. “A Comparison of Classical and Deep Reinforcement Learning Methods for HVAC Control.” In *20th IEEE International Conference on Ubiquitous Intelligence and Computing (UIC)*. 2023.

Refereed Workshops

Thomas Jiralerspong, Berton Earnshaw, Jason Hartford, Yoshua Bengio, Luca Scimeca. “Shaping Inductive Bias in Diffusion Models through Frequency-Based Noise Control” In *The ICLR Workshop on Deep Generative Models in Machine Learning: Theory, Principle and Efficacy (DeLTa)*. 2025.

Marco Jiralerspong, **Thomas Jiralerspong**, Vedant Shah, Dhanya Sridhar, Gauthier Gidel. “General Causal Imputation via Synthetic Interventions.” In *The Causal Representation Learning Workshop at NeurIPS*. 2024.

Thomas Jiralerspong*, Xiaoyin Chen*, Yash More, Vedant Shah, Yoshua Bengio. “Efficient Causal Graph Discovery Using Large Language Models.” In *How Far Are We From AGI? Workshop at ICLR*. 2024.

Thomas Jiralerspong*, Flemming Kondrup*, Doina Precup, Khimya Khetarpal. “Forecaster: Towards Temporally Abstract Tree-Search Planning from Pixels.” In *Seventh Workshop on Generalization in Planning at NeurIPS*. 2023.

Flemming Kondrup*, **Thomas Jiralerspong***, Elaine Lau, Nathan de Lara, Jacob Shkrob, My Duc Tran, Doina Precup, Sumana Basu. “Deep Conservative Reinforcement Learning for Personalization of Mechanical Ventilation Treatment.” In *The Multi-disciplinary Conference on Reinforcement Learning and Decision Making (RLDM)*. 2022.

Research Experience

Anthropic

Research Fellow

March 2024 - Present

Project: Mechanistic interpretability of emergent misalignment models

Occam AI

LLM Research Intern

Jun 2024 - Present

Projects: Optimization of interactions between network of LLM agents, automated SQL query generation using LLMs

Waabi

Deep Learning Research Intern

Jun 2023 – Aug 2023

Mentored by Kelvin Wong and Chris Zhang

Project: Realistic and controllable traffic simulation using a transformer based variational autoencoder

Reasoning and Learning Lab, Mila/McGill University

Research Intern

Jan 2022 – Aug 2023

* Equal Contribution

Supervised by Prof. Doina Precup

Project: Model-based reinforcement learning with affordance aware tree-search planning directly from pixels

Learning in Neural Circuits Lab, Mila/McGill University

Research Intern

Sep 2022 – Aug 2023

Supervised by Prof. Blake Richards

Project: Contrastive learning to discover critical states for reinforcement learning in sparse reward environments

Vector Institute for A.I.

Machine Learning Research Intern

Sep 2022 – Dec 2022

Mentored by John Willes and Marshall Wang

Project: Model-based reinforcement learning for HVAC control

Project X, Machine Learning Research Competition

Co-leader of McGill's Team

Jun 2021 – Feb 2022

Received the highest score out of 25 submitted papers

Project: Deep offline conservative reinforcement learning for mechanical ventilation treatment

Industry
Experience

Amazon Web Services (AWS) – S3 Team

Software Development Engineer Intern

May 2022 – Jul 2022

Project: JavaScript/Python tool to automate the Incremental Backup recovery system for AWS S3 (stores ~14 trillion objects)

Square Enix

Software Development Intern

May 2021 – Aug 2021

Project: Localization system to allow a MOBA game to be translated into over 10 languages

Expedia

Software Development Intern

Jun 2019 – Aug 2019

Project: React/TypeScript tool to identify which elements of a webpage are broken and conveniently display them to developers

Teaching

Université de Montréal

Teaching Assistant, Representation Learning

2023

McGill A.I. Society

Organizer/Teaching Assistant, Accelerated Intro to ML

2021 – 2023

McGill University

Teaching Assistant, Software Systems

2021 – 2022

Guest Lecturer, Theory of Machine Learning

2022

Honors

Vanier Canada Graduate Scholarship (150000\$)

2025

	FRQNT Master's Scholarship (40000\$) (Rank #1 among all applicants in category)	
	2024	
	Arbour Scholarship (7500\$)	2024
	Hydro-Québec Excellence Scholarship (10000\$)	2024
	Chosen to attend the 10th Heidelberg Laureate Forum	2023
	NSERC Canada Graduate Scholarship (17500\$)	2023
	University of Montreal Master's Scholarship (5000\$)	2023
	McGill Mobility Bursary for Exchanges (6000\$)	2022
	Winner of UofT AI's Project X competition (25000\$)	2022
	J.W. McConnell Major Entrance Scholarship (9000\$)	2020 – 2022
	CIBPA Foundation Bursary (1000\$, 2500\$, 1000\$)	2021, 2022, 2023
	Marianopolis College Valedictorian	2020
	Governor General of Canada's Academic Medal	2020
Invited Talks	Canadian Undergraduate Conference on AI (CUCAI)	2022
	University of Toronto AI Conference	2022
	McGill AI Society Learnathon	2022
Professional Activities	Mila	
	Chairman of Lab Representatives	2023 – Present
	Chairman of Social Committee	2023 – Present
	Executive Member of Recruitment Committee	2023 – Present
	McGill AI Society	
	Senior Advisor	2023 – Present
	Technical Project Manager	2021 – 2023
	Montreal AI & Neuroscience Conference	
	Organizer – Introduction to deep learning with PyTorch workshop	2022
	McGill NeuroTech	
	Machine Learning Developer	2021 – 2022
	McGill Robotics	
	Software Developer	2020 – 2021
Languages	Native: English, French	
	Advanced: Italian, Spanish	
	Beginner: Mandarin, Japanese	
Skills	Programming Languages: Python, Java, JavaScript, R, C, C++, C#, OCaml, SQL, HTML, CSS	
	Machine Learning Libraries: PyTorch, TensorFlow, Keras, Pandas, NumPy, Matplotlib	
	Other: L ^A T _E X, Slurm, Jupyter Notebooks, Perforce, GitHub, Jira, Unity	

Press

SciLogs. Nina Beier. Jan 24, 2024. [What Do Food and Research Have in Common? More Than You Might Think.](#)

The McGill Tribune. Mikaela Shadick. March 15, 2022. [Six McGill undergrads win UofT international artificial intelligence competition.](#)

McGill Reporter. Richard Deschamps. March 1, 2022. [Undergrad team uses machine learning to create a better hospital ventilator.](#)

Advanced Coursework

Université de Montréal

Representation Learning
Reinforcement Learning & Optimal Control
Scaling Laws
Causal Inference & Machine Learning
Probabilistic Graphical Models

McGill University

Reinforcement Learning
Brain Inspired Artificial Intelligence
Honours Math for Machine Learning
Probabilistic Programming
Network Science

National University of Singapore

Quantum Computing
Information Theory