Yung-Sung Chuang

✉ yungsung@mit.edu | 🏠 people.csail.mit.edu/yungsung | ⬛ voidism | 🐦 YungSungChuang | 🎓 Google Scholar
link to this file (latest): bit.ly/yungsung

Education

Massachusetts Institute of Technology (MIT) *Sep. 2021 - PRESENT*

PH.D. STUDENT IN COMPUTER SCIENCE

- **GPA: 5.0/5.0** (up-to-date)
- Computer Science & Artificial Intelligence Laboratory (CSAIL), Advisor: Dr. James Glass

National Taiwan University (NTU) *Sep. 2016 - Jun. 2020*

B.S. IN ELECTRICAL ENGINEERING

- **GPA:** overall: **4.18/4.30**, major: **4.22/4.30**, last 60: **4.27/4.30**, graduate ranking: **15/177 (8%)**
- **Honors:** Presidential Award * 4 (S '18, S '19, F '19, S '20), Irving T. Ho Memorial Scholarship (F '18, F '19)
- **Selected Courses:** Data Structure and Programming 2017 Fall (**A+**), Algorithms 2019 Spring (**A+**),
  Machine Learning 2018 Spring (**A+**), Digital Speech Processing 2018 Fall (**A+**),
  Deep Learning for Computer Vision 2019 Spring (**A+**), Computer Architecture 2019 Fall (**A+**)

Research Experiences

Spoken Language Systems Group, CSAIL, MIT, *Advisor: Dr. James Glass* *Sep. 2021 - PRESENT*

RESEARCH ASSISTANT

- Researching on **Passage Retrieval** for open-domain question answering. *[ACL'23]*
- Researching on **Equivariant Contrastive Learning** for unsupervised sentence embedding. *[NAACL'22]*
- Built a comprehensive and reproducible benchmark for **Self-supervised Speech Representation Learning**. *[Interspeech'21]*

Microsoft, Redmond, *Advisor: Dr. Pengcheng He & Dr. Yujia Xie* *Jun. 2023 - PRESENT*

RESEARCH INTERN

- Researching on **Hallucination Reduction** in text generation for open-source **Large Language Models**. *[arXiv'23]*

MIT-IBM Watson AI Lab, Cambridge, *Advisor: Dr. Yang Zhang & Prof. Yoon Kim & Prof. Shiyu Chang* *Jun. 2022 - Sep. 2022*

RESEARCH INTERN

- Researching on **Cross-Lingual transfer** for low-resource languages with unseen scripts.

Speech Processing Lab, NTU, *Advisor: Prof. Hung-yi Lee & Lin-shan Lee* *Aug. 2018 - Aug. 2021*

UNDERGRADUATE RESEARCHER

- Researched on **Non-autoregressive Speech-to-Text Translation**. *[Findings@ACL'21]*
- Researched on **End-to-end Spoken Language Understanding**. *[ICASSP'21]*
- Researched on **Cross-Lingual Zero-shot Transfer** tasks. *[arXiv Link]*
- Researched on **Language Model Pre-training** for speech and text to solve **Spoken Question Answering** tasks. *[Interspeech'20]*
  - Receive **Travel Grant** in Interspeech 2020 with this paper.
- Researched on **Text Style Transfer** with CycleGAN architecture *[Github Link]* and delete-insert-based Transformer. *[Github Link]*

Machine Intelligence and Understanding Lab, NTU, *Advisor: Prof. Yun-Nung (Vivian) Chen* *Feb. 2019 - Aug. 2020*

UNDERGRADUATE RESEARCHER

- Researched on **Knowledge Distillation** to improve **Lifelong Learning** for language tasks. *[EMNLP'20]*
  - Won **Appier Best Application Award** in **2020 NTU CSIE Undergrad Special Research Exhibition**.
- Researched on **Cycle Consistency and Duality of NLU and NLG** to improve both NLU/NLG tasks. *[Findings@EMNLP'20]*
- Researched on Generating Conclusions from Medical RCT Papers. *[LOUHI 2019@EMNLP]*
  - Won the **2nd place** and **Appier 1st prize** in **2019 NTU CSIE Undergrad Special Research Exhibition**.

Intelligent Agent Systems Lab, Academia Sinica, *Advisor: Prof. Wen-Lian Hsu* *Jul. 2018 - Feb. 2019*

RESEARCH INTERN AND RESEARCH ASSISTANT

- Built a supervised **Accurate Collocation Parsing System** with state-of-the-art deep learning methods. *[Github Link]*
- Developed fully-unsupervised methods to find collocation pairs in a large corpus with Word2Vec technique. *[Github Link]*

Selected Publications († indicates equal contribution.)

[1] **Yung-Sung Chuang**, Yujia Xie, Hongyin Luo, Yoon Kim, James Glass, Pengcheng He "DoLa: Decoding by Contrasting Layers Improves Factuality in Large Language Models". *In submission, arXiv 2023*.

[2] Hongyin Luo, **Yung-Sung Chuang**, Yuan Gong, Tianhua Zhang, Yoon Kim, Xixin Wu, Danny Fox, Helen Meng, James Glass "SAIL: Search-Augmented Instruction Learning". *To appear in The 2023 Conference on Empirical Methods in Natural Language Processing*.

