IJCNLP 2011
The Fifth International Joint Conference on Natural Language Processing

November 8-13, 2011
Shangri-La Hotel
Chiang Mai, Thailand
We wish to thank our sponsors

Gold Sponsors

Google
www.google.com

Baidu
www.baidu.com

The Office of Naval Research (ONR)

The Asian Office of Aerospace Research and Development (AOARD)

Silver Sponsors

Microsoft
Microsoft Corporation

Bronze Sponsors

COLIPS
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

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
Chaiyong Eurviriyanukul, 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
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 Bodenstab, Ester Boldriní, 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 Christodouloupoulos, Kenneth Church, Alex Clark, Shay Cohen, Trevor Cohn, Gao Cong, Marta R. Costa-jussa, Paul Crook, Montse Cuadros, Ronan Cummins
Robert Damper, 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, Aranpta 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 Emm
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 Fürstenau
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, Xiaojian Jiang, Yun Huang, Zhongqiang Huang
Francisco Iacobelli, Diana Inkpen, Aminul Islam, Ruben Izquierdo
Heng Ji, Sittichai Jiampojaman, Hongfei Jiang, Wenbin Jiang, Xing Jiang, Cai Jie, Rong Jin, Richard Johansson, Hammin Jung
# Table of Contents

## Part A: Full Papers

*Analyzing the Dynamics of Research by Extracting Key Aspects of Scientific Papers*
Sonal Gupta and Christopher Manning .................................................. 1

*Dependency-directed Tree Kernel-based Protein-Protein Interaction Extraction from Biomedical Literature*
Longhua Qian and Guodong Zhou ...................................................... 10

*Learning Logical Structures of Paragraphs in Legal Articles*
Ngo Xuan Bach, Nguyen Le Minh, Tran Thi Oanh and Akira Shimazu ............. 20

*Extracting Pre-ordering Rules from Predicate-Argument Structures*
Xianchao Wu, Katsuhiro Sudoh, Kevin Duh, Hajime Tsukada and Masaaki Nagata .......... 29

*Context-Sensitive Syntactic Source-Reordering by Statistical Transduction*
Maxim Khalilov and Khalil Sima’an .................................................. 38

*Discriminative Phrase-based Lexicalized Reordering Models using Weighted Reordering Graphs*
Wang Ling, João Graça, David Martins de Matos, Isabel Trancoso and Alan W Black ........ 47

*Active Learning Strategies for Support Vector Machines, Application to Temporal Relation Classification*
Seyed Abolghasem Mirroshandel, Gholamreza Ghassem-Sani and Alexis Nasr ............ 56

*A Fast Accurate Two-stage Training Algorithm for L1-regularized CRFs with Heuristic Line Search Strategy*
Jinlong Zhou, Xipeng Qiu and Xuanjing Huang ....................................... 65

*Automatic Topic Model Adaptation for Sentiment Analysis in Structured Domains*
Geoffrey Levine and Gerald DeJong .................................................. 75

*Multi-modal Reference Resolution in Situated Dialogue by Integrating Linguistic and Extra-Linguistic Clues*
Ryu Iida, Masaaki Yasuhara and Takenobu Tokunaga .................................. 84

*Single and multi-objective optimization for feature selection in anaphora resolution*
Sriparna Saha, Asif Ekbal, Olga Uryupina and Massimo Poesio .......................... 93

*A Unified Event Coreference Resolution by Integrating Multiple Resolvers*
Bin Chen, Jian Su, Sinno Jialin Pan and Chew Lim Tan ................................ 102

*Handling verb phrase morphology in highly inflected Indian languages for Machine Translation*
Ankur Gandhe, Rashmi Gangadharaiah, Karthik Visweswariah and Ananthakrishnan Ramanathan 111

*Japanese Pronunciation Prediction as Phrasal Statistical Machine Translation*
Jun Hatori and Hisami Suzuki .......................................................... 120

