Lecture Notes in Artificial Intelligence 11221

Subseries of Lecture Notes in Computer Science

More information about this series at http://www.springer.com/series/1244

Maosong Sun · Ting Liu
Xiaojie Wang · Zhiyuan Liu
Yang Liu (Eds.)

Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data

17th China National Conference, CCL 2018
and 6th International Symposium, NLP-NABD 2018
Changsha, China, October 19–21, 2018
Proceedings

Springer

Editors
Maosong Sun
Tsinghua University
Beijing, China

Ting Liu
Harbin Institute of Technology
Harbin, China

Xiaojie Wang
Beijing University of Posts
and Telecommunications
Beijing, China

Zhiyuan Liu
Tsinghua University
Beijing, China

Yang Liu
Tsinghua University
Beijing, China

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Artificial Intelligence
ISBN 978-3-030-01715-6 ISBN 978-3-030-01716-3 (eBook)
https://doi.org/10.1007/978-3-030-01716-3

Library of Congress Control Number: 2018956472

LNCS Sublibrary: SL7 – Artificial Intelligence

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Welcome to the proceedings of the 17th China National Conference on Computational Linguistics (17th CCL) and the 6th International Symposium on Natural Language Processing Based on Naturally Annotated Big Data (6th NLP-NABD). The conference and symposium were hosted by Changsha University of Science and Technology located in Changsha City, Hunan Province, China.

CCL is an annual conference (bi-annual before 2013) that started in 1991. It is the flagship conference of the Chinese Information Processing Society of China (CIPS), which is the largest NLP scholar and expert community in China. CCL is a premier nation-wide forum for disseminating new scholarly and technological work in computational linguistics, with a major emphasis on computer processing of the languages in China such as Mandarin, Tibetan, Mongolian, and Uyghur.

Affiliated with the 17th CCL, the 6th International Symposium on Natural Language Processing Based on Naturally Annotated Big Data (NLP-NABD) covered all the topics of NLP, with particular focus on methodologies and techniques relating to naturally annotated big data. In contrast to manually annotated data such as treebanks that are constructed for specific NLP tasks, naturally annotated data come into existence through users' normal activities, such as writing, conversation, and interactions on the Web. Although the original purposes of these data typically were unrelated to NLP, they can nonetheless be purposefully exploited by computational linguists to acquire linguistic knowledge. For example, punctuation marks in Chinese text can help word boundaries identification, social tags in social media can provide signals for keyword extraction, and categories listed in Wikipedia can benefit text classification. The natural annotation can be explicit, as in the aforementioned examples, or implicit, as in Hearst patterns (e.g., "Beijing and other cities" implies "Beijing is a city"). This symposium focuses on numerous research challenges ranging from very-large-scale unsupervised/semi-supervised machine leaning (deep learning, for instance) of naturally annotated big data to integration of the learned resources and models with existing handcrafted "core" resources and "core" language computing models. NLP-NABD 2018 was supported by the National Key Basic Research Program of China (i.e., "973" Program) "Theory and Methods for Cyber-Physical-Human Space Oriented Web Chinese Information Processing" under grant no.2014CB340500 and the Major Project of the National Social Science Foundation of China under grant no. 13&ZD190.

The Program Committee selected 102 papers (69 Chinese papers and 33 English papers) out of 277 submissions from China, Hong Kong (region), Singapore, and USA for publication. The acceptance rate is 36.82%. The 33 English papers cover the following topics:

- Semantics (3)
- Machine translation (6)
- Knowledge graph and information extraction (7)
- Linguistic resource annotation and evaluation (2)

- Information retrieval and question answering (4)
- Text classification and summarization (5)
- Social computing and sentiment analysis (2)
- NLP applications (4)

The final program for the 17th CCL and the sixth NLP-NABD was the result of a great deal of work by many dedicated colleagues. We want to thank, first of all, the authors who submitted their papers, and thus contributed to the creation of the high-quality program that allowed us to look forward to an exciting joint conference. We are deeply indebted to all the Program Committee members for providing high-quality and insightful reviews under a tight schedule. We are extremely grateful to the sponsors of the conference. Finally, we extend a special word of thanks to all the colleagues of the Organizing Committee and secretariat for their hard work in organizing the conference, and to Springer for their assistance in publishing the proceedings in due time.

