PACLIC 23

Proceedings of the 23rd Pacific Asia Conference on Language, Information and Computation

Volume 1

Edited by Olivia Kwong

3 – 5 December 2009
City University of Hong Kong

© 2009 City University of Hong Kong and PACLIC Steering Committee

All rights reserved. No part of this publication may be reproduced, digitized, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, Internet or otherwise, without the prior permission of the publisher.

Copyright of contributed papers reserved by respective authors.

ISBN 978-962-442-319-8

Published by City University of Hong Kong Press

For the **Department of Chinese, Translation and Linguistics**
City University of Hong Kong
Tat Chee Avenue
Kowloon
Hong Kong

Preface: Conference Chair

Hong Kong, and, in particular, the City University of Hong Kong are pleased to welcome PACLIC again for the third time: the first was in 1995 and the second in 2001. Since our last PACLIC in Hong Kong, *Language, Information and Computation*, the principal concerns of PACLIC have increasingly impacted on human life and society, ranging from research underlying search engines to work on speech technology and mobile phone access. Interests in this field have blossomed considerably, with many truly global and regional conferences having been held in the Pacific Rim. They include ACL and IJCNLP (International Joint Conference on Natural Language Processing), the flagship conference of the recently established Asian Federation of Natural Language Processing (AFNLP), of which PACLIC is a member.

The 145 submissions to the 23rd PACLIC have more than doubled the number of submissions in each of the recent conferences, thereby showing a tremendous expansion of interests. It also shows the drawing power of PACLIC as a successful regional conference, which is reflected by the 50th place it holds out of 701 conferences in the area of AI and Machine Learning according to the Computer Science Conference Ranking Website. The acceptance rate of 40% for oral presentation and 25% for poster session continues the tradition of rigorous review process for PACLIC.

We are very grateful that the 23rd PACLIC has brought together a hundred presenters, as well as a large group of invited speakers with diverse background: Dr. Keh-Jiann Chen (Academia Sinica, Taipei), Professor Dafydd Gibbon (Universität Bielefeld), Professor Noriko Kando (National Institute of Informatics, Tokyo), Dr. Anna Korhonen (University of Cambridge), Professor Ik-Hwan Lee (Sangmyung University), Professor Helen Meng (Chinese University of Hong Kong), and Dr. Ming Zhou (Microsoft Research Asia, Beijing).

The success of PACLIC has come about as a result of cultivation by many scholars over nearly a quarter of a century. They include Professor Akira Ikeya from Tokyo and Professor Byung-Soo Park from Seoul. For the 23rd PACLIC, we are especially pleased to see a special co-located event: A Dialogue between Professor Sydney Lamb and Professor William Wang. The two distinguished scholars have been keen observers and influential contributors on *Language, Information and Computation* since the 1960's when a paradigmatic shift in their related studies became evident with the advent of the computer and the additional availability of new frontiers in languages and linguistics.

The 23rd PACLIC is grateful for the support and work of Professor Jae-Woong Choe, Chair of the PACLIC Steering Committee, Professor Chu-Ren Huang, the Conference Co-Chair, Professor Jonathan Webster, the Program Committee Advisor, and Dr. Olivia Kwong, under whom the Program Committee members and the entire Local Organization Team have attended to the conference preparations with dedication, and taken care of even the finest logistic details with devotion. I hope this will be a memorable occasion for the PACLIC participants as well as for the colleagues and students in the Department of Chinese, Translation & Linguistics and Language Information Sciences Research Centre of the City University of Hong Kong.

Benjamin K. T'sou
November 2009

Preface: Program Committee Chair

The number and breadth of submissions to PACLIC conferences as well as the diversity of authorship have been on an upward trend since the inauguration of the series in 1995. However it expands, PACLIC is certainly proud to remain one of the few which manage to maintain a healthily balanced and symbiotic co-existence of formal linguists and computational linguists, and continue to emphasize the synergy of theoretical analysis and computer processing of language among researchers in the Pacific Asia region.

