



# Lei Sha

## Curriculum Vitae

### Education & Work Experience

- 2013.09-present **PhD student of Natural Language Processing**, *Peking University*.
- 2017.07-present **Intern of System**, *Microsoft Research Asia*.  
Working on Dialogue system
- 2014.11-2015.6 **Intern of Knowledge Mining**, *Microsoft Research Asia*.  
Working on Event Extraction
- 2009.09-2013.06 **Bachelor of Computer Science**, *Beijing University of Technology*.  
GPA – 3.9
- 2006.09-2009.06 **High School**, *The Experimental High School attached to Beijing Normal University*.

### Publications

- 2016 [1] Reading and Thinking: Re-read LSTM Unit for Textual Entailment Recognition. **Lei Sha**, Sujian Li, Baobao Chang and Zhifang Sui. In Proceedings of Coling 2016 Conference, 2016. (Coling 2016).
- 2016 [2] Capturing Argument Connection for Chinese Semantic Role Labeling. **Lei Sha**, Sujian Li, Baobao Chang and Zhifang Sui. In Proceedings of EMNLP 2016 Conference, 2016. (EMNLP 2016).
- 2016 [3] RBPB: Regularization-Based Pattern Balancing Method for Event Extraction. **Lei Sha**, Chin-Yew Lin, Sujian Li, Baobao Chang and Zhifang Sui. In Proceedings of ACL 2016 Conference, 2016. (ACL 2016).
- 2016 [4] Joint Learning Templates and Slots for Event Schema Induction. **Lei Sha**, Sujian Li, Baobao Chang and Zhifang Sui. In Proceedings of NAACL 2016 Conference, 2016. (NAACL 2016).
- 2015 [5] Recognizing Textual Entailment Using Probabilistic Inference. **Lei Sha**, Sujian Li, Tingsong Jiang, Baobao Chang, and Zhifang Sui. In Proceedings of the EMNLP 2015 Conference, 2015. (EMNLP 2015).
- 2017 [6] Syntax Aware LSTM Model for Chinese Semantic Role Labeling. Feng Qian, **Lei Sha**, Baobao Chang, Lu-chen Liu, Ming Zhang. In Proceedings of EMNLP 2017 Conference, 2017. (EMNLP 2017).
- 2017 [7] A Progressive Learning Approach to Chinese SRL Using Heterogeneous Data. Qiaolin Xia, **Lei Sha**, Baobao Chang, Zhifang Sui. In Proceedings of ACL 2017 Conference, 2017. (ACL 2017).
- 2017 [8] Attentive Interactive Neural Networks for Answer Selection in Community Question Answering. Xiaodong Zhang, **Lei Sha**, Sujian Li, Houfeng Wang. In Proceedings of AAAI 2017 Conference, 2017. (AAAI 2017).
- 2016 [9] Encoding Temporal Information for Time-Aware Link Prediction. Tingsong Jiang, Tianyu Liu, Tao Ge, **Lei Sha**, Sujian Li, Baobao Chang and Zhifang Sui. In Proceedings of EMNLP 2016 Conference, 2016. (EMNLP 2016).
- 2016 [10] Towards Time-Aware Knowledge Graph Completion. Tingsong Jiang, Tianyu Liu, Tao Ge, **Lei Sha**, Sujian Li, Baobao Chang and Zhifang Sui. In Proceedings of Coling 2016 Conference, 2016. (Coling 2016).

- 2015 [11] Multi-label Text Categorization with Joint Learning Predictions-as-Features Method. Li Li, Houfeng Wang, **Lei Sha**, Xu Sun, Baobao Chang, Shi Zhao. In Proceedings of EMNLP 2015 Conference, 2015. (EMNLP 2015)
- 2014 [12] Event Schema Induction Based on Relational Co-occurrence over Multiple Documents. Tingsong Jiang, **Lei Sha** and Zhifang Sui. In Proceedings of Natural Language Processing and Chinese Computing, 2014. (NLPCC 2014).

## Awards

- 2016 Merit Student, May 4th Scholarship
- 2015 "Stars of Tomorrow" in MSRA, Peking University PhD Fellowship
- 2014 The 3<sup>rd</sup> Prize in the ACM contest of Peking University
- 2012 The 1<sup>st</sup> Prize in the 3<sup>rd</sup> "Lanqiao Cup" National Software Professional Design and Business Competition,
- 2011 The National 2<sup>nd</sup> Prize in the China Undergraduate Mathematical Contest in Modeling, National Scholarship(**The only one** in my grade)
- 2010 The 2<sup>nd</sup> Prize in Chinese Mathematical Contest(Beijing division), Beichuan Scholarship(one of the two in my grade)

## Experience

### Projects

- 2017.5– **Table-to-text.**
- 2017.7
  - We use sequence-to-sequence LSTM-RNN model to generate brief summary from Wikipedia infobox.
  - Apart from the content-based addressing, we add link-based addressing to make our model an order-sensitive model. This propose increased the BLEU and ROUGE score significantly.
- 2017.2– **Kensho Demo.**
- 2017.5
  - We use rule-based method to turn natural language query to SPARQL.
  - We use SPARQL query to seek for answers from knowledge base.
- 2017.2– **Automatically Design Neural Network Architecture.**
- 2017.5
  - We use reinforcement learning to decide the layer type and layer size to add in each decision step.
  - The reward is set as the  $F_1$  value on validation set.
  - We use beam search to prune those architecture with low reward.
- 2014.11– **Event Extraction.**
- 2015.5
  - Read papers about event extraction
  - Implement the new ideas and write a paper "RBPB: Regularization-Based Pattern Balancing Method for Event Extraction."
- 2014.03– **Event Schema Induction.**
- 2014.5
  - Read papers about event schema Induction
  - Implement the new ideas and write a paper "Joint Learning Templates and Slots for Event Schema Induction"

## Computer skills

- Intermediate HTML,  $\text{\LaTeX}$ , Linux, Microsoft Windows
- Advanced PYTHON, JAVA, C/C++, MATLAB
- Advanced THEANO, TENSORFLOW

## Interests

- Table Tennis
- Badminton
- Running
- Swimming
- Photographing
- Playing Piano