[3] **Yung-Sung Chuang**, Wei Fang, Shang-Wen Li, Wen-tau Yih, James Glass "Expand, Rerank, and Retrieve: Query Reranking for Open-Domain Question Answering". *In Findings of The 61st Annual Meeting of the Association for Computational Linguistics (ACL) 2023*.

[4] Tianhua Zhang, Jiaxin Ge, Hongyin Luo, **Yung-Sung Chuang**, Mingye Gao, Yuan Gong, Xixin Wu, Yoon Kim, Helen Meng, James Glass "Natural Language Embedded Programs for Hybrid Language Symbolic Reasoning". *In submission, arXiv 2023*.

[5] **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". *In Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) 2022*.

[6] Andrew Rouditchenko, **Yung-Sung Chuang**, Nina Shvetsova, Samuel Thomas, Rogerio Feris, Brian Kingsbury, Leonid Karlinsky, David Harwath, Hilde Kuehne, James Glass "C2KD: Cross-Lingual Cross-Modal Knowledge Distillation for Multilingual Text-Video Retrieval". *In 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2023*.

[7] Guan-Ting Lin, **Yung-Sung Chuang**, Ho-Lam Chung, Shu-wen Yang, Hsuan-Jui Chen, Shang-Wen Li, Abdelrahman Mohamed, Hung-yi Lee, Lin-shan Lee "DUAL: Textless Spoken Question Answering with Speech Discrete Unit Adaptive Learning". *Interspeech 2022*.

[8] Shu-wen Yang, Po-Han Chi[†], **Yung-Sung Chuang**[†], Cheng-I Jeff Lai[†], Kushal Lakhotia[†], Yist Y. Lin[†], Andy T. Liu[†], Jiatong Shi[†], Xuankai Chang, Guan-Ting Lin, Tzu-Hsien Huang, Wei-Cheng Tseng, Ko-tik Lee, Da-Rong Liu, Zili Huang, Shuyan Dong, Shang-Wen Li, Shinji Watanabe, Abdelrahman Mohamed, Hung-Yi Lee. "SUPERB: Speech processing Universal PERformance Benchmark". *In Interspeech 2021*.

[9] Shun-Po Chuang[†], **Yung-Sung Chuang**[†], Chih-Chiang Chang[†], Hung-Yi Lee. "Investigating the Reordering Capability in CTC-based Non-Autoregressive End-to-End Speech Translation". *In ACL-IJCNLP 2021: Findings*.

[10] **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". *In The 5th Workshop on Online Abuse and Harms at The Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (WOAH@ ACL-IJCNLP)* 2021.

[11] Cheng-I Lai, **Yung-Sung Chuang**, Hung-Yi Lee, Shang-Wen Li, James Glass. "Semi-Supervised Spoken Language Understanding via Self-Supervised Speech and Language Model Pretraining". *In ICASSP 2021*.

[12] **Yung-Sung Chuang**, Shang-Yu Su, Yun-Nung Chen. "Lifelong Language Knowledge Distillation". *In EMNLP 2020*.

[13] Shang-Yu Su[†], **Yung-Sung Chuang**[†], Yun-Nung Chen. "Dual Inference for Improving Language Understanding and Generation". *In Findings of EMNLP 2020*.

[14] **Yung-Sung Chuang**, Chi-Liang Liu, Hung-Yi Lee, Lin-shan Lee. "SpeechBERT: An Audio-and-text Jointly Learned Language Model for End-to-end Spoken Question Answering". *In Proc. Interspeech 2020*.

[15] Alexander Te-Wei Shieh[†], **Yung-Sung Chuang**[†], Shang-Yu Su, Yun-Nung Chen. "Towards Understanding of Medical Randomized Controlled Trials by Conclusion Generation". In *Proceedings of the 10th International Workshop on Health Text Mining and Information Analysis at EMNLP 2019*

Teachings

Tutorial Speaker at AACL-IJCNLP 2022 *Nov. 2022*

- Co-hosted a tutorial: **Recent Advances in Pre-trained Language Models: Why Do They Work and How to Use Them** with Prof. Hung-yi Lee and Cheng-Han Chiang at AACL-IJCNLP 2022. [*Slide Link*] [*YouTube Link*]
- Introduced the topic of **Contrastive Learning for Pre-trained Language Models**.

Teaching Assistant on Deep Learning for Human Language Processing 2020 Spring Course *Mar. 2020 - Jun. 2020*

- Designed homework 1 on **End-to-end Speech Recognition** topic. [*Slide Link*] [*YouTube Link*]
- Introduced the topic of **Non-Autoregressive Sequence Generation**.

Honors & Awards

Reviewers, NeurIPS (2021, 2022, 2023), ICLR (2021, 2022, 2023, 2024), ICML (2022, 2023), AAAI (2023, 2024), EMNLP (2022), ICASSP (2021, 2022, 2023), Interspeech (2021) *Apr. 2021 - PRESENT*

Presidential Award - for top 5% students (4 times), Electrical Engineering Dept. at NTU *Spring '18, Spring '19, Fall '19, Spring '20*

Irving T. Ho Memorial Scholarship (2 times), EECS at NTU *Fall '18, Fall '19*

Travel Grant, INTERSPEECH 2020 conference *Sep. 2020*