Such strengths are evident from the 145 papers submitted to PACLIC 23 from 20 countries or regions: 82.8% from 9 regions in Asia, 8.3% from 7 regions in Europe, 5.5% from the United States and Canada, 2.8% from the Middle East, and 0.7% from Australia. We would like to thank all authors for their submissions. Of these, 40% were accepted for regular paper presentations and an additional 25% for poster presentations, constituting a diverse technical program. I must say a big thankyou to our professional Program Committee, including the six regional Co-Chairs: Dr. Shu-Kai Hsieh, Dr. Donghong Ji, Prof. Yae-Sheik Lee, Prof. Rachel Edita Roxas, Prof. Kei Yoshimoto, and Dr. Min Zhang, 73 supportive committee members and 34 helpful additional reviewers, for their competence and hard work, and most importantly, prompt feedback.

PACLIC 23 will span three days from 3rd to 5th December 2009, featuring invited talks, plenary talks, regular paper presentations, and a poster session. We would like to thank the invited and plenary speakers, who will enlighten us on a wide range of topics: Dr. Keh-Jiann Chen and Dr. Anna Korhonen on lexical semantics, Prof. Dafydd Gibbon and Prof. Helen Meng on speech technology, Prof. Noriko Kando on system evaluation, Prof. Ik-Hwan Lee on formal semantics, and Dr. Ming Zhou on Chinese couplet and quatrain generation. PACLIC 23 is also pleased to co-locate with a lecture series on *LANGUAGE: Its Neurocognitive Bases*, which will climax with a dialogue between Prof. Sydney Lamb and Prof. William Wang on 5th December.

My job would have been impossible without the help of many individuals. I am very much indebted to Prof. Benjamin T'sou, the Conference Chair, for his long-term unfailing support and particularly his unswayed confidence in my role for this conference. I am also most grateful to Prof. Chu-Ren Huang, the Conference Co-Chair, for his significant help in promoting the conference. My heart-felt thanks also goes to Prof. Jonathan Webster, the Program Committee Advisor, for his advice and generous departmental support. I would also like to thank the PACLIC Steering Committee, particularly Prof. Jae-Woong Choe, the Committee Chair, and Prof. Yasunari Harada, the Digital Archivist; as well as previous PACLIC organizers, including Prof. Rachel Edita Roxas, Prof. Hee-Rahk Chae, and Prof. Tingting He, for sharing their experience. The concerted effort of the devoted and resourceful Local Organization Team is the major force putting the whole conference in shape, and the credit goes to Tom Lai, the co-convenor, Caesar Lun, Ian Chow, Ernest Lam, Kin-Tat Ko, Toby Hong, Sadie Chan, Alpha Ho, and all colleagues who have contributed in one way or another behind the scene.

Last but not least, as a participant, your presence makes our conference special! Thank you for making a difference to PACLIC 23! Enjoy our program and have a nice time in Hong Kong!

Olivia O.Y. Kwong
November 2009

Organizers

PACLIC Steering Committee

Jae-Woong Choe, Korea University
Yasunari Harada, Waseda University
Chu-Ren Huang, Hong Kong Polytechnic University
Kim Teng Lua, Chinese and Oriental Languages Information Processing Society
Maosong Sun, Tsinghua University
Benjamin T'sou, City University of Hong Kong

Conference Chair

Benjamin T'sou, City University of Hong Kong

Conference Co-Chair

Chu-Ren Huang, Hong Kong Polytechnic University

Program Committee

Advisor

Jonathan Webster, City University of Hong Kong

Chair

Olivia Kwong, City University of Hong Kong

Co-Chairs

Shu-Kai Hsieh, National Taiwan Normal University
Donghong Ji, Wuhan University
Yae-Sheik Lee, Kyungpook National University
Rachel Edita Roxas, De La Salle University-Manila
Kei Yoshimoto, Tohoku University
Min Zhang, Institute for Infocomm Research

Members

Timothy Baldwin, University of Melbourne
Francis Bond, Nanyang Technological University
Claire Cardie, Cornell University
Hee-Rahk Chae, Hankuk University of Foreign Studies
Jason Chang, National Tsing Hua University
Jing-Shin Chang, National Chi-Nan University
Eng Siong Chng, Nanyang Technological University
Siaw-Fong Chung, National Chengchi University
Beatrice Daille, University of Nantes
Alex Fang, City University of Hong Kong
Guohong Fu, Heilongjiang University
Helena Hong Gao, Nanyang Technological University
Zhao-Ming Gao, National Taiwan University

