

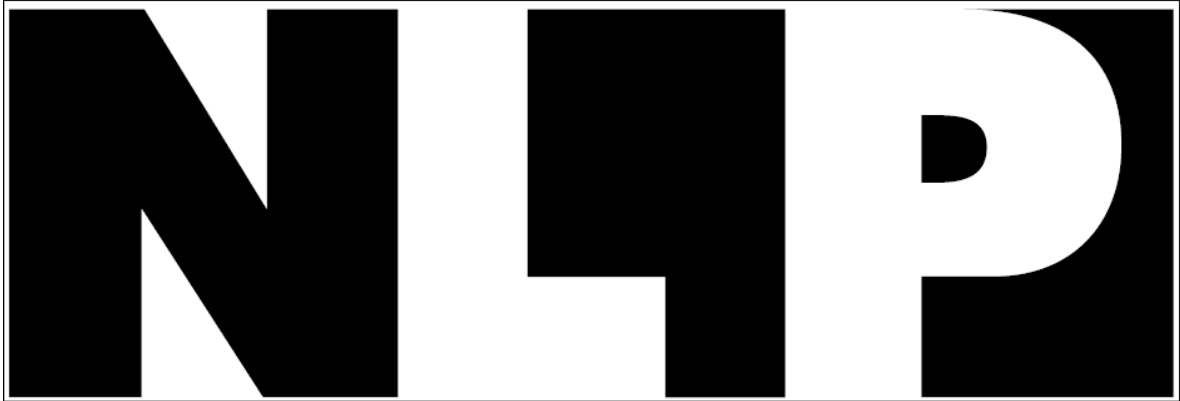
Sixth International Joint Conference on
Natural Language Processing



**Proceedings of the Seventh SIGHAN Workshop on
Chinese Language Processing**

We wish to thank our sponsors and supporters!

Platinum Sponsors



www.anlp.jp

Silver Sponsors



www.google.com

Bronze Sponsors



www.rakuten.com

Supporters



**NAGOYA CONVENTION
& VISITORS BUREAU**

Nagoya Convention & Visitors Bureau

We wish to thank our organizers!

Organizers



[Asian Federation of Natural Language Processing \(AFNLP\)](#)



[Toyohashi University of Technology](#)

©2013 Asian Federation of Natural Language Processing

ISBN 978-4-9907348-5-5

Preface

Welcome to the Seventh SIGHAN Workshop on Chinese Language Processing! Sponsored by the Association for Computational Linguistics (ACL) Special Interest Group on Chinese Language Processing (SIGHAN), this year's SIGHAN-7 workshop is being held in Nagoya, Japan, on October 14, 2013, and is co-located with IJCNLP 2013. The workshop program includes a keynote speech, research paper presentations and a Chinese Spelling Check Bake-off. We hope that these events will encourage the participation of researchers and bring them together to share ideas and developments in various aspects of Chinese language processing.

We are honored to welcome as our distinguished speaker Dr. Keh-Jiann Chen (Research Fellow, Academia Sinica, Taiwan). Dr. Chen will be speaking on "Lexical Semantics of Chinese Language". We would also like to thank Shih-Hung Wu, Chao-Lin Liu and Lung-Hao Lee for their great efforts in organizing the Chinese Spelling Check Bake-off which will feature seventeen teams from China, Japan, Singapore, Taiwan and United Kingdom, and is expected to further the development of more accurate Chinese spelling checkers.

Finally, we would like to thank all authors for their submissions. We appreciate your active participation and support to ensure a smooth and successful conference. The publication of these papers represents the joint effort of many researchers, and we are grateful to the efforts of the review committee for their work, and to the SIGHAN committee for their continuing support.

We wish all a rewarding and eye-opening time at the workshop.

Liang-Chih Yu
Yuen-Hsien Tseng
Jingbo Zhu
Fuji Ren
SIGHAN-7 Workshop Co-Chairs

Organizers

SIGHAN Committee:

Hsin-Hsi Chen, National Taiwan University
Chengqing Zhong, Chinese Academy of Science
Gina-Anne Levow, University of Washington
Ming Zhou, Microsoft Research Asia

Workshop Co-Organizers:

Liang-Chih Yu, Yuan Ze University
Yuen-Hsien Tseng, National Taiwan Normal University
Jingbo Zhu, Northeastern University
Fuji Ren, The University of Tokushima

Bake-off Co-Organizers:

Shih-Hung Wu, Chaoyang University of Technology
Chao-Lin Liu, National Chengchi University
Lung-Hao Lee, National Taiwan University

Steering Committee:

