

Maosong Sun · Xuanjing Huang
Hongfei Lin · Zhiyuan Liu
Yang Liu (Eds.)

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

15th China National Conference, CCL 2016, and
4th International Symposium, NLP-NABD 2016
Yantai, China, October 15–16, 2016
Proceedings

Contents

Semantics

- Improving Chinese Semantic Role Labeling with English Proposition Bank . . . 3
Tianshi Li, Qi Li, and BaoBao Chang
- Transition-Based Chinese Semantic Dependency Graph Parsing 12
Yuxuan Wang, Jiang Guo, Wanxiang Che, and Ting Liu
- Improved Graph-Based Dependency Parsing via Hierarchical
LSTM Networks 25
Wenhui Wang and Baobao Chang

Machine Translation

- Error Analysis of English-Chinese Machine Translation 35
Fei Fang, Shili Ge, and Rou Song
- I Can Guess What You Mean: A Monolingual Query Enhancement
for Machine Translation 50
Chenxi Pang, Hai Zhao, and Zhongyi Li
- Keeping the Meanings of the Source Text: An Introduction
to Yes Translate 64
Xiaoheng Zhang
- Sentence Alignment Method Based on Maximum Entropy Model
Using Anchor Sentences 76
Chao Che, Wenwen Guo, and Jianxin Zhang
- Using Collaborative Training Method to Build Vietnamese
Dependency Treebank 86
Guoke Qiu, Jianyi Guo, Zhengtao Yu, Yantuan Xian, and Cunli Mao

Multilinguality in NLP

- A Novel Approach to Improve the Mongolian Language Model
Using Intermediate Characters 103
Xiaofei Yan, Feilong Bao, Hongxi Wei, and Xiangdong Su
- Improved Joint Kazakh POS Tagging and Chunking 114
Hao Wu and Gulila Altenbek

Coping with Problems of Unicoded Traditional Mongolian. 125
Boli Wang, Xiaodong Shi, and Yidong Chen

Tibetan Person Attributes Extraction Based on BP Neural Network. 132
Lili Guo and Yuan Sun

Semi-supervised Learning for Mongolian Morphological Segmentation 143
Zhenxin Yang, Miao Li, Lei Chen, Weihui Zeng, Yi Gao, and Sha Fu

Investigation and Use of Methods for Defining the Extends of Similarity
of Kazakh Language Sentences. 153
*Unzila Kamanur, Altynbek Sharipbay, Gulila Altenbek,
Gulmira Bekmanova, and Lena Zhetkenbay*

Knowledge Graph and Information Extraction

Recognizing Biomedical Named Entities Based on the Sentence
Vector/Twin Word Embeddings Conditioned Bidirectional LSTM. 165
Lishuang Li, Liuke Jin, Yuxin Jiang, and Degen Huang

Definition Extraction with LSTM Recurrent Neural Networks. 177
SiLiang Li, Bin Xu, and Tong Lee Chung

Event Extraction via Bidirectional Long Short-Term Memory
Tensor Neural Networks 190
Yubo Chen, Shulin Liu, Shizhu He, Kang Liu, and Jun Zhao

Chinese Hedge Scope Detection Based on Structure and Semantic
Information 204
*Huiwei Zhou, Junli Xu, Yunlong Yang, Huijie Deng, Long Chen,
and Degen Huang*

Combining Event-Level and Cross-Event Semantic Information
for Event-Oriented Relation Classification by SCNN. 216
Siyuan Ding, Yu Hong, Shanshan Zhu, Jianmin Yao, and Qiaoming Zhu

Linguistic Resource Annotation and Evaluation

The Constitution of a Fine-Grained Opinion Annotated Corpus on Weibo . . . 227
Liao Jian, Li Yang, and Wang Suge

The Construction of a Customized Medical Corpus for Assisting Chinese
Clinicians in English Research Article Writing 241
Xiaowen Wang, Yong Gao, and Tianyong Hao

Information Retrieval and Question Answering

- Topic-Sentiment Mining from Multiple Text Collections 255
Qifeng Zhu and Fang Li
- A New Focus Strategy for Efficient Dialog Management 267
Xinqi Bao, Yunfang Wu, and Xueqiang Lv

Text Classification and Summarization

- Recognizing Textual Entailment via Multi-task Knowledge Assisted LSTM . . . 285
Lei Sha, Sujian Li, Baobao Chang, and Zhifang Sui
- Multilingual Multi-document Summarization with Enhanced
 hLDA Features 299
Taiwen Huang, Lei Li, and Yazhao Zhang
- News Abridgement Algorithm Based on Word Alignment
 and Syntactic Parsing 313
*Min Yu, Huaping Zhang, Yu Zhang, Yang Qiao, Zhonghua Zhao,
 and Yueying He*
- A Hierarchical LSTM Model for Joint Tasks 324
Qianrong Zhou, Liyun Wen, Xiaojie Wang, Long Ma, and Yue Wang
- Enhancing Neural Disfluency Detection with Hand-Crafted Features 336
Shaolei Wang, Wanxiang Che, Yijia Liu, and Ting Liu

Social Computing and Sentiment Analysis

- Active Learning for Age Regression in Social Media. 351
Jing Chen, Shoushan Li, Bin Dai, and Guodong Zhou
- A Novel Approach for Discovering Local Community Structure
 in Networks 363
Jinglian Liu, Daling Wang, Weiwei Zhao, Shi Feng, and Yifei Zhang
- Identifying Suspected Cybermob on Tieba 375
Shumin Shi, Xinyu Zhou, Meng Zhao, and Heyan Huang
- Chinese Sentiment Analysis Exploiting Heterogeneous Segmentations 387
Da Pan, Meishan Zhang, and Guohong Fu
- Towards Scalable Emotion Classification in Microblog Based on Noisy
 Training Data 399
Minglei Li, Qin Lu, Lin Gui, and Yunfei Long

NLP Applications

A Bootstrapping Approach to Symptom Entity Extraction on Chinese Electronic Medical Records	413
<i>Tianyi Qin and Yi Guan</i>	
Automatic Naming of Speakers in Video via Name-Face Mapping	424
<i>Zhixin Liu, Cheng Jin, Yuejie Zhang, and Tao Zhang</i>	
Image Tag Recommendation via Deep Cross-Modal Correlation Mining	437
<i>Xingmeng Zhang, Cheng Jin, Yuejie Zhang, and Tao Zhang</i>	
Is Local Window Essential for Neural Network Based Chinese Word Segmentation?	450
<i>Jinchao Zhang, Fandong Meng, Mingxuan Wang, Daqi Zheng, Wenbin Jiang, and Qun Liu</i>	
Author Index	459