



IJCNLP 2011

**The Fifth International Joint Conference on Natural
Language Processing**

**November 8-13, 2011
Shangri-La Hotel
Chiang Mai, Thailand**



IJCNLP 2011

**Proceedings of
the Fifth International Joint Conference on Natural
Language Processing**

November 8 – 13, 2011
Chiang Mai, Thailand

We wish to thank our sponsors

Gold Sponsors



www.google.com



www.baidu.com



[The Office of Naval Research \(ONR\)](#)



[The Asian Office of Aerospace Research and Development \(AOARD\)](#)



[Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong](#)

Silver Sponsors



[Microsoft Corporation](#)

Bronze Sponsors



[Chinese and Oriental Languages Information Processing Society \(COLIPS\)](#)

Supporter



[Thailand Convention and Exhibition Bureau \(TCEB\)](#)

We wish to thank our sponsors

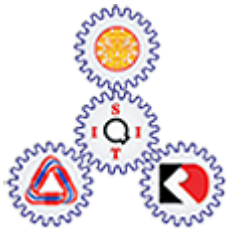
Organizers



[Asian Federation of Natural Language Processing \(AFNLP\)](#)



[National Electronics and Computer Technology Center \(NECTEC\), Thailand](#)



[Sirindhorn International Institute of Technology \(SIIT\), Thailand](#)



[Rajamangala University of Technology Lanna \(RMUTL\), Thailand](#)



[Maejo University, Thailand](#)



[Chiang Mai University \(CMU\), Thailand](#)

©2011 Asian Federation of Natural Language Processing

ISBN 978-974-466-564-5

FOREWORD

IJCNLP2011, where Anna(s) meet the king(s) for sharing knowledge in natural language

IJCNLP2011 is held in Chiang Mai. It is a historic city situated in the northern part of Thailand. Organizing the conference in this part of Asia made us think of the classic movie “The King and I” (1956), where King Mongkut of Siam invited Anna Leonowens an Anglo-Indian school teacher to Siam to teach his family English. Similar to the movie, IJCNLP2011 brings together scientists and practitioners from the East and West in pursuit of the knowledge of natural language processing (NLP).

Virach, Hitoshi and I compiled this passage collaboratively online using our own iPads. Despite us being physically apart, in Thailand, Japan and Hong Kong respectively, our collaborative editorial work went smoothly with virtually no distance. The increasing popularity of smart handheld devices, such as iPhones and iPads has practically made the world flat. The hurdles and boundaries between people have effectively been lifted enabling friends and relatives over the globe to keep in close contact with each other. We use email, blog, facebook and twitter regularly and ubiquitously for communications. Non-traditional they may be, the languages for communication over these channels are natural as they are used by the netizens (human) for information exchange. Processing of these natural languages is inevitably unconventional and the task is challenging, which requires much innovation. For this reason, NLP is a key research area both in the industry and in universities worldwide. Therefore, it is not surprising that we have received over 500 submissions from different countries around the world in this year’s IJCNLP. This number is in fact the largest in the history of the conference series.

Organizing a conference of the scale of IJCNLP2011 (with over 300 participants) is never easy. We worked closely as a team in the past ten months. It is really not easy for us to express our gratitude to any one individual. The names of the hard working conference officers, the track chairs, the workshop chairs, the tutors as well as the reviewers are enlisted in the proceedings. We owe everyone a billion. Without their hard work IJCNLP2011 would never have reached this stage. So please help me praise and thank them when you meet them in the conference.

Chiang Mai is a cultural city full of history and traditions, with many famous attractions such as its melodious colloquial language, Lanna style of clothing, mellow taste of food, etc. During the conference period, we will experience the “Loi Krathong Festival” where people float krathong (floating basket) on a river to pay respect to the spirit of the waters. IJCNLP2011 in November Chiang Mai is unique. It coincides with the unforgettable Lanna Festival. Locally known as “Yi Peng”, the festival will bring to you a memorable cultural experience. You will witness a multitude of Lanna-style sky lanterns (khom loi, literally “floating lanterns”) gently rising in the air. These lanterns resemble large flocks of giant fluorescent jellyfish gracefully floating by through the sky. Honestly, these attractions are just too good to be missed.

Dear friends and colleagues of the world NLP communities, honorable guests of Chiang Mai, we are glad to see you in IJCNLP2011. We hope you find the technical program useful to your research and can discover something insightful at the end. And before closing, as one often said “seeing is believing”, we urge you to spare some time after the conference to explore and to enjoy the city.

Ka Poon Kap (thank you)

Kam-Fai Wong, General Chair, The Chinese University of Hong Kong (CUHK), China

Virach Sornlertlamvanich, Organization Co-Chair, National Electronics and Computer Technology Center (NECTEC), Thailand

Hitoshi Isahara, Organization Co-Chair, Toyohashi University of Technology, Japan

November 7, 2011

PREFACE

As the flagship conference of the Asian Federation of Natural Language Processing (AFNLP), IJCNLP has now rapidly grown into a renowned international event. IJCNLP 2011 covers a broad spectrum of technical areas related to natural language processing. The conference includes full papers, short papers, demonstrations, a student research workshop, as well as pre- and post-conference tutorials and workshops.

This year, we received a record 478 valid paper submissions, which is well beyond our initial expectations. This represents an increasing interest of research on NLP and the growing reputation of IJCNLP as an international event. The 478 submissions include 385 full-paper submissions and 93 short-paper submissions from more than 40 countries. Specifically, approximately 61% of the papers are from 16 countries and areas in Asia Pacific, 22% from 16 countries in Europe, 14% from the United States and Canada; we also have 2% of the papers from the Middle East and Africa, and 1% from South America.

We would like to thank all the authors for submitting papers to IJCNLP 2011. The significant increase in the number of submissions and the wide range of demographic areas represent a rapid growth of our field. We would also like to thank the 22 area chairs and 474 program committee members for writing over 1400 reviews and meta-reviews and for paving the way for the final paper selection. Of all 478 submissions, a total of 176 papers were accepted, representing a healthy 36% acceptance rate. The accepted papers are comprised of 149 full papers (8+ pages), of which 107 are presented orally and 42 as posters, and 27 short papers (4+ pages) where 25 are presented orally and 2 as posters. 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 be possible.

We are delighted to have invited three strategic keynote speakers addressing different application aspects of NLP for the Web in IJCNLP2011. Mathew Lease will talk about “crowdsourcing”, which is a trendy and effective means to perform a task that requires hundreds/thousands of people, such as corpus tagging. Wai Lam will present the latest techniques for information extraction, which is essential for today’s Internet business. And last but not the least, Mengqiu Wang, Vice President of Baidu, the largest Internet search company in China, will share with us the recent trends in search and social network technologies and how NLP techniques can be applied to improve performance in the real world. These speeches will surely be informative and enlightening to the audience leading to many innovative research ideas. We are excited about it and are looking forward to them. Best paper awards will be announced in the last session of the conference as well.

We thank General Chair Kam-Fai Wong, the Local Arrangements Committee headed by Virach Sornlertlamvanich and Hitoshi Isahara, and the AFNLP Conference Coordination Committee chaired by Yuji Matsumoto, for their help and advice. Thanks to Min Zhang and Sudeshria Sarkar, the Publication Co-Chairs for putting the proceedings together, and all the other committee chairs for their work.

We hope that you enjoy the conference!

Haifeng Wang, Baidu

David Yarowsky, Johns Hopkins University

November 7, 2011

Honorary Conference Chair

Chaoyong Eurviriyankul, Rajamangala University of Technology Lanna, Thailand
Chongrak Polprasert, Sirindhorn International Institute of Technology, Thailand
Thaweesak Koanantakool, NSTDA, Thailand

General Chair

Kam-Fai Wong, The Chinese University of Hong Kong, China

Program Co-Chairs:

Haifeng Wang, Baidu, China
David Yarowsky, John Hopkins University, USA

Organisation Co-Chairs:

Virach Sornlertlamvanich, NECTEC, Thailand
Hitoshi Isahara, Toyohashi University of Technology, Japan

Workshop Co-Chairs:

Sivaji Bandyopadhyay, Jadavpur University, India
Jong Park, KAIST, Korea
Noriko Kando, NII, Japan

Tutorial Co-Chairs:

Kentaro Inui, Tohoku University, Japan
Wei Gao, The Chinese University of Hong Kong, China
Dawei Song, Robert Gordon University, UK

Demonstration Co-Chairs:

Ken Church, Johns Hopkins University, USA
Yunqing Xia, Tsinghua University, China

Publication Co-Chairs:

Min Zhang, I2R, Singapore
Sudeshna Sarkar, IIT Kharagpur, India

Finance Co-Chairs:

Vilas Wuwongse, AIT, Thailand
Gary Lee, POSTECH, Korea

Sponsorship Co-Chairs:

Asanee Kawtrakul, Kasetsart University, Thailand
Methinee Sirikrai, NECTEC, Thailand
Hiromi Nakaiwa, NTT, Japan

Publicity Committee:

Steven Bird, University of Melbourne, Australia
Le Sun, CIPS, China
Kevin Knight, USC, USA
Nicoletta Calzolari, Istituto di Linguistica Computazionale del CNR, Italy
Thanaruk Theeramunkong, SIIT, Thailand

Webmasters:

Swit Phuvipadawat, Tokyo Institute of Technology, Japan
Wirat Chinnan, SIIT, Thailand

