ACL-IJCNLP 2015

The 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing

Proceedings of The 2nd Workshop on Natural Language Processing Techniques for Educational Applications

July 31, 2015
Beijing, China

©2015 The Association for Computational Linguistics
and The Asian Federation of Natural Language Processing

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-941643-70-9

Preface

Welcome to the 2nd Workshop on Natural Language Processing Techniques for Educational Applications (NLP-TEA-2), with a Shared Task on Chinese Grammatical Error Diagnosis (CGED).

The development of Natural Language Processing (NLP) has advanced to a level that affects the research landscape of many academic domains and has practical applications in many industrial sectors. On the other hand, educational environment has also been improved to impact the world society, such as the emergence of MOOCs (Massive Open Online Courses). With these trends, this workshop focuses on the NLP techniques applied to the educational environment. Research issues in this direction have gained more and more attention, examples including the activities like the workshops on Innovative Use of NLP for Building Educational Applications since 2005 and educational data mining conferences since 2008.

This is the second workshop held in the Asian area, with the first one NLP-TEA-1 being held in conjunction with the 22nd International Conference in Computer Education (ICCE 2014) from Nov. 30 to Dec. 4, 2014 in Japan. This year, we continue to promote this research line by holding the workshop in conjunction with the 2015 ACL-IJCNLP conference and also holding the second shared task on Chinese Grammatical Error Diagnosis. During this short period between the first and second workshop, we still receive 14 valid submissions for regular session, each of which was reviewed by three experts, and have 9 teams participating in the shared task, with 6 of them submitting their testing results. In total, there are 6 oral papers and 12 posters accepted. We also organize a keynote speech session from the industrial sector in this workshop.

Overall, we would like to promote this line of research and benefit the participants of the workshop and the shared task.

Workshop Chairs
Hsin-Hsi Chen, National Taiwan University
Yuen-Hsien Tseng, National Taiwan Normal University
Yuji Matsumoto, Nara Institute of Science and Technology
Lung Hsiang Wong, Nanyang Technological University

Organizers

Workshop Organizers:

Hsin-Hsi Chen, National Taiwan University
Yuen-Hsien Tseng, National Taiwan Normal University
Yuji Matsumoto, Nara Institute of Science and Technology
Lung Hsiang Wong, Nanyang Technological University

Shared Task Organizers:

Lung-Hao Lee, National Taiwan Normal University
Liang-Chih Yu, Yuan Ze University
Li-Ping Chang, National Taiwan Normal University

Program Committee:

Yuki Arase, Osaka University
Aoife Cahill, Educational Testing Services
Li-Ping Chang, National Taiwan Normal University
Mariano Felice, Cambridge University
Dongfei Feng, Google Inc.
Trude Heift, Simon Fraser University
Mamoru Komachi, Tokyo Metropolitan University
Lun-Wei Ku, Academia Sicina
Lung-Hao Lee, National Taiwan Normal University
Xiaofei Lu, Pennsylvania State University
Detmar Meurers, University of Tubingen
Ildiko Pilan, University of Gothenburg
Benno Stein, Bauhaus-University Weimar
Yukio Tono, Tokyo University of Foreign Studies
Elena Volodina, University of Gothenburg
Houfeng Wang, Peking University
Jinhua Xiong, Chinese Academy of Science
Jui-Feng Yeh, National Chiayi University
Liang-Chih Yu, Yuan Ze University
Marcos Zampieri, Saarland University
Torsten Zesch, University of Duisburg-Essen
Hui Zhang, Facebook Inc.

Table of Contents

Overview of the NLP-TEA 2015 Shared Task for Chinese Grammatical Error Diagnosis
Lung-Hao Lee, Liang-Chih Yu and Li-Ping Chang 1

Chinese Grammatical Error Diagnosis by Conditional Random Fields
Shih-Hung Wu, Po-Lin Chen, Liang-Pu Chen, Ping-Che Yang and Ren-Dar Yang 7

NTOU Chinese Grammar Checker for CGED Shared Task
Chuan-Jie Lin and Shao-Heng Chen 15

Collocation Assistant for Learners of Japanese as a Second Language
Lis Pereira and Yuji Matsumoto 20

Semi-automatic Generation of Multiple-Choice Tests from Mentions of Semantic Relations
Renlong Ai, Sebastian Krause, Walter Kasper, Feiyu Xu and Hans Uszkoreit 26

Interactive Second Language Learning from News Websites
Tao Chen, Naijia Zheng, Yue Zhao, Muthu Kumar Chandrasekaran and Min-Yen Kan 34

Bilingual Keyword Extraction and its Educational Application
Chung-Chi Huang, Mei-Hua Chen and Ping-Che Yang 43

Annotating Entailment Relations for Shortanswer Questions
Simon Ostermann, Andrea Horbach and Manfred Pinkal 49

An Automated Scoring Tool for Korean Supply-type Items Based on Semi-Supervised Learning
Minah Cheon, Hyeong-Won Seo, Jae-Hoon Kim, Eun-Hee Noh, Kyung-Hee Sung and EunYong Lim 59

A System for Generating Multiple Choice Questions: With a Novel Approach for Sentence Selection
Mukta Majumder and Sujan Kumar Saha 64

The "News Web Easy" news service as a resource for teaching and learning Japanese: An assessment of the comprehension difficulty of Japanese sentence-end expressions
Hideki Tanaka, Tadashi Kumano and Isao Goto 73

