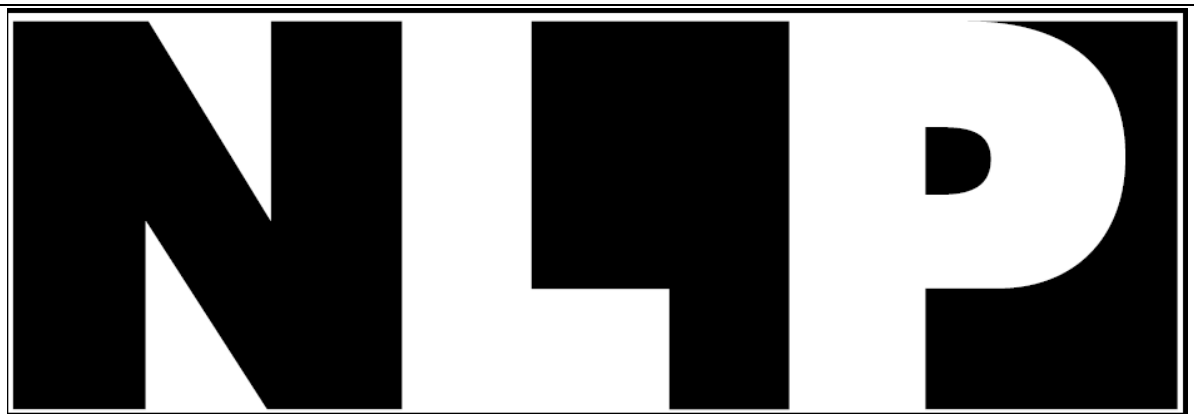
Sixth International Joint Conference on
Natural Language Processing

Proceedings of the Main Conference

We wish to thank our sponsors and supporters!

Platinum Sponsors



www.anlp.jp

Silver Sponsors

www.google.com

Bronze Sponsors

www.rakuten.com

Supporters

NAGOYA CONVENTION & VISITORS BUREAU

Nagoya Convention & Visitors Bureau

We wish to thank our organizers!

Organizers

Asian Federation of Natural Language Processing (AFNLP)

Toyohashi University of Technology

©2013 Asian Federation of Natural Language Processing

ISBN 978-4-9907348-0-0

FOREWORD

Welcome to the 6th International Joint Conference on Natural Language Processing (IJCNLP) in Nagoya, Japan. IJCNLP was initiated in 2004 by The Asian Federation of Natural Language Processing (AFNLP) with the major goal to provide a platform for researchers and professionals from around the world to share their experiences related to natural language processing and computational linguistics. In the past years, IJCNLPs were held in 5 different places: Hainan Island (2004), Jeju Island (2005), Hyderabad (2008), Singapore (2009), and Chiang Mai (2011). This year the 6th IJCNLP is held in Nagoya Congress Center on October 14-18, 2013. The conference covers a broad spectrum of technical areas related to natural language and computation. Besides main conference, the program includes 2 keynote speeches, 3 tutorials, 12 demonstrations, and 7 workshops.

We would like to thank many people who contribute their efforts to IJCNLP 2013. Program chairs Ruslan Mitkov and Jong C. Park select a strong set of papers and organize a wonderful program. PC coordinators Jin-Woo Chung and Isabel Duran support authors and PC committee a stable submission and review platform. Workshop chairs Naoaki Okazaki and Scott Wen-tau Yih organize 7 nice pre-conference and post-conference workshops. Tutorial chairs Vincent Ng and Satoshi Sekine choose 3 very good tutorials. Demo chairs Hang Li and Kentaro Torisawa recommend 12 demonstrations. Sponsorship chair Hiromi Nakaiwa designs sponsor packages and finds financial supports. We thank all the sponsors. Publicity chairs Gareth Jones, Gary Geunbae Lee, Diego Mollá-Aliod, Chengqing Zong and Stajner Sanja help circulate the conference information and promote the conference. We would like to express our special thanks to publication chairs Jing Jiang and Lun-Wei Ku. They bore two babies during the organization of the conference. After the hard work, they deliver an excellent proceeding to the participants. Finally, we are very thankful to those people who dedicate their time and energy to IJCNLP 2013, but are not mentioned in the above. Without them, we would not have had a successful conference.

Hsin-Hsi Chen, General Chair, National Taiwan University, Taiwan
Hitoshi Isahara, Organization Chair, Toyohashi University of Technology, Japan

October 15, 2013

PREFACE

As the flagship conference of the Asian Federation of Natural Language Processing (AFNLP), IJCNLP continues to establish itself as a highly influential international event. IJCNLP 2013 covers a broad spectrum of technical areas related to Natural Language Processing. The conference includes regular papers, short papers, poster papers, and system demonstrations, as well as pre- and post-conference tutorials and workshops.

This year, we received 363 paper submissions, which is not as many as the record submissions in the previous conference (e.g., 478 in 2011) but still quite admirable, considering the tough competition for good papers this year, with EMNLP 2013 and RANLP 2013, and the fact that many organizations have restricted their spending in the current economic climate. This represents increasing interest in research on NLP and the growing reputation of IJCNLP as an international event. The 363 submissions include 235 regular, 86 short, and 42 poster paper submissions from more than 37 countries. In particular, approximately 63% of the papers are from 14 countries and areas in Asia Pacific, 18% from 14 countries in Europe, 14% from the United States and Canada; in addition, 4% of the papers are from the Middle East and Africa, and 1% come from South America.

We would like to thank all the authors for submitting papers to IJCNLP 2013. The significant increase in the number of submissions, the topics covered and the wide range of demographic areas represent a rapid and steady growth of our field and hold promise for a bright future. We would also like to thank the 23 area chairs and 439 program committee members for writing over 1078 reviews and meta-reviews and for paving the way for the final paper selection. Of all 363 submissions, a total of 88 papers were accepted as regular papers, representing a healthy 24.4% acceptance rate. Additional 56 papers were accepted as short papers, which, together with regular papers, represent a 39.8% acceptance rate. In addition, 74 papers were accepted as poster papers. Due to various reasons, some authors of accepted papers chose to withdraw their submissions afterwards. As a result, we have 85 regular papers (23.4% acceptance rate), 53 short papers (38.0% acceptance rate), and 62 poster papers. All the regular and short papers are presented orally, and all the poster papers are presented in the plenary poster session. We are extremely grateful to the area chairs and program committee members for all their hard work, without which the preparation of this program would not have been possible. The help of PC coordinators is also much appreciated.

We are delighted to have two keynote speakers addressing different aspects of NLP in IJCNLP 2013. Hwee Tou Ng will present a talk about improving students' writing with automated grammatical error correction, including the review of recent research and advances in grammatical error correction. Roberto Navigli will present a talk about BabelNet 2.0, a very large multilingual semantic network that covers 50 languages and provides both lexicographic and encyclopedic knowledge for all the open-class parts of speech. These plenary talks will surely be not only informative but also enlightening to the audience, leading to many innovative research ideas. We would like to thank General Chair Hsin-Hsi Chen, the Local Arrangements Committee headed by Hitoshi Isahara, and the AFNLP Conference Coordination Committee chaired by Yuji Matsumoto, for their help and advice. Thanks to Jing Jiang and Lun-Wei Ku, the Publication Committee Chairs, for putting the proceedings together, and all the other committee chairs for their great work.

We hope that you enjoy the conference!

Ruslan Mitkov, University of Wolverhampton, England, United Kingdom
Jong C. Park, Korea Advanced Institute of Science and Technology, Republic of Korea
IJCNLP 2013 Program Committee Chairs

October 15, 2013

General Chair

Hsin-Hsi Chen, National Taiwan University, Taiwan

Program Committee Chairs

Ruslan Mitkov, University of Wolverhampton, UK
Jong C. Park, KAIST, Korea

Local Organization Committee Chair

Hitoshi Isahara, Toyohashi University of Technology, Japan

Workshop Committee Chairs

Naoaki Okazaki, Tohoku University, Japan
Scott Wen-tau Yih, Microsoft Research, USA

Tutorial Chairs

Vincent Ng, The University of Texas at Dallas, USA
Satoshi Sekine, New York University, USA

Demo Chairs

Hang Li, Huawei Technologies Co., China
Kentaro Torisawa, NICT, Japan

Sponsorship Committee Chair

Hiromi Nakaiwa, NTT, Japan

Publication Committee Chairs

Jing Jiang, Singapore Management University, Singapore
Lun-Wei Ku, Academia Sinica, Taiwan

Finance Committee Chairs

Masayuki Okabe, Toyohashi University of Technology, Japan
Masatoshi Tsuchiya, Toyohashi University of Technology, Japan

Publicity Committee Chairs

Gareth Jones, Dublin City University, Ireland
Gary Geunbae Lee, POSTECH, Korea
Diego Mollá-Aliod, Macquarie University, Australia
Chengqing Zong, Chinese Academy of Sciences, China

PC Coodinators

Jin-Woo Chung, KAIST, Korea
Isabel Duran, University of Wolverhampton, UK

Area Chairs

Phonology and Morphology
Mans Hulden, University of Arizona, USA

Syntax and Semantics
Mary Dalrymple, University of Oxford, UK

Pragmatics and Discourse
Joey Frazee, University of Texas, USA

Dialogue and Dialogue Systems
Gary Geunbae Lee, POSTECH, South Korea

Language Resources
Key-Sun Choi, KAIST, South Korea
Doaa Samy, Cairo University, Egypt

Statistical and ML Language Models
Fumiyo Fukumoto, University of Yamanashi, Japan
Leonor Becerra Bonache, Universitat Rovira i Virgili, Spain

POS Tagging and Parsing
Sandra Kuebler, Indiana University, USA
Yusuke Miyao, National Institute of Informatics, Japan

Semantic Processing
Alessandro Moschitti, University of Trento, Italy
Idan Szpektor, Yahoo! Research

Information Extraction
Nigel Collier, National Institute of Informatics, Japan
Jin-Dong Kim, Database Center for Life Science, Japan

Text Summarisation
Inderjeet Mani, Yahoo! Labs, Sunnyvale, USA
Helen Meng, Chinese University of Hong Kong, Hong Kong

Information Retrieval and QA
Qiaozhu Mei, University of Michigan, USA
Iustin Dornescu University of Wolverhampton, UK

Text Mining
Wai Lam, Chinese University of Hong Kong, Hong Kong

Opinion Mining
Alfonso Urena, University of Jaen, Spain

NLP for Educational Applications
Jin-Dong Kim, Database Center for Life Science, Japan
Nigel Collier, National Institute of Informatics, Japan

Recent NLP Applications
Constantin Orasan, University of Wolverhampton, UK

Machine Translation
Dekai Wu, The Hong Kong University of Science and Technology, Hong Kong
Young-suk Lee, IBM, USA