Area Chairs:**Discourse, Dialogue and Pragmatics**

David Schlangen, The University of Potsdam, Germany

Generation /Summarization

Xiaojun Wan, Peking University, China

Information Extraction

Wenjie Li, The Hong Kong Polytechnic University, Hong Kong

Information Retrieval

Gareth Jones, Dublin City University, Ireland

Language Resource

Eneko Agirre, University of the Basque Country, Spain

Machine Translation

David Chiang, USC-ISI, USA

Min Zhang, Institute for Infocomm Research, Singapore

Hua Wu, Baidu, China

Phonology/morphology, POS tagging and chunking, Word Segmentation

Richard Sproat, Oregon Health & Science University, USA

Gary Lee, Pohang University of Science and Technology, Korea

Question Answering

Jun Zhao, Institute of Automation, Chinese Academy of Sciences, China

Semantics

Pushpak Bhattacharyya, Indian Institute of Technology, India

Hinrich Schuetze, University of Stuttgart, Germany

Sentiment Analysis, Opinion Mining and Text Classification

Rafael Banchs, Institute for Infocomm Research, Singapore

Theresa Wilson, Johns Hopkins University, USA

Spoken Language Processing

Chung-Hsien Wu, National Cheng Kung University, Taiwan

Statistical and ML Methods

Miles Osborne, The University of Edinburgh, UK

David Smith, University of Massachusetts Amherst, USA

Syntax and Parsing

Stephen Clark, University of Cambridge, UK
Yusuke Miyao, National Institute of Informatics, Japan

Text Mining and NLP Applications

Juanzi Li, Tsinghua University, China
Patrick Pantel, Microsoft Research, USA

Reviewers

Ahmed Abbasi, Omri Abend, Akiko Aizawa, Ahmet Aker, Enrique Alfonseca, DAUD ALI, Ben Allison, Robin Aly, Alina Andreevskaia, Masayuki Asahara, Ai Azuma

Jing Bai, Alexandra Balahur, Timothy Baldwin, Kalika Bali, Carmen Banea, Srinivas Bangalore, Mohit Bansal, Marco Bbaroni, Roberto Basili, Timo Baumann, Emily Bender, Shane Bergsma, Pushpak Bhattacharyya, Dan Bikel, Wang Bin, Lexi Birch, Michael Bloodgood, Phil Blunsom, Nate Bodestab, Ester Boldrini, Gemma Boleda, Danushka Bollegala, Luc Boruta, Stefan Bott, Chris Brew, Sam Brody, Julian Brooke, Paul Buitelaar, Miriam Butt

Aoife Cahill, Li Cai, Yi Cai, Nicoletta Calzolari, Jaime Carbonell, Marine Carpuat, John Carroll, Paula Carvalho, Suleyman Cetintas, Debasri Chakrabarti, Nate Chambers, Niladri Chatterjee, Wanxiang Che, Berlin Chen, Boxing Chen, Chia-Ping Chen, Hsin-Hsi Chen, Wenliang Chen, Ying Chen, Yufeng Chen, Pu-Jen Cheng, Colin Cherry, Jackie Chi KiCheung, Key-Sun Choi, Monojit Choudhury, Christos Christodoulopoulos, Kenneth Church, Alex Clark, Shay Cohen, Trevor Cohn, Gao Cong, Marta R. Costa-jussa, Paul Crook, Montse Cuadros, Ronan Cummins

Robert Dampier, Kareem Darwish, Dipanjan Das, Niladri Dash, Adrià de Gispert, Daniel de Kok, Eric De La Clergerie, Stijn De Saeger, Steve DeNeefe, Pascal Denis, Ann Devitt, Arantza Diaz de Ilarraza, Anne Diekema, Markus Dreyer, Rebecca Dridan, Jinhua Du, Xiangyu Duan, Amit Dubey, Kevin Duh, Chris Dyer, Michal Dziemianko

Jacob Eisenstein, Michael Elhadad, Micha Elsner, Martin Emms

Angela Fahrni, Hui Fang, Yi Fang, Li Fangtao, Christiane Fellbaum, Raquel Fernandez, Colum Foley, Jennifer Foster, Timothy Fowler, Stella Frank, Guohong Fu, Atsushi Fujii, Kotaro Funakoshi, Hagen Fürstenaue

Matthias Galle, Michael Gamon, Michaela Geierhos, Eugenie Giesbrecht, Alastair Gill, Roxana Girju, Bruno Golenia, Carlos Gomez-Rodriguez, Zhengxian Gong, Matt Gormley, Amit Goyal, João Graça, Jens Grivolla, Iryna Gurevych

Stephanie Haas, Barry Haddow, Eva Hajicova, David Hall, Keith Hall, Xianpei Han, Kazuo Hara, Donna Harman, Kazi Hasan, Chikara Hashimoto, Koiti Hasida, Eva Hasler, Samer Hassan, Claudia Hauff, Xiaodong He, Yulan He, Zhongjun He, Carlos Henriquez, Tsutomu Hirao, Hieu Hoang, Tracy Holloway King, Matthew Honnibal, Mark Hopkins, Meishan Hu, Chien-Lin Huang, Fei Huang, Minlie Huang, Ruizhang Huang, Xiaojiang Huang, Xuanjing Huang, Yun Huang, Zhongqiang Huang

Francisco Iacobelli, Diana Inkpen, Aminul Islam, Ruben Izquierdo

Heng Ji, Sittichai Jiampojarn, Hongfei Jiang, Wenbin Jiang, Xing Jiang, Cai Jie, Rong Jin, Richard Johansson, Hanmin Jung

Sarvnaz Karimi, Daisuke Kawahara, Jun'ichi Kazama, Liadh Kelly, Maxim Khalilov, Mitesh Khapra, Adam Kilgarriff, Byeongchang Kim, Irwin King, Alistair Knott, Philipp Koehn, Rob Koeling, Oskar Kohonen, Mamoru Komachi, Grzegorz Kondrak, Fang Kong, Valia Kordoni, Lili Kotlerman, Zornitsa Kozareva, Wessel Kraaij, Parton Kristen, Lun-Wei Ku, Sudip Kumar Naskar, June-Jei Kuo, Kow Kuroda, Sadao KUROHASHI, Kui-Lam Kwok, Han Kyoung-Soo

Sobha Lalitha Devi, Wai Lam, Joel Lang, Jun Lang, Matt Lease, Cheongjae Lee, Jung-Tae Lee, Sungjin Lee, Tan Lee, Russell Lee-goldman, Alessandro Lenci, Johannes Leveling, Abby Levenberg, Gina-Anne Levow, Baoli Li, Daifeng Li, Haizhou Li, linlin li, Mu Li, Qing Li, Shoushan Li, Sujian Li, Yunyao Li, Shasha Liao, Yuan-Fu Liao, Chin-Yew Lin, Pierre Lison, Ken Litkowski, Marina Litvak, Bing Liu, Fei Liu, Feifan Liu, Kang Liu, Pengyuan Liu, Qun Liu, Shui Liu, Xiaohua Liu, Yang Liu (UT Dallas), Yang Liu (ICT CAS), Yi Liu, Ying Liu, Yiqun Liu, Zhanyi Liu, Hector Llorens, Elena Lloret, Wai-Kit Lo, QIU Long, Adam Lopez, Yajuan Lu

Bin Ma, Yanjun Ma, Walid Magdy, OKUMURA Manabu, Suresh Manandhar, Maria Antonia Marti, David Martinez, Andre Martins, Yuji Matsumoto, Yutaka Matsuo, Takuya Matsuzaki, Mike Maxwell, Jonathan May, Diana McCarthy, David McClosky, Ryan McDonald, Paul McNamee, Beata Megyesi, Donald Metzler, Haitao Mi, Lukas Michelbacher, Dipti Mishra Sharma, Mandar Mitra, Daichi Mochihashi, Saif Mohammed, Behrang Mohit, Karo Moilanen, Christian Monson, Paul Morarescu, Jin'ichi Murakami, Sung Hyon Myaeng

Seung-Hoon Na, Masaaki Nagata, Mikio Nakano, Preslav Nakov, Jason Naradowsky, Vivi Nastase, Roberto Navigli, Mark-Jan Nederhof, Ani Nenkova, Vincent Ng, Truc-Vien T. Nguyen, Eric Nichols, Tadashi Nomoto, Scott Nowson, Andreas Nuernberger, Pierre Nugues

Diarmuid O Seaghdha, Brendan O'Connor, Neil O'Hare, Stephan Oepen, Kemal Oflazer, Kemal Oflazer, Alice Oh, Naoaki Okazaki, Constantin Orasan, Arantxa Otegi, Myle Ott, Jahna Otterbacher, You Ouyang

Alexandre Patry, Soma Paul, Adam Pease, Ted Peders, Wei Peng, Gerald Penn, Sasa Petrovic, Christian Pietsch, Juan Pino, Matt Post, John Prager, Daniel Preotiuc, Matthew Purver

Vahed Qazvinian, Guang Qiu, Chris Quirk

Altaf Rahman, Ganesh Ramakrishnan, Karthik Raman, AnanthakrishnRamanathan, Sujith Ravi, Bunescu Razvan, Jonathon Read, Marta Recasens, Jeremy Reffin, Roi Reichart, Jason Riesa, Verena Rieser, Arndt Riestler, Stefan Riezler, German Rigau, Laura Rimell, Carlos Rodriguez, Kepa Rodriguez, Robert Ross, Michael Roth, Sasha Rush

Kenji Sagae, Benoît Sagot, Agnes Sandor, Anoop Sarkar, Sudeshna Sarkar, Ryohei Sasano, Roser Sauri, Helmut Schmid, Satoshi Sekine, Arulmozi Selvaraj, Pavel Serdyukov, Gao Sheng, Masashi Shimbo, Darla Shockley, Luo Si, Khalil Sima'an, Ben Snyder, Ruihua Song, Young-In Song, Sebastian Spiegler, Valentin Spitkovsky, Caroline Sporleder, Manfred Stede, Mark Steedman, Mark Stevenson, Nicola Stokes, Veselin Stoyanov, Michael Strube, Jian Su, Keh-Yih Su, Zhifang Sui, Aixin Sun, Jun Sun, Weiwei Sun, Mihai Surdeanu

Oscar Tackstrom, Hiroya Takamura, Jianhua Tao, Joel Tetreault, Stefan Thater, Jörg Tiedemann, Ivan Titov, Takenobu Tokunaga, Kentaro Torisawa, Lamia Tounsi, Kristina Toutanova, Roy Tromble, Reut Tsarfaty, Yuen-Hsien Tseng, Hajime Tsukada

