

CONTACT INFORMATION	One Apple Park Way Cupertino, CA 95014	znculee@gmail.com https://znculee.github.io
EDUCATION	The Chinese University of Hong Kong , Shatin N.T., Hong Kong Ph.D. in Electronic Engineering Advisor: Prof. Max Q.-H. Meng Thesis: Research on Word Alignment and Context Modeling for NMT	Aug 2015 – Oct 2019
	Northeastern University , Shenyang, P.R. China B.E. in Automation Graduated with University Honors. Cumulative GPA: 3.78 / 4.0 (Ranked 3 / 270)	Sep 2011 – Jul 2015
EXPERIENCE	Apple , California, United States Machine Learning Research ICT5 in AI/ML Machine Learning Research ICT4 in AI/ML I am a core developer of Apple's translation models, driving both the deployment of cutting-edge research into future products and the continuous performance improvement of current production models. My work includes training/deploying production translation models, and building end-to-end automation pipelines for data cleaning, augmentation, synthesis, distillation, preprocessing, training, and evaluation. I designed an internal orchestration system and a distributed, plug-and-play data processing framework that powers large-scale experimentation. I also drive the architecture of Apple's reinforcement-learning framework for machine translation. Products: <ul style="list-style-type: none">• Live translation in Messages, FaceTime, Phone, and AirPods• Translate app, System-wide, Safari, and Siri translation	Oct 2025 – present Aug 2022 – Sep 2025
	Baidu Research , California, United States Senior Research Scientist in Institute of Deep Learning Advisor: Prof. Liang Huang Research: <ul style="list-style-type: none">• A³T: Alignment-Aware Acoustic and Text Pretraining for Speech Synthesis and Editing [1] It reconstructs masked acoustic signals with text input and acoustic-text alignment during training, then the pretrained model can generate high quality reconstructed spectrogram, which can be applied to the speech editing and unseen speaker TTS directly.• PaddleSpeech: An Easy-to-Use All-in-One Speech Toolkit [2] I led the Paddle NLP team to demonstrate an all-in-one speech toolkit and wrote a paper that won the best demo award at NAACL. PaddleSpeech provides an easy-to-use command-line interface and a clear code structure. It also achieves state-of-the-art performance on various speech datasets and implements the most popular methods.	Nov 2021 – Jun 2022

- Prefix-finetuning for Simultaneous Machine Translation

It proposes an efficient method to enable the full-sentence NMT model to achieve state-of-the-art simultaneous translation performance by a quick finetuning on an optimized prefix-to-prefix dataset. It also proposes an efficient and simple method for boosting lagging and translation quality trade-off by alleviating the train-test mismatch.

The Ohio State University, Ohio, United States

Postdoc in Linguistics Funded by Facebook

Nov 2019 – Oct 2021

Advisor: Prof. Michael White

Research:

- Building Adaptive Acceptability Classifiers for Neural NLG [3]

It proposes a novel framework to train models to classify acceptability of responses generated by natural language generation (NLG) models, improving upon existing sentence transformation and model-based approaches.

- Leveraging Large Pretrained Models for WebNLG 2020 [4]

It reports experiments on finetuning large pretrained models to realize resource description framework (RDF) triples to natural language.

- Self-Training for Compositional Neural NLG in Task-Oriented Dialogue [5, 6]

It shows that self-training enhanced with constrained decoding yields large gains in data efficiency on datasets that employ compositional meaning representations. In particular, the experiments indicate that self-training with constrained decoding can enable sequence-to-sequence models to achieve satisfactory quality using vanilla decoding with 5 to 10 times less data than with ordinary supervised baseline; moreover, by leveraging pretrained models, data efficiency can be increased further to 50 times. The end result is an approach that makes it possible to achieve acceptable performance on compositional NLG tasks using hundreds rather than tens of thousands of training samples.

- Neural NLG for Methodius [7, 8]

It shows that discourse relation relations significantly improve NLG when data is limited.

Tencent AI Lab, Shenzhen, P.R. China

Research Intern in the NLP Group

Jul 2017 – May 2019

Advisor: Dr. Lemao Liu

Research:

- Regularized Context Gates on Transformer for Machine Translation [9]

It proposes a gate mechanism to control source and target contexts for the state of the art translation system. In addition, it proposes an effective method to regularize the context gates to reduce the bias of context gates learned from scratch.

- On the Word Alignment from Neural Machine Translation [10]

It reveals attention may not capture word alignment and proposes two effective methods to induce word alignment from general NMT models. It also shows NMT indeed learn excellent word alignment regarding its own translation although its alignment regarding reference seems weaker than statistical aligners. Finally, it demonstrates word alignment errors have adverse effect on translation quality.

- Target Foresight based Attention for Neural Machine Translation [11, 12]
It proposes a novel attention mechanism by predicting the target foresight word with an auxiliary network to enhance the attention model and achieves significant improvement both on translation and alignment performance.
- Understanding and Improving Hidden Representations for Neural Machine Translation [13]
It proposes methods to regularize hidden representations and interpret what hidden representations learn from translation.