Reviewers

Muhammad Abdul-Mageed, Amjad Abu-Jbara, Hani AbuSalem, Karteek Addanki, Khurshid Ahmad, Akiko Aizawa, Ahmet Aker, Iñaki Alegria, Hadi Amiri, Dana Angluin, Kenji Araki, Eiji Aramaki, Ash Asudeh, Sören Auer, Nguyen Bach, Amit Bagga, Alexandra Balahur, Rafael E. Banchs, Yang Bao, Roberto Basili, Cosmin Bejan, Gemma Bel Enguix, Núria Bel, Lidong Bing, Phil Blunsom, Igor Boguslavsky, Bernd Bohnet, Ilaria Bordino, Stefano Borgo, Johan Bos, Houda Bouamor, Chris Brew, Christopher Brewster, Christopher Brown, Miriam Butt, Aoife Cahill, Lynne Cahill, Nicoletta Calzolari, Erik Cambria, Burcu Can, Marie Candito, Marine Carpuat, Fabio Celli, Daniel Cer, Hutchatai Chanlekha, Wanxiang Che, Berlin Chen, Boxing Chen, Wenliang Chen, Ying Chen, Pu-Jen Cheng, Colin Cherry, David Chiang, Han-Cheol Cho, Jinho D. Choi, Miranda Chong, Khalid Choukri, Janara Christensen, Grzegorz Chrupała, Kevin Cohen, Trevor Cohn, Nigel Collier, Kevyn Collins-Thompson, Gao Cong, John Conroy, Mike Conway, Justin Cope, Bonaventura Coppola, Anna Corazza, Josep Maria Crego, Dan Cristea, Danilo Croce, Andras Csomai, Iria da Cunha, Kareem Darwish, Amitava Das, Adrià de Gispert, Gerard de Melo, Marie-Catherine de Marneffe, Thierry Declerck, Adrian Horia Dediu, Rodolfo Delmonte, Dina Demner-Fushman, Barbara Di Eugenio, Giorgio Maria Di Nunzio, Mona Diab, Markus Dickinson, Marco Dinarelli, Liviu Dinu, Quang Do, Son Doan, Isabel Durán-Muñoz, Ismail El Maarouf, Michael Elhadad, Ossama Emam, Tomaž Erjavec, Gülşen Eryiğit, Maxine Eskenasi, Andrea Esuli, Richard Evans, James Fan, Richárd Farkas, Óscar Ferrández, Corina Forascu, Jennifer Foster, Wei Gao, Claire Gardent, Anna Lisa Gentile, Matthew Gerber, Alec Go, Koldo Gojenola Galletebeitia, José Miguel Goñi-Menoyo, Edward Grefenstette, Jiafeng Guo, Iryna Gurevych, Francisco Guzman, Amaury Habrard, Keith Hall, Olivier Hamon, Dan Han, Xianpei Han, Laura Hasler, Hany Hassan, Jiyin He, Yulan He, Jeffrey Heinz, James Henderson, Sanjika Hewavitharana, Tsutomu Hirao, Hieu Hoang, Matthew Honnibal, Veronique Hoste, Dirk Hovy, Fei Huang, Zhongqiang Huang, Young-Sook Hwang, Nancy Ide, Adrian Iftene, Radu Ion, Hitoshi Isahara, Guillaume Jacquet, Minwoo Jeong, Sittichai Jiampojamarn, Wenbin Jiang, Yunliang Jiang, Maria Dolores Jimenez Lopez, Richard Johansson, Kristiina Jokinen, Senay Kafkas, Min-Yen Kan, Nattiya Kanhabua, Ron Kaplan, Makoto Kato, Mitesh M. Khapra, Adam Kilgarriff, Byeongchang Kim, Hong Kook Kim, Jee-Hyub Kim, Jung-Jae Kim, Seokhwan Kim, Soo-Min Kim, Sungchul Kim, Sun Kim, Woosung Kim, Tracy Holloway King, Philipp Koehn, Oskar Kohonen, Oleksandr Kolomiyets, Mamoru Komachi, Natalia Konstantinova, Hahn Koo, Valia Kordoni, Ioannis Korkontzelos, Alexander Kotov, Zornitsa Kozareva, Udo Kruschwitz, Jonas Kuhn, Sadao Kurohashi, Sobha Lalitha Devi, Wai Lam, Guy Lapalme, Jey Han Lau, Raymond Lau, Alberto Lavelli, Ho-Joon Lee, Jae Sung Lee, Sungjin Lee, Young-Suk Lee, Alessandro Lenci, Yves Lepage, Gina-Anne Levow, Si Li, Maria Liakata, Chin-Yew Lin, Christina Lioma, Bing Liu, Haibin Liu, Kang Liu, Qiaoling Liu, Qun

Liu, Yang Liu, Yiqun Liu, Zhiyuan Liu, Avishay Livne, Elena Lloret, Chi-kiu Lo, Christoph Lofi, Oier Lopez de Lacalle, Ramon Lopez-Cozar, Bin Lu, Yue Lu, Zhiyong Lu, Giorgio Magri, Wolfgang Maier, Suresh Manandhar, Inderjeet Mani, Maite Martin, David Martinez, Paloma Martínez, Patricio Martinez-Barco, Eugenio Martínez-Cámara, Yuji Matsumoto, Yutaka Matsuo, Suguru Matsuyoshi, Takuya Matsuzaki, Evgeny Matusov, Arne Mauser, Mark Maybury, Diana McCarthy, John Philip McCrae, Beata Megyesi, Yashar Mehdad, Edgar Meij, Wolfgang Menzel, Farid Meziane, Haitao Mi, Bonan Min, Hye-Jin Min, Wolfgang Minker, Shachar Mirkin, Teruhisa Misu, Makoto Miwa, Yusuke Miyao, Manuel Montes, Alessandro Moschitti, Arjun Mukherjee, Rafael Muñoz-Guillena, Louise Mycock, Satoshi Nakamura, Preslav Nakov, Jason Naradowsky, Vivi Nastase, Costanza Navarretta, Roberto Navigli, Jun Ping Ng, Vincent Ng, ThuyLinh Nguyen, Truc-Vien T. Nguyen, Hitoshi Nishikawa, Joakim Nivre, Josef Novak, Kemal Oflazer, Jong-Hoon Oh, Kousaku Okubo, Petya Osenova, Alexander Osherenko, Shiyan Ou, Arzucan Özgür, Georgios Paltoglou, Rebecca J. Passonneau, Jong Park, Jungyeul Park, Siddharth Patwardhan, Michael Paul, Ted Pedersen, Viktor Pekar, Jose Manuel Perea-Ortega, Wim Peters, Michael Piotrowski, Emily Pitler, Barbara Plank, Natalia Ponomareva, Simone Paolo Ponzetto, Maja Popović, Christopher Potts, Sameer Pradhan, John Prager, Sampo Pyysalo, Vahed Qazvinian, Yao Qian, Tao Qin, Sriram Raghavan, Ganesh Ramakrishnan, Owen Rambow, Ines Rehbein, Luz Rello, German Rigau, Fabio Rinaldi, Horacio Rodriguez, Laurent Romary, Paolo Rosso, Bogdan Sacaleanu, Markus Saers, Kenji Sagae, Horacio Saggion, Benoît Sagot, Patrick Saint-Dizier, Agnes Sandor, Marina Santini, Diana Santos, Anoop Sarkar, Kenji Satou, Helmut Schmid, Lane Schwartz, Djamé Seddah, Kazuhiro Seki, Satoshi Sekine, Hendra Setiawan, Aliaksei Severyn, Khaled Shaalan, Libin Shen, Wade Shen, Hiroyuki Shindo, Hiroyuki Shinnou, Eyal Shnarch, Khalil Sima'an, Kiril Simov, Anders Søgaard, Edward Stabler, Sanja Štajner, Efstathios Stamatatos, Pontus Stenetorp, Matthew Stone, Michael Strube, Jian Su, Keh-Yih Su, Kazunari Sugiyama, Ang Sun, Yizhou Sun, Jun Suzuki, Yoshimi Suzuki, Stan Szpakowicz, Hiroya Takamura, Chenhao Tan, Liling Tan, Yee Fan Tan, Ivan Titov, Noriko Tomuro, Sara Tonelli, Lamia Tounsi, Harald Trost, Jose A. Troyano, Manos Tsagkias, Richard Tzong-Han Tsai, Yuen-Hsien Tseng, Yoshimasa Tsuruoka, Kateryna Tymoshenko, Mike Unwalla, Olga Uryupina, Takehito Utsuro, Kees van Deemter, Josef van Genabith, Gertjan van Noord, Menno van Zaanen, Adriano Veloso, Yannick Versley, Karin Verspoor, Jose Vicedo, Laure Vieu, Clare Voss, Vinod Vydiswaran, Henning Wachsmuth, Hui Wan, Xiaojun Wan, Haifeng Wang, Hongning Wang, Hsin-Min Wang, Lidan Wang, Yue Wang, Yu Wang, Houfeng Wang, Leo Wanner, Taro Watanabe, Bonnie Webber, William Webber, Deyi Xiong, Peng Xu, Yasunori Yamamoto, Anssi Yli-Jyrä, Naoki Yoshinaga, Heng Yu, Bei Yu, Fabio Massimo Zanzotto, Duo Zhang, Joy Ying Zhang, Lanbo Zhang, Min Zhang, Qi Zhang, Yi Zhang, Yue Zhang, Ziqi Zhang, Bing Zhao, Le Zhao, Tiejun Zhao, Desislava Zhekova, Jing Zheng, Guodong Zhou, Xiaodan Zhu, Heike Zinsmeister, Michael Zock

Table of Contents

Semi-Supervised Answer Extraction from Discussion Forums
Rose Catherine, Rashmi Gangadharaiah, Karthik Visweswariah and Dinesh Raghu 1

WordTopic-MultiRank: A New Method for Automatic Keyphrase Extraction
Fan Zhang, Lian'en Huang and Bo Peng 10

Towards Contextual Healthiness Classification of Food Items - A Linguistic Approach
Michael Wiegand and Dietrich Klakow 19

Learning a Replacement Model for Query Segmentation with Consistency in Search Logs
Wei Zhang, Yunbo Cao, Chin-Yew Lin, Jian Su and Chew-Lim Tan 28

Precise Information Retrieval Exploiting Predicate-Argument Structures
Daisuke Kawahara, Keiji Shinzato, Tomohide Shibata and Sadao Kurohashi 37

Global Model for Hierarchical Multi-Label Text Classification
Yugo Murawaki 46

(Pre-)Annotation of Topic-Focus Articulation in Prague Czech-English Dependency Treebank
Jiří Mírovský, Kateřina Rysová, Magdaléna Rysová and Eva Hajičová 55

