Sixth SIGHAN Workshop on Chinese Language Processing

(IJCNLP 2008)

Hyderabad, India
11-12 January 2008
# Table of Contents

Preface ................................................................. iii
Organizers ............................................................ v
Workshop Program .................................................. vii

*An Example-based Decoder for Spoken Language Machine Translation*
   Zhou-Jun Li, Wen-Han Chao and Yue-Xin Chen ........................................... 1

*Automatic Extraction of English-Chinese Transliteration Pairs using Dynamic Window and Tokenizer*
   Chengguo Jin, Seung-Hoon Na, Dong-Il Kim and Jong-Hyeok Lee ......................... 9

*Mining Transliterations from Web Query Results: An Incremental Approach*
   Jin-Sheu Kuo, Haizhou Li and Chih-Lung Lin ................................................. 16

*An Effective Hybrid Machine Learning Approach for Coreference Resolution*
   Feiliang Ren and Jingbo Zhu ................................................................. 24

*Use of Event Types for Temporal Relation Identification in Chinese Text*
   Yuchang Cheng, Masayuki Asahara and Yuji Matsumoto .................................. 31

*Chinese Word Sense Disambiguation with PageRank and HowNet*
   Jinghua Wang, Jianyi Liu and Ping Zhang .................................................... 39

*Stochastic Dependency Parsing Based on A* Admissible Search*
   Bor-shen Lin ........................................................................ 45

*Analyzing Chinese Synthetic Words with Tree-based Information and a Survey on Chinese Morphologically Derived Words*
   Jia Lu, Masayuki Asahara and Yuji Matsumoto .............................................. 53

*Which Performs Better on In-Vocabulary Word Segmentation: Based on Word or Character?*
   Zhenxing Wang, Changning Huang and Jingbo Zhu ....................................... 61

   Guangjin Jin and Xiao Chen ................................................................. 69

*A Two-Stage Approach to Chinese Part-of-Speech Tagging*
   Aitao Chen, Ya Zhang and Gordon Sun .................................................... 82

*NOKIA Research Center Beijing Chinese Word Segmentation System for the SIGHAN Bakeoff 2007*
   Jiang Li, Rile Hu, Guohua Zhang, Yuezhong Tang, Zhanjiang Song and Xia Wang .... 86
Description of the NCU Chinese Word Segmentation and Part-of-Speech Tagging for SIGHAN Bakeoff 2007
    Yu-Chieh Wu, Jie-Chi Yang and Yue-Shi Lee .......................................................... 161

CRF-based Hybrid Model for Word Segmentation, NER and even POS Tagging
    Zhiting Xu, Xian Qian, Yuejie Zhang and Yaqian Zhou ........................................ 167

CRFs-Based Named Entity Recognition Incorporated with Heuristic Entity List Searching
    Fan Yang, Jun Zhao and Bo Zou ................................................................. 171

A Chinese Word Segmentation System Based on Cascade Model
    Jianfeng Zhang, Jiaheng Zheng, Hu Zhang and Hongye Tan ............................... 175

Achilles: NiCT/ATR Chinese Morphological Analyzer for the Fourth Sighan Bakeoff
    Ruiqiang Zhang and Eiichiro Sumita ............................................................. 178

Author Index ............................................................................................................. 183