Yi Guan, Harbin Institute of Technology
Yasunari Harada, Waseda University
Munpyo Hong, Sungkyunkwan University
Xuanjing Huang, Fudan University
Kiyoshi Ishikawa, Hosei University
Jong-Bok Kim, Kyung Hee University
Young-Gil Kim, Electronics and Telecommunications Research Institute
Chunyu Kit, City University of Hong Kong
Valia Kordoni, Saarland University
Anna Korhonen, University of Cambridge
Paul Law, City University of Hong Kong
Chungmin Lee, Seoul National University
Hae-yun Lee, Hankuk University of Foreign Studies
Alessandro Lenci, University of Pisa
Gina-Anne Levow, University of Manchester
Haizhou Li, Institute for Infocomm Research
Wenjie Li, Hong Kong Polytechnic University
Qun Liu, Chinese Academy of Sciences
Ting Liu, Harbin Institute of Technology
Qin Lu, Hong Kong Polytechnic University
Qing Ma, Ryukoku University
Yuji Matsumoto, Nara Institute of Science and Technology
Yoshiki Mori, University of Tokyo
Vincent Ng, University of Texas at Dallas
Grace Ngai, Hong Kong Polytechnic University
Toshiyuki Ogihara, University of Washington
Ryo Otoguro, Waseda University
Haihua Pan, City University of Hong Kong
Jong C. Park, Korea Advanced Institute of Science and Technology
Laurent Prévot, Université de Provence
Long Qiu, Institute for Infocomm Research
Rajeev Sangal, International Institute of Information Technology-Hyderabad
Sachiko Shudo, Waseda University
Pornsiri Singhapreecha, Thammasat University
Virach Sornlertlamvanich, Thai Computational Linguistics Laboratory, NICT
I-Wen Su, National Taiwan University
Jian Su, Institute for Infocomm Research
Keh-Yih Su, Behavior Design Corporation
Zhifang Sui, Peking University
Le Sun, Chinese Academy of Sciences
Takenobu Tokunaga, Tokyo Institute of Technology
Shu-Chuan Tseng, Academia Sinica
Aline Villavicencio, Federal University of Rio Grande do Sul
Haifeng Wang, Toshiba (China) R&D Center
Houfeng Wang, Peking University
Hui Wang, National University of Singapore
Kam-Fai Wong, Chinese University of Hong Kong
Dekai Wu, Hong Kong University of Science and Technology
Jiun-Shiung Wu, National Chiayi University
Fei Xia, University of Washington

Nianwen Xue, Brandeis University
Jae-Il Yeom, Hongik University
Yujie Zhang, National Institute of Information and Communications Technology
Jun Zhao, Chinese Academy of Sciences
Tiejun Zhao, Harbin Institute of Technology
Guo Dong Zhou, Soochow University
Qiang Zhou, Tsinghua University
Jingbo Zhu, Northeastern University
Michael Zock, Laboratoire d'Informatique Fondamentale de Marseille, C.N.R.S.
Chengqing Zong, Chinese Academy of Sciences

Additional Reviewers

Baobao Chang, Peking University
Wanxiang Che, Harbin Institute of Technology
Charibeth Cheng, De La Salle University-Manila
Shi Feng, Northeastern University
Wei Gao, Chinese University of Hong Kong
Yun Huang, Institute for Infocomm Research
Wenbin Jiang, Chinese Academy of Sciences
Naoko Kinoshita, Waseda University
Lian Hau Lee, Institute for Infocomm Research
Binyang Li, Chinese University of Hong Kong
Sujian Li, Peking University
Hua Shen Liang, Institute for Infocomm Research
Nathalie Rose Lim, De La Salle University-Manila
Lemao Liu, Harbin Institute of Technology
Yang Liu, Peking University
Yongliang Ma, Harbin Institute of Technology
Dalos Miguel, Saint Louis University, Philippines
Yoon-Hyeng Noh, Electronics and Telecommunications Research Institute
You Ouyang, Hong Kong Polytechnic University
Young-Ae Seo, Electronics and Telecommunications Research Institute
Yan Song, City University of Hong Kong
Jun Sun, Institute for Infocomm Research
Masanori Suzuki, Pearson Knowledge Technologies and Waseda University
Thuy Vu, Institute for Infocomm Research
Billy Wong, City University of Hong Kong
Yunqing Xia, Tsinghua University
Chong Xu, Harbin Institute of Technology
Ge Xu, Peking University
Ruifeng Xu, City University of Hong Kong
Muyun Yang, Harbin Institute of Technology
Seong-Il Yang, Electronics and Telecommunications Research Institute
Hui Zhang, Institute for Infocomm Research
Hai Zhao, City University of Hong Kong
Yu Zhou, Chinese Academy of Sciences