Animacy Acquisition Using Morphological Case
Riyaz Ahmad Bhat and Dipti Misra Sharma 64

The Complexity of Math Problems – Linguistic, or Computational?
Takuya Matsuzaki, Hidenao Iwane, Hirokazu Anai and Noriko Arai 73

Hybrid Models for Lexical Acquisition of Correlated Styles
Julian Brooke and Graeme Hirst 82

Introducing the Prague Discourse Treebank 1.0
Lucie Poláková, Jiří Mírovský, Anna Nedoluzhko, Pavlína Jínová, Šárka Zikánová and Eva Hajičová 91

Multilingual Mention Detection for Coreference Resolution
Olga Uryupina and Alessandro Moschitti 100

A Weakly Supervised Bayesian Model for Violence Detection in Social Media
Amparo Elizabeth Cano Basave, Yulan He, Kang Liu and Jun Zhao 109

Detecting Spammers in Community Question Answering
Zhuoye Ding, Yeyun Gong, Yaqian Zhou, Qi Zhang and Xuanjing Huang 118

Chinese Informal Word Normalization: an Experimental Study
Aobo Wang, Min-Yen Kan, Daniel Andrade, Takashi Onishi and Kai Ishikawa 127

Feature Selection Using a Semantic Hierarchy for Event Recognition and Type Classification
Yoonjae Jeong and Sung-Hyon Myaeng 136

Romanization-based Approach to Morphological Analysis in Korean SMS Text Processing
Youngsam Kim and Hyopil Shin 145

Efficient Word Lattice Generation for Joint Word Segmentation and POS Tagging in Japanese
Nobuhiro Kaji and Masaru Kitsuregawa 153

A Simple Approach to Unknown Word Processing in Japanese Morphological Analysis
Ryohei Sasano, Sadao Kurohashi and Manabu Okumura 162

Chinese Word Segmentation by Mining Maximized Substrings
Mo Shen, Daisuke Kawahara and Sadao Kurohashi 171

Capturing Long-distance Dependencies in Sequence Models: A Case Study of Chinese Part-of-speech Tagging
Weiwei Sun, Xiaochang Peng and Xiaojun Wan 180

Exploring Semantic Information in Hindi WordNet for Hindi Dependency Parsing
Sambhav Jain, Naman Jain, Aniruddha Tammewar, Riyaz Ahmad Bhat and Dipti Sharma 189

Towards Robust Cross-Domain Domain Adaptation for Part-of-Speech Tagging
Tobias Schnabel and Hinrich Schütze 198

Dependency Parsing for Identifying Hungarian Light Verb Constructions
Veronika Vincze, János Zsibrita and István Nagy T. 207

Written Dialog and Social Power: Manifestations of Different Types of Power in Dialog Behavior
Vinodkumar Prabhakaran and Owen Rambow 216

Evaluation of the Scusi? Spoken Language Interpretation System – A Case Study
Thomas Kleinbauer, Ingrid Zukerman and Su Nam Kim 225

A Noisy Channel Approach to Error Correction in Spoken Referring Expressions
Su Nam Kim, Ingrid Zukerman, Thomas Kleinbauer and Farshid Zavareh 234

Natural Language Query Refinement for Problem Resolution from Crowd-Sourced Semi-Structured Data
Rashmi Gangadharaiah and Balakrishnan Narayanaswamy 243

Ensemble Triangulation for Statistical Machine Translation
Majid Razmara and Anoop Sarkar 252

Robust Transliteration Mining from Comparable Corpora with Bilingual Topic Models
John Richardson, Toshiaki Nakazawa and Sadao Kurohashi 261

SuMT: A Framework of Summarization and MT
Houda Bouamor, Behrang Mohit and Kemal Oflazer 270

Tuning SMT with a Large Number of Features via Online Feature Grouping
Lemao Liu, Tiejun Zhao, Taro Watanabe and Eiichiro Sumita 279

Multimodal Comparable Corpora as Resources for Extracting Parallel Data: Parallel Phrases Extraction
Haithem Afli, Loïc Barrault and Holger Schwenk 286

Bootstrapping Large-scale Named Entities using URL-Text Hybrid Patterns
Chao Zhang, Shiqi Zhao and Haifeng Wang 293

Feature-Rich Segment-Based News Event Detection on Twitter
Yanxia Qin, Yue Zhang, Min Zhang and Dequan Zheng 302

Building Chinese Event Type Paradigm Based on Trigger Clustering
Xiao Ding, Bing Qin and Ting Liu 311

Chinese Named Entity Abbreviation Generation Using First-Order Logic
Huan Chen, Qi Zhang, Jin Qian and Xuanjing Huang 320

Full-coverage Identification of English Light Verb Constructions
István Nagy T., Veronika Vincze and Richárd Farkas 329

Detecting Deceptive Opinions with Profile Compatibility
Vanessa Wei Feng and Graeme Hirst 338

Behind the Times: Detecting Epoch Changes using Large Corpora
Octavian Popescu and Carlo Strapparava 347

How Noisy Social Media Text, How Diffrnt Social Media Sources?
Timothy Baldwin, Paul Cook, Marco Lui, Andrew MacKinlay and Li Wang 356

Who Had the Upper Hand? Ranking Participants of Interactions Based on Their Relative Power
Vinodkumar Prabhakaran, Ajita John and Dorée D. Seligmann 365

Readability Indices for Automatic Evaluation of Text Simplification Systems: A Feasibility Study for Spanish
Sanja Štajner and Horacio Saggion 374

Weasels, Hedges and Peacocks: Discourse-level Uncertainty in Wikipedia Articles
Veronika Vincze 383

Automatically Developing a Fine-grained Arabic Named Entity Corpus and Gazetteer by utilizing Wikipedia
Fahd Alotaibi and Mark Lee 392

Ranking Translation Candidates Acquired from Comparable Corpora
Rima Harastani, Béatrice Daille and Emmanuel Morin 401

Using the Semantic-Syntactic Interface for Reliable Arabic Modality Annotation
Rania Al-Sabbagh, Jana Diesner and Roxana Girju 410

Mapping Rules for Building a Tunisian Dialect Lexicon and Generating Corpora
Rahma Boujelbane, Mariem Ellouze Khemekhem and Lamia Hadrich Belguith 419

Hypothesis Refinement Using Agreement Constraints in Machine Translation
Ankur Gandhe and Rashmi Gangadharaiah 429

Scalable Variational Inference for Extracting Hierarchical Phrase-based Translation Rules
Baskaran Sankaran, Gholamreza Haffari and Anoop Sarkar 438

A Topic-Triggered Language Model for Statistical Machine Translation
Heng Yu, Jinsong Su, Yajuan Lv and Qun Liu 447

Reserved Self-training: A Semi-supervised Sentiment Classification Method for Chinese Microblogs
Zhiguang Liu, Xishuang Dong, Yi Guan and Jinfeng Yang 455

Enhancing Lexicon-Based Review Classification by Merging and Revising Sentiment Dictionaries
Heeryon Cho, Jong-Seok Lee and Songkuk Kim 463

Exploring the Effects of Word Roots for Arabic Sentiment Analysis
Shereen Oraby, Yasser El-Sonbaty and Mohamad Abou El-Nasr 471

Topical Key Concept Extraction from Folksonomy
Han Xue, Bing Qin, Ting Liu and Chao Xiang 480

Uncovering Distributional Differences between Synonyms and Antonyms in a Word Space Model
Silke Scheible, Sabine Schulte im Walde and Sylvia Springorum 489

Multilingual Word Sense Disambiguation Using Wikipedia
Bharath Dandala, Rada Mihalcea and Razvan Bunescu 498

Semantic v.s. Positions: Utilizing Balanced Proximity in Language Model Smoothing for Information Retrieval
Rui Yan, Han Jiang, Mirella Lapata, Shou-De Lin, Xueqiang Lv and Xiaoming Li 507

An Unsupervised Parameter Estimation Algorithm for a Generative Dependency N-gram Language Model
Chenchen Ding and Mikio Yamamoto 516

Learning a Product of Experts with Elitist Lasso
Mengqiu Wang and Christopher D. Manning 525

Learning Efficient Information Extraction on Heterogeneous Texts
Henning Wachsmuth, Benno Stein and Gregor Engels 534

TopicRank: Graph-Based Topic Ranking for Keyphrase Extraction
Adrien Bougouin, Florian Boudin and Béatrice Daille 543

Understanding the Semantic Intent of Natural Language Query
Juan Xu, Qi Zhang and Xuanjing Huang 552

Sentiment Classification for Movie Reviews in Chinese Using Parsing-based Methods
Wen-Juan Hou and Chuang-Ping Chang 561

Sentiment Aggregation using ConceptNet Ontology
Subhabrata Mukherjee and Sachindra Joshi 570

Detecting Cyberbullying Entries on Informal School Websites Based on Category Relevance Maximization
Taisei Nitta, Fumito Masui, Michal Ptaszynski, Yasutomo Kimura, Rafal Rzepka and Kenji Araki 579

A Lexicon-based Investigation of Research Issues in Japanese Factuality Analysis
Kazuya Narita, Junta Mizuno and Kentaro Inui 587

A Hierarchical Semantics-Aware Distributional Similarity Scheme
Shuqi Sun, Ke Sun, Shiqi Zhao, Haifeng Wang, Muyun Yang and Sheng Li 596

Labeled Alignment for Recognizing Textual Entailment
Xiaolin Wang, Hai Zhao and Bao-Liang Lu 605

Context-Based Chinese Word Segmentation using SVM Machine-Learning Algorithm without Dictionary Support
Chia-ming Lee and Chien-Kang Huang 614

A Common Case of Jekyll and Hyde: The Synergistic Effect of Using Divided Source Training Data for Feature Augmentation
Yan Song and Fei Xia 623

Detecting Polysemy in Hard and Soft Cluster Analyses of German Preposition Vector Spaces
Sylvia Springorum, Sabine Schulte im Walde and Jason Utt 632

Generalized Abbreviation Prediction with Negative Full Forms and Its Application on Improving Chinese Web Search
Xu Sun, Wenjie Li, Fanqi Meng and Houfeng Wang 641

Prosody-Based Unsupervised Speech Summarization with Two-Layer Mutually Reinforced Random Walk
Sujay Kumar Jauhar, Yun-Nung Chen and Florian Metze 648

Mining the Gaps: Towards Polynomial Summarization
Marina Litvak and Natalia Vanetik 655

Detecting Domain Dedicated Polar Words
Raksha Sharma and Pushpak Bhattacharyya 661

Can I Hear You? Sentiment Analysis on Medical Forums
Tanveer Ali, David Schramm, Marina Sokolova and Diana Inkpen 667

Construction of Emotional Lexicon Using Potts Model
Braja Gopal Patra, Hiroya Takamura, Dipankar Das, Manabu Okumura and Sivaji Bandyopadhyay 674

