

# Hongyin Luo

32 Vassar St.  
Cambridge, MA 02139  
<https://www.csail.mit.edu/person/hongyin-luo>

Phone: +1 (617) 253-3049  
Email: [hyluo@mit.edu](mailto:hyluo@mit.edu)  
Alt: [hongyin.luo@icloud.com](mailto:hongyin.luo@icloud.com)

## PARTICULARS

---

### WORK

MIT Computer Science and Artificial Intelligence Laboratory Postdoctoral Associate	Cambridge MA, USA <i>July 2022 - Now</i>
---	---

### EDUCATION

Massachusetts Institute of Technology Ph. D. in Computer Science <i>Minor in Cognitive Science</i>	Cambridge MA, USA <i>May 2022</i>
--	--------------------------------------

Massachusetts Institute of Technology S. M. in Computer Science	Cambridge MA, USA <i>May 2019</i>
--	--------------------------------------

Tsinghua University B. Eng. in Computer Science and Technology	Beijing, China <i>June 2016</i>
---	------------------------------------

### RESEARCH INTERESTS

My research interests span the areas of natural language understanding, generation, and the social impacts of language models. I have a specific interest in grounding language models on formal reasoning engines to improve their trustworthiness, efficiency, and problem solving ability.

### DISSERTATION

Title: “Self-training for Natural Language Processing”  
Advisor: James R. Glass

My thesis develops a self-training framework to adapt efficient language models on different downstream tasks without using any human-annotated training data. This approach make small language models outperform 500 times larger GPT models on a wide range of standard natural language processing benchmarks.

### Research Highlights

---

- IA générative: quelle technologie après ChatGPT? *Le Journal du Net*, July 2023. [\(link\)](#)
- The Cost of Doing AI Business. *PYMNTS*, July 2023. [\(link\)](#)
- Search engines don’t always help chatbots generate accurate answers. *The Register*, June 2023. [\(link\)](#)

- Bigger is Not Always Better. *Analytics India Magazine*, June 2023. ([link](#))
- MIT, 자가학습 AI 공개... 성능최대 500 배향상. *ZDNet Korea*, June 2023. ([link](#))
- The Little Language Model That Could. *Hackster.io*, June 2023. ([link](#))
- MIT researchers make language models scalable self-learners. *MIT News*, June 2023. ([link](#))
- MIT researchers develop self-learning language models that outperform larger counterparts. *VentureBeats*, June 2023. ([link](#))
- Текстовую нейросеть научили думать» еще лучше, чтобы избавиться от расизма и сексизма. *Nauka TV*, March 2023. ([link](#))
- Large language models are biased. Can logic help save them? *MIT News*, March 2023. ([link](#))

## WORK EXPERIENCE

---

- **Postdoctoral Associate, MIT CSAIL**, July 2022 - Now.  
Leading research projects on grounded and reasoning language models, mentoring junior Ph.D., master, and undergraduate students in the research group.
- **Lecturer, MIT Momentum AI**, September 2021 - January 2022.  
Offering machine learning and natural language procession lectures for minority and refugee students.
- **Research Intern, Amazon AWS AI**, May 2021 - August 2021.  
Built self-trained, prefix-tuning systems for document QA toolkit of AWS.
- **Research Intern, Google Brain**, May 2019 - August 2019.  
Developed simultaneous translation system of Google that is applied in the Tokyo Olympics.

## TEACHING EXPERIENCE

---

- **MIT Kaufman Teaching Certificate**, Fall 2022, MIT Teaching + Learning Lab.
- **Lecturer**, Fall 2021, MIT Momentum AI.
- **Teaching Assistant**. 6.864 Advanced Natural Language Processing, Prof. Jacob Andreas and Prof. James R. Glass, Spring and Fall 2020, MIT.
- **Teaching Assistant and Project Mentor**. 6.862 Applied Machine Learning, Prof. Stefanie Jegelka, Spring 2019, MIT.

## PUBLICATIONS

---

### Preprints

1. Junmo Kang, **Hongyin Luo**, Yada Zhu, James Glass, David Cox, Alan Ritter, Rogerio Feris, Leonid Karlinsky “Self-Specialization: Uncovering Latent Expertise within Large Language Models,” 2023.
2. Jiaxin Ge, **Hongyin Luo**, Siyuan Qian, Yulu Gan, Jie Fu, Shanghang Zhan, “Chain of Thought Prompt Tuning in Vision Language Models,” 2023.