We thank the Program and Organizing Committees for helping to make the conference successful, and we hope all the participants enjoyed a remarkable visit to Changsha, a historical and beautiful city in South China.

August 2018

Maosong Sun
Ting Liu
Xiaojie Wang
Randy Goebel
Heng Ji

Organization

General Chairs

Sheng Li	Harbin Institute of Technology, China
Changning Huang	Tsinghua University, China
Kaiying Liu	Shanxi University, China

Program Committee

17th CCL Program Chairs

Maosong Sun	Tsinghua University, China
Ting Liu	Harbin Institute of Technology, China
Xiaojie Wang	Beijing University of Posts and Telecommunications, China

17th CCL Area Co-chairs

Linguistics and Cognitive Science

Weiguang Qu	Nanjing Normal University, China
Yulin Yuan	Peking University, China

Fundamental Theory and Methods of Computational Linguistics

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

Information Retrieval and Question Answering

Jun Xu	Institute of Computing Technology, CAS, China
Aixin Sun	Nanyang Technological University, Singapore

Text Classification and Summarization

Xipeng Qiu	Fudan University, China
Xiaodan Zhu	Queen's University, Canada

Knowledge Graph and Information Extraction

Xianpei Han	Institute of Software, CAS, China
Ni Lao	Google Inc., USA

Machine Translation

Shujian Huang	Nanjing University, China
Haitao Mi	Ant Financial, USA

Minority Language Information Processing

Xiaobing Zhao	Minzu University of China, China
Celimuge Wu	The University of Electro-Communications, Japan

Language Resource and Evaluation

Muyun Yang	Harbin Institute of Technology, China
Rui Wang	National Institute of Information and Communications Technology, NICT, Japan

Social Computing and Sentiment Analysis

Jia Jia	Tsinghua University, China
Yanyan Zhao	Harbin Institute of Technology, China

NLP Applications

Qingcai Chen	Harbin Institute of Technology, Shenzhen, China
Cui Tao	The University of Texas Health Science Center at Houston, USA

6th NLP-NABD Program Chairs

Maosong Sun	Tsinghua University, China
Randy Goebel	University of Alberta, Canada
Heng Ji	Rensselaer Polytechnic Institute, USA

17th CCL and 6th NLP-NABD Local Arrangements Chairs

Feng Li	Changsha University of Science and Technology, China
Jin Wang	Changsha University of Science and Technology, China
Yang Liu	Tsinghua University, China

17th CCL and 6th NLP-NABD Evaluation Chairs

Ting Liu	Harbin Institute of Technology, China
Wei Song	Capital Normal University, China

17th CCL and 6th NLP-NABD Publication Chairs

Zhiyuan Liu	Tsinghua University, China
Shizhu He	Institute of Automation, CAS, China

17th CCL and 6th NLP-NABD Publicity Chairs

Ruifeng Xu	Harbin Institute of Technology, Shenzhen, China
William Wang	University of California, Santa Barbara, USA

17th CCL and 6th NLP-NABD Advance Lectures Chairs

Xiaojun Wan	Peking University, China
Jiafeng Guo	Institute of Computing Technology, CAS, China

17th CCL and 6th NLP-NABD Sponsorship Chairs

Qi Zhang	Fudan University, China
Jiajun Zhang	Institute of Automation, CAS, China

17th CCL and 6th NLP-NABD System Demonstration Chairs

Tong Xiao	Northeastern University, China
Binyang Li	University of International Relations, China

17th CCL and 6th NLP-NABD Student Seminar Chairs

Yang Feng	Institute of Computing Technology, CAS, China
Dong Yu	Beijing Language and Culture University, China

17th CCL and 6th NLP-NABD Organizers

Chinese Information Processing Society of China

Tsinghua University

Changsha University of Science & Technology, China

Publishers

Springer

Lecture Notes in Artificial Intelligence, Springer

中国中文信息学会会刊 中文信息学报 ZHONGWEN XINXI XUEBAO

Journal of Chinese Information Processing

《中国科学》杂志社 Science China Press

Science China

清华大学学报（自然科学版） Journal of Tsinghua University (Science and Technology)

Journal of Tsinghua University (Science and Technology)

Sponsoring Institutions

Platinum

犀语科技 AlphaInsight

Alex&Masen 雅乐美森

Gold

Silver

Bronze

Contents

Text Classification and Summarization

Social Computing and Sentiment Analysis

NLP Applications