Suicidal Tendencies: The Automatic Classification of Suicidal and Non-Suicidal Lyricists Using NLP
Matthew Mulholland and Joanne Quinn 680

Unsupervised Word Class Induction for Under-resourced Languages: A Case Study on Indonesian
Meladel Mistica, Jey Han Lau and Timothy Baldwin 685

An Efficient Active Learning Framework for New Relation Types
Lisheng Fu and Ralph Grishman 692

Parsing Dependency Paths to Identify Event-Argument Relations
Seung-Cheol Baek and Jong Park 699

Augmentable Paraphrase Extraction Framework
MeiHua Chen, YiChun Chen, ShihTing Huang and Jason S. Chang 706

Automatic Prediction of Evidence-based Recommendations via Sentence-level Polarity Classification
Abeed Sarker, Diego Mollá-Aliod and Cécile Paris 712

Clustering Microtext Streams for Event Identification
Jie Yin 719

Automatic Corpora Construction for Text Classification
Dandan Wang, Qingcai Chen, Xiaolong Wang and Bingyang Yu 726

Learning to Generate Diversified Query Interpretations using Biconvex Optimization
Ramakrishna Bairi, Ambha A and Ganesh Ramakrishnan 733

Learning Based Approaches for Vietnamese Question Classification Using Keywords Extraction from the Web
Dang Tran, Cuong Chu, Son Pham and Minh Nguyen 740

Detecting Bot-Answerable Questions in Ubuntu Chat
David Uthus and David Aha 747

Alignment-based Annotation of Proofreading Texts toward Professional Writing Assistance
Ngan Nguyen and Yusuke Miyao 753

Toward Automatic Processing of English Metalanguage
Shomir Wilson 760

On the Effectiveness of Using Syntactic and Shallow Semantic Tree Kernels for Automatic Assessment of Essays
Yllias Chali and Sadid A. Hasan 767

Little by Little: Semi Supervised Stemming through Stem Set Minimization
Vasudevan N and Pushpak Bhattacharyya 774

What Information is Helpful for Dependency Based Semantic Role Labeling
Yanyan Luo, Kevin Duh and Yuji Matsumoto 781

Classifying Taxonomic Relations between Pairs of Wikipedia Articles
Or Biran and Kathleen McKeown 788

A Rule System for Chinese Time Entity Recognition by Comprehensive Linguistic Study
Hongzhi Xu and Chu-Ren Huang 795

Financial Sentiment Analysis for Risk Prediction
Chuan-Ju Wang, Ming-Feng Tsai, Tse Liu and Chin-Ting Chang 802

Sense Disambiguation: From Natural Language Words to Mathematical Terms
Minh-Quoc Nghiem, Giovanni Yoko Kristianto, Goran Topic and Akiko Aizawa 809

Adapting a State-of-the-art Anaphora Resolution System for Resource-poor Language
Utpal Sikdar, Asif Ekbal, Sriparna Saha, Olga Uryupina and Massimo Poesio 815

Chinese Event Coreference Resolution: Understanding the State of the Art
Chen Chen and Vincent Ng 822

A Two-Step Named Entity Recognizer for Open-Domain Search Queries
Andreas Eiselt and Alejandro Figueroa 829

A Comparison of Centrality Measures for Graph-Based Keyphrase Extraction
Florian Boudin 834

Translating Chinese Unknown Words by Automatically Acquired Templates
Ming-Hong Bai, Yu-Ming Hsieh, Keh-Jiann Chen and Jason S. Chang 839

Multilingual Lexicon Bootstrapping - Improving a Lexicon Induction System Using a Parallel Corpus
Patrick Ziering, Lonneke van der Plas and Hinrich Schütze 844

Mining Japanese Compound Words and Their Pronunciations from Web Pages and Tweets
Xianchao Wu 849

A Factoid Question Answering System Using Answer Pattern Matching
Nagehan Pala Er and Ilyas Cicekli 854

Chinese Short Text Classification Based on Domain Knowledge
Xiao Feng, Yang Shen, Chengyong Liu, Wei Liang and Shuwu Zhang 859

Applying Graph-based Keyword Extraction to Document Retrieval
Youngsam Kim, Munhyong Kim, Andrew Cattle, Julia Otmakhova, Suzi Park and Hyopil Shin 864

Semi-supervised Classification of Twitter Messages for Organization Name Disambiguation
Shu Zhang, Jianwei Wu, Dequan Zheng, Yao Meng and Hao Yu 869

Word in a Dictionary is used by Numerous Users
Eiji Aramaki, Sachiko Maskawa, Mai Miyabe, Mizuki Morita and Sachi Yasuda 874

Extracting Evaluative Conditions from Online Reviews: Toward Enhancing Opinion Mining
Yuki Nakayama and Atsushi Fujii 878

Cognate Production using Character-based Machine Translation
Lisa Beinborn, Torsten Zesch and Iryna Gurevych 883

An Empirical Study of Combing Multiple Models in Bengali Question Classification
Somnath Banerjee and Sivaji Bandyopadhyay 892

A Two-Stage Classifier for Sentiment Analysis
Dai Quoc Nguyen, Dat Quoc Nguyen and Son Bao Pham 897

Exploiting User Search Sessions for the Semantic Categorization of Question-like Informational Search Queries
Alejandro Figueroa and Guenter Neumann 902

Influence of Part-of-Speech and Phrasal Category Universal Tag-set in Tree-to-Tree Translation Models
Francisco Oliveira, Derek F. Wong, Lidia S. Chao, Liang Tian and Liangye He 907

Interest Analysis using PageRank and Social Interaction Content
Chung-chi Huang and Lun-Wei Ku 912

Time Series Topic Modeling and Bursty Topic Detection of Correlated News and Twitter
Daichi Koike, Yusuke Takahashi, Takehito Utsuro, Masaharu Yoshioka and Noriko Kando 917

A Distant Supervision Approach for Identifying Perspectives in Unstructured User-Generated Text
Attapol Thamrongrattanarit, Colin Pollock, Benjamin Goldenberg and Jason Fennell 922

An Approach of Hybrid Hierarchical Structure for Word Similarity Computing by HowNet
Jiangming Liu, Jinan Xu and Yujie Zhang 927

Extracting Causes of Emotions from Text
Alena Neviarouskaya and Masaki Aono 932

Automated Grammar Correction Using Hierarchical Phrase-Based Statistical Machine Translation
Bibek Behera and Pushpak Bhattacharyya 937

Finding Dependency Parsing Limits over a Large Spanish Corpus
Muntsa Padró, Miguel Ballesteros, Héctor Martínez and Bernd Bohnet 942

High Quality Dependency Selection from Automatic Parses
Gongye Jin, Daisuke Kawahara and Sadao Kurohashi 947

Building Specialized Bilingual Lexicons Using Word Sense Disambiguation
Dhouha Bouamor, Nasredine Semmar and Pierre Zweigenbaum 952

Predicate Argument Structure Analysis using Partially Annotated Corpora
Koichiro Yoshino, Shinsuke Mori and Tatsuya Kawahara 957

Statistical Dialogue Management using Intention Dependency Graph
Koichiro Yoshino, Shinji Watanabe, Jonathan Le Roux and John R. Hershey 962

Repairing Incorrect Translation with Examples
Junguo Zhu, Muyun Yang, Sheng Li and Tiejun Zhao 967

Phrase-based Parallel Fragments Extraction from Comparable Corpora
Xiaoyin Fu, Wei Wei, Shixiang Lu, Zhenbiao Chen and Bo Xu 972

A Hybrid Approach for Anaphora Resolution in Hindi
Praveen Dakwale, Vandan Mujadia and Dipti M Sharma 977

Structure Cognizant Pseudo Relevance Feedback
Arjun Atreya V, Yogesh Kakde, Pushpak Bhattacharyya and Ganesh Ramakrishnan 982

Cross-Domain Answer Ranking using Importance Sampling
Anders Johannsen and Anders Søgaard 987

Morphological Analysis of Tunisian Dialect
Inès Zribi, Mariem Ellouze Khemakhem and Lamia Hadrich Belguith 992

Disambiguating Explicit Discourse Connectives without Oracles
Anders Johannsen and Anders Søgaard 997

Updating Rare Term Vector Replacement
Tobias Berka and Marian Vajteršic 1002

Statistical Morphological Analyzer for Hindi
Deepak Kumar Malladi and Prashanth Mannem 1007

Induction of Root and Pattern Lexicon for Unsupervised Morphological Analysis of Arabic
Bilal Khaliq and John Carrol 1012

Using Shallow Semantic Parsing and Relation Extraction for Finding Contradiction in Text
Minh Quang Nhat Pham, Minh Le Nguyen and Akira Shimazu 1017

Using Transliteration of Proper Names from Arabic to Latin Script to Improve English-Arabic Word Alignment
Nasredine Semmar and Houda Saadane 1022

A Semi-Supervised Method for Arabic Word Sense Disambiguation Using a Weighted Directed Graph
Laroussi Merhbene, Anis Zouaghi and Mounir Zrigui 1027

Incremental Segmentation and Decoding Strategies for Simultaneous Translation
Mahsa Yarmohammadi, Vivek Kumar Rangarajan Sridhar, Srinivas Bangalore and Baskaran Sankaran 1032

Two Case Studies on Translating Pronouns in a Deep Syntax Framework
Michal Novák, Zdenek Zabokrtsky and Anna Nedoluzhko 1037

Bootstrapping Phrase-based Statistical Machine Translation via WSD Integration
Hien Vu Huy, Phuong-Thai Nguyen, Tung-Lam Nguyen and M.L Nguyen 1042

Orthographic and Morphological Processing for Persian-to-English Statistical Machine Translation
Mohammad Sadegh Rasooli, Ahmed El Kholy and Nizar Habash 1047

Interoperability between Service Composition and Processing Pipeline: Case Study on the Language Grid and UIMA
Trang Mai Xuan, Yohei Murakami, Donghui Lin and Toru Ishida 1052

Improving Calculation of Contextual Similarity for Constructing a Bilingual Dictionary via a Third Language
Takashi Tsunakawa, Yosuke Yamamoto and Hiroyuki Kaji 1057

Two-Stage Pre-ordering for Japanese-to-English Statistical Machine Translation
Sho Hoshino, Yusuke Miyao, Katsuhito Sudoh and Masaaki Nagata 1062

Grammatical Error Correction Using Feature Selection and Confidence Tuning
Yang Xiang, Yaoyun Zhang, Xiaolong Wang, Chongqiang Wei, Wen Zheng, Xiaoqiang Zhou, Yuxiu Hu and Yang Qin 1067

An Online Algorithm for Learning over Constrained Latent Representations using Multiple Views
Ann Clifton, Max Whitney and Anoop Sarkar 1072