Local Organizing Committee

Convenors

Tom Lai, City University of Hong Kong
Olivia Kwong, City University of Hong Kong

Members

Caesar Lun, City University of Hong Kong
Ian Chow, City University of Hong Kong

Secretary

Ernest Lam, City University of Hong Kong

Webmaster

Kin-Tat Ko, City University of Hong Kong

Graphic Design

Toby Hong, City University of Hong Kong

Editorial Assistants

Sadie Chan, City University of Hong Kong
Alpha Ho, City University of Hong Kong

Table of Contents

Preface: Conference Chair iii

Preface: Program Committee Chair v

Organizers vii

Invited and Plenary Talks

A Step toward Compositional Semantics: E-HowNet a Lexical Semantic Representation System
Keh-Jiann Chen and Shu-Ling Huang 1

Gesture Theory is Linguistics: On Modelling Multimodality as Prosody
Dafydd Gibbon 9

Automatic Lexical Classification – Balancing between Machine Learning and Linguistics
Anna Korhonen 19

Resultatives as Causal Relations between Events
Ik-Hwan Lee 29

Developing Speech Recognition and Synthesis Technologies to Support Computer-Aided Pronunciation Training for Chinese Learners of English
Helen Meng 40

Generating Chinese Couplets and Quatrain Using a Statistical Approach
Ming Zhou, Long Jiang and Jing He 43

Regular Papers

Capturing Lexical Variation in MT Evaluation Using Automatically Built Sense-Cluster Inventories
Marianna Apidianaki, Yifan He and Andy Way 53

Dependency Grammar Based English Subject-Verb Agreement Evaluation
Dongfeng Cai, Yonghua Hu, Xuelei Miao and Yan Song 63

WikiSense: Supersense Tagging of Wikipedia Named Entities Based WordNet
Joseph Chang, Richard Tzong-Han Tsai and Jason S. Chang 72

An Integrated Approach to Heterogeneous Data for Information Extraction
Ying Chen, Sophia Y. M. Lee and Chu-Ren Huang 82

Are Emotions Enumerable or Decomposable? And its Implications for Emotion Processing
Ying Chen, Sophia Y. M. Lee and Chu-Ren Huang 92

Scope and Anaphoric Links in Dynamic Discourse Representation Theory
So-Woo Chung and Jungmin Lee 101

Coupling an Annotated Corpus and a Morphosyntactic Lexicon for State-of-the-Art POS Tagging with Less Human Effort
Pascal Denis and Benoît Sagot 110

Voted Approach for Part of Speech Tagging in Bengali
Asif Ekbal, Md. Hasanuzzaman and Sivaji Bandyopadhyay 120

Adjective Density as a Text Formality Characteristic for Automatic Text Classification: A Study Based on the British National Corpus
Alex Chengyu Fang and Jing Cao 130

Correcting Errors Using the Framework of Argumentation: Towards Generating Argumentative Correction Propositions from Error Annotation Schemas
Marie Garnier, Arnaud Rykner and Patrick Saint-Dizier 140

Shallow Semantic Parsing of Persian Sentences
Azadeh Kamel Ghalibaf, Saeed Rahati and Azam Estaji 150

LogisticLDA: Regularizing Latent Dirichlet Allocation by Logistic Regression
Jia-Cheng Guo, Bao-Liang Lu, Zhiwei Li and Lei Zhang 160