*Comparing Two Techniques for Learning Transliteration Models Using a Parallel Corpus*
Hassan Sajjad, Nadir Durrani, Helmut Schmid and Alexander Fraser ................. 129
A Semantic-Specific Model for Chinese Named Entity Translation
Yufeng Chen and Chengqing Zong .......................................................... 138

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 ................. 147

Modality Specific Meta Features for Authorship Attribution in Web Forum Posts
Thamar Solorio, Sangita Pillay, Sindhu Raghavan and Manuel Montes-Gomez .............. 156

Keyphrase Extraction from Online News Using Binary Integer Programming
Zhuoye Ding, Qi Zhang and Xuanjing Huang .............................................. 165

Improving Related Entity Finding via Incorporating Homepages and Recognizing Fine-grained Entities
Youzheng Wu, Chiori Hori, Hisashi Kawai and Hideki Kashioka ............................ 174

Enhancing Active Learning for Semantic Role Labeling via Compressed Dependency Trees
Chenhua Chen, Alexis Palmer and Caroline Sporleder ........................................ 183

Semantic Role Labeling Without Treebanks?
Stephen Boxwell, Chris Brew, Jason Baldridge, Dennis Mehay and Sujith Ravi ............. 192

Japanese Predicate Argument Structure Analysis Exploiting Argument Position and Type
Yuta Hayashibe, Mamoru Komachi and Yuji Matsumoto ...................................... 201

An Empirical Study on Compositionality in Compound Nouns
Siva Reddy, Diana McCarthy and Suresh Manandhar ......................................... 210

Feature-Rich Log-Linear Lexical Model for Latent Variable PCFG Grammars
Zhongqiang Huang and Mary Harper ................................................................ 219

Improving Dependency Parsing with Fined-Grained Features
Guangyou Zhou, Li Cai, Kang Liu and Jun Zhao .................................................. 228

Natural Language Programming Using Class Sequential Rules
Cohan Sujay Carlos ......................................................................................... 237

Treeblazing: Using External Treebanks to Filter Parse Forests for Parse Selection and Treebanking
Andrew MacKinlay, Rebecca Dridan, Dan Flickinger, Stephan Oepen and Timothy Baldwin ... 246

Cross-Language Entity Linking
Paul McNamee, James Mayfield, Dawn Lawrie, Douglas Oard and David Doermann ........ 255

Generating Chinese Named Entity Data from a Parallel Corpus
Ruiji Fu, Bing Qin and Ting Liu ........................................................................ 264

Learning the Latent Topics for Question Retrieval in Community QA
Li Cai, Guangyou Zhou, Kang Liu and Jun Zhao ................................................ 273

Identifying Event Descriptions using Co-training with Online News Summaries
William Yang Wang, Kapil Thadani and Kathleen McKeown ................................. 282

Automatic Labeling of Voiced Consonants for Morphological Analysis of Modern Japanese Literature
Teruaki Oka, Mamoru Komachi, Toshinobu Ogiso and Yuji Matsumoto ................. 292
S3 - Statistical Sandhi Splitting
   Abhiram Natarajan and Eugene Charniak .................................................. 301

Improving Chinese Word Segmentation and POS Tagging with Semi-supervised Methods Using Large Auto-Analyzed Data
   You Wang, Jun’ichi Kazama, Yoshimasa Tsuruoka, Wenliang Chen, Yujie Zhang and Kentaro Torisawa .................................................. 309

CODACT: Towards Identifying Orthographic Variants in Dialectal Arabic
   Pradeep Dasigi and Mona Diab ................................................................. 318

Enhancing the HL-SOT Approach to Sentiment Analysis via a Localized Feature Selection Framework
   Wei and Jon Atle Gulla ................................................................. 327

Fine-Grained Sentiment Analysis with Structural Features
   Cacilia Zirn, Mathias Niepert, Heiner Stuckenschmidt and Michael Strube ................................. 336