Synonym Acquisition Using Bilingual Comparable Corpora
Daniel Andrade, Masaaki Tsuchida, Takashi Onishi and Kai Ishikawa 1077

Exploring Verb Frames for Sentence Simplification in Hindi
Ankush Soni, Sambhav Jain and Dipti Misra Sharma 1082

Dirichlet Processes for Joint Learning of Morphology and PoS Tags
Burcu Can and Suresh Manandhar 1087

Parser Accuracy in Quality Estimation of Machine Translation: A Tree Kernel Approach
Rasoul Samad Zadeh Kaljahi, Jennifer Foster, Raphael Rubino, Johann Roturier and Fred Hollowood 1092

Attribute Relation Extraction from Template-inconsistent Semi-structured Text by Leveraging Site-level Knowledge
Yang Liu, Fang Liu, Siwei Lai, Kang Liu, Guangyou Zhou and Jun Zhao 1097

Optimum Parameter Selection for K.L.D. Based Authorship Attribution in Gujarati
Parth Mehta and Prasenjit Majumder 1102

Modeling User Leniency and Product Popularity for Sentiment Classification
Wenliang Gao, Naoki Yoshinaga, Nobuhiro Kaji and Masaru Kitsuregawa 1107

A Generalized LCS Algorithm and Its Application to Corpus Alignment
Jin-Dong Kim 1112

Semantic Naïve Bayes Classifier for Document Classification
How Jing, Yu Tsao, Kuan-Yu Chen and Hsin-Min Wang 1117

Cluster-based Web Summarization
Yves Petinot, Kathleen McKeown and Kapil Thadani 1124

Automated Activity Recognition in Clinical Documents
Camilo Thorne, Marco Montali, Diego Calvanese, Elena Cardillo and Claudio Eccher 1129

Large-Scale Text Collection for Unwritten Languages
Florian R. Hanke and Steven Bird 1134

A Self-learning Template Approach for Recognizing Named Entities from Web Text
Qian Liu, Bingyang Liu, Dayong Wu, Yue Liu and Xueqi Cheng 1139

Accurate Parallel Fragment Extraction from Quasi–Comparable Corpora using Alignment Model and Translation Lexicon
Chenhui Chu, Toshiaki Nakazawa and Sadao Kurohashi 1144

Meta-level Statistical Machine Translation
Sajad Ebrahimi, Kourosh Meshgi, Shahram Khadivi and Mohammad Ebrahim Shiri Ahmad Abady 1151

Bayesian Induction of Bracketing Inversion Transduction Grammars
Markus Saers and Dekai Wu 1158

Estimating the Quality of Translated User-Generated Content
Raphael Rubino, Jennifer Foster, Rasoul Samad Zadeh Kaljahi, Johann Roturier and Fred Hollowood 1167

Selective Combination of Pivot and Direct Statistical Machine Translation Models
Ahmed El Kholy, Nizar Habash, Gregor Leusch, Evgeny Matusov and Hassan Sawaf 1174

Multiword Expressions in the Context of Statistical Machine Translation
Mahmoud Ghoneim and Mona Diab 1181

Uncertainty Detection for Natural Language Watermarking
György Szarvas and Iryna Gurevych 1188

KySS 1.0: a Framework for Automatic Evaluation of Chinese Input Method Engines
Zhongye Jia and Hai Zhao 1195

Automatic Extraction of Social Networks from Literary Text: A Case Study on Alice in Wonderland
Apoorv Agarwal, Anup Kotalwar and Owen Rambow 1202

Using the Web to Train a Mobile Device Oriented Japanese Input Method Editor
Xianchao Wu, Rixin Xiao and Xiaoxin Chen 1209

A Novel Approach Towards Incorporating Context Processing Capabilities in NLIDB System
Arjun Akula, Rajeev Sangal and Radhika Mamidi 1216

Iterative Development and Evaluation of a Social Conversational Agent
Annika Silvervarg and Arne Jönsson 1223

A Hybrid Morphological Disambiguation System for Turkish
Mucahid Kutlu and Ilyas Cicekli 1230

A Dynamic Confusion Score for Dependency Arc Labels
Sambhav Jain and Bhasha Agrawal 1237

Increasing the Quality and Quantity of Source Language Data for Unsupervised Cross-Lingual POS Tagging
Long Duong, Paul Cook, Steven Bird and Pavel Pecina 1243

Towards the Annotation of Penn TreeBank with Information Structure
Bernd Bohnet, Alicia Burga and Leo Wanner 1250

Constituency and Dependency Relationship from a Tree Adjoining Grammar and Abstract Categorial Grammars Perspective
Aleksandre Maskharashvili and Sylvain Pogodalla 1257

Named Entity Extraction using Information Distance
Sangameshwar Patil, Sachin Pawar and Girish Palshikar 1264

Feature-based Neural Language Model and Chinese Word Segmentation
Mairgup Mansur, Wenzhe Pei and Baobao Chang 1271

Human-Computer Interactive Chinese Word Segmentation: An Adaptive Dirichlet Process Mixture Model Approach
Tongfei Chen, Xiaojun Zou, Weimeng Zhu and Junfeng Hu 1278

Effect of Non-linear Deep Architecture in Sequence Labeling
Mengqiu Wang and Christopher D. Manning 1285

Case Study of Model Adaptation: Transfer Learning and Online Learning
Kenji Imamura 1292

Source and Translation Classification using Most Frequent Words
Zahurul Islam and Armin Hoenen 1299

Comparison of Algorithmic and Human Assessments of Sentence Similarity
John Mersch and R. Raymond Lang 1306

Effective Selectional Restrictions for Unsupervised Relation Extraction
Alan Akbik, Larysa Visengeriyeva, Johannes Kirschnick and Alexander Löser 1312

Bootstrapping Semantic Lexicons for Technical Domains
Patrick Ziering, Lonneke van der Plas and Hinrich Schütze 1321

Long-Distance Time-Event Relation Extraction
Alessandro Moschitti, Siddharth Patwardhan and Chris Welty 1330

Unsupervised Extraction of Attributes and Their Values from Product Description
Keiji Shinzato and Satoshi Sekine 1339

Stance Classification of Ideological Debates: Data, Models, Features, and Constraints
Kazi Saidul Hasan and Vincent Ng 1348

University Entrance Examinations as a Benchmark Resource for NLP-based Problem Solving
Yusuke Miyao and Ai Kawazoe 1357

Linguistically Aware Coreference Evaluation Metrics
Chen Chen and Vincent Ng 1366

An Empirical Assessment of Contemporary Online Media in Ad-Hoc Corpus Creation for Social Events
Kanika Narang, Seema Nagar, Sameep Mehta, L V Subramaniam and Kuntal Dey 1375

Diagnosing Causes of Reading Difficulty using Bayesian Networks
Pascual Martínez-Gómez and Akiko Aizawa 1383

Word Co-occurrence Counts Prediction for Bilingual Terminology Extraction from Comparable Corpora
Amir Hazem and Emmanuel Morin 1392

Measuring the Effect of Discourse Relations on Blog Summarization
Shamima Mithun and Leila Kosseim 1401

Supervised Sentence Fusion with Single-Stage Inference
Kapil Thadani and Kathleen McKeown 1410

Detecting and Correcting Learner Korean Particle Omission Errors
Ross Israel, Markus Dickinson and Sun-Hee Lee 1419

Automatic Identification of Learners' Language Background Based on Their Writing in Czech
Katsiaryna Aharodnik, Marco Chang, Anna Feldman and Jirka Hana 1428

Conference Program

October 15, 2013 (Tuesday)

09:00-09:20 Opening

09:20-10:20 Keynote Speech - Hwee Tou Ng (National University of Singapore)

10:20-10:50 Coffee Break

10:50-12:05 Regular Papers

Information Extraction I

10:50–11:15 *Semi-Supervised Answer Extraction from Discussion Forums*
Rose Catherine, Rashmi Gangadharaiah, Karthik Visweswariah and Dinesh Raghu

11:15–11:40 *WordTopic-MultiRank: A New Method for Automatic Keyphrase Extraction*
Fan Zhang, Lian'en Huang and Bo Peng

11:40–12:05 *Towards Contextual Healthiness Classification of Food Items - A Linguistic Approach*
Michael Wiegand and Dietrich Klakow

Information Retrieval I

10:50–11:15 *Learning a Replacement Model for Query Segmentation with Consistency in Search Logs*
Wei Zhang, Yunbo Cao, Chin-Yew Lin, Jian Su and Chew-Lim Tan

11:15–11:40 *Precise Information Retrieval Exploiting Predicate-Argument Structures*
Daisuke Kawahara, Keiji Shinzato, Tomohide Shibata and Sadao Kurohashi

11:40–12:05 *Global Model for Hierarchical Multi-Label Text Classification*
Yugo Murawaki

October 15, 2013 (Tuesday) (continued)

Syntax and Semantics

10:50–11:15 *(Pre-)Annotation of Topic-Focus Articulation in Prague Czech-English Dependency Treebank*
Jiří Mírovský, Kateřina Rysová, Magdaléna Rysová and Eva Hajičová

11:15–11:40 *Animacy Acquisition Using Morphological Case*
Riyaz Ahmad Bhat and Dipti Misra Sharma

11:40–12:05 *The Complexity of Math Problems – Linguistic, or Computational?*
Takuya Matsuzaki, Hidenao Iwane, Hirokazu Anai and Noriko Arai

Pragmatics and Discourse

10:50–11:15 *Hybrid Models for Lexical Acquisition of Correlated Styles*
Julian Brooke and Graeme Hirst

11:15–11:40 *Introducing the Prague Discourse Treebank 1.0*
Lucie Poláková, Jiří Mírovský, Anna Nedoluzhko, Pavlína Jínová, Šárka Zikánová and Eva Hajičová

11:40–12:05 *Multilingual Mention Detection for Coreference Resolution*
Olga Uryupina and Alessandro Moschitti

12:05-13:30 Lunch

13:30-15:10 Regular Papers

October 15, 2013 (Tuesday) (continued)

Text Mining

13:30–13:55 *A Weakly Supervised Bayesian Model for Violence Detection in Social Media*
Amparo Elizabeth Cano Basave, Yulan He, Kang Liu and Jun Zhao

13:55–14:20 *Detecting Spammers in Community Question Answering*
Zhuoye Ding, Yeyun Gong, Yaqian Zhou, Qi Zhang and Xuanjing Huang

14:20–14:45 *Chinese Informal Word Normalization: an Experimental Study*
Aobo Wang, Min-Yen Kan, Daniel Andrade, Takashi Onishi and Kai Ishikawa

14:45–15:10 *Feature Selection Using a Semantic Hierarchy for Event Recognition and Type Classification*
Yoonjae Jeong and Sung-Hyon Myaeng

Phonology and Morphology

13:30–13:55 *Romanization-based Approach to Morphological Analysis in Korean SMS Text Processing*
Youngsam Kim and Hyopil Shin

13:55–14:20 *Efficient Word Lattice Generation for Joint Word Segmentation and POS Tagging in Japanese*
Nobuhiro Kaji and Masaru Kitsuregawa

14:20–14:45 *A Simple Approach to Unknown Word Processing in Japanese Morphological Analysis*
Ryohei Sasano, Sadao Kurohashi and Manabu Okumura

14:45–15:10 *Chinese Word Segmentation by Mining Maximized Substrings*
Mo Shen, Daisuke Kawahara and Sadao Kurohashi

October 15, 2013 (Tuesday) (continued)

POS Tagging and Parsing

13:30–13:55 *Capturing Long-distance Dependencies in Sequence Models: A Case Study of Chinese Part-of-speech Tagging*
Weiwei Sun, Xiaochang Peng and Xiaojun Wan

13:55–14:20 *Exploring Semantic Information in Hindi WordNet for Hindi Dependency Parsing*
Sambhav Jain, Naman Jain, Aniruddha Tammewar, Riyaz Ahmad Bhat and Dipti Sharma

14:20–14:45 *Towards Robust Cross-Domain Domain Adaptation for Part-of-Speech Tagging*
Tobias Schnabel and Hinrich Schütze

14:45–15:10 *Dependency Parsing for Identifying Hungarian Light Verb Constructions*
Veronika Vincze, János Zsibrita and István Nagy T.

