EDWARD HU

ai@edwardj.hu | @edwardjhu

EDUCATION

Université de Montréal, Montréal, QC, Canada — Jan 2022 – Present
Doctor of Philosophy in Computer Science
- Advised by Yoshua Bengio
- Cumulative GPA: 4.00/4.00

Johns Hopkins University, Baltimore, MD — Class of 2019
Bachelor of Science in Cognitive Science, Computer Science
- Advised by Benjamin Van Durme
- Cumulative GPA: 3.96/4.00
- Member of Upsilon Pi Epsilon, Omega Psi

NOTABLE PUBLICATIONS

- ***LoRA*: Low-Rank Adaptation of Large Language Models**
  *Edward J. Hu**, Y Shen*, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, W. Chen — ICLR 2022
- ***µTransfer*: Tuning Large Neural Networks via Zero-Shot Hyperparameter Transfer**
  G. Yang*, *Edward J. Hu**, I. Babuschkin, S. Sidor, X. Liu, D. Farhi, N. Ryder,
  J. Pachocki, W. Chen, J. Gao — NeurIPS 2021
- ***µP:* Feature Learning in Infinite-Width Neural Networks**
  Greg Yang, *Edward J. Hu* — ICML 2021

OPEN-SOURCE REPOSITORIES

- **lora**: efficient model adaptation (3,700+ stars)
  Maintainer github.com/microsoft/lora/
- **mup:** principled model parametrization for hyperparameter transfer (800+ stars)
  Maintainer github.com/microsoft/mup/

RESEARCH EXPERIENCE

Mila, Université de Montréal — Jan 2022 – Present
Montreal, Quebec, Canada
PhD Student
- Research principled approaches to robust reasoning
- Contributor to the theories of generative flow networks and creator of GFlowNet-EM

Microsoft Research AI / Microsoft Azure AI, — Sept 2019 – Dec 2021
Microsoft Corporation, Redmond, WA
AI Resident / Researcher
- Conducted research on the fundamentals of deep learning, principled approaches to large-scale machine learning, and its practical deployment
- Drove early research integration of OpenAI GPT-3 as one of the 8 members of the Microsoft-OpenAI partnership onsite engineering team
- Led research collaborations with CMU and OpenAI, resulting in two publications, one at ICLR 2022 and the other at NeurIPS 2021
- Generated 2 U.S. patents as the lead inventor and published 5 papers at major conferences including a best paper for the Trustworthy ML workshop at ICLR 2020

Center for Language and Speech Processing, Jan 2018 – Aug 2019
Johns Hopkins University, Baltimore, MD
Research Assistant

- Conducted research on paraphrase generation and monolingual rewriting with applications in data augmentation and plagiarism detection
- Built ParaBank, the largest English paraphrase dataset at the time with more than 4 billion generated tokens
- Developed an lexically-constrained decoding algorithm 5 times more efficient than the best prior approach while being more accurate
- Implemented features including improved lexically-constrained decoding and decoding by sampling in AWS Sockeye

ONLINE ARTICLES

- **µTransfer: A technique for hyperparameter tuning of enormous neural networks**
  Edward Hu, Greg Yang, Jianfeng Gao — Microsoft Research Blog **(Link)**
- **On infinitely wide neural networks that exhibit feature learning**
  Edward Hu, Greg Yang — Microsoft Research Blog **(Link)**

SELECTED PUBLICATIONS

- **GFlowNet-EM for Learning Compositional Latent Variable Models**
  *Edward J. Hu**, N. Malkin*, M. Jain, K.E. Everett, A. Graikos, Y. Bengio — ICML 2023
- **GFlowNet Foundations** — JMLR 2023
  Yoshua Bengio*, Salem Lahlou*, Tristan Deleu*, *Edward Hu*, Mo Tiwari, Emmanuel Bengio
- **LoRA: Low-Rank Adaptation of Large Language Models**
  *Edward J. Hu**, Y. Shen*, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, W. Chen — ICLR 2022
- **Tuning Large Neural Networks via Zero-Shot Hyperparameter Transfer**
  G. Yang*, *Edward J. Hu**, I. Babuschkin, S. Sidor, X. Liu, D. Farhi, N. Ryder, J. Pachocki, W. Chen, J. Gao — NeurIPS 2021
- **Feature Learning in Infinite-Width Neural Networks**
  Greg Yang, *Edward J. Hu* — ICML 2021
- **Improved Image Wasserstein Attacks and Defenses** *(Best Paper)*
  Edward J. Hu, Adith Swaminathan, Hadi Salman, Greg Yang — ICLR 2020 Workshop
- **Randomized Smoothing of All Shapes and Sizes**
  G. Yang, T. Duan, *Edward J. Hu*, H. Salman, I. Razenshteyn, J. Li — ICML 2020
- **Large-scale, Diverse, Paraphrastic Bitexts via Sampling and Clustering**
  Edward J. Hu, A. Singh, N Holzenberger, M. Post, B. Van Durme — CoNLL 2019
- **Improved Lexically-Constrained Decoding for Translation and Monolingual Rewriting** — NAACL 2019
  Edward J. Hu, H. Khayrallah, R. Culkin, P. Xia, T. Chen, M. Post, B. Van Durme
- **ParaBank: Monolingual Bitext Generation and Sentential Paraphrasing via Lexically-constrained Neural Machine Translation** — AAAI 2019
  Edward J. Hu, Rachel Rudinger, Matt Post, Benjamin Van Durme
- **Towards a Unified Natural Language Inference Framework to Evaluate Sentence Representations** — EMNLP 2018
  A. Poliak, A. Haldar, R. Rudinger, *Edward J. Hu*, E. Pavlick, A. S. White, B. Van Durme

HONORS / SCHOLARSHIPS

Graduate fellowship in AI 2023, CA$10,000	Université de Montréal
Nominated for Apple Scholars in AI/ML 2023	Université de Montréal
Best paper award, US$1,000	ICLR 2022 Workshop
Departmental Honors in Cog. Sci., Comp. Sci.	Johns Hopkins University

Updated on 5/9/2023