Predicting Opinion Dependency Relations for Opinion Analysis
   Lun-Wei Ku, Ting-Hao Huang and Hsin-Hsi Chen ........................................ 345

Detecting and Blocking False Sentiment Propagation
   Hye-Jin Min and Jong C. Park ................................................................. 354

Efficient induction of probabilistic word classes with LDA
   Grzegorz Chrupala ................................................................. 363

Quality-biased Ranking of Short Texts in Microblogging Services
   Minlie Huang, Yi Yang and Xiaoyan Zhu .................................................. 373

Labeling Unlabeled Data using Cross-Language Guided Clustering
   Sachindra Joshi, Danish Contractor and Sumit Negi ..................................... 383

Extracting Relation Descriptors with Conditional Random Fields
   Yaliang Li, Jing Jiang, Hai Leong Chieu and Kian Ming A. Chai .................... 392

Attribute Extraction from Synthetic Web Search Queries
   Marius Pasca ................................................................. 401

Japanese Abbreviation Expansion with Query and Clickthrough Logs
   Kei Uchiumi, Mamoru Komachi, Keigo Machinaga, Toshiyuki Maezawa, Toshinori Satou and Yoshinori Kobayashi ........................................... 410

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 ........ 420

Crawling Back and Forth: Using Back and Out Links to Locate Bilingual Sites
   Luciano Barbosa, Srinivas Bangalore and Vivek Kumar Rangarajan Sridhar ........ 429

Grammar Induction from Text Using Small Syntactic Prototypes
   Prachya Boonkwan and Mark Steedman .................................................. 438