Dialogue and Dialogue Systems

13:30–13:55 *Written Dialog and Social Power: Manifestations of Different Types of Power in Dialog Behavior*
Vinodkumar Prabhakaran and Owen Rambow

13:55–14:20 *Evaluation of the Scusi? Spoken Language Interpretation System – A Case Study*
Thomas Kleinbauer, Ingrid Zukerman and Su Nam Kim

14:20–14:45 *A Noisy Channel Approach to Error Correction in Spoken Referring Expressions*
Su Nam Kim, Ingrid Zukerman, Thomas Kleinbauer and Farshid Zavareh

14:45–15:10 *Natural Language Query Refinement for Problem Resolution from Crowd-Sourced Semi-Structured Data*
Rashmi Gangadharaiah and Balakrishnan Narayanaswamy

15:10-15:40 Coffee Break

15:40-17:45 Regular Papers

October 15, 2013 (Tuesday) (continued)

Machine Translation I

15:40–16:05 *Ensemble Triangulation for Statistical Machine Translation*
Majid Razmara and Anoop Sarkar

16:05–16:30 *Robust Transliteration Mining from Comparable Corpora with Bilingual Topic Models*
John Richardson, Toshiaki Nakazawa and Sadao Kurohashi

16:30–16:55 *SuMT: A Framework of Summarization and MT*
Houda Bouamor, Behrang Mohit and Kemal Oflazer

16:55–17:15 *Tuning SMT with a Large Number of Features via Online Feature Grouping*
Lemao Liu, Tiejun Zhao, Taro Watanabe and Eiichiro Sumita

17:15–17:35 *Multimodal Comparable Corpora as Resources for Extracting Parallel Data: Parallel Phrases Extraction*
Haithem Afli, Loïc Barrault and Holger Schwenk

Information Extraction II

15:40–16:05 *Bootstrapping Large-scale Named Entities using URL-Text Hybrid Patterns*
Chao Zhang, Shiqi Zhao and Haifeng Wang

16:05–16:30 *Feature-Rich Segment-Based News Event Detection on Twitter*
Yanxia Qin, Yue Zhang, Min Zhang and Dequan Zheng

16:30–16:55 *Building Chinese Event Type Paradigm Based on Trigger Clustering*
Xiao Ding, Bing Qin and Ting Liu

16:55–17:20 *Chinese Named Entity Abbreviation Generation Using First-Order Logic*
Huan Chen, Qi Zhang, Jin Qian and Xuanjing Huang

17:20–17:45 *Full-coverage Identification of English Light Verb Constructions*
István Nagy T., Veronika Vincze and Richárd Farkas

October 15, 2013 (Tuesday) (continued)

Recent NLP Applications I

15:40–16:05 *Detecting Deceptive Opinions with Profile Compatibility*
Vanessa Wei Feng and Graeme Hirst

16:05–16:30 *Behind the Times: Detecting Epoch Changes using Large Corpora*
Octavian Popescu and Carlo Strapparava

16:30–16:55 *How Noisy Social Media Text, How Diffrnt Social Media Sources?*
Timothy Baldwin, Paul Cook, Marco Lui, Andrew MacKinlay and Li Wang

16:55–17:20 *Who Had the Upper Hand? Ranking Participants of Interactions Based on Their Relative Power*
Vinodkumar Prabhakaran, Ajita John and Dorée D. Seligmann

17:20–17:45 *Readability Indices for Automatic Evaluation of Text Simplification Systems: A Feasibility Study for Spanish*
Sanja Štajner and Horacio Saggion

Language Resources I

15:40–16:05 *Weasels, Hedges and Peacocks: Discourse-level Uncertainty in Wikipedia Articles*
Veronika Vincze

16:05–16:30 *Automatically Developing a Fine-grained Arabic Named Entity Corpus and Gazetteer by utilizing Wikipedia*
Fahd Alotaibi and Mark Lee

16:30–16:55 *Ranking Translation Candidates Acquired from Comparable Corpora*
Rima Harastani, Béatrice Daille and Emmanuel Morin

16:55–17:20 *Using the Semantic-Syntactic Interface for Reliable Arabic Modality Annotation*
Rania Al-Sabbagh, Jana Diesner and Roxana Girju

17:20–17:45 *Mapping Rules for Building a Tunisian Dialect Lexicon and Generating Corpora*
Rahma Boujelbane, Mariem Ellouze Khemekhem and Lamia Hadrich Belguith

19:30-21:30 Reception (ANA Crowne Plaza Hotel Grand Court Nagoya)

October 16, 2013 (Wednesday)

09:00-10:15 Regular Papers

Machine Translation II

09:00–09:25 *Hypothesis Refinement Using Agreement Constraints in Machine Translation*
Ankur Gandhe and Rashmi Gangadharaiah

09:25–09:50 *Scalable Variational Inference for Extracting Hierarchical Phrase-based Translation Rules*
Baskaran Sankaran, Gholamreza Haffari and Anoop Sarkar

09:50–10:15 *A Topic-Triggered Language Model for Statistical Machine Translation*
Heng Yu, Jinsong Su, Yajuan Lv and Qun Liu

Opnion Mining I

09:00–09:25 *Reserved Self-training: A Semi-supervised Sentiment Classification Method for Chinese Microblogs*
Zhiguang Liu, Xishuang Dong, Yi Guan and Jinfeng Yang

09:25–09:50 *Enhancing Lexicon-Based Review Classification by Merging and Revising Sentiment Dictionaries*
Heeryon Cho, Jong-Seok Lee and Songkuk Kim

09:50–10:15 *Exploring the Effects of Word Roots for Arabic Sentiment Analysis*
Shereen Oraby, Yasser El-Sonbaty and Mohamad Abou El-Nasr

October 16, 2013 (Wednesday) (continued)

Semantic Processing I

09:00–09:25 *Topical Key Concept Extraction from Folksonomy*
Han Xue, Bing Qin, Ting Liu and Chao Xiang

09:25–09:50 *Uncovering Distributional Differences between Synonyms and Antonyms in a Word Space Model*
Silke Scheible, Sabine Schulte im Walde and Sylvia Springorum

09:50–10:15 *Multilingual Word Sense Disambiguation Using Wikipedia*
Bharath Dandala, Rada Mihalcea and Razvan Bunescu

Statistical and ML Language Modeling I

09:00–09:25 *Semantic v.s. Positions: Utilizing Balanced Proximity in Language Model Smoothing for Information Retrieval*
Rui Yan, Han Jiang, Mirella Lapata, Shou-De Lin, Xueqiang Lv and Xiaoming Li

09:25–09:50 *An Unsupervised Parameter Estimation Algorithm for a Generative Dependency N-gram Language Model*
Chenchen Ding and Mikio Yamamoto

09:50–10:15 *Learning a Product of Experts with Elitist Lasso*
Mengqiu Wang and Christopher D. Manning

10:15-10:45 Coffee Break

10:45-12:00 Regular Papers

October 16, 2013 (Wednesday) (continued)

Information Extraction III / Question Answering

10:45–11:10 *Learning Efficient Information Extraction on Heterogeneous Texts*
Henning Wachsmuth, Benno Stein and Gregor Engels

11:10–11:35 *TopicRank: Graph-Based Topic Ranking for Keyphrase Extraction*
Adrien Bougouin, Florian Boudin and Béatrice Daille

11:35–12:00 *Understanding the Semantic Intent of Natural Language Query*
Juan Xu, Qi Zhang and Xuanjing Huang

Opnion Mining II

10:45–11:10 *Sentiment Classification for Movie Reviews in Chinese Using Parsing-based Methods*
Wen-Juan Hou and Chuang-Ping Chang

11:10–11:35 *Sentiment Aggregation using ConceptNet Ontology*
Subhabrata Mukherjee and Sachindra Joshi

11:35–12:00 *Detecting Cyberbullying Entries on Informal School Websites Based on Category Relevance Maximization*
Taisei Nitta, Fumito Masui, Michal Ptaszynski, Yasutomo Kimura, Rafal Rzepka and Kenji Araki

Semantic Processing II

10:45–11:10 *A Lexicon-based Investigation of Research Issues in Japanese Factuality Analysis*
Kazuya Narita, Junta Mizuno and Kentaro Inui

11:10–11:35 *A Hierarchical Semantics-Aware Distributional Similarity Scheme*
Shuqi Sun, Ke Sun, Shiqi Zhao, Haifeng Wang, Muyun Yang and Sheng Li

11:35–12:00 *Labeled Alignment for Recognizing Textual Entailment*
Xiaolin Wang, Hai Zhao and Bao-Liang Lu

October 16, 2013 (Wednesday) (continued)

Statistical and ML Language Modeling II

10:45–11:10 *Context-Based Chinese Word Segmentation using SVM Machine-Learning Algorithm without Dictionary Support*
Chia-ming Lee and Chien-Kang Huang

11:10–11:35 *A Common Case of Jekyll and Hyde: The Synergistic Effect of Using Divided Source Training Data for Feature Augmentation*
Yan Song and Fei Xia

11:35–12:00 *Detecting Polysemy in Hard and Soft Cluster Analyses of German Preposition Vector Spaces*
Sylvia Springorum, Sabine Schulte im Walde and Jason Utt