Christina Unger, Takehito Utsuro

Antal van den Bosch, Gertjan van Noord, Vasudeva Varma, Silvia Vazquez, Tony Veale, Olga Vechtomova, Sriram Venkatapathy, Yannick Versley, Jesus Vilares, Sami Virpioja, Andreas Vlachos, Piek Vossen

Stephen Wan, Bin Wang, Bo Wang, Dingding Wang, Hsin-Min Wang, Ting Wang, Wei Wang, Zhichun Wang, Taro Watanabe, Yotaro Watanabe, Bonnie Webber, Furu Wei, Richard Wicentowski, Shuly Wintner, Kristian Woodsend, Gang Wu, Zhiyong Wu

Yunqing Xia, Tong Xiao, Xin Xin, Deyi Xiong, Qiu Xipeng, Jun Xu, Ruifeng Xu

Christopher Yang, Grace Yang, Muyun Yang, Yuhang Yang, Zi Yang, Benajiba Yassine, Mark Yatskar, Patrick Ye, Jui-Feng Yeh, Ainur Yessenalina, Scott Wen-tauYih, Bei Yu, Hong Yu

Taras Zagibalov, Benat Zaporain, Alessandra Zarccone, Duo Zhang, Hao Zhang, Jiajun Zhang, Jing Zhang, Lanbo Zhang, Lei Zhang, Min Zhang, Qi Zhang, Yi Zhang (UCSC), Yi Zhang (DFKI), Yue

Zhang, Shiqi Zhao, Tiejun Zhao, Haitao Zheng, Zhi Zhong, Bowen Zhou, Dong Zhou, GuoDong
Zhou, Qiang Zhou, Yu Zhou, Muhua Zhu, Xiaodan Zhu, Chengqing Zong

Table of Contents

Part A: Full Papers

<i>Analyzing the Dynamics of Research by Extracting Key Aspects of Scientific Papers</i> Sonal Gupta and Christopher Manning	1
<i>Dependency-directed Tree Kernel-based Protein-Protein Interaction Extraction from Biomedical Literature</i> Longhua Qian and Guodong Zhou	10
<i>Learning Logical Structures of Paragraphs in Legal Articles</i> Ngo Xuan Bach, Nguyen Le Minh, Tran Thi Oanh and Akira Shimazu	20
<i>Extracting Pre-ordering Rules from Predicate-Argument Structures</i> Xianchao Wu, Katsuhito Sudoh, Kevin Duh, Hajime Tsukada and Masaaki Nagata	29
<i>Context-Sensitive Syntactic Source-Reordering by Statistical Transduction</i> Maxim Khalilov and Khalil Sima'an	38
<i>Discriminative Phrase-based Lexicalized Reordering Models using Weighted Reordering Graphs</i> Wang Ling, João Graça, David Martins de Matos, Isabel Trancoso and Alan W Black	47
<i>Active Learning Strategies for Support Vector Machines, Application to Temporal Relation Classification</i> Seyed Abolghasem Mirroshandel, Gholamreza Ghassem-Sani and Alexis Nasr	56
<i>A Fast Accurate Two-stage Training Algorithm for L1-regularized CRFs with Heuristic Line Search Strategy</i> Jinlong Zhou, Xipeng Qiu and Xuanjing Huang	65
<i>Automatic Topic Model Adaptation for Sentiment Analysis in Structured Domains</i> Geoffrey Levine and Gerald DeJong	75
<i>Multi-modal Reference Resolution in Situated Dialogue by Integrating Linguistic and Extra-Linguistic Clues</i> Ryu Iida, Masaaki Yasuhara and Takenobu Tokunaga	84
<i>Single and multi-objective optimization for feature selection in anaphora resolution</i> Sriparna Saha, Asif Ekbal, Olga Uryupina and Massimo Poesio	93
<i>A Unified Event Coreference Resolution by Integrating Multiple Resolvers</i> Bin Chen, Jian Su, Sinno Jialin Pan and Chew Lim Tan	102
<i>Handling verb phrase morphology in highly inflected Indian languages for Machine Translation</i> Ankur Gandhe, Rashmi Gangadharaiah, Karthik Visweswariah and Ananthakrishnan Ramanathan	111
<i>Japanese Pronunciation Prediction as Phrasal Statistical Machine Translation</i> Jun Hatori and Hisami Suzuki	120
<i>Comparing Two Techniques for Learning Transliteration Models Using a Parallel Corpus</i> Hassan Sajjad, Nadir Durrani, Helmut Schmid and Alexander Fraser	129

<i>A Semantic-Specific Model for Chinese Named Entity Translation</i> Yufeng Chen and Chengqing Zong	138
<i>Mining Revision Log of Language Learning SNS for Automated Japanese Error Correction of Second Language Learners</i> Tomoya Mizumoto, Mamoru Komachi, Masaaki Nagata and Yuji Matsumoto	147
<i>Modality Specific Meta Features for Authorship Attribution in Web Forum Posts</i> Thamar Solorio, Sangita Pillay, Sindhu Raghavan and Manuel Montes-Gomez	156
<i>Keyphrase Extraction from Online News Using Binary Integer Programming</i> Zhuoye Ding, Qi Zhang and Xuanjing Huang	165
<i>Improving Related Entity Finding via Incorporating Homepages and Recognizing Fine-grained Entities</i> Youzheng Wu, Chiori Hori, Hisashi Kawai and Hideki Kashioka	174
<i>Enhancing Active Learning for Semantic Role Labeling via Compressed Dependency Trees</i> Chenhua Chen, Alexis Palmer and Caroline Sporleder	183
<i>Semantic Role Labeling Without Treebanks?</i> Stephen Boxwell, Chris Brew, Jason Baldrige, Dennis Mehay and Sujith Ravi	192
<i>Japanese Predicate Argument Structure Analysis Exploiting Argument Position and Type</i> Yuta Hayashibe, Mamoru Komachi and Yuji Matsumoto	201
<i>An Empirical Study on Compositionality in Compound Nouns</i> Siva Reddy, Diana McCarthy and Suresh Manandhar	210
<i>Feature-Rich Log-Linear Lexical Model for Latent Variable PCFG Grammars</i> Zhongqiang Huang and Mary Harper	219
<i>Improving Dependency Parsing with Fined-Grained Features</i> Guangyou Zhou, Li Cai, Kang Liu and Jun Zhao	228
<i>Natural Language Programming Using Class Sequential Rules</i> Cohan Sujay Carlos	237
<i>Treeblazing: Using External Treebanks to Filter Parse Forests for Parse Selection and Treebanking</i> Andrew MacKinlay, Rebecca Dridan, Dan Flickinger, Stephan Oepen and Timothy Baldwin ..	246
<i>Cross-Language Entity Linking</i> Paul McNamee, James Mayfield, Dawn Lawrie, Douglas Oard and David Doermann	255
<i>Generating Chinese Named Entity Data from a Parallel Corpus</i> Ruiji Fu, Bing Qin and Ting Liu	264
<i>Learning the Latent Topics for Question Retrieval in Community QA</i> Li Cai, Guangyou Zhou, Kang Liu and Jun Zhao	273
<i>Identifying Event Descriptions using Co-training with Online News Summaries</i> William Yang Wang, Kapil Thadani and Kathleen McKeown	282
<i>Automatic Labeling of Voiced Consonants for Morphological Analysis of Modern Japanese Literature</i> Teruaki Oka, Mamoru Komachi, Toshinobu Ogiso and Yuji Matsumoto	292

<i>S³ - Statistical Sandhi Splitting</i>	
Abhiram Natarajan and Eugene Charniak	301
<i>Improving Chinese Word Segmentation and POS Tagging with Semi-supervised Methods Using Large Auto-Analyzed Data</i>	
Yiou Wang, Jun'ichi Kazama, Yoshimasa Tsuruoka, Wenliang Chen, Yujie Zhang and Kentaro Torisawa	309
<i>CODACT: Towards Identifying Orthographic Variants in Dialectal Arabic</i>	
Pradeep Dasigi and Mona Diab	318
<i>Enhancing the HL-SOT Approach to Sentiment Analysis via a Localized Feature Selection Framework</i>	
Wei and Jon Atle Gulla	327
<i>Fine-Grained Sentiment Analysis with Structural Features</i>	
Cäcilia Zirn, Mathias Niepert, Heiner Stuckenschmidt and Michael Strube	336
<i>Predicting Opinion Dependency Relations for Opinion Analysis</i>	
Lun-Wei Ku, Ting-Hao Huang and Hsin-Hsi Chen	345
<i>Detecting and Blocking False Sentiment Propagation</i>	
Hye-Jin Min and Jong C. Park	354
<i>Efficient induction of probabilistic word classes with LDA</i>	
Grzegorz Chrupala	363
<i>Quality-biased Ranking of Short Texts in Microblogging Services</i>	
Minlie Huang, Yi Yang and Xiaoyan Zhu	373
<i>Labeling Unlabeled Data using Cross-Language Guided Clustering</i>	
Sachindra Joshi, Danish Contractor and Sumit Negi	383
<i>Extracting Relation Descriptors with Conditional Random Fields</i>	
Yaliang Li, Jing Jiang, Hai Leong Chieu and Kian Ming A. Chai	392
<i>Attribute Extraction from Synthetic Web Search Queries</i>	
Marius Pasca	401
<i>Japanese Abbreviation Expansion with Query and Clickthrough Logs</i>	
Kei Uchiumi, Mamoru Komachi, Keigo Machinaga, Toshiyuki Maezawa, Toshinori Satou and Yoshinori Kobayashi	410
<i>Mining Parallel Documents Using Low Bandwidth and High Precision CLIR from the Heterogeneous Web</i>	
Simon Shi, Pascale Fung, Emmanuel Prochasson, Chi-kiu Lo and Dekai Wu	420
<i>Crawling Back and Forth: Using Back and Out Links to Locate Bilingual Sites</i>	
Luciano Barbosa, Srinivas Bangalore and Vivek Kumar Rangarajan Sridhar	429
<i>Grammar Induction from Text Using Small Syntactic Prototypes</i>	
Prachya Boonkwan and Mark Steedman	438
<i>Transferring Syntactic Relations from English to Hindi Using Alignments on Local Word Groups</i>	
Aswarth Dara, Prashanth Mannem, Hemanth Sagar Bayyarapu and Avinesh PVS	447