## PAPERS

3. Tianhua Zhang\*, Jiaxin Ge\*, **Hongyin Luo\***, Yung-Sung Chuang, Yuan Gong, Yoon Kim, Xixin Wu, Helen Meng, James Glass, “Natural Language Embedded Programs for Hybrid Language Symbolic Reasoning,” NAACL 2024.
4. Zhaorun Chen, Zhuokai Zhao, **Hongyin Luo**, Huaxiu Yao, Bo Li, Jiawei Zhou, “HALC: Object Hallucination Reduction via Adaptive Focal-Contrast Decoding”, ICLR 2024 R2-FM Workshop.
5. Yung-Sung Chuang, Yujia Xie, **Hongyin Luo**, Yoon Kim, James Glass, Pengcheng He, “Dola: Decoding by contrasting layers improves factuality in large language models,” ICLR 2024.
6. Yuan Gong, **Hongyin Luo**, Alexander H. Liu, Leonid Karlinsky, James Glass “Listen, Think, and Understand,” ASRU 2023.
7. **Hongyin Luo\***, Tianhua Zhang\*, Yung-Sung Chuang, Yuan Gong, Yoon Kim, Xixin Wu, Helen Meng, James Glass “Search-Augmented Instruction Learning,” EMNLP 2023.
8. Jiaxin Ge\*, **Hongyin Luo\***, Yoon Kim, James Glass, “Entailment as Robust Self-Learner,” ACL 2023.
9. Yuan Gong, Alexander H Liu, **Hongyin Luo**, Leonid Karlinsky, James Glass, “Joint Audio and Speech Understanding,” ASRU 2023.
10. Tianhua Zhang, Liping Tang, Wei Fang, **Hongyin Luo**, Xixin Wu, Helen Meng, James Glass, “ConvrGX: Recognition, Generation, and Extraction for Self-trained Conversational Question Answering,” Doc2dial Workshop at ACL 2023.
11. **Hongyin Luo**, James Glass, “Logic against bias: Textual entailment mitigates stereotypical sentence reasoning,” EACL 2023.
12. **Hongyin Luo**, Shang-Wen Li, Mingye Gao, Seunghak Yu, James Glass, “Cooperative Self-training of Machine Reading Comprehension,” NAACL 2022.
13. Yung-Sung Chuang, Rumen Dangovski, **Hongyin Luo**, Yang Zhang, Shiyu Chang, Marin Soljačić, Shang-Wen Li, Wen-tau Yih, Yoon Kim, James Glass, “DiffCSE: Difference-based contrastive learning for sentence embeddings,” NAACL 2022.
14. **Hongyin Luo**, James Glass, Garima Lalwani, Yi Zhang, Shang-Wen Li, “Joint Retrieval-Extraction Training for Evidence-Aware Dialog Response Selection,” EACL 2021.
15. Yung-Sung Chuang, Mingye Gao, **Hongyin Luo**, James Glass, Hung-yi Lee, Yun-Nung Chen, Shang-Wen Li, “Mitigating biases in toxic language detection through invariant rationalization,” WOH workshop at EMNLP 2021.
16. **Hongyin Luo**, Shang-Wen Li, James Glass, “Knowledge grounded conversational symptom detection with graph memory networks,” EMNLP 2021.
17. **Hongyin Luo**, Shuyan Dong, Yung-Sung Chuang, Shang-Wen Li, “Meta-learning for downstream aware and agnostic pretraining,” MetaNLP workshop at EMNLP 2020.
18. **Hongyin Luo**, Shang-Wen Li, James Glass, “Prototypical q networks for automatic conversational diagnosis and few-shot new disease adaption,” Interspeech 2020.

19. **Hongyin Luo**, Lan Jiang, Yonatan Belinkov, James Glass, “Improving neural language models by segmenting, attending, and predicting the future,” ACL 2019.
20. **Hongyin Luo**, Mitra Mohtarami, James Glass, Karthik Krishnamurthy, Brigitte Richardson, “Integrating Video Retrieval and Moment Detection in a Unified Corpus for Video Question Answering,” Interspeech 2019.
21. **Hongyin Luo**, James Glass, “Learning word representations with cross-sentence dependency for end-to-end co-reference resolution,” EMNLP 2018.
22. Jie Fu, **Hongyin Luo**, Jiashi Feng, Kian Hsiang Low, Tat-Seng Chua, “DrMAD: distilling reverse-mode automatic differentiation for optimizing hyperparameters of deep neural networks,” IJCAI 2016.
23. **Hongyin Luo**, Zhiyuan Liu, Huanbo Luan, Maosong Sun, “Online learning of interpretable word embeddings,” EMNLP 2015.