12:00-13:10 Lunch

13:10-15:30 Short Papers

Recent NLP Applications / Text Summarization / Opinion Mining

13:10–13:30 *Generalized Abbreviation Prediction with Negative Full Forms and Its Application on Improving Chinese Web Search*
Xu Sun, Wenjie Li, Fanqi Meng and Houfeng Wang

13:30–13:50 *Prosody-Based Unsupervised Speech Summarization with Two-Layer Mutually Reinforced Random Walk*
Sujay Kumar Jauhar, Yun-Nung Chen and Florian Metze

13:50–14:10 *Mining the Gaps: Towards Polynomial Summarization*
Marina Litvak and Natalia Vanetik

14:10–14:30 *Detecting Domain Dedicated Polar Words*
Raksha Sharma and Pushpak Bhattacharyya

14:30–14:50 *Can I Hear You? Sentiment Analysis on Medical Forums*
Tanveer Ali, David Schramm, Marina Sokolova and Diana Inkpen

14:50–15:10 *Construction of Emotional Lexicon Using Potts Model*
Braja Gopal Patra, Hiroya Takamura, Dipankar Das, Manabu Okumura and Sivaji Bandyopadhyay

15:10–15:30 *Suicidal Tendencies: The Automatic Classification of Suicidal and Non-Suicidal Lyricists Using NLP*
Matthew Mulholland and Joanne Quinn

October 16, 2013 (Wednesday) (continued)

Language Resources / Information Extraction / Text Mining

13:10–13:30 *Unsupervised Word Class Induction for Under-resourced Languages: A Case Study on Indonesian*
Meladel Mistica, Jey Han Lau and Timothy Baldwin

13:30–13:50 *An Efficient Active Learning Framework for New Relation Types*
Lisheng Fu and Ralph Grishman

13:50–14:10 *Parsing Dependency Paths to Identify Event-Argument Relations*
Seung-Cheol Baek and Jong Park

14:10–14:30 *Augmentable Paraphrase Extraction Framework*
MeiHua Chen, YiChun Chen, ShihTing Huang and Jason S. Chang

14:30–14:50 *Automatic Prediction of Evidence-based Recommendations via Sentence-level Polarity Classification*
Abeed Sarker, Diego Mollá-Aliod and Cécile Paris

14:50–15:10 *Clustering Microtext Streams for Event Identification*
Jie Yin

15:10–15:30 *Automatic Corpora Construction for Text Classification*
Dandan Wang, Qingcai Chen, Xiaolong Wang and Bingyang Yu

Information Retrieval / NLP for Educational Applications / Phonology and Morphology

13:10–13:30 *Learning to Generate Diversified Query Interpretations using Biconvex Optimization*
Ramakrishna Bairi, Ambha A and Ganesh Ramakrishnan

13:30–13:50 *Learning Based Approaches for Vietnamese Question Classification Using Keywords Extraction from the Web*
Dang Tran, Cuong Chu, Son Pham and Minh Nguyen

13:50–14:10 *Detecting Bot-Answerable Questions in Ubuntu Chat*
David Uthus and David Aha

14:10–14:30 *Alignment-based Annotation of Proofreading Texts toward Professional Writing Assistance*
Ngan Nguyen and Yusuke Miyao

October 16, 2013 (Wednesday) (continued)

14:30–14:50 *Toward Automatic Processing of English Metalanguage*
Shomir Wilson

14:50–15:10 *On the Effectiveness of Using Syntactic and Shallow Semantic Tree Kernels for Automatic Assessment of Essays*
Yllias Chali and Sadid A. Hasan

15:10–15:30 *Little by Little: Semi Supervised Stemming through Stem Set Minimization*
Vasudevan N and Pushpak Bhattacharyya

Semantic Processing / Pragmatics and Discourse

13:10–13:30 *What Information is Helpful for Dependency Based Semantic Role Labeling*
Yanyan Luo, Kevin Duh and Yuji Matsumoto

13:30–13:50 *Classifying Taxonomic Relations between Pairs of Wikipedia Articles*
Or Biran and Kathleen McKeown

13:50–14:10 *A Rule System for Chinese Time Entity Recognition by Comprehensive Linguistic Study*
Hongzhi Xu and Chu-Ren Huang

14:10–14:30 *Financial Sentiment Analysis for Risk Prediction*
Chuan-Ju Wang, Ming-Feng Tsai, Tse Liu and Chin-Ting Chang

14:30–14:50 *Sense Disambiguation: From Natural Language Words to Mathematical Terms*
Minh-Quoc Nghiem, Giovanni Yoko Kristianto, Goran Topic and Akiko Aizawa

14:50–15:10 *Adapting a State-of-the-art Anaphora Resolution System for Resource-poor Language*
Utpal Sikdar, Asif Ekbal, Sriparna Saha, Olga Uryupina and Massimo Poesio

15:10–15:30 *Chinese Event Coreference Resolution: Understanding the State of the Art*
Chen Chen and Vincent Ng

15:30-17:00 Poster Presentations and System Demonstrations

A Two-Step Named Entity Recognizer for Open-Domain Search Queries
Andreas Eiselt and Alejandro Figueroa

October 16, 2013 (Wednesday) (continued)

A Comparison of Centrality Measures for Graph-Based Keyphrase Extraction
Florian Boudin

Translating Chinese Unknown Words by Automatically Acquired Templates
Ming-Hong Bai, Yu-Ming Hsieh, Keh-Jiann Chen and Jason S. Chang

Multilingual Lexicon Bootstrapping - Improving a Lexicon Induction System Using a Parallel Corpus
Patrick Ziering, Lonneke van der Plas and Hinrich Schütze

Mining Japanese Compound Words and Their Pronunciations from Web Pages and Tweets
Xianchao Wu

A Factoid Question Answering System Using Answer Pattern Matching
Nagehan Pala Er and Ilyas Cicekli

Chinese Short Text Classification Based on Domain Knowledge
Xiao Feng, Yang Shen, Chengyong Liu, Wei Liang and Shuwu Zhang

Applying Graph-based Keyword Extraction to Document Retrieval
Youngsam Kim, Munhyong Kim, Andrew Cattle, Julia Otmakhova, Suzi Park and Hyopil Shin

Semi-supervised Classification of Twitter Messages for Organization Name Disambiguation
Shu Zhang, Jianwei Wu, Dequan Zheng, Yao Meng and Hao Yu

Word in a Dictionary is used by Numerous Users
Eiji Aramaki, Sachiko Maskawa, Mai Miyabe, Mizuki Morita and Sachi Yasuda

Extracting Evaluative Conditions from Online Reviews: Toward Enhancing Opinion Mining
Yuki Nakayama and Atsushi Fujii

Cognate Production using Character-based Machine Translation
Lisa Beinborn, Torsten Zesch and Iryna Gurevych

An Empirical Study of Combing Multiple Models in Bengali Question Classification
Somnath Banerjee and Sivaji Bandyopadhyay

October 16, 2013 (Wednesday) (continued)

A Two-Stage Classifier for Sentiment Analysis
Dai Quoc Nguyen, Dat Quoc Nguyen and Son Bao Pham

Exploiting User Search Sessions for the Semantic Categorization of Question-like Informational Search Queries
Alejandro Figueroa and Guenter Neumann

Influence of Part-of-Speech and Phrasal Category Universal Tag-set in Tree-to-Tree Translation Models
Francisco Oliveira, Derek F. Wong, Lidia S. Chao, Liang Tian and Liangye He

Interest Analysis using PageRank and Social Interaction Content
Chung-chi Huang and Lun-Wei Ku

Time Series Topic Modeling and Bursty Topic Detection of Correlated News and Twitter
Daichi Koike, Yusuke Takahashi, Takehito Utsuro, Masaharu Yoshioka and Noriko Kando

A Distant Supervision Approach for Identifying Perspectives in Unstructured User-Generated Text
Attapol Thamrongrattanarit, Colin Pollock, Benjamin Goldenberg and Jason Fennell

An Approach of Hybrid Hierarchical Structure for Word Similarity Computing by HowNet
Jiangming Liu, Jinan Xu and Yujie Zhang

Extracting Causes of Emotions from Text
Alena Neviarouskaya and Masaki Aono

Automated Grammar Correction Using Hierarchical Phrase-Based Statistical Machine Translation
Bibek Behera and Pushpak Bhattacharyya

Finding Dependency Parsing Limits over a Large Spanish Corpus
Muntsa Padró, Miguel Ballesteros, Héctor Martínez and Bernd Bohnet

High Quality Dependency Selection from Automatic Parses
Gongye Jin, Daisuke Kawahara and Sadao Kurohashi

Building Specialized Bilingual Lexicons Using Word Sense Disambiguation
Dhouha Bouamor, Nasredine Semmar and Pierre Zweigenbaum

October 16, 2013 (Wednesday) (continued)

Predicate Argument Structure Analysis using Partially Annotated Corpora
Koichiro Yoshino, Shinsuke Mori and Tatsuya Kawahara

Statistical Dialogue Management using Intention Dependency Graph
Koichiro Yoshino, Shinji Watanabe, Jonathan Le Roux and John R. Hershey

Repairing Incorrect Translation with Examples
Junguo Zhu, Muyun Yang, Sheng Li and Tiejun Zhao

Phrase-based Parallel Fragments Extraction from Comparable Corpora
Xiaoyin Fu, Wei Wei, Shixiang Lu, Zhenbiao Chen and Bo Xu

A Hybrid Approach for Anaphora Resolution in Hindi
Praveen Dakwale, Vandan Mujadia and Dipti M Sharma

Structure Cognizant Pseudo Relevance Feedback
Arjun Atreya V, Yogesh Kakde, Pushpak Bhattacharyya and Ganesh Ramakrishnan

Cross-Domain Answer Ranking using Importance Sampling
Anders Johannsen and Anders Søgaard

Morphological Analysis of Tunisian Dialect
Inès Zribi, Mariem Ellouze Khemakhem and Lamia Hadrich Belguith

Disambiguating Explicit Discourse Connectives without Oracles
Anders Johannsen and Anders Søgaard

Updating Rare Term Vector Replacement
Tobias Berka and Marian Vajteršic

Statistical Morphological Analyzer for Hindi
Deepak Kumar Malladi and Prashanth Mannem

Induction of Root and Pattern Lexicon for Unsupervised Morphological Analysis of Arabic
Bilal Khaliq and John Carrol

October 16, 2013 (Wednesday) (continued)

Using Shallow Semantic Parsing and Relation Extraction for Finding Contradiction in Text
Minh Quang Nhat Pham, Minh Le Nguyen and Akira Shimazu

Using Transliteration of Proper Names from Arabic to Latin Script to Improve English-Arabic Word Alignment
Nasredine Semmar and Houda Saadane

A Semi-Supervised Method for Arabic Word Sense Disambiguation Using a Weighted Directed Graph
Laroussi Merhbene, Anis Zouaghi and Mounir Zrigui