Dependency Relations as Source Context in Phrase-Based SMT
Rejwanul Haque, Sudip Kumar Naskar, Antal van den Bosch and Andy Way 170

Multi-Task Learning in Conditional Random Fields for Chunking in Shallow Semantic Parsing
Saike He, Xiaojie Wang, Yuan Dong, Taozheng Zhang and Xue Bai 180

Query-Focused Multi-Document Summarization Using Co-Training Based Semi-Supervised Learning
Po Hu, Donghong Ji, Hai Wang and Chong Teng 190

Review Classification Using Semantic Features and Run-Time Weighting
Chung-chi Huang, Meng-chiech Lee, Zhe-nan Lin and Jason S. Chang 200

A Key Word Analysis of English Intensifying Adverbs in Male and Female Speech in ICE-GB
Jieun Jeon and Jae-Woong Choe 210

Gender and Animacy Knowledge Discovery from Web-Scale N-Grams for Unsupervised Person Mention Detection
Heng Ji and Dekang Lin 220

Layer-Based Dependency Parsing
Ping Jian and Chengqing Zong 230

Mediatory Summary Generation: Summary-Passage Extraction for Information Credibility on the Web
Koichi Kaneko, Hideyuki Shibuki, Masahiro Nakano, Rintaro Miyazaki, Madoka Ishioroshi and Tatsunori Mori 240

An Experimental Syntactic Study of Binding: A Case Study of Korean Long-Distance Anaphor caki
Ji-Hye Kim and James H. Yoon 250

Method of Extracting Is-A and Part-Of Relations Using Pattern Pairs in Mass Corpus
Se-Jong Kim, Yong-Hun Lee and Jong-Hyeok Lee 260

On the Scope Interaction of Japanese Indefinites: An Epsilon Calculus Approach
Masahiro Kobayashi and Hiroaki Nakamura 269

Pattern Lattice as a Model for Linguistic Knowledge and Performance
Kow Kuroda 278

Chinese WordNet Domains: Bootstrapping Chinese WordNet with Semantic Domain Labels
Lung-Hao Lee, Yu-Ting Yu and Chu-Ren Huang 288

Sentiment Classification Considering Negation and Contrast Transition
Shoushan Li and Chu-Ren Huang 297

Modal Verbs for the Advice Move in Advice Columns
Ying-shu Liao and Ting-gen Liao 307

Approach to Selecting Best Development Set for Phrase-Based Statistical Machine Translation
Peng Liu, Yu Zhou and Chengqing Zong 315

Using Extra-Linguistic Material for Mandarin-French Verbal Constructions Comparison
Pierre Magistry, Laurent Prévot, Hintat Cheung, Chien-yun Shiao, Yann Desalle and Bruno Gaume 325

Improving Unsegmented Dialogue Turns Annotation with N-gram Transducers
Carlos-D. Martínez-Hinarejos, Vicent Tamarit and José-Miguel Benedí 335

Cause Event Representations for Happiness and Surprise
Sophia Yat Mei Lee, Ying Chen and Chu-Ren Huang 345

Using Tree Kernels for Classifying Temporal Relations between Events
Seyed Abolghasem Mirroshandel, Gholamreza Ghassem-Sani and Mahdy Khayyamian 355

How Language Processing Constrains (Computational) Natural Language Processing: A Cognitive Perspective
Prakash Mondal 365

Extending Bilingual WordNet via Hierarchical Word Translation Classification
Tzu-yi Nien, Tsun Ku, Chung-chi Huang, Mei-hua Chen and Jason S. Chang 375

Extended GL and Japanese Postposition No
Sumiyo Nishiguchi 385

Note on Japanese Epistemic Verb Constructions: A Surface-Compositional Analysis
Akira Ohtani and Mark Steedman 395

On the So-Called Thematic Use of Wa: Reconsideration and Reconciliation
David Y. Oshima 405

Chinese Semantic Class Learning from Web Based on Concept-Level Characteristics
Wenbo Pang, Xiaozhong Fan, Jiangde Yu and Yuxiang Jia 415

On Pseudogapping in HPSG
Dong-woo Park 425

Resolving Multiple Wh-Fragments: A Syntax-Pragmatics Approach
Sang-Hee Park 435

A Lexicalized Tree Adjoining Grammar for Thai
Siripong Potisuk 444

Incorporate Web Search Technology to Solve Out-of-Vocabulary Words in Chinese Word Segmentation
Wei Qiao and Maosong Sun 454

A Hybrid Model for Sense Guessing of Chinese Unknown Words
Likun Qiu, Kai Zhao and Changjian Hu 464

Collaborative Summarization: When Collaborative Filtering Meets Document Summarization
Yang Qu and Qunxiu Chen 474

Finding Answers to Definition Questions Using Web Knowledge Bases
Han Ren, Donghong Ji, Jing Wan and Chong Teng 484

Incorporating Statistical Information of Lexical Dependency into a Rule-Based Parser
Yoon-Hyung Roh, Ki-Young Lee and Young-Gil Kim 493

An Automated Thematic Role Labeler and Generalizer for Filipino Verb Arguments
Briane Paul Samson, Bianca Pamela Alcera, Ed Oswald Go, Czarina Meg Gonzales and Nathalie Rose Lim 501

Hybrid N-gram Probability Estimation in Morphologically Rich Languages
Hyopil Shin and Hyunjo You 511

Summarization Approaches Based on Document Probability Distributions
Sandeep Sripada and Jagadeesh Jagarlamudi 521

Chinese Function Tag Labeling
Weiwei Sun and Zhifang Sui 530

What L2 Processing Strategy Reveals about the Prototypical Relationship: A Case of Japanese Modal Markers of Possibility
Mizuho Tamaji 540

HMM Expanded to Multiple Interleaved Chains as a Model for Word Sense Disambiguation
Denis Turdakov and Dmitry Lizorkin 549

Passage Retrieval Using Answer Type Profiles in Question Answering
Surya Ganesh Veeravalli and Vasudeva Varma 559

A Continuum-Based Approach for Tightness Analysis of Chinese Semantic Units
Ying Xu, Christoph Ringlstetter and Randy Goebel 569

A Framework for Effectively Integrating Hard and Soft Syntactic Rules into Phrase Based Translation
Jiajun Zhang and Chengqing Zong 579

A Bootstrapping Method for Finer-Grained Opinion Mining Using Graph Model
Shu Zhang, Yingju Xia, Yao Meng and Hao Yu 589

SESS: A Self-Supervised and Syntax-Based Method for Sentiment Classification
Weishi Zhang, Kai Zhao, Likun Qiu and Changjian Hu 596

Poster Papers

Summarizing Opinions in Blog Threads
Alexandra Balahur, Mijail Kabadjov, Josef Steinberger, Ralf Steinberger and Andrés Montoyo 606

Constraint Based Hybrid Approach to Parsing Indian Languages
Akshar Bharati, Samar Husain, Meher Vijay, Kalyan Deepak, Dipti Misra Sharma and Rajeev Sangal 614

Interpolated PLSI for Learning Plausible Verb Arguments
Hiram Calvo, Kentaro Inui and Yuji Matsumoto 622

An EM Algorithm for Context-Based Searching and Disambiguation with Application to Synonym Term Alignment
Jing-Shin Chang and Shih-Jay Chiou 630

VocabAnalyzer: A Referred Word List Analyzing Tool with Keyword, Concordancing and N-gram Functions
Siaw-Fong Chung, F.Y. August Chao and Yi-Chen Hsieh 638

Building Online Corpora of Philippine Languages
Shirley N. Dita, Rachel Edita O. Roxas and Paul Inventado 646

Machine Translation Using Automatically Inferred Construction-Based Correspondence and Language Models
Shimon Edelman and Zach Solan 654

Latin Etymologies as Features on BNC Text Categorization
Alex Chengyu Fang, Wanyin Li and Nancy Ide 662

Experiments on Domain Adaptation for English–Hindi SMT
Rejwanul Haque, Sudip Kumar Naskar, Josef van Genabith and Andy Way 670

Extraction of English Ditransitive Constructions Using Formal Concept Analysis
Yoichiro Hasebe and Kow Kuroda 678