Berlin Chen, National Taiwan Normal University
Keh-Jiann Chen, Academia Sinica
Sin-Horng Chen, National Chiao Tung University
Eduard Hovy, Carnegie Mellon University
Haizhou Li, Institute for Infocomm Research
Chao-Lin Liu, National Chengchi University
Hwee Tou Ng, National University of Singapore
Jianyun Nie, University of Montreal
Wen-Lian Hsu, Academia Sinica
Martha Palmer, University of Colorado Boulder
Jian Su, Institute for Infocomm Research
Keh-Yih Su, Behavior Design Corporation
Hsin-Min Wang, Academia Sinica
Kam Fai Wong, Chinese University of Hong Kong
Chung-Hsien Wu, National Chen Kung University
Guodong Zhou, Soochow University

Program Committee:

Chia-Hui Chang, National Central University
Chien-Liang Chen, Academia Sinica
Kuan-hua Chen, National Taiwan University
Minghui Dong, Institute of Infocomm Research
Donghui Feng, Google Inc.
Zhao-Ming Gao, National Taiwan University
Xungjing Huang, Fudan University
Chunyu Kit, City University of Hong Kong

Olivia Kwong, City University of Hong Kong
Lung-Hao Lee, National Taiwan University
Jun-Lin Lin, Yuan-Ze University
Chao-Hong Liu, National Chen Kung University
Cheng-Jye Luh, Yuan-Ze University
Weiyun Ma, Columbia University
Houfeng Wang, Peking University
Jia-Ching Wang, National Central University
Xiangli Wang, Japan Patent Information Organization
Derek F. Wong, University of Macau
Nianwen Xue, Brandeis University
Chin-Sheng Yang, Yuan-Ze University
Jui-Feng Yeh, National ChiaYi University
Min Zhang, Tsinghua University

Table of Contents

| | |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <i>Keynote Speech: Lexical Semantics of Chinese Language</i> Keh-Jiann Chen | 1 |
| <i>Can MDL Improve Unsupervised Chinese Word Segmentation?</i> Pierre Magistry and Benoît Sagot | 2 |
| <i>Deep Context-Free Grammar for Chinese with Broad-Coverage</i> Xiangli Wang, Yi Zhang, Yusuke Miyao, Takuya Matsuzaki and Junichi Tsujii | 11 |
| <i>Lexical Representation and Classification of Eventive Verbs - Polarity and Interaction between Process and State</i> Shu-Ling Huang, Yu-Ming Hsieh, Su-Chu Lin and Keh-Jiann Chen | 20 |
| <i>Response Generation Based on Hierarchical Semantic Structure with POMDP Re-ranking for Conversational Dialogue Systems</i> Jui-Feng Yeh and Yuan-Cheng Chu | 29 |
| <i>Chinese Spelling Check Evaluation at SIGHAN Bake-off 2013</i> Shih-Hung Wu, Chao-Lin Liu and Lung-Hao Lee | 35 |
| <i>Chinese Word Spelling Correction Based on N-gram Ranked Inverted Index List</i> Jui-Feng Yeh, Sheng-Feng Li, Mei-Rong Wu, Wen-Yi Chen and Mao-Chuan Su | 43 |
| <i>Chinese Spelling Checker Based on Statistical Machine Translation</i> Hsun-wen Chiu, Jian-cheng Wu and Jason S. Chang | 49 |
| <i>A Hybrid Chinese Spelling Correction Using Language Model and Statistical Machine Translation with Reranking</i> Xiaodong Liu, Kevin Cheng, Yanyan Luo, Kevin Duh and Yuji Matsumoto | 54 |
| <i>Introduction to CKIP Chinese Spelling Check System for SIGHAN Bakeoff 2013 Evaluation</i> Yu-Ming Hsieh, Ming-Hong Bai and Keh-Jiann Chen | 59 |
| <i>Automatic Chinese Confusion Words Extraction Using Conditional Random Fields and the Web</i> Chun-Hung Wang, Jason S. Chang and Jian-Cheng Wu | 64 |
| <i>Conditional Random Field-based Parser and Language Model for Traditional Chinese Spelling Checker</i> Yih-Ru Wang, Yuan-Fu Liao, Yeh-Kuang Wu and Liang-Chun Chang | 69 |
| <i>A Maximum Entropy Approach to Chinese Spelling Check</i> Dongxu Han and Baobao Chang | 74 |
| <i>A Study of Language Modeling for Chinese Spelling Check</i> Kuan-Yu Chen, Hung-Shin Lee, Chung-Han Lee, Hsin-Min Wang and Hsin-Hsi Chen | 79 |
| <i>Description of HLJU Chinese Spelling Checker for SIGHAN Bakeoff 2013</i> Yu He and Guohong Fu | 84 |
| <i>Graph Model for Chinese Spell Checking</i> Zhongye Jia, Peilu Wang and Hai Zhao | 88 |