Incremental Segmentation and Decoding Strategies for Simultaneous Translation
Mahsa Yarmohammadi, Vivek Kumar Rangarajan Sridhar, Srinivas Bangalore and Baskaran Sankaran

Two Case Studies on Translating Pronouns in a Deep Syntax Framework
Michal Novák, Zdenek Zabokrtsky and Anna Nedoluzhko

Bootstrapping Phrase-based Statistical Machine Translation via WSD Integration
Hien Vu Huy, Phuong-Thai Nguyen, Tung-Lam Nguyen and M.L Nguyen

Orthographic and Morphological Processing for Persian-to-English Statistical Machine Translation
Mohammad Sadegh Rasooli, Ahmed El Kholy and Nizar Habash

Interoperability between Service Composition and Processing Pipeline: Case Study on the Language Grid and UIMA
Trang Mai Xuan, Yohei Murakami, Donghui Lin and Toru Ishida

Improving Calculation of Contextual Similarity for Constructing a Bilingual Dictionary via a Third Language
Takashi Tsunakawa, Yosuke Yamamoto and Hiroyuki Kaji

Two-Stage Pre-ordering for Japanese-to-English Statistical Machine Translation
Sho Hoshino, Yusuke Miyao, Katsuhito Sudoh and Masaaki Nagata

Grammatical Error Correction Using Feature Selection and Confidence Tuning
Yang Xiang, Yaoyun Zhang, Xiaolong Wang, Chongqiang Wei, Wen Zheng, Xiaoqiang Zhou, Yuxiu Hu and Yang Qin

An Online Algorithm for Learning over Constrained Latent Representations using Multiple Views
Ann Clifton, Max Whitney and Anoop Sarkar

October 16, 2013 (Wednesday) (continued)

Synonym Acquisition Using Bilingual Comparable Corpora
Daniel Andrade, Masaaki Tsuchida, Takashi Onishi and Kai Ishikawa

Exploring Verb Frames for Sentence Simplification in Hindi
Ankush Soni, Sambhav Jain and Dipti Misra Sharma

Dirichlet Processes for Joint Learning of Morphology and PoS Tags
Burcu Can and Suresh Manandhar

Parser Accuracy in Quality Estimation of Machine Translation: A Tree Kernel Approach
Rasoul Samad Zadeh Kaljahi, Jennifer Foster, Raphael Rubino, Johann Roturier and Fred Hollowood

Attribute Relation Extraction from Template-inconsistent Semi-structured Text by Leveraging Site-level Knowledge
Yang Liu, Fang Liu, Siwei Lai, Kang Liu, Guangyou Zhou and Jun Zhao

Optimum Parameter Selection for K.L.D. Based Authorship Attribution in Gujarati
Parth Mehta and Prasenjit Majumder

Modeling User Leniency and Product Popularity for Sentiment Classification
Wenliang Gao, Naoki Yoshinaga, Nobuhiro Kaji and Masaru Kitsuregawa

A Generalized LCS Algorithm and Its Application to Corpus Alignment
Jin-Dong Kim

Semantic Naïve Bayes Classifier for Document Classification
How Jing, Yu Tsao, Kuan-Yu Chen and Hsin-Min Wang

Cluster-based Web Summarization
Yves Petinot, Kathleen McKeown and Kapil Thadani

Automated Activity Recognition in Clinical Documents
Camilo Thorne, Marco Montali, Diego Calvanese, Elena Cardillo and Claudio Eccher

Large-Scale Text Collection for Unwritten Languages
Florian R. Hanke and Steven Bird

October 16, 2013 (Wednesday) (continued)

A Self-learning Template Approach for Recognizing Named Entities from Web Text
Qian Liu, Bingyang Liu, Dayong Wu, Yue Liu and Xueqi Cheng

18:00-21:00 Banquet (Port of Nagoya Public Aquarium)

October 17, 2013 (Thursday)

09:00-10:00 Keynote Speech - Roberto Navigli (Sapienza University of Rome)

10:00-10:30 Coffee Break

10:30-12:30 Short Papers

Machine Translation

10:30–10:50 *Accurate Parallel Fragment Extraction from Quasi–Comparable Corpora using Alignment Model and Translation Lexicon*
Chenhui Chu, Toshiaki Nakazawa and Sadao Kurohashi

10:50–11:10 *Meta-level Statistical Machine Translation*
Sajad Ebrahimi, Kourosh Meshgi, Shahram Khadivi and Mohammad Ebrahim Shiri Ahmad Abady

11:10–11:30 *Bayesian Induction of Bracketing Inversion Transduction Grammars*
Markus Saers and Dekai Wu

11:30–11:50 *Estimating the Quality of Translated User-Generated Content*
Raphael Rubino, Jennifer Foster, Rasoul Samad Zadeh Kaljahi, Johann Roturier and Fred Hollowood

11:50–12:10 *Selective Combination of Pivot and Direct Statistical Machine Translation Models*
Ahmed El Kholy, Nizar Habash, Gregor Leusch, Evgeny Matusov and Hassan Sawaf

12:10–12:30 *Multiword Expressions in the Context of Statistical Machine Translation*
Mahmoud Ghoneim and Mona Diab

October 17, 2013 (Thursday) (continued)

Recent NLP Applications / Dialogue and Dialogue Systems

10:30–10:50 *Uncertainty Detection for Natural Language Watermarking*
György Szarvas and Iryna Gurevych

10:50–11:10 *KySS 1.0: a Framework for Automatic Evaluation of Chinese Input Method Engines*
Zhongye Jia and Hai Zhao

11:10–11:30 *Automatic Extraction of Social Networks from Literary Text: A Case Study on Alice in Wonderland*
Apoorv Agarwal, Anup Kotalwar and Owen Rambow

11:30–11:50 *Using the Web to Train a Mobile Device Oriented Japanese Input Method Editor*
Xianchao Wu, Rixin Xiao and Xiaoxin Chen

11:50–12:10 *A Novel Approach Towards Incorporating Context Processing Capabilities in NLIDB System*
Arjun Akula, Rajeev Sangal and Radhika Mamidi

12:10–12:30 *Iterative Development and Evaluation of a Social Conversational Agent*
Annika Silvervarg and Arne Jönsson

POS Tagging and Parsing / Syntax and Semantics / Information Extraction

10:30–10:50 *A Hybrid Morphological Disambiguation System for Turkish*
Mucahid Kutlu and Ilyas Cicekli

10:50–11:10 *A Dynamic Confusion Score for Dependency Arc Labels*
Sambhav Jain and Bhasha Agrawal

11:10–11:30 *Increasing the Quality and Quantity of Source Language Data for Unsupervised Cross-Lingual POS Tagging*
Long Duong, Paul Cook, Steven Bird and Pavel Pecina

11:30–11:50 *Towards the Annotation of Penn TreeBank with Information Structure*
Bernd Bohnet, Alicia Burga and Leo Wanner

11:50–12:10 *Constituency and Dependency Relationship from a Tree Adjoining Grammar and Abstract Categorial Grammars Perspective*
Aleksandre Maskharashvili and Sylvain Pogodalla

October 17, 2013 (Thursday) (continued)

12:10–12:30 *Named Entity Extraction using Information Distance*
Sangameshwar Patil, Sachin Pawar and Girish Palshikar

Statistical and ML Language Modeling

10:30–10:50 *Feature-based Neural Language Model and Chinese Word Segmentation*
Mairgup Mansur, Wenzhe Pei and Baobao Chang

10:50–11:10 *Human-Computer Interactive Chinese Word Segmentation: An Adaptive Dirichlet Process Mixture Model Approach*
Tongfei Chen, Xiaojun Zou, Weimeng Zhu and Junfeng Hu

11:10–11:30 *Effect of Non-linear Deep Architecture in Sequence Labeling*
Mengqiu Wang and Christopher D. Manning

11:30–11:50 *Case Study of Model Adaptation: Transfer Learning and Online Learning*
Kenji Imamura

11:50–12:10 *Source and Translation Classification using Most Frequent Words*
Zahurul Islam and Armin Hoenen

12:10–12:30 *Comparison of Algorithmic and Human Assessments of Sentence Similarity*
John Mersch and R. Raymond Lang

12:30-14:00 Lunch

14:00-16:05 Regular Papers

October 17, 2013 (Thursday) (continued)

Information Extraction IV

14:05–14:30 *Effective Selectional Restrictions for Unsupervised Relation Extraction*
Alan Akbik, Larysa Visengeriyeva, Johannes Kirschnick and Alexander Löser

14:30–14:55 *Bootstrapping Semantic Lexicons for Technical Domains*
Patrick Ziering, Lonneke van der Plas and Hinrich Schütze

14:55–15:20 *Long-Distance Time-Event Relation Extraction*
Alessandro Moschitti, Siddharth Patwardhan and Chris Welty

15:20–15:45 *Unsupervised Extraction of Attributes and Their Values from Product Description*
Keiji Shinzato and Satoshi Sekine

15:45–16:05 *Stance Classification of Ideological Debates: Data, Models, Features, and Constraints*
Kazi Saidul Hasan and Vincent Ng

Language Resources II / Recent NLP Applications II

14:05–14:30 *University Entrance Examinations as a Benchmark Resource for NLP-based Problem Solving*
Yusuke Miyao and Ai Kawazoe

14:30–14:55 *Linguistically Aware Coreference Evaluation Metrics*
Chen Chen and Vincent Ng

14:55–15:20 *An Empirical Assessment of Contemporary Online Media in Ad-Hoc Corpus Creation for Social Events*
Kanika Narang, Seema Nagar, Sameep Mehta, L V Subramaniam and Kuntal Dey

15:20–15:45 *Diagnosing Causes of Reading Difficulty using Bayesian Networks*
Pascual Martínez-Gómez and Akiko Aizawa

15:45–16:05 *Word Co-occurrence Counts Prediction for Bilingual Terminology Extraction from Comparable Corpora*
Amir Hazem and Emmanuel Morin

October 17, 2013 (Thursday) (continued)

Text Summarization / NLP for Educational Applications

14:05–14:30 *Measuring the Effect of Discourse Relations on Blog Summarization*
Shamima Mithun and Leila Kosseim

14:30–14:55 *Supervised Sentence Fusion with Single-Stage Inference*
Kapil Thadani and Kathleen McKeown

14:55–15:20 *Detecting and Correcting Learner Korean Particle Omission Errors*
Ross Israel, Markus Dickinson and Sun-Hee Lee

15:20–15:45 *Automatic Identification of Learners' Language Background Based on Their Writing in Czech*
Katsiaryna Aharodnik, Marco Chang, Anna Feldman and Jirka Hana

16:05-16:35 Coffee Break

16:35-17:15 Best Papers

17:15-17:25 Future Conferences

17:25-17:35 Closing