### Software and Application

1. **LangCode**. A jupyter & Colab interface for generating natural language embedded program (NLEP).
2. **Anchoring**. A platform for generating large language model applications with natural language instructions, powered by neuro-symbolic augmented LLMs.
3. **SAIL-7b**. Search engine augmented large language model with entailment based retrieval denoising.
4. **UniLC**. A unified factual and moral language assessment toolkit.
5. **EntST**. Entailment self-training toolkit with simple pseudo-label editing (SimPLE) algorithm. Boosting small pretrained 350M-parameter language models to outperform GPT3-175B performance.
6. **RGX**. A framework for automatic question-answer data generation and self-training.

### TALKS

---

#### INVITED TALKS

1. “Program Interfaces Grounded, Transparent, and Reasoning AI” *Chinese University of Hong Kong*, Hong King, Marck 2024
2. “Specialized Large Language Models”, *Quantiphi*, Cambridge MA, USA, October 2023.
3. “SAIL-7b: Search Augmented Instruction Learning LLM”,
  - *Sanofi*, Cambridge MA, USA, August 2023.
  - *Amazon*, Sunnyvale CA, USA, July 2023.
  - *MassMutual*, Boston MA, USA, June 2023.
4. “Entailment as Robust Self-learner”, *IBM Watson AI Lab*, Cambridge MA, USA, May 2023.
5. “Entailment Self-training”, *Embodied Intelligence Seminar*, Cambridge MA, USA, October 2022.
6. “Cooperative Self-training for Machine Reading Comprehension”,
  - *MIT CSAIL Annual Meeting*, Cambridge MA, USA, June 2022.

- *Embodied Intelligence Seminar*, Cambridge MA, USA, March 2022.
  - *Itaú Unibanco*, State of São Paulo, Brazil, February 2022.
  - *Chinese University of Hong Kong*, Hong Kong SAR, China, January 2022.
  - *Amazon Alexa*, Sunnyvale CA, USA, September 2021.
7. “Belief State Generation for Misinformation Detection”, *Defence Science and Technology Agency (DSTA)*, Singapore, January 2021

## CONFERENCE TALKS

8. “Search Augmented Instruction Learning”, *EMNLP 2023*, Singapore, December 2023.
9. “Entailment as Robust Self-learner”, *ACL 2023*, Toronto ON, Canada, July 2023.
10. “Logic Against Bias: Textual Entailment Mitigates Stereotypical Sentence Reasoning”, *EACL 2023*, Dubrovnik, Croatia, May. 2023.
11. “Cooperative Self-training of Machine Reading Comprehension”, *NAACL 2022*, Seattle WA, USA, July 2022.
12. “Knowledge Grounded Conversational Symptom Detection with Graph Memory Networks”, *EMNLP 2021*, Punta Cana, Dominican Republic, November 2021.
13. “Joint Retrieval-Extraction Training for Evidence-Aware Dialog Response Selection”, *Interspeech 2021*, Shanghai, China, May 2021.
14. “Prototypical Q Networks for Automatic Conversational Diagnosis and Few-shot New Disease Adaption”, *Interspeech 2020*, Virtual, October 2020.
15. “Integrating Video Retrieval and Moment Detection in a Unified Corpus for Video Question Answering”, *Interspeech 2019*, Graz, Austria, September 2019.
16. “Improving Neural Language Models by Segmenting, Attending, and Predicting the Future”, *ACL 2019*, Florence, Italy, July 2019.
17. “Learning Word Representations with Cross-sentence Dependency for End-to-end Co-reference Resolution”, *EMNLP 2018*, Brussels, Belgium, November 2018.
18. “Online Learning of Interpretable Word Embeddings”, *EMNLP 2015*, Lisbon, Portugal, September 2015.

## PROFESSIONAL ACTIVITIES

---

- Reviewer (Journals) - *Springer Artificial Intelligence Review (AIRE)*, *ACL Rolling Review (ARR)*.
- Reviewer (Conferences) - *COLING (2018, 2021, 2022)*, *Neurips (2022)*, *ACL (2018, 2019, 2020, 2021, 2022, 2023)*, *EMNLP (2018, 2019, 2020, 2021, 2022, 2023)*, *NAACL (2021, 2022)*, *EACL (2022, 2023)*, *IJCAI-PRICAI (2020)*, *WNGT (2020)*, *PRMI (2020)*, *WOHA (2022, 2023)*.
- Mentor - *Tianhua Zhang (Ph.D. student at Chinese University of Hong Kong)*, *Mark Razanau*, *Luc Gaiskell*, *Kimberly Wong (undergraduate students at MIT)*, *Jiaxin Ge (undergraduate student at Peking University)*.

**LANGUAGES**

---

Proficient in Chinese and English.