Grammatical Error Correction Considering Multi-word Expressions
Tomoya Mizumoto, Masato Mita and Yuji Matsumoto 82

Salinlahi III: An Intelligent Tutoring System for Filipino Heritage Language Learners
Ralph Vincent Regalado, Michael Louie Boñon, Nadine Chua, Rene Rose Piñera and Shannen Rose Dela Cruz 87

Using Finite State Transducers for Helping Foreign Language Learning
Hasan Kaya and Gülşen Eryiğit 94

Chinese Grammatical Error Diagnosis Using Ensemble Learning
Yang Xiang, Xiaolong Wang, Wenying Han and Qinghua Hong 99

Condition Random Fields-based Grammatical Error Detection for Chinese as Second Language
Jui-Feng Yeh, Chan Kun Yeh, Kai-Hsiang Yu, Ya-Ting Li and Wan-Ling Tsai 105

Improving Chinese Grammatical Error Correction with Corpus Augmentation and Hierarchical Phrase-based Statistical Machine Translation
Yinchen Zhao, Mamoru Komachi and Hiroshi Ishikawa 111

Chinese Grammatical Error Diagnosis System Based on Hybrid Model
Xiupeng Wu, Peijie Huang, Jundong Wang, Qingwen Guo, Yuhong Xu and Chuping Chen 117

Workshop Program

Friday, July 31, 2015

09:30–09:40 Opening Ceremony

09:40–10:30 Invited Speech

09:40–10:30 *Big Data-Based Automatic Essay Scoring Service — www.pigai.org*
Zhang Yu, Beijing Cikuu Corp.

10:30–11:00 Coffee Break

11:00–12:00 Shared Task Session

11:00–11:20 *Overview of the NLP-TEA 2015 Shared Task for Chinese Grammatical Error Diagnosis*
Lung-Hao Lee, Liang-Chih Yu and Li-Ping Chang

11:20–11:40 *Chinese Grammatical Error Diagnosis by Conditional Random Fields*
Shih-Hung Wu, Po-Lin Chen, Liang-Pu Chen, Ping-Che Yang and Ren-Dar Yang

11:40–12:00 *NTOU Chinese Grammar Checker for CGED Shared Task*
Chuan-Jie Lin and Shao-Heng Chen

12:00–14:00 Lunch

Friday, July 31, 2015 (continued)

14:00–15:30 Regular Paper Session

14:00–14:30 *Collocation Assistant for Learners of Japanese as a Second Language*
Lis Pereira and Yuji Matsumoto

14:30–15:00 *Semi-automatic Generation of Multiple-Choice Tests from Mentions of Semantic Relations*
Renlong Ai, Sebastian Krause, Walter Kasper, Feiyu Xu and Hans Uszkoreit

15:00–15:30 *Interactive Second Language Learning from News Websites*
Tao Chen, Naijia Zheng, Yue Zhao, Muthu Kumar Chandrasekaran and Min-Yen Kan

15:30–16:00 Coffee Break

16:00–17:00 Poster Session

16:00–16:05 *Bilingual Keyword Extraction and its Educational Application*
Chung-Chi Huang, Mei-Hua Chen and Ping-Che Yang

16:05–16:10 *Annotating Entailment Relations for Shortanswer Questions*
Simon Ostermann, Andrea Horbach and Manfred Pinkal

16:10–16:15 *An Automated Scoring Tool for Korean Supply-type Items Based on Semi-Supervised Learning*
Minah Cheon, Hyeong-Won Seo, Jae-Hoon Kim, Eun-Hee Noh, Kyung-Hee Sung and EunYong Lim

16:15–16:20 *A System for Generating Multiple Choice Questions: With a Novel Approach for Sentence Selection*
Mukta Majumder and Sujan Kumar Saha

16:20–16:25 *The "News Web Easy" news service as a resource for teaching and learning Japanese: An assessment of the comprehension difficulty of Japanese sentence-end expressions*
Hideki Tanaka, Tadashi Kumano and Isao Goto

16:25–16:30 *Grammatical Error Correction Considering Multi-word Expressions*
Tomoya Mizumoto, Masato Mita and Yuji Matsumoto

16:30–16:35 *Salinlahi III: An Intelligent Tutoring System for Filipino Heritage Language Learners*
Ralph Vincent Regalado, Michael Louie Boñon, Nadine Chua, Rene Rose Piñera and Shannen Rose Dela Cruz

Friday, July 31, 2015 (continued)

16:35–16:40 *Using Finite State Transducers for Helping Foreign Language Learning*
Hasan Kaya and Gülşen Eryiğit

16:40–16:45 *Chinese Grammatical Error Diagnosis Using Ensemble Learning*
Yang Xiang, Xiaolong Wang, Wenying Han and Qinghua Hong

16:45–16:50 *Condition Random Fields-based Grammatical Error Detection for Chinese as Second Language*
Jui-Feng Yeh, Chan Kun Yeh, Kai-Hsiang Yu, Ya-Ting Li and Wan-Ling Tsai

16:50–16:55 *Improving Chinese Grammatical Error Correction with Corpus Augmentation and Hierarchical Phrase-based Statistical Machine Translation*
Yinchen Zhao, Mamoru Komachi and Hiroshi Ishikawa

16:55–17:00 *Chinese Grammatical Error Diagnosis System Based on Hybrid Model*
Xiupeng Wu, Peijie Huang, Jundong Wang, Qingwen Guo, Yuhong Xu and Chuping Chen

17:00–17:10 Closing Remarks