<i>Generative Modeling of Coordination by Factoring Parallelism and Selectional Preferences</i> Daisuke Kawahara and Sadao Kurohashi	456
<i>Syntactic Parsing for Ranking-Based Coreference Resolution</i> Altat Rahman and Vincent Ng	465
<i>TriS: A Statistical Sentence Simplifier with Log-linear Models and Margin-based Discriminative Training</i> Nguyen Bach, Qin Gao, Stephan Vogel and Alex Waibel	474
<i>Social Summarization via Automatically Discovered Social Context</i> Po Hu, Cheng Sun, Longfei Wu, Donghong Ji and Chong Teng	483
<i>Simultaneous Clustering and Noise Detection for Theme-based Summarization</i> Xiaoyan Cai, Renxian Zhang, Dehong Gao and Wenjie Li	491
<i>Extractive Summarization Method for Contact Center Dialogues based on Call Logs</i> Akihiro Tamura, Kai Ishikawa, Masahiro Saikou and Masaaki Tsuchida	500
<i>Indexing Spoken Documents with Hierarchical Semantic Structures: Semantic Tree-to-string Alignment Models</i> Xiaodan Zhu, Colin Cherry and Gerald Penn	509
<i>Structured and Extended Named Entity Evaluation in Automatic Speech Transcriptions</i> Olivier Galibert, Sophie Rosset, Cyril Grouin, Pierre Zweigenbaum and Ludovic Quintard	518
<i>Normalising Audio Transcriptions for Unwritten Languages</i> Adel Foda and Steven Bird	527
<i>Similarity Based Language Model Construction for Voice Activated Open-Domain Question Answering</i> Istvan Varga, Kiyonori Ohtake, Kentaro Torisawa, Stijn De Saeger, Teruhisa Misu, Shigeki Matsuda and Jun'ichi Kazama	536
<i>The application of chordal graphs to inferring phylogenetic trees of languages</i> Jessica Enright and Grzegorz Kondrak	545
<i>Cross-domain Feature Selection for Language Identification</i> Marco Lui and Timothy Baldwin	553
<i>A Wikipedia-LDA Model for Entity Linking with Batch Size Changing Instance Selection</i> Wei Zhang, Jian Su and Chew-Lim Tan	562
<i>Discovering Latent Concepts and Exploiting Ontological Features for Semantic Text Search</i> Vuong M. Ngo and Tru H. Cao	571
<i>CLGVSM: Adapting Generalized Vector Space Model to Cross-lingual Document Clustering</i> Guoyu Tang, Yunqing Xia, Min Zhang, Haizhou Li and Fang Zheng	580
<i>Thread Cleaning and Merging for Microblog Topic Detection</i> Jianfeng Zhang, Yunqing Xia, Bin Ma, Jianmin Yao and Yu Hong	589
<i>Training a BN-based user model for dialogue simulation with missing data</i> Stéphane Rossignol, Olivier Pietquin and Michel Iannotto	598
<i>Automatic identification of general and specific sentences by leveraging discourse annotations</i> Annie Louis and Ani Nenkova	605

<i>A POS-based Ensemble Model for Cross-domain Sentiment Classification</i> Rui Xia and Chengqing Zong	614
<i>Ensemble-style Self-training on Citation Classification</i> Cailing Dong and Ulrich Schäfer	623
<i>Back to the Roots of Genres: Text Classification by Language Function</i> Henning Wachsmuth and Kathrin Bujna	632
<i>Transductive Minimum Error Rate Training for Statistical Machine Translation</i> Yinggong Zhao, Shujie Liu, Yangsheng Ji, Jiajun Chen and Guodong Zhou	641
<i>Distributed Minimum Error Rate Training of SMT using Particle Swarm Optimization</i> Jun Suzuki, Kevin Duh and Masaaki Nagata	649
<i>Going Beyond Word Cooccurrences in Global Lexical Selection for Statistical Machine Translation using a Multilayer Perceptron</i> Alexandre Patry and Philippe Langlais	658
<i>System Combination Using Discriminative Cross-Adaptation</i> Jacob Devlin, Antti-Veikko Rosti, Sankaranarayanan Ananthakrishnan and Spyros Matsoukas	667
<i>Word Sense Disambiguation by Combining Labeled Data Expansion and Semi-Supervised Learning Method</i> Sanae Fujita and Akinori Fujino	676
<i>Combining ConceptNet and WordNet for Word Sense Disambiguation</i> Junpeng Chen and Juan Liu	686
<i>It Takes Two to Tango: A Bilingual Unsupervised Approach for Estimating Sense Distributions using Expectation Maximization</i> Mitesh M Khapra, Salil Joshi and Pushpak Bhattacharyya	695
<i>Dynamic and Static Prototype Vectors for Semantic Composition</i> Siva Reddy, Ioannis Klapaftis, Diana McCarthy and Suresh Manandhar	705
<i>Using Prediction from Sentential Scope to Build a Pseudo Co-Testing Learner for Event Extraction</i> Shasha Liao and Ralph Grishman	714
<i>Text Segmentation and Graph-based Method for Template Filling in Information Extraction</i> Ludovic Jean-Louis, Romaric Besançon and Olivier Ferret	723
<i>Joint Distant and Direct Supervision for Relation Extraction</i> Truc-Vien T. Nguyen and Alessandro Moschitti	732
<i>A Cross-lingual Annotation Projection-based Self-supervision Approach for Open Information Extraction</i> Seokhwan Kim, Minwoo Jeong, Jonghoon Lee and Gary Geunbae Lee	741
<i>Exploring Difficulties in Parsing Imperatives and Questions</i> Tadayoshi Hara, Takuya Matsuzaki, Yusuke Miyao and Jun'ichi Tsujii	749
<i>A Discriminative Approach to Japanese Zero Anaphora Resolution with Large-scale Lexicalized Case Frames</i> Ryohei Sasano and Sadao Kurohashi	758

<i>An Empirical Comparison of Unknown Word Prediction Methods</i> Kostadin Cholakov, Gertjan van Noord, Valia Kordoni and Yi Zhang	767
<i>Training Dependency Parsers from Partially Annotated Corpora</i> Daniel Flannery, Yusuke Miayo, Graham Neubig and Shinsuke Mori	776
<i>A Breadth-First Representation for Tree Matching in Large Scale Forest-Based Translation</i> Sumukh Ghodke, Steven Bird and Rui Zhang	785
<i>Bayesian Subtree Alignment Model based on Dependency Trees</i> Toshiaki Nakazawa and Sadao Kurohashi	794
<i>Enriching SMT Training Data via Paraphrasing</i> Wei He, Shiqi Zhao, Haifeng Wang and Ting Liu	803
<i>Translation Quality Indicators for Pivot-based Statistical MT</i> Michael Paul and Eiichiro Sumita	811
<i>Source Error-Projection for Sample Selection in Phrase-Based SMT for Resource-Poor Languages</i> Sankaranarayanan Ananthakrishnan, Shiv Vitaladevuni, Rohit Prasad and Prem Natarajan	819
<i>A Named Entity Recognition Method based on Decomposition and Concatenation of Word Chunks</i> Tomoya Iwakura, Hiroya Takamura and Manabu Okumura	828
<i>Extract Chinese Unknown Words from a Large-scale Corpus Using Morphological and Distributional Evidences</i> Kaixu Zhang, Ruining Wang, Ping Xue and Maosong Sun	837
<i>Entity Disambiguation Using a Markov-Logic Network</i> Hong-Jie Dai, Richard Tzong-Han Tsai and Wen-Lian Hsu	846
<i>Named Entity Recognition in Chinese News Comments on the Web</i> Xiaojun Wan, Liang Zong, Xiaojiang Huang, Tengfei Ma, Houping Jia, Yuqian Wu and Jianguo Xiao	856
<i>Clustering Semantically Equivalent Words into Cognate Sets in Multilingual Lists</i> Bradley Hauer and Grzegorz Kondrak	865
<i>Extending WordNet with Hypernyms and Siblings Acquired from Wikipedia</i> Ichiro Yamada, Jong-Hoon Oh, Chikara Hashimoto, Kentaro Torisawa, Jun'ichi Kazama, Stijn De Saeger and Takuya Kawada	874
<i>What Psycholinguists Know About Chemistry: Aligning Wiktionary and WordNet for Increased Domain Coverage</i> Christian M. Meyer and Iryna Gurevych	883
<i>From News to Comment: Resources and Benchmarks for Parsing the Language of Web 2.0</i> Jennifer Foster, Ozlem Cetinoglu, Joachim Wagner, Joseph Le Roux, Joakim Nivre, Deirdre Hogan and Josef van Genabith	893
<i>Toward Finding Semantic Relations not Written in a Single Sentence: An Inference Method using Auto-Discovered Rules</i> Masaaki Tsuchida, Kentaro Torisawa, Stijn De Saeger, Jong Hoon Oh, Jun'ichi Kazama, Chikara Hashimoto and Hayato Ohwada	902

<i>Fleshing it out: A Supervised Approach to MWE-token and MWE-type Classification</i> Richard Fothergill and Timothy Baldwin	911
<i>Identification of relations between answers with global constraints for Community-based Question Answering services</i> Hikaru Yokono, Takaaki Hasegawa, Genichiro Kikui and Manabu Okumura	920
<i>Automatically Generating Questions from Queries for Community-based Question Answering</i> Shiqi Zhao, Haifeng Wang, Chao Li, Ting Liu and Yi Guan	929
<i>Question classification based on an extended class sequential rule model</i> Zijing Hui, Juan Liu and Lumei Ouyang	938
<i>K2Q: Generating Natural Language Questions from Keywords with User Refinements</i> Zhicheng Zheng, Xiance Si, Edward Chang and Xiaoyan Zhu	947
<i>Answering Complex Questions via Exploiting Social Q&A Collection</i> Youzheng Wu, Chiori Hori, Hisashi Kawai and Hideki Kashioka	956
<i>Safety Information Mining — What can NLP do in a disaster—</i> Graham Neubig, Yuichiroh Matsubayashi, Masato Hagiwara and Koji Murakami	965
<i>A Character-Level Machine Translation Approach for Normalization of SMS Abbreviations</i> Deana Pennell and Yang Liu	974
<i>Using Text Reviews for Product Entity Completion</i> Mrinmaya Sachan, Tanveer Faruque, L. V. Subramaniam and Mukesh Mohania	983
<i>Mining bilingual topic hierarchies from unaligned text</i> Sumit Negi	992
<i>Efficient Near-Duplicate Detection for Q&A Forum</i> Yan Wu, Qi Zhang and Xuanjing Huang	1001
<i>A Graph-based Method for Entity Linking</i> Yuhang Guo, Wanxiang Che, Ting Liu and Sheng Li	1010
<i>Harvesting Related Entities with a Search Engine</i> Shuqi Sun, Shiqi Zhao, Muyun Yang, Haifeng Wang and Sheng Li	1019
<i>Acquiring Strongly-related Events using Predicate-argument Co-occurring Statistics and Case Frames</i> Tomohide Shibata and Sadao Kurohashi	1028
<i>Relevance Feedback using Latent Information</i> Jun Harashima and Sadao Kurohashi	1037
<i>Passage Retrieval for Information Extraction using Distant Supervision</i> Wei Xu, Ralph Grishman and Le Zhao	1046
<i>Using Context Inference to Improve Sentence Ordering for Multi-document Summarization</i> Peifeng Li, Guangxi Deng and Qiaoming Zhu	1055
<i>Enhancing extraction based summarization with outside word space</i> Christian Smith and Arne Jönsson	1062

<i>Shallow Discourse Parsing with Conditional Random Fields</i> Sucheta Ghosh, Richard Johansson, Giuseppe Riccardi and Sara Tonelli	1071
<i>Relational Lasso —An Improved Method Using the Relations Among Features—</i> Kotaro Kitagawa and Kumiko Tanaka-Ishii	1080
<i>Enhance Top-down method with Meta-Classification for Very Large-scale Hierarchical Classification</i> Xiao-lin Wang, Hai Zhao and Bao-Liang Lu	1089
<i>Using Syntactic and Shallow Semantic Kernels to Improve Multi-Modality Manifold-Ranking for Topic-Focused Multi-Document Summarization</i> Yllias Chali, Sadid A. Hasan and Kaisar Imam	1098
<i>Automatic Determination of a Domain Adaptation Method for Word Sense Disambiguation Using Decision Tree Learning</i> Kanako Komiya and Manabu Okumura	1107
<i>Learning from Chinese-English Parallel Data for Chinese Tense Prediction</i> Feifan Liu, Fei Liu and Yang Liu	1116
<i>Jointly Extracting Japanese Predicate-Argument Relation with Markov Logic</i> Katsumasa Yoshikawa, Masayuki Asahara and Yuji Matsumoto	1125
<i>Word Meaning in Context: A Simple and Effective Vector Model</i> Stefan Thater, Hagen Fürstenaу and Manfred Pinkal	1134
<i>Automatic Analysis of Semantic Coherence in Academic Abstracts Written in Portuguese</i> Vinícius Mourão Alves de Souza and Valéria Delisandra Feltrim	1144
<i>Sentence Subjectivity Detection with Weakly-Supervised Learning</i> Chenghua Lin, Yulan He and Richard Everson	1153
<i>Opinion Expression Mining by Exploiting Keyphrase Extraction</i> Gábor Berend	1162
<i>Extracting Resource Terms for Sentiment Analysis</i> Lei Zhang and Bing Liu	1171
<i>Towards Context-Based Subjectivity Analysis</i> Farah Benamara, Baptiste Chardon, Yannick Mathieu and Vladimir Popescu	1180
<i>Compression Methods by Code Mapping and Code Dividing for Chinese Dictionary Stored in a Double-Array Trie</i> Huidan Liu, Minghua Nuo, Longlong Ma, Jian Wu and Yeping He	1189
<i>Functional Elements and POS Categories</i> Qiuye Zhao and Mitch Marcus	1198
<i>Joint Alignment and Artificial Data Generation: An Empirical Study of Pivot-based Machine Transliteration</i> Min Zhang, Xiangyu Duan, Ming Liu, Yunqing Xia and Haizhou Li	1207
<i>Incremental Joint POS Tagging and Dependency Parsing in Chinese</i> Jun Hatori, Takuya Matsuzaki, Yusuke Miyao and Jun'ichi Tsujii	1216

<i>Extending the adverbial coverage of a NLP oriented resource for French</i> Elsa Tolone and Stavroula Voyatzi	1225
<i>Linguistic Phenomena, Analyses, and Representations: Understanding Conversion between Treebanks</i> Rajesh Bhatt, Owen Rambow and Fei Xia	1234
<i>Automatic Transformation of the Thai Categorical Grammar Treebank to Dependency Trees</i> Christian Rishøj, Taneth Ruangrajitpakorn, Prachya Boonkwan and Thepchai Supnithi	1243
<i>Parse Reranking Based on Higher-Order Lexical Dependencies</i> Zhiguo Wang and Chengqing Zong	1251
<i>Improving Part-of-speech Tagging for Context-free Parsing</i> Xiao Chen and Chunyu Kit	1260
<i>Models Cascade for Tree-Structured Named Entity Detection</i> Marco Dinarelli and Sophie Rosset	1269
<i>Clausal parsing helps data-driven dependency parsing: Experiments with Hindi</i> Samar Husain, Phani Gadde, Joakim Nivre and Rajeev Sangal	1279
<i>Word-reordering for Statistical Machine Translation Using Trigram Language Model</i> Jing He and Hongyu Liang	1288
<i>Extracting Hierarchical Rules from a Weighted Alignment Matrix</i> Zhaopeng Tu, Yang Liu, Qun Liu and Shouxun Lin	1294
<i>Integration of Reduplicated Multiword Expressions and Named Entities in a Phrase Based Statistical Machine Translation System</i> Thoudam Doren Singh and Sivaji Bandyopadhyay	1304
<i>Regularizing Mono- and Bi-Word Models for Word Alignment</i> Thomas Schoenemann	1313
<i>Parametric Weighting of Parallel Data for Statistical Machine Translation</i> Kashif Shah, Loïc Barrault and Holger Schwenk	1323
<i>An Effective and Robust Framework for Transliteration Exploration</i> EA-EE JAN, Niyu Ge, Shih-Hsiang Lin and Berlin Chen	1332

Part B: Short Papers

<i>An Evaluation of Alternative Strategies for Implementing Dialogue Policies Using Statistical Classification and Hand-Authored Rules</i>	
David DeVault, Anton Leuski and Kenji Sagae	1341
<i>Reducing Asymmetry between language-pairs to Improve Alignment and Translation Quality</i>	
Rashmi Gangadharaiah	1346
<i>Clause-Based Reordering Constraints to Improve Statistical Machine Translation</i>	
Ananthakrishnan Ramanathan, Pushpak Bhattacharyya, Karthik Visweswariah, Kushal Ladha and Ankur Gandhe	1351
<i>Generalized Minimum Bayes Risk System Combination</i>	
Kevin Duh, Katsuhito Sudoh, Xianchao Wu, Hajime Tsukada and Masaaki Nagata	1356
<i>Enhancing scarce-resource language translation through pivot combinations</i>	
Marta R. Costa-jussà, Carlos Henríquez and Rafael E. Banchs	1361
<i>A Baseline System for Chinese Near-Synonym Choice</i>	
Liang-Chih Yu, Wei-Nan Chien and Shih-Ting Chen	1366
<i>Cluster Labelling based on Concepts in a Machine-Readable Dictionary</i>	
Fumiyo Fukumoto and Yoshimi Suzuki	1371
<i>Text Patterns and Compression Models for Semantic Class Learning</i>	
Chung-Yao Chuang, Yi-Hsun Lee and Wen-Lian Hsu	1376
<i>Potts Model on the Case Fillers for Word Sense Disambiguation</i>	
Hiroya Takamura and Manabu Okumura	1382
<i>Improving Word Sense Induction by Exploiting Semantic Relevance</i>	
Zhenzhong Zhang and Le Sun	1387
<i>Predicting Word Clipping with Latent Semantic Analysis</i>	
Julian Brooke, Tong Wang and Graeme Hirst	1392
<i>A Semantic Relatedness Measure Based on Combined Encyclopedic, Ontological and Collocational Knowledge</i>	
Yannis Haralambous and Vitaly Klyuev	1397
<i>Going Beyond Text: A Hybrid Image-Text Approach for Measuring Word Relatedness</i>	
Chee Wee Leong and Rada Mihalcea	1403
<i>Domain Independent Model for Product Attribute Extraction from User Reviews using Wikipedia</i>	
Sudheer Kovelamudi, Sethu Ramalingam, Arpit Sood and Vasudeva Varma	1408
<i>Finding Problem Solving Threads in Online Forum</i>	
Zhonghua Qu and Yang Liu	1413

<i>Compiling Learner Corpus Data of Linguistic Output and Language Processing in Speaking, Listening, Writing, and Reading</i>	
Katsunori Kotani, Takehiko Yoshimi, Hiroaki Nanjo and Hitoshi Isahara	1418
<i>Mining the Sentiment Expectation of Nouns Using Bootstrapping Method</i>	
Miaomiao Wen and Yunfang Wu	1423
<i>An Analysis of Questions in a Q&A Site Resubmitted Based on Indications of Unclear Points of Original Questions</i>	
Masahiro Kojima, Yasuhiko Watanabe and Yoshihiro Okada	1428
<i>Diversifying Information Needs in Results of Question Retrieval</i>	
Yaoyun Zhang, Xiaolong Wang, Xuan Wang, Ruifeng Xu, Jun Xu and ShiXi Fan	1432
<i>Beyond Normalization: Pragmatics of Word Form in Text Messages</i>	
Tyler Baldwin and Joyce Chai	1437
<i>Chinese Discourse Relation Recognition</i>	
Hen-Hsen Huang and Hsin-Hsi Chen	1442
<i>Improving Chinese POS Tagging with Dependency Parsing</i>	
Zhenghua Li, Wanxiang Che and Ting Liu	1447
<i>Exploring self training for Hindi dependency parsing</i>	
Rahul Goutam and Bharat Ram Ambati	1452
<i>Reduction of Search Space to Annotate Monolingual Corpora</i>	
Prajol Shrestha, Christine Jacquin and Beatrice Daille	1457
<i>Toward a Parallel Corpus of Spoken Cantonese and Written Chinese</i>	
John Lee	1462
<i>Query Expansion for IR using Knowledge-Based Relatedness</i>	
Arantxa Otegi, Xabier Arregi and Eneko Agirre	1467
<i>Word Sense Disambiguation Corpora Acquisition via Confirmation Code</i>	
Wanxiang Che and Ting Liu	1472

Conference Program

November 9

(9:00-9:20) Welcome

(9:20-10:20) Keynote – Mengqiu Wang (Baidu)

(10:20 - 10:50) Coffee Break

(10:50-12:05) Text Mining I

Analyzing the Dynamics of Research by Extracting Key Aspects of Scientific Papers
Sonal Gupta and Christopher Manning

Dependency-directed Tree Kernel-based Protein-Protein Interaction Extraction from Biomedical Literature
Longhua Qian and Guodong Zhou

Learning Logical Structures of Paragraphs in Legal Articles
Ngo Xuan Bach, Nguyen Le Minh, Tran Thi Oanh and Akira Shimazu

(10:50-12:05) Machine Translation I

Extracting Pre-ordering Rules from Predicate-Argument Structures
Xianchao Wu, Katsuhito Sudoh, Kevin Duh, Hajime Tsukada and Masaaki Nagata

Context-Sensitive Syntactic Source-Reordering by Statistical Transduction
Maxim Khalilov and Khalil Sima'an

Discriminative Phrase-based Lexicalized Reordering Models using Weighted Re-ordering Graphs
Wang Ling, João Graça, David Martins de Matos, Isabel Trancoso and Alan W Black

November 9 (continued)

(10:50-12:05) Machine Learning I

Active Learning Strategies for Support Vector Machines, Application to Temporal Relation Classification

Seyed Abolghasem Mirroshandel, Gholamreza Ghassem-Sani and Alexis Nasr

A Fast Accurate Two-stage Training Algorithm for L1-regularized CRFs with Heuristic Line Search Strategy

Jinlong Zhou, Xipeng Qiu and Xuanjing Huang

Automatic Topic Model Adaptation for Sentiment Analysis in Structured Domains

Geoffrey Levine and Gerald DeJong

(10:50-12:05) Discourse / Dialog I

Multi-modal Reference Resolution in Situated Dialogue by Integrating Linguistic and Extra-Linguistic Clues

Ryu Iida, Masaaki Yasuhara and Takenobu Tokunaga

Single and multi-objective optimization for feature selection in anaphora resolution

Sriparna Saha, Asif Ekbal, Olga Uryupina and Massimo Poesio

A Unified Event Coreference Resolution by Integrating Multiple Resolvers

Bin Chen, Jian Su, Sinno Jialin Pan and Chew Lim Tan

(12:05-13:30) Lunch

(13:30-15:10) Machine Translation II

Handling verb phrase morphology in highly inflected Indian languages for Machine Translation

Ankur Gandhe, Rashmi Gangadharaiah, Karthik Visweswariah and Ananthkrishnan Ramanathan

Japanese Pronunciation Prediction as Phrasal Statistical Machine Translation

Jun Hatori and Hisami Suzuki

Comparing Two Techniques for Learning Transliteration Models Using a Parallel Corpus

Hassan Sajjad, Nadir Durrani, Helmut Schmid and Alexander Fraser

A Semantic-Specific Model for Chinese Named Entity Translation

Yufeng Chen and Chengqing Zong

November 9 (continued)

(13:30-15:10) Text Mining II

Mining Revision Log of Language Learning SNS for Automated Japanese Error Correction of Second Language Learners

Tomoya Mizumoto, Mamoru Komachi, Masaaki Nagata and Yuji Matsumoto

Modality Specific Meta Features for Authorship Attribution in Web Forum Posts

Thamar Solorio, Sangita Pillay, Sindhu Raghavan and Manuel Montes-Gomez

Keyphrase Extraction from Online News Using Binary Integer Programming

Zhuoye Ding, Qi Zhang and Xuanjing Huang

Improving Related Entity Finding via Incorporating Homepages and Recognizing Fine-grained Entities

Youzheng Wu, Chiori Hori, Hisashi Kawai and Hideki Kashioka

(13:30-15:10) Semantics I

Enhancing Active Learning for Semantic Role Labeling via Compressed Dependency Trees

Chenhua Chen, Alexis Palmer and Caroline Sporleder

Semantic Role Labeling Without Treebanks?

Stephen Boxwell, Chris Brew, Jason Baldridge, Dennis Mehay and Sujith Ravi

Japanese Predicate Argument Structure Analysis Exploiting Argument Position and Type

Yuta Hayashibe, Mamoru Komachi and Yuji Matsumoto

An Empirical Study on Compositionality in Compound Nouns

Siva Reddy, Diana McCarthy and Suresh Manandhar

November 9 (continued)

(13:30-15:10) Syntax I

Feature-Rich Log-Linear Lexical Model for Latent Variable PCFG Grammars
Zhongqiang Huang and Mary Harper

Improving Dependency Parsing with Fined-Grained Features
Guangyou Zhou, Li Cai, Kang Liu and Jun Zhao

Natural Language Programming Using Class Sequential Rules
Cohan Sujay Carlos

Treeblazing: Using External Treebanks to Filter Parse Forests for Parse Selection and Treebanking
Andrew MacKinlay, Rebecca Dridan, Dan Flickinger, Stephan Oepen and Timothy Baldwin

(15:10-15:30) Coffee Break

(15:30-17:10) Text Mining / Information Extraction

Cross-Language Entity Linking
Paul McNamee, James Mayfield, Dawn Lawrie, Douglas Oard and David Doermann

Generating Chinese Named Entity Data from a Parallel Corpus
Ruiji Fu, Bing Qin and Ting Liu

Learning the Latent Topics for Question Retrieval in Community QA
Li Cai, Guangyou Zhou, Kang Liu and Jun Zhao

Identifying Event Descriptions using Co-training with Online News Summaries
William Yang Wang, Kapil Thadani and Kathleen McKeown

November 9 (continued)

(15:30-17:10) Phonology / Morphology I

Automatic Labeling of Voiced Consonants for Morphological Analysis of Modern Japanese Literature

Teruaki Oka, Mamoru Komachi, Toshinobu Ogiso and Yuji Matsumoto

S³ - Statistical Sandhi Splitting

Abhiram Natarajan and Eugene Charniak

Improving Chinese Word Segmentation and POS Tagging with Semi-supervised Methods Using Large Auto-Analyzed Data

Yiou Wang, Jun'ichi Kazama, Yoshimasa Tsuruoka, Wenliang Chen, Yujie Zhang and Kentaro Torisawa

CODACT: Towards Identifying Orthographic Variants in Dialectal Arabic

Pradeep Dasigi and Mona Diab

(15:30-17:10) Sentiment I

Enhancing the HL-SOT Approach to Sentiment Analysis via a Localized Feature Selection Framework

Wei and Jon Atle Gulla

Fine-Grained Sentiment Analysis with Structural Features

Cäcilia Zirn, Mathias Niepert, Heiner Stuckenschmidt and Michael Strube

Predicting Opinion Dependency Relations for Opinion Analysis

Lun-Wei Ku, Ting-Hao Huang and Hsin-Hsi Chen

Detecting and Blocking False Sentiment Propagation

Hye-Jin Min and Jong C. Park

November 9 (continued)

(15:30-17:10) Machine Learning II

Efficient induction of probabilistic word classes with LDA

Grzegorz Chrupala

Quality-biased Ranking of Short Texts in Microblogging Services

Minlie Huang, Yi Yang and Xiaoyan Zhu

Labeling Unlabeled Data using Cross-Language Guided Clustering

Sachindra Joshi, Danish Contractor and Sumit Negi

Extracting Relation Descriptors with Conditional Random Fields

Yaliang Li, Jing Jiang, Hai Leong Chieu and Kian Ming A. Chai

(18:00-21:00) Poster/Demo (Pre Function Level 1) + Reception

November 10

(9:00-10:00) Keynote – Matt Lease (University of Texas)

(10:00 - 10:30) Coffee Break

(10:30 - 12:10) Information Retrieval I

Attribute Extraction from Synthetic Web Search Queries

Marius Pasca

Japanese Abbreviation Expansion with Query and Clickthrough Logs

Kei Uchiumi, Mamoru Komachi, Keigo Machinaga, Toshiyuki Maezawa, Toshinori Satou and Yoshinori Kobayashi

Mining Parallel Documents Using Low Bandwidth and High Precision CLIR from the Heterogeneous Web

Simon Shi, Pascale Fung, Emmanuel Prochasson, Chi-kiu Lo and Dekai Wu

Crawling Back and Forth: Using Back and Out Links to Locate Bilingual Sites

Luciano Barbosa, Srinivas Bangalore and Vivek Kumar Rangarajan Sridhar

November 10 (continued)

(10:30 - 12:10) Syntax II

Grammar Induction from Text Using Small Syntactic Prototypes

Prachya Boonkwan and Mark Steedman

Transferring Syntactic Relations from English to Hindi Using Alignments on Local Word Groups

Aswarth Dara, Prashanth Mannem, Hemanth Sagar Bayyarapu and Avinesh PVS

Generative Modeling of Coordination by Factoring Parallelism and Selectional Preferences

Daisuke Kawahara and Sadao Kurohashi

Syntactic Parsing for Ranking-Based Coreference Resolution

Altaf Rahman and Vincent Ng

(10:30 - 12:10) Generation / Summarization

TriS: A Statistical Sentence Simplifier with Log-linear Models and Margin-based Discriminative Training

Nguyen Bach, Qin Gao, Stephan Vogel and Alex Waibel

Social Summarization via Automatically Discovered Social Context

Po Hu, Cheng Sun, Longfei Wu, Donghong Ji and Chong Teng

Simultaneous Clustering and Noise Detection for Theme-based Summarization

Xiaoyan Cai, Renxian Zhang, Dehong Gao and Wenjie Li

Extractive Summarization Method for Contact Center Dialogues based on Call Logs

Akihiro Tamura, Kai Ishikawa, Masahiro Saikou and Masaaki Tsuchida

November 10 (continued)

(10:30 - 12:10) Spoken Language Processing

Indexing Spoken Documents with Hierarchical Semantic Structures: Semantic Tree-to-string Alignment Models

Xiaodan Zhu, Colin Cherry and Gerald Penn

Structured and Extended Named Entity Evaluation in Automatic Speech Transcriptions

Olivier Galibert, Sophie Rosset, Cyril Grouin, Pierre Zweigenbaum and Ludovic Quintard

Normalising Audio Transcriptions for Unwritten Languages

Adel Foda and Steven Bird

Similarity Based Language Model Construction for Voice Activated Open-Domain Question Answering

Istvan Varga, Kiyonori Ohtake, Kentaro Torisawa, Stijn De Saeger, Teruhisa Misu, Shigeki Matsuda and Jun'ichi Kazama

(12:10-13:55) Lunch

(13:55-15:10) Machine Learning III

The application of chordal graphs to inferring phylogenetic trees of languages

Jessica Enright and Grzegorz Kondrak

Cross-domain Feature Selection for Language Identification

Marco Lui and Timothy Baldwin

A Wikipedia-LDA Model for Entity Linking with Batch Size Changing Instance Selection

Wei Zhang, Jian Su and Chew-Lim Tan

November 10 (continued)

(13:55-15:10) Information Retrieval II

Discovering Latent Concepts and Exploiting Ontological Features for Semantic Text Search

Vuong M. Ngo and Tru H. Cao

CLGVSM: Adapting Generalized Vector Space Model to Cross-lingual Document Clustering

Guoyu Tang, Yunqing Xia, Min Zhang, Haizhou Li and Fang Zheng

Thread Cleaning and Merging for Microblog Topic Detection

Jianfeng Zhang, Yunqing Xia, Bin Ma, Jianmin Yao and Yu Hong

(13:55-15:10) Dialog/Discourse II

Training a BN-based user model for dialogue simulation with missing data

Stéphane Rossignol, Olivier Pietquin and Michel Ianotto

Automatic identification of general and specific sentences by leveraging discourse annotations

Annie Louis and Ani Nenkova

An Evaluation of Alternative Strategies for Implementing Dialogue Policies Using Statistical Classification and Hand-Authored Rules

David DeVault, Anton Leuski and Kenji Sagae

(13:55-15:10) Sentiment II

A POS-based Ensemble Model for Cross-domain Sentiment Classification

Rui Xia and Chengqing Zong

Ensemble-style Self-training on Citation Classification

Cailing Dong and Ulrich Schäfer

Back to the Roots of Genres: Text Classification by Language Function

Henning Wachsmuth and Kathrin Bujna

November 10 (continued)

(15:10-15:30) Coffee Break

(15:30-17:10) Machine Translation III

Transductive Minimum Error Rate Training for Statistical Machine Translation

Yinggong Zhao, Shujie Liu, Yangsheng Ji, Jiajun Chen and Guodong Zhou

Distributed Minimum Error Rate Training of SMT using Particle Swarm Optimization

Jun Suzuki, Kevin Duh and Masaaki Nagata

Going Beyond Word Cooccurrences in Global Lexical Selection for Statistical Machine Translation using a Multilayer Perceptron

Alexandre Patry and Philippe Langlais

System Combination Using Discriminative Cross-Adaptation

Jacob Devlin, Antti-Veikko Rosti, Sankaranarayanan Ananthakrishnan and Spyros Matsoukas

(15:30-17:10) Semantics II

Word Sense Disambiguation by Combining Labeled Data Expansion and Semi-Supervised Learning Method

Sanae Fujita and Akinori Fujino

Combining ConceptNet and WordNet for Word Sense Disambiguation

Junpeng Chen and Juan Liu

It Takes Two to Tango: A Bilingual Unsupervised Approach for Estimating Sense Distributions using Expectation Maximization

Mitesh M Khapra, Salil Joshi and Pushpak Bhattacharyya

Dynamic and Static Prototype Vectors for Semantic Composition

Siva Reddy, Ioannis Klapaftis, Diana McCarthy and Suresh Manandhar

November 10 (continued)

(15:30-17:10) Information Extraction

Using Prediction from Sentential Scope to Build a Pseudo Co-Testing Learner for Event Extraction

Shasha Liao and Ralph Grishman

Text Segmentation and Graph-based Method for Template Filling in Information Extraction

Ludovic Jean-Louis, Romaric Besançon and Olivier Ferret

Joint Distant and Direct Supervision for Relation Extraction

Truc-Vien T. Nguyen and Alessandro Moschitti

A Cross-lingual Annotation Projection-based Self-supervision Approach for Open Information Extraction

Seokhwan Kim, Minwoo Jeong, Jonghoon Lee and Gary Geunbae Lee

(15:30-17:10) Syntax III

Exploring Difficulties in Parsing Imperatives and Questions

Tadayoshi Hara, Takuya Matsuzaki, Yusuke Miyao and Jun'ichi Tsujii

A Discriminative Approach to Japanese Zero Anaphora Resolution with Large-scale Lexicalized Case Frames

Ryohei Sasano and Sadao Kurohashi

An Empirical Comparison of Unknown Word Prediction Methods

Kostadin Cholakov, Gertjan van Noord, Valia Kordoni and Yi Zhang

Training Dependency Parsers from Partially Annotated Corpora

Daniel Flannery, Yusuke Miyao, Graham Neubig and Shinsuke Mori

(19:00-21:00) Banquet

November 11

(9:00-10:00) Keynote – Wai Lam (Chinese University of Hong Kong)

(10:00-10:30) Coffee Break

(10:30-12:10) Machine Translation / Machine Learning (short papers)

Reducing Asymmetry between language-pairs to Improve Alignment and Translation Quality

Rashmi Gangadharaiah

Clause-Based Reordering Constraints to Improve Statistical Machine Translation

Ananthakrishnan Ramanathan, Pushpak Bhattacharyya, Karthik Visweswariah, Kushal Ladha and Ankur Gandhe

Generalized Minimum Bayes Risk System Combination

Kevin Duh, Katsuhito Sudoh, Xianchao Wu, Hajime Tsukada and Masaaki Nagata

Enhancing scarce-resource language translation through pivot combinations

Marta R. Costa-jussà, Carlos Henríquez and Rafael E. Banchs

A Baseline System for Chinese Near-Synonym Choice

Liang-Chih Yu, Wei-Nan Chien and Shih-Ting Chen

Cluster Labelling based on Concepts in a Machine-Readable Dictionary

Fumiyo Fukumoto and Yoshimi Suzuki

November 11 (continued)

(10:30-12:10) Semantics (short papers)

Text Patterns and Compression Models for Semantic Class Learning

Chung-Yao Chuang, Yi-Hsun Lee and Wen-Lian Hsu

Potts Model on the Case Fillers for Word Sense Disambiguation

Hiroya Takamura and Manabu Okumura

Improving Word Sense Induction by Exploiting Semantic Relevance

Zhenzhong Zhang and Le Sun

Predicting Word Clipping with Latent Semantic Analysis

Julian Brooke, Tong Wang and Graeme Hirst

(10:30-12:10) Text Mining / Question Answering (short papers)

Domain Independent Model for Product Attribute Extraction from User Reviews using Wikipedia

Sudheer Kovelamudi, Sethu Ramalingam, Arpit Sood and Vasudeva Varma

Finding Problem Solving Threads in Online Forum

Zhonghua Qu and Yang Liu

Compiling Learner Corpus Data of Linguistic Output and Language Processing in Speaking, Listening, Writing, and Reading

Katsunori Kotani, Takehiko Yoshimi, Hiroaki Nanjo and Hitoshi Isahara

Mining the Sentiment Expectation of Nouns Using Bootstrapping Method

Miaomiao Wen and Yunfang Wu

An Analysis of Questions in a Q&A Site Resubmitted Based on Indications of Unclear Points of Original Questions

Masahiro Kojima, Yasuhiko Watanabe and Yoshihiro Okada

Diversifying Information Needs in Results of Question Retrieval

Yaoyun Zhang, Xiaolong Wang, Xuan Wang, Ruifeng Xu, Jun Xu and ShiXi Fan

November 11 (continued)

(10:30-12:10) Other NLP topics (short papers)

Beyond Normalization: Pragmatics of Word Form in Text Messages
Tyler Baldwin and Joyce Chai

Chinese Discourse Relation Recognition
Hen-Hsen Huang and Hsin-Hsi Chen

Improving Chinese POS Tagging with Dependency Parsing
Zhenghua Li, Wanxiang Che and Ting Liu

Exploring self training for Hindi dependency parsing
Rahul Goutam and Bharat Ram Ambati

Reduction of Search Space to Annotate Monolingual Corpora
Prajol Shrestha, Christine Jacquin and Beatrice Daille

Toward a Parallel Corpus of Spoken Cantonese and Written Chinese
John Lee

(12:00-13:30) Lunch

(13:30-15:35) Machine Translation IV

A Breadth-First Representation for Tree Matching in Large Scale Forest-Based Translation
Sumukh Ghodke, Steven Bird and Rui Zhang

Bayesian Subtree Alignment Model based on Dependency Trees
Toshiaki Nakazawa and Sadao Kurohashi

Enriching SMT Training Data via Paraphrasing
Wei He, Shiqi Zhao, Haifeng Wang and Ting Liu

Translation Quality Indicators for Pivot-based Statistical MT
Michael Paul and Eiichiro Sumita

Source Error-Projection for Sample Selection in Phrase-Based SMT for Resource-Poor Languages
Sankaranarayanan Ananthkrishnan, Shiv Vitaladevuni, Rohit Prasad and Prem Natarajan

November 11 (continued)

(13:30-15:35) Morphology / Named-Entity Recognition

A Named Entity Recognition Method based on Decomposition and Concatenation of Word Chunks

Tomoya Iwakura, Hiroya Takamura and Manabu Okumura

Extract Chinese Unknown Words from a Large-scale Corpus Using Morphological and Distributional Evidences

Kaixu Zhang, Ruining Wang, Ping Xue and Maosong Sun

Entity Disambiguation Using a Markov-Logic Network

Hong-Jie Dai, Richard Tzong-Han Tsai and Wen-Lian Hsu

Named Entity Recognition in Chinese News Comments on the Web

Xiaojun Wan, Liang Zong, Xiaojiang Huang, Tengfei Ma, Houping Jia, Yuqian Wu and Jianguo Xiao

Clustering Semantically Equivalent Words into Cognate Sets in Multilingual Lists

Bradley Hauer and Grzegorz Kondrak

(13:30-15:35) Resources

Extending WordNet with Hypernyms and Siblings Acquired from Wikipedia

Ichiro Yamada, Jong-Hoon Oh, Chikara Hashimoto, Kentaro Torisawa, Jun'ichi Kazama, Stijn De Saeger and Takuya Kawada

What Psycholinguists Know About Chemistry: Aligning Wiktionary and WordNet for Increased Domain Coverage

Christian M. Meyer and Iryna Gurevych

From News to Comment: Resources and Benchmarks for Parsing the Language of Web 2.0

Jennifer Foster, Ozlem Cetinoglu, Joachim Wagner, Joseph Le Roux, Joakim Nivre, Deirdre Hogan and Josef van Genabith

Toward Finding Semantic Relations not Written in a Single Sentence: An Inference Method using Auto-Discovered Rules

Masaaki Tsuchida, Kentaro Torisawa, Stijn De Saeger, Jong Hoon Oh, Jun'ichi Kazama, Chikara Hashimoto and Hayato Ohwada

Fleshing it out: A Supervised Approach to MWE-token and MWE-type Classification

Richard Fothergill and Timothy Baldwin

November 11 (continued)

(13:30-15:35) Question Answering

Identification of relations between answers with global constraints for Community-based Question Answering services

Hikaru Yokono, Takaaki Hasegawa, Genichiro Kikui and Manabu Okumura

Automatically Generating Questions from Queries for Community-based Question Answering

Shiqi Zhao, Haifeng Wang, Chao Li, Ting Liu and Yi Guan

Question classification based on an extended class sequential rule model

Zijing Hui, Juan Liu and Lumei Ouyang

K2Q: Generating Natural Language Questions from Keywords with User Refinements

Zhicheng Zheng, Xiance Si, Edward Chang and Xiaoyan Zhu

Answering Complex Questions via Exploiting Social Q&A Collection

Youzheng Wu, Chiori Hori, Hisashi Kawai and Hideki Kashioka

(15:35-16:05) Coffee Break

(16:05-17:15) Best paper session (2 plenary papers announced during this session)

(17:15-17:25) Presentation of Future Conferences (including ACL/IJCNLP 2012)

(17:25-17:35) Closing Ceremony

November 9

(18:00-21:00) Poster Session

Text Mining

Safety Information Mining — What can NLP do in a disaster—

Graham Neubig, Yuichiroh Matsubayashi, Masato Hagiwara and Koji Murakami

A Character-Level Machine Translation Approach for Normalization of SMS Abbreviations

Deana Pennell and Yang Liu

Using Text Reviews for Product Entity Completion

Mrinmaya Sachan, Tanveer Faruque, L. V. Subramaniam and Mukesh Mohania

Mining bilingual topic hierarchies from unaligned text

Sumit Negi

Efficient Near-Duplicate Detection for Q&A Forum

Yan Wu, Qi Zhang and Xuanjing Huang

Information Extraction

A Graph-based Method for Entity Linking

Yuhang Guo, Wanxiang Che, Ting Liu and Sheng Li

Harvesting Related Entities with a Search Engine

Shuqi Sun, Shiqi Zhao, Muyun Yang, Haifeng Wang and Sheng Li

Acquiring Strongly-related Events using Predicate-argument Co-occurring Statistics and Case Frames

Tomohide Shibata and Sadao Kurohashi

November 9 (continued)

Information Retrieval

Relevance Feedback using Latent Information

Jun Harashima and Sadao Kurohashi

Passage Retrieval for Information Extraction using Distant Supervision

Wei Xu, Ralph Grishman and Le Zhao

Query Expansion for IR using Knowledge-Based Relatedness

Arantxa Otegi, Xabier Arregi and Eneko Agirre

Summarization

Using Context Inference to Improve Sentence Ordering for Multi-document Summarization

Peifeng Li, Guangxi Deng and Qiaoming Zhu

Enhancing extraction based summarization with outside word space

Christian Smith and Arne Jönsson

Discourse

Shallow Discourse Parsing with Conditional Random Fields

Sucheta Ghosh, Richard Johansson, Giuseppe Riccardi and Sara Tonelli

Machine Learning

Relational Lasso —An Improved Method Using the Relations Among Features—

Kotaro Kitagawa and Kumiko Tanaka-Ishii

Enhance Top-down method with Meta-Classification for Very Large-scale Hierarchical Classification

Xiao-lin Wang, Hai Zhao and Bao-Liang Lu

Using Syntactic and Shallow Semantic Kernels to Improve Multi-Modality Manifold-Ranking for Topic-Focused Multi-Document Summarization

Yllias Chali, Sadid A. Hasan and Kaisar Imam

Automatic Determination of a Domain Adaptation Method for Word Sense Disambiguation Using Decision Tree Learning

Kanako Komiya and Manabu Okumura

Learning from Chinese-English Parallel Data for Chinese Tense Prediction

Feifan Liu, Fei Liu and Yang Liu

November 9 (continued)

Semantics

Jointly Extracting Japanese Predicate-Argument Relation with Markov Logic
Katsumasa Yoshikawa, Masayuki Asahara and Yuji Matsumoto

Word Meaning in Context: A Simple and Effective Vector Model
Stefan Thater, Hagen Fürstenau and Manfred Pinkal

Automatic Analysis of Semantic Coherence in Academic Abstracts Written in Portuguese
Vinícius Mourão Alves de Souza and Valéria Delisandra Feltrim

Word Sense Disambiguation Corpora Acquisition via Confirmation Code
Wanxiang Che and Ting Liu

Sentiment

Sentence Subjectivity Detection with Weakly-Supervised Learning
Chenghua Lin, Yulan He and Richard Everson

Opinion Expression Mining by Exploiting Keyphrase Extraction
Gábor Berend

Extracting Resource Terms for Sentiment Analysis
Lei Zhang and Bing Liu

Towards Context-Based Subjectivity Analysis
Farah Benamara, Baptiste Chardon, Yannick Mathieu and Vladimir Popescu

Phonology/Morphology

Compression Methods by Code Mapping and Code Dividing for Chinese Dictionary Stored in a Double-Array Trie
Huidan Liu, Minghua Nuo, Longlong Ma, Jian Wu and Yeping He

Functional Elements and POS Categories
Qiuye Zhao and Mitch Marcus

Joint Alignment and Artificial Data Generation: An Empirical Study of Pivot-based Machine Transliteration
Min Zhang, Xiangyu Duan, Ming Liu, Yunqing Xia and Haizhou Li

November 9 (continued)

Incremental Joint POS Tagging and Dependency Parsing in Chinese

Jun Hatori, Takuya Matsuzaki, Yusuke Miyao and Jun'ichi Tsujii

Resources

Extending the adverbial coverage of a NLP oriented resource for French

Elsa Tolone and Stavroula Voyatzi

Linguistic Phenomena, Analyses, and Representations: Understanding Conversion between Treebanks

Rajesh Bhatt, Owen Rambow and Fei Xia

Syntax

Automatic Transformation of the Thai Categorical Grammar Treebank to Dependency Trees

Christian Rishøj, Taneth Ruangrajitpakorn, Prachya Boonkwan and Thepchai Supnithi

Parse Reranking Based on Higher-Order Lexical Dependencies

Zhiguo Wang and Chengqing Zong

Improving Part-of-speech Tagging for Context-free Parsing

Xiao Chen and Chunyu Kit

Models Cascade for Tree-Structured Named Entity Detection

Marco Dinarelli and Sophie Rosset

Clausal parsing helps data-driven dependency parsing: Experiments with Hindi

Samar Husain, Phani Gadde, Joakim Nivre and Rajeev Sangal

November 9 (continued)

Machine Translation

Word-reordering for Statistical Machine Translation Using Trigram Language Model
Jing He and Hongyu Liang

Extracting Hierarchical Rules from a Weighted Alignment Matrix
Zhaopeng Tu, Yang Liu, Qun Liu and Shouxun Lin

Integration of Reduplicated Multiword Expressions and Named Entities in a Phrase Based Statistical Machine Translation System
Thoudam Doren Singh and Sivaji Bandyopadhyay

Regularizing Mono- and Bi-Word Models for Word Alignment
Thomas Schoenemann

Parametric Weighting of Parallel Data for Statistical Machine Translation
Kashif Shah, Loïc Barrault and Holger Schwenk

An Effective and Robust Framework for Transliteration Exploration
EA-EE JAN, Niyu Ge, Shih-Hsiang Lin and Berlin Chen

