

SIGHAN-9

**The 9th SIGHAN Workshop on Chinese Language Processing**

**Proceedings of the SIGHAN-9 Workshop**

December 1, 2017

Taipei, Taiwan

©2017 Asian Federation of Natural Language Processing

ISBN 978-1-948087-09-4

## Introduction

The 9th SIGHAN Workshop on Chinese Language Processing (SIGHAN-9) at IJCNLP 2017, December 01, 2017, Taiwan

Growing interest in Chinese language processing is leading to the development of linguistic corpora and automatic tools. As more resources have become available recently, it is crucial to create a platform that allows easy exchange of information and data and the comparison of different approaches to various NLP tasks. Among about 20 SIGs within the Association for Computational Linguistics, SIGHAN (Special Interest Group on Chinese Language Processing) provides an umbrella for researchers in industry and academia working in various aspects of Chinese language processing. SIGHAN workshop provides an international forum for both researchers and practitioners to report their latest innovations, summarize state-of-the-art in the field, as well as exchange ideas, results, and visions in all aspects of Chinese language processing.

The previous SIGHAN workshop series were started in 2002. SIGHAN-1 was held in conjunction with COLING-02 (Taipei, Taiwan); SIGHAN-2 was held in conjunction with ACL-03 (Sapporo, Japan); SIGHAN-3 was held in conjunction with ACL-04 (Barcelona, Spain); SIGHAN-4 was held in conjunction with IJCNLP-05 (Jeju Island, Korea); SIGHAN-5 was held in conjunction with ACL-COLING-06 (Sydney, Australia); SIGHAN-6 was held in conjunction with IJCNLP-08 (Hyderabad, India); SIGHAN-7 was held in conjunction with IJCNLP-13 (Nagoya, Japan) and SIGHAN-8 was held in conjunction with ACL-IJCNLP 2015 (Beijing, China).

The purpose of the SIGHAN-9 is to identify challenging problems facing the development of Chinese language processing systems, and to shape future research directions through the publication of high quality, applied and theoretical research findings. We encourage submissions of papers on all topics in the general areas related to computational linguistics and speech/text processing of Chinese natural languages.



**Organizers:**

Yue Zhang, Singapore University of Technology and Design  
Zhifang Sui, Peking University

**Program Committee:**

Jie Yang, Singapore University of Technology and Design  
Jiangming Liu, University of Edinburgh  
Yu Yuan, Singapore University of Technology and Design  
Yuze Gao, Singapore University of Technology and Design  
Yanxia Qin, Harbin Institute of Technology  
Shuailong Liang, Singapore University of Technology and Design  
Deng Cai, Shanghai Jiao Tong University  
Chenhua Chen, Singapore University of Technology and Design

**Invited Speaker:**

Yue Zhang, Singapore University of Technology and Design  
Wanxiang Che, Harbin Institute of Technology



## Table of Contents

<i>Group Linguistic Bias Aware Neural Response Generation</i>	
Jianan Wang, Xin Wang, Fang Li, Zhen Xu, Zhuoran Wang and Baoxun Wang .....	1
<i>Neural Regularized Domain Adaptation for Chinese Word Segmentation</i>	
Zuyi Bao, Si Li, Weiran XU and Sheng GAO .....	11
<i>The Sentimental Value of Chinese Sub-Character Components</i>	
Yassine Benajiba, Or Biran, Zhiliang Weng, Yong Zhang and Jin Sun .....	21
<i>Chinese Answer Extraction Based on POS Tree and Genetic Algorithm</i>	
Shuihua Li, Xiaoming Zhang and Zhoujun Li .....	30
<i>Learning from Parenthetical Sentences for Term Translation in Machine Translation</i>	
Guoping Huang, Jiajun Zhang, Yu Zhou and Chengqing Zong .....	37





## Conference Program

[8:00-9:00] *Chinese Segmentation and POS-tagging Leveraging Heterogenous Corpora*

Yue Zhang (**Invited talk**)

[9:00-9:25] *Group Linguistic Bias Aware Neural Response Generation*

Jianan Wang, Xin Wang, Fang Li, Zhen Xu, Zhuoran Wang and Baoxun Wang

[9:25-9:50] *Neural Regularized Domain Adaptation for Chinese Word Segmentation*

Zuyi Bao, Si Li, Weiran XU and Sheng GAO

[9:50-10:15] *The Sentimental Value of Chinese Sub-Character Components*

Yassine Benajiba, Or Biran, Zhiliang Weng, Yong Zhang and Jin Sun

[10:15-11:00] *Coffe break*

[11:00-12:00] *Syntactic and Semantic Parsing*

Wanxiang Che (**Invited talk**)

[12:00-12:25] *Chinese Answer Extraction Based on POS Tree and Genetic Algorithm*

Shuihua Li, Xiaoming Zhang and Zhoujun Li

[12:25-12:50] *Learning from Parenthetical Sentences for Term Translation in Machine Translation*

Guoping Huang, Jiajun Zhang, Yu Zhou and Chengqing Zong