Department of Electronic Engineering, The Chinese University of Hong Kong
Ph.D. Student in Robotics and Perception Laboratory Aug 2015 – Jul 2019

Advisor: Prof. Max Q.-H. Meng

Research:

- Efficient Object Search With Belief Road Map Using Mobile Robot [14]
It formulates the object search problem as a Partially Observable Markov Decision Process utilizing semantic information.
- Cuffless Blood Pressure Estimation using Recurrent Neural Network [15]
Considering electrocardiogram (ECG) and photoplethysmography (PPG) signals are time series, the feed-forward neural networks have difficulties to handle this type of data. From these observations, a novel structure of recurrent neural network based on long short-term memory has been proposed to learn the mapping from ECG and PPG to blood pressure.
- Algorithms for Reducing Motion Artifacts in Photoplethysmography Signal [16]
The project aims at removing motion artifacts in PPG signal captured by wearable sensors so as to enhance the accuracy and credibility of continuous blood pressure estimation. Statistical signal processing techniques such as independent component analysis, dimensionality reduction and matrix factorization have been involved in the project.
- Robotic Path Planning with Obstacles [17]
The project takes obstacles' information into rapidly-exploring random tree algorithm, making the two dimensional robotic path planning more efficient.

Teaching:

- Teaching Assistant of Medical Robotics Jan 2016 – May 2016
Teaching kinematics of robot manipulator and fundamentals of robot control in surgical robotics, and introducing the novel robotic applications in medical care.
- Teaching Assistant of Biomedical Modeling Sep 2015 – Dec 2016
Teaching basic concepts and fundamental techniques of mathematical modeling, and how the methodologies are used in biomedical modeling via various physiological case studies.

PUBLICATIONS

[GOOGLE SCHOLAR]

- [1] He Bai, Renjie Zheng, Junkun Chen, **Xintong Li**, Mingbo Ma, and Liang Huang. A³T: Alignment-Aware Acoustic and Text Pretraining for Speech Synthesis and Editing. In *ICML*, 2022. (To appear).
- [2] Hui Zhang, Tian Yuan, Junkun Chen, **Xintong Li**, Renjie Zheng, Yuxin Huang, Xiaojie Chen, Enlei Gong, Zeyu Chen, Xiaoguang Hu, Dianhai Yu, Yanjun Ma, and Liang Huang. PaddleSpeech: An Easy-to-Use All-in-One Speech Toolkit. In *NAACL Demo Track*, 2022. (To appear).

- [3] Soumya Batra, Shashank Jain, Peyman Heidari, Ankit Arun, Catharine Youngs, **Xintong Li**, Pinar Dommez, Shawn Mei, Shiunzu Kuo, Vikas Bhardwaj, Anuj Kumar, and Michael White. **Building Adaptive Acceptability Classifiers for Neural NLG**. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 682–697, Online and Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics.
- [4] **Xintong Li**, Aleksandre Maskharashvili, Symon Jory Stevens-Guille, and Michael White. **Leveraging Large Pretrained Models for WebNLG 2020**. In *Proceedings of the 3rd International Workshop on Natural Language Generation from the Semantic Web (WebNLG+)*, pages 117–124, Dublin, Ireland (Virtual), 12 2020. Association for Computational Linguistics.
- [5] **Xintong Li** and Michael White. **Self-Training for Compositional Neural NLG**, 2020. The third annual West Coast NLP Summit (WeCNLP), poster session.
- [6] **Xintong Li**, Symon Stevens-Guille, Aleksandre Maskharashvili, and Michael White. **Self-Training for Compositional Neural NLG in Task-Oriented Dialogue**. In *Proceedings of the 14th International Conference on Natural Language Generation*, pages 87–102, Aberdeen, Scotland, UK, August 2021. Association for Computational Linguistics.
- [7] Symon Stevens-Guille, Aleksandre Maskharashvili, Amy Isard, **Xintong Li**, and Michael White. **Neural NLG for Methodius: From RST Meaning Representations to Texts**. In *Proceedings of the 13th International Conference on Natural Language Generation*, pages 306–315, Dublin, Ireland, December 2020. Association for Computational Linguistics.
- [8] Aleksandre Maskharashvili, Symon Stevens-Guille, **Xintong Li**, and Michael White. **Neural Methodius Revisited: Do Discourse Relations Help with Pre-Trained Models Too?** In *Proceedings of the 14th International Conference on Natural Language Generation*, pages 12–23, Aberdeen, Scotland, UK, August 2021. Association for Computational Linguistics.
- [9] **Xintong Li**, Lemao Liu, Rui Wang, Guoping Huang, and Max Meng. **Regularized Context Gates on Transformer for Machine Translation**. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8555–8562, Online, July 2020. Association for Computational Linguistics.
- [10] **Xintong Li**, Guanlin Li, Lemao Liu, Max Meng, and Shuming Shi. **On the Word Alignment from Neural Machine Translation**. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 1293–1303, Florence, Italy, July 2019. Association for Computational Linguistics.
- [11] **Xintong Li**, Lemao Liu, Zhaopeng Tu, Shuming Shi, and Max Meng. **Target Foresight Based Attention for Neural Machine Translation**. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 1380–1390, New Orleans, Louisiana, June 2018. Association for Computational Linguistics.
- [12] **Xintong Li**, Lemao Liu, Zhaopeng Tu, Guanlin Li, Shuming Shi, and Max Q.-H. Meng. **Attending From Foresight: A Novel Attention Mechanism for Neural Machine Translation**. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 29:2606–2616, 2021.
- [13] Guanlin Li, Lemao Liu, **Xintong Li**, Conghui Zhu, Tiejun Zhao, and Shuming Shi. **Understanding and Improving Hidden Representations for Neural Machine Translation**. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*,

- Volume 1 (Long and Short Papers)*, pages 466–477, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics.
- [14] Chaoqun Wang, Jiyu Cheng, Jiankun Wang, **Xintong Li**, and Max Q-H Meng. Efficient Object Search With Belief Road Map Using Mobile Robot. *IEEE Robotics and Automation Letters*, 3(4):3081–3088, 2018.
 - [15] Frank P.-W. Lo, **Charles X.-T. Li**, Jiankun Wang, Jiyu Cheng, and Max Q.-H. Meng. Continuous Systolic and Diastolic Blood Pressure Estimation utilizing Long Short-Memory Network. In *Engineering in Medicine and Biology Society (EMBC), 2017 IEEE 39th Annual International Conference of the*. IEEE, 2017.
 - [16] Frank P.-W. Lo, **Charles X.-T. Li**, Jiankun Wang, and Max Q.-H. Meng. Motion Artifact Reduction in PPG Signals based on Periodic Component Factorization. In *Engineering in Medicine and Biology Society (EMBC), 2017 IEEE 39th Annual International Conference of the*. IEEE, 2017.
 - [17] Jiankun Wang, **Xintong Li**, and Max Q-H Meng. An improved rrt algorithm incorporating obstacle boundary information. In *2016 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pages 625–630. IEEE, 2016.

ACTIVITIES

- Poster presentation on the 14th International Conference on Natural Language Generation Sep 2021
- Oral presentation on the WebNLG workshop of 13th International Conference on Natural Language Generation Dec 2020
- Poster presentation and being volunteer on the 57th Annual Meeting of the Association for Computational Linguistics Jul 2019
- Local Chair of 15th International Conference on Information and Automation Aug 2018
- Poster presentation on the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies Jun 2018
- Oral presentation on 39th International Conference of the IEEE Engineering in Medicine and Biology Society Jul 2017

SERVICES

Area Chair ARR

Conference Reviewer
ACL, EMNLP, ACL-IJCNLP, COLING, WebNLG+, AAAI, EACL, NAACL-HLT

Journal Reviewer
NEUCOM

AWARDS & HONORS

- Best Demo Award at NAACL Jul 2022
- Postgraduate Studentships from CUHK 2015 – 2019
- Best Innovation Award of Innovation and Entrepreneurship Competition in Greater China Nov 2016
- Excellent Final Year Project Award Highest score final year project in the Department of Automation. Jun 2015
- Meritorious Winner of the American Mathematical Contest in Modeling Feb 2015

- National Scholarship
For the first 0.2% undergraduates in China 2014
- Outstanding Award of the “Challenge Cup”
[National Academic Competition](#) May 2013
- [Baosteel Merit Scholarship](#) (Grand Prize, ranked 5 / 50) 2013
For the first 50 Chinese engineering undergraduates and postgraduates.
- The Excellent Undergraduate of Shenyang
For the first undergraduate in the [ISE college](#). 2012 – 2014
- The First Prize of Scholarship of [Northeastern University](#) 2011 – 2015
For the first 3% undergraduates in the Department of Automation.

LANGUAGES

- Mandarin: Native language
- English: Fluent (speaking, reading, writing)
- Cantonese: Fluent (speaking, reading, writing)
- Latin: Intermediate (reading); basic (speaking, writing)
- Japanese: basic (speaking, reading, writing)
- Spanish: basic (speaking, reading, writing)

SKILLS

Programming Languages

Python, Rust, C, C++, Bash,
Typescript, Javascript, HTML, CSS,
Vimscript, Lua, AWK,
Matlab, x86 Assembly

Deep Learning Platforms

[Pytorch](#), [Jax](#), [HuggingFace](#),
[Tensorflow](#), [Keras](#), [TFLearn](#), [Theano](#)

Distributed Computing

[Spark](#), [Dask](#), [Ray](#)

Others

[LATEX](#), [ROS](#), [OpenCV](#), [Qt](#), [CUDA](#)

[Last update on November 30, 2025]