Yijia Liu

Ph.D. candidate, *Language analysis* group in HIT-SCIR
oneplus.lau@gmail.com

RESEARCH INTEREST

Natural language processing, Chinese word segmentation, parsing and machine learning. My supervisor is Wanxiang Che.

EDUCATION

Ph.D. candidate, Harbin Institute of Technology — 2014.9 - present
Major: Computer Science

Visiting student, University of Washington — 2016.10 - 2017.9
Supervisor: Noah A. Smith

M.S., Harbin Institute of Technology — 2012.9 - 2014.7
Major: Computer Science, Score: 79.14 (rank top 10%)

B.E., Harbin Institute of Technology — 2008.9 - 2012.7
Major: Computer Science, Score: 88.7 (rank top 8%)

PUBLICATION

Yijia Liu, Wanxiang Che, Huaipeng Zhao, Bing Qin, and Ting Liu. 2018. Knowledge Distilling for Search-based Structured Prediction. (to appear) In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics* (ACL).

Yijia Liu, Yi Zhu, Wanxiang Che, Bing Qin, Nathan Schneder, and Noah A. Smith. 2018. Parsing Tweets into Universal Dependency. (to appear) In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (NAACL).

Wanxiang Che, **Yijia Liu**, Yuxuan Wang, Bo Zheng, and Ting Liu. 2018. Towards Better UD Parsing: Deep Contextualized Word Embeddings, Ensemble, and Treebank Concatenation. (to appear) In *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies* (CoNLL).

Yutai Hou, **Yijia Liu**, Wanxiang Che, and Ting Liu. 2018. Sequence-to-Sequence Data Augmentation for Dialogue Language Understanding. (to appear) In *Proceedings of the 27th International Conference on Computational Linguistics* (COLING).

Haoyang Wen, **Yijia Liu**, Wanxiang Che, Libo Qin, and Ting Liu, 2018. Sequence-to-Sequence Learning for Task-oriented Dialogue with Dialogue State Representation. (to appear) In *Proceedings of the 27th International Conference on Computational Linguistics* (COLING).

Yijia Liu, Wanxiang Che, Jiang Guo, Bing Qin, and Ting Liu. 2016. Exploring Segment Representations for Neural Segmentation Models. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence* (IJCAI).

Yijia Liu, Wanxiang Che, Bing Qin, and Ting Liu. 2016. HC-search for Incremental Parsing. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence* (IJCAI).

Yijia Liu, Yue Zhang, Wanxiang Che, and Ting Liu. 2015. Transition-Based Syntactic Linearization. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics* (NAACL).

Yijia Liu, Yue Zhang, Wanxiang Che, and Ting Liu. 2014. Domain Adaptation for CRF-based Chinese Word Segmentation using Free Annotations, In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing* (EMNLP).

Yijia Liu, Wanxiang Che, and Ting Liu. 2013. Enhancing chinese word segmentation with character clustering. In *Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data* (CCL).

Yijia Liu, Meishan Zhang, Wanxiang Che, Ting Liu, and Yihe Deng. 2012. Micro blogs Oriented Word Segmentation System. In *Proceedings of the Second CIPS-SIGHAN Joint Conference on Chinese Language Processing*.

Meishan Zhang, Wanxiang Che, **Yijia Liu**, Zhenghua Li, and Ting Liu. 2012. HIT dependency parsing: Bootstrap aggregating heterogeneous parsers. In *Notes of the First Workshop on Syntactic Analysis of Non-Canonical Language* (SANCL).

PATTERNS

Wanxiang Che, **Yijia Liu**, Ting Liu, and Yanyan Zhao. An Incremental Training Method for Domain Adaptation in Chinese Word Segmentation, CN201510604035.0.

PROJECTS

CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies 2018.4 - 2018.6

Our system (HIT-SCIR) was ranked **first** out of 26 submissions according to LAS.

- contributed the ideas of using deep contextualized word embeddings and ensemble.
- developed the treebank concatenation strategies.

Language Technology Platform (LTP) 2013.6 - present

LTP is a software package that provides Chinese NLP pipeline and web service API.

- one of the developers and the major maintainer of LTP.
- developed 4 modules including Chinese word segmentation, part-of-speech tagging, NER and dependency parsing in a perceptron algorithm framework.
- developed the RESTful API and contributed to the development of website (`http://ltp-cloud.com`).

SERVICES

Conference Reviewer/Secondary Reviewer: ACL 2014, 2018, CCL 2015-2017, NLPCC 2015-2017, NAACL 2016, IJCAI 2016, SemEval 2016.

EMPLOYMENT

Research Assistance, SUTD. 2013.10 - 2014.10

worked with Dr. Yue Zhang, on statistical machine translation, Chinese tagging and transition based dependency parsing.

Intern Researcher and Developer, Baidu Inc., NLP Department. 2011.7 - 2011.11

implemented query template extraction toolkit and built a python extension for baidu wordseg library.

AWARDS

First Class Award in HeiLongJiang Provincial Science and Technology Prizes:
The Language Technology Platform and its Applications 2016.10
Hua Wei Scholarship (for graduate student) 2016.9
The National Scholarship for graduate students 2013.9
2010 ACM/ICPC Asia Regional Contest Hangzhou Onsite, Silver Medal 2010.10
Hua Wei Scholarship (for undergraduate student) 2010.9