| | |
|-------------------------------------------------------------------------------------------------------------------------|-----|
| <i>Sinica-IASL Chinese spelling check system at Sighan-7</i> | |
| Ting-Hao Yang, Yu-Lun Hsieh, Yu-Hsuan Chen, Michael Tsang, Cheng-Wei Shih and Wen-lian Hsu | 93 |
| <i>Automatic Detection and Correction for Chinese Misspelled Words Using Phonological and Orthographic Similarities</i> | |
| Tao-Hsing Chang, Hsueh-Chih Chen, Yuen-Hsien Tseng and Jian-Liang Zheng | 97 |
| <i>NTOU Chinese Spelling Check System in SIGHAN Bake-off 2013</i> | |
| Chuan-Jie Lin and Wei-Cheng Chu | 102 |
| <i>Candidate Scoring Using Web-Based Measure for Chinese Spelling Error Correction</i> | |
| Liang-Chih Yu, Chao-Hong Liu and Chung-Hsien Wu | 108 |

Workshop Program

Monday, October 14, 2013

09:30 09:40 Opening

09:40 10:30 *Keynote Speech: Lexical Semantics of Chinese Language*
Keh-Jiann Chen

10:30 10:50 Break

Oral Session 1: Chinese Language Processing

10:50 11:15 *Can MDL Improve Unsupervised Chinese Word Segmentation?*
Pierre Magistry and Benoît Sagot

11:15 11:40 *Deep Context-Free Grammar for Chinese with Broad-Coverage*
Xiangli Wang, Yi Zhang, Yusuke Miyao, Takuya Matsuzaki and Junichi Tsujii

11:40 12:05 *Lexical Representation and Classification of Eventive Verbs - Polarity and Interaction between Process and State*
Shu-Ling Huang, Yu-Ming Hsieh, Su-Chu Lin and Keh-Jiann Chen

12:05 12:30 *Response Generation Based on Hierarchical Semantic Structure with POMDP Re-ranking for Conversational Dialogue Systems*
Jui-Feng Yeh and Yuan-Cheng Chu

12:30 13:30 Lunch

Oral Session 2: Chinese Spelling Check Bake-off

13:30 13:50 *Chinese Spelling Check Evaluation at SIGHAN Bake-off 2013*
Shih-Hung Wu, Chao-Lin Liu and Lung-Hao Lee

13:50 14:10 *Chinese Word Spelling Correction Based on N-gram Ranked Inverted Index List*
Jui-Feng Yeh, Sheng-Feng Li, Mei-Rong Wu, Wen-Yi Chen and Mao-Chuan Su

14:10 14:30 *Chinese Spelling Checker Based on Statistical Machine Translation*
Hsun-wen Chiu, Jian-cheng Wu and Jason S. Chang

Monday, October 14, 2013 (continued)

14:30 14:50 *A Hybrid Chinese Spelling Correction Using Language Model and Statistical Machine Translation with Reranking*

Xiaodong Liu, Kevin Cheng, Yanyan Luo, Kevin Duh and Yuji Matsumoto

14:50 15:10 *Introduction to CKIP Chinese Spelling Check System for SIGHAN Bakeoff 2013 Evaluation*

Yu-Ming Hsieh, Ming-Hong Bai and Keh-Jiann Chen

15:10 15:30 Break

15:30 16:20 **Poster Session**

Automatic Chinese Confusion Words Extraction Using Conditional Random Fields and the Web

Chun-Hung Wang, Jason S. Chang and Jian-Cheng Wu

Conditional Random Field-based Parser and Language Model for Traditional Chinese Spelling Checker

Yih-Ru Wang, Yuan-Fu Liao, Yeh-Kuang Wu and Liang-Chun Chang

A Maximum Entropy Approach to Chinese Spelling Check

Dongxu Han and Baobao Chang

A Study of Language Modeling for Chinese Spelling Check

Kuan-Yu Chen, Hung-Shin Lee, Chung-Han Lee, Hsin-Min Wang and Hsin-Hsi Chen

Description of HLJU Chinese Spelling Checker for SIGHAN Bakeoff 2013

Yu He and Guohong Fu

Graph Model for Chinese Spell Checking

Zhongye Jia, Peilu Wang and Hai Zhao

Sinica-IASL Chinese spelling check system at Sighan-7

Ting-Hao Yang, Yu-Lun Hsieh, Yu-Hsuan Chen, Michael Tsang, Cheng-Wei Shih and Wen-lian Hsu

Automatic Detection and Correction for Chinese Misspelled Words Using Phonological and Orthographic Similarities

Tao-Hsing Chang, Hsueh-Chih Chen, Yuen-Hsien Tseng and Jian-Liang Zheng

NTOU Chinese Spelling Check System in SIGHAN Bake-off 2013

Chuan-Jie Lin and Wei-Cheng Chu

Monday, October 14, 2013 (continued)

Candidate Scoring Using Web-Based Measure for Chinese Spelling Error Correction
Liang-Chih Yu, Chao-Hong Liu and Chung-Hsien Wu

16:20 16:30 Closing