Bridging the Gap between Graph Modeling and Developmental Psycholinguistics: An Experiment on Measuring Lexical Proximity in Chinese Semantic Space
Shu-Kai Hsieh, Chun-Han Chang, Ivy Kuo, Hintat Cheung, Chu-Ren Huang and Bruno Gaume 686

Effective Use of Chinese Structural Auxiliaries for Chinese Parsing
Yun Jin, Qing Li, Yingshun Wu and Young-Gil Kim 694

Towards Conceptual Indexing of the Blogosphere through Wikipedia Topic Hierarchy
Mariko Kawaba, Daisuke Yokomoto, Hiroyuki Nakasaki, Takehito Utsuro and Tomohiro Fukuhara 702

Syntactic Category Prediction for Improving Translation Quality in English-Korean Machine Translation
Sung-Dong Kim 710

Customizing an English-Korean Machine Translation System for Patent/Technical Documents Translation
Oh-Woog Kwon, Sung-Kwon Choi, Ki-Young Lee, Yoon-Hyung Roh and Young-Gil Kim 718

Word Boundary Decision with CRF for Chinese Word Segmentation
Shoushan Li and Chu-Ren Huang 726

Extracting Keyphrases from Chinese News Articles Using TextRank and Query Log Knowledge
Weiming Liang, Chang-Ning Huang, Mu Li and Bao-Liang Lu 733

Modeling the Relationship among Linguistic Typological Features with Hierarchical Dirichlet Process
Chu-Cheng Lin, Yu-Chun Wang and Richard Tzong-Han Tsai 741

Document Re-ranking via Wikipedia Articles for Definition/Biography Type Questions
Maofu Liu, Fang Fang and Donghong Ji 748

Towards Bilingual Term Extraction in Comparable Patents
Bin Lu and Benjamin K. Tsou 755

Factors Affecting Part-of-Speech Tagging for Tagalog
Erlyn Manguilimotan and Yuji Matsumoto 763

GuideLink: A Corpus Annotation System that Integrates the Management of Annotation Guidelines
Kenta Oouchida, Jin-Dong Kim, Toshihisa Takagi and Jun'ichi Tsujii 771

Using Wikipedia for Hierarchical Finer Categorization of Named Entities
Aasish Pappu 779

A Novel Method of Sentence Ordering Based on Support Vector Machine
Gongfu Peng, Yanxiang He, Ye Tian, Yingsheng Tian and Weidong Wen 787

Supertagging with Factorial Hidden Markov Models
Srivatsan Ramanujam and Jason Baldridge 795

Identifying and Utilizing the Class of Monosemous Japanese Functional Expressions in Machine Translation
Akiko Sakamoto, Taiji Nagasaka, Takehito Utsuro and Suguru Matsuyoshi 803

Named Entity Recognition for Manipuri Using Support Vector Machine
Thoudam Doren Singh, Kishorjit Nongmeikapam, Asif Ekbal and Sivaji Bandyopadhyay 811

BaseNP Supersense Tagging for Japanese Texts
Hirotoshi Taira, Sen Yoshida and Masaaki Nagata 819

Which is More Suitable for Chinese Word Segmentation, the Generative Model or the Discriminative One?
Kun Wang, Chengqing Zong and Keh-Yih Su 827

Design of Chinese HPSG Framework for Data-Driven Parsing
Xiangli Wang, Shunya Iwasawa, Yusuke Miyao, Takuya Matsuzaki, Kun Yu and Jun'ichi Tsujii 835

Rule-based Korean Grapheme to Phoneme Conversion Using Sound Patterns
Yu-Chun Wang and Richard Tzong-Han Tsai 843

Discovery of Dependency Tree Patterns for Relation Extraction
Hongzhi Xu, Changjian Hu and Guoyang Shen 851

Generalizable Features Help Semantic Role Labeling
Li Yang 859

Utilizing Features of Verbs in Statistical Zero Pronoun Resolution for Japanese Speech
Sen Yoshida and Masaaki Nagata 867

Towards Establishing a Hierarchy in the Japanese Sentence Structure
Kei Yoshimoto, Chidori Nakamura and Alastair Butler 875

Author Index 883