Transferring Syntactic Relations from English to Hindi Using Alignments on Local Word Groups
   Aswarth Dara, Prashanth Mannem, Hemanth Sagar Bayyarapu and Avinesh PVS .......... 447
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generative Modeling of Coordination by Factoring Parallelism and Selectional Preferences</td>
<td>456</td>
</tr>
<tr>
<td>Daisuke Kawahara and Sadao Kurohashi</td>
<td></td>
</tr>
<tr>
<td>Syntactic Parsing for Ranking-Based Coreference Resolution</td>
<td>465</td>
</tr>
<tr>
<td>Altaf Rahman and Vincent Ng</td>
<td></td>
</tr>
<tr>
<td>TriS: A Statistical Sentence Simulator with Log-linear Models and Margin-based Discriminative Training</td>
<td>474</td>
</tr>
<tr>
<td>Nguyen Bach, Qin Gao, Stephan Vogel and Alex Waibel</td>
<td></td>
</tr>
<tr>
<td>Social Summarization via Automatically Discovered Social Context</td>
<td>483</td>
</tr>
<tr>
<td>Po Hu, Cheng Sun, Longfei Wu, Donghong Ji and Chong Teng</td>
<td></td>
</tr>
<tr>
<td>Simultaneous Clustering and Noise Detection for Theme-based Summarization</td>
<td>491</td>
</tr>
<tr>
<td>Xiaoyan Cai, Renxian Zhang, Dehong Gao and Wenjie Li</td>
<td></td>
</tr>
<tr>
<td>Extractive Summarization Method for Contact Center Dialogues based on Call Logs</td>
<td>500</td>
</tr>
<tr>
<td>Akihiro Tamura, Kai Ishikawa, Masahiro Saikou and Masaaki Tsuchida</td>
<td></td>
</tr>
<tr>
<td>Indexing Spoken Documents with Hierarchical Semantic Structures: Semantic Tree-to-string Alignment Models</td>
<td>509</td>
</tr>
<tr>
<td>Xiaodan Zhu, Colin Cherry and Gerald Penn</td>
<td></td>
</tr>
<tr>
<td>Structured and Extended Named Entity Evaluation in Automatic Speech Transcriptions</td>
<td>518</td>
</tr>
<tr>
<td>Olivier Galibert, Sophie Rosset, Cyril Grouin, Pierre Zweigenbaum and Ludovic Quintard</td>
<td></td>
</tr>
<tr>
<td>Normalising Audio Transcriptions for Unwritten Languages</td>
<td>527</td>
</tr>
<tr>
<td>Adel Foda and Steven Bird</td>
<td></td>
</tr>
<tr>
<td>Similarity Based Language Model Construction for Voice Activated Open-Domain Question Answering</td>
<td>536</td>
</tr>
<tr>
<td>Istvan Varga, Kiyonori Ohtake, Kentaro Torisawa, Stijn De Saeger, Teruhisa Misu, Shigeki Matsuda and Jun`ichi Kazama</td>
<td></td>
</tr>
<tr>
<td>The application of chordal graphs to inferring phylogenetic trees of languages</td>
<td>545</td>
</tr>
<tr>
<td>Jessica Enright and Grzegorz Kondrak</td>
<td></td>
</tr>
<tr>
<td>Cross-domain Feature Selection for Language Identification</td>
<td>553</td>
</tr>
<tr>
<td>Marco Lui and Timothy Baldwin</td>
<td></td>
</tr>
<tr>
<td>A Wikipedia-LDA Model for Entity Linking with Batch Size Changing Instance Selection</td>
<td>562</td>
</tr>
<tr>
<td>Wei Zhang, Jian Su and Chew-Lim Tan</td>
<td></td>
</tr>
<tr>
<td>Discovering Latent Concepts and Exploiting Ontological Features for Semantic Text Search</td>
<td>571</td>
</tr>
<tr>
<td>Vuong M. Ngo and Tru H. Cao</td>
<td></td>
</tr>
<tr>
<td>CLGVSM: Adapting Generalized Vector Space Model to Cross-lingual Document Clustering</td>
<td>580</td>
</tr>
<tr>
<td>Guoyu Tang, Yunqing Xia, Min Zhang, Haizhou Li and Fang Zheng</td>
<td></td>
</tr>
<tr>
<td>Thread Cleaning and Merging for Microblog Topic Detection</td>
<td>589</td>
</tr>
<tr>
<td>Jianfeng Zhang, Yunqing Xia, Bin Ma, Jianmin Yao and Yu Hong</td>
<td></td>
</tr>
<tr>
<td>Training a BN-based user model for dialogue simulation with missing data</td>
<td>598</td>
</tr>
<tr>
<td>Stéphane Rossignol, Olivier Pietquin and Michel Ianotto</td>
<td></td>
</tr>
<tr>
<td>Automatic identification of general and specific sentences by leveraging discourse annotations</td>
<td>605</td>
</tr>
<tr>
<td>Annie Louis and Ani Nenkova</td>
<td></td>
</tr>
</tbody>
</table>
A POS-based Ensemble Model for Cross-domain Sentiment Classification
Rui Xia and Chengqing Zong ................................................................. 614

Ensemble-style Self-training on Citation Classification
Cailing Dong and Ulrich Schäfer .......................................................... 623

Back to the Roots of Genres: Text Classification by Language Function
Henning Wachsmuth and Kathrin Bujna ................................................... 632

Transductive Minimum Error Rate Training for Statistical Machine Translation
Yinggong Zhao, Shujie Liu, Yangsheng Ji, Jiajun Chen and Guodong Zhou ................. 641

Distributed Minimum Error Rate Training of SMT using Particle Swarm Optimization
Jun Suzuki, Kevin Duh and Masaaki Nagata ............................................. 649

Going Beyond Word Cooccurrences in Global Lexical Selection for Statistical Machine Translation using a Multilayer Perceptron
Alexandre Patry and Philippe Langlais .................................................... 658

System Combination Using Discriminative Cross-Adaptation
Jacob Devlin, Antti-Veikko Rosti, Sankaranarayanan Ananthakrishnan and Spyros Matsoukas. 667

Word Sense Disambiguation by Combining Labeled Data Expansion and Semi-Supervised Learning Method
Sanae Fujita and Akinori Fujino ............................................................. 676

Combining ConceptNet and WordNet for Word Sense Disambiguation
Junpeng Chen and Juan Liu ................................................................... 686

It Takes Two to Tango: A Bilingual Unsupervised Approach for Estimating Sense Distributions using Expectation Maximization
Mitesh M Khapra, Salil Joshi and Pushpak Bhattacharyya .............................. 695

Dynamic and Static Prototype Vectors for Semantic Composition
Siva Reddy, Ioannis Klapaftis, Diana McCarthy and Suresh Manandhar .................. 705

Using Prediction from Sentential Scope to Build a Pseudo Co-testing Learner for Event Extraction
Shasha Liao and Ralph Grishman ........................................................... 714

Text Segmentation and Graph-based Method for Template Filling in Information Extraction
Ludovic Jean-Louis, Romaric Besançon and Olivier Ferret .............................. 723

Joint Distant and Direct Supervision for Relation Extraction
Truc-Vien T. Nguyen and Alessandro Moschitti ........................................... 732

A Cross-lingual Annotation Projection-based Self-supervision Approach for Open Information Extraction
Seokhwan Kim, Minwoo Jeong, Jonghoon Lee and Gary Geunbae Lee ...................... 741

Exploring Difficulties in Parsing Imperatives and Questions
Tadayoshi Hara, Takuya Matsuzaki, Yusuke Miyao and Jun’ichi Tsujii .................... 749

A Discriminative Approach to Japanese Zero Anaphora Resolution with Large-scale Lexicalized Case Frames
Ryohei Sasano and Sadao Kurohashi ...................................................... 758
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>An Empirical Comparison of Unknown Word Prediction Methods</td>
<td>Kostadin Cholakov, Gertjan van Noord, Valia Kordoni and Yi Zhang</td>
<td>767</td>
</tr>
<tr>
<td>Training Dependency Parsers from Partially Annotated Corpora</td>
<td>Daniel Flannery, Yusuke Miayo, Graham Neubig and Shinsuke Mori</td>
<td>776</td>
</tr>
<tr>
<td>A Breadth-First Representation for Tree Matching in Large Scale Forest-Based Translation</td>
<td>Sumukh Ghodke, Steven Bird and Rui Zhang</td>
<td>785</td>
</tr>
<tr>
<td>Bayesian Subtree Alignment Model based on Dependency Trees</td>
<td>Toshiaki Nakazawa and Sadao Kurohashi</td>
<td>794</td>
</tr>
<tr>
<td>Enriching SMT Training Data via Paraphrasing</td>
<td>Wei He, Shiqi Zhao, Haifeng Wang and Ting Liu</td>
<td>803</td>
</tr>
<tr>
<td>Translation Quality Indicators for Pivot-based Statistical MT</td>
<td>Michael Paul and Eiichiro Sumita</td>
<td>811</td>
</tr>
<tr>
<td>Source Error-Projection for Sample Selection in Phrase-Based SMT for Resource-Poor Languages</td>
<td>Sankaranarayanan Ananthakrishnan, Shiv Vitaladevuni, Rohit Prasad and Prem Natarajan</td>
<td>819</td>
</tr>
<tr>
<td>A Named Entity Recognition Method based on Decomposition and Concatenation of Word Chunks</td>
<td>Tomoya Iwakura, Hiroya Takamura and Manabu Okumura</td>
<td>828</td>
</tr>
<tr>
<td>Extract Chinese Unknown Words from a Large-scale Corpus Using Morphological and Distributional Evidences</td>
<td>Kaixu Zhang, Ruining Wang, Ping Xue and Maosong Sun</td>
<td>837</td>
</tr>
<tr>
<td>Entity Disambiguation Using a Markov-Logic Network</td>
<td>Hong-Jie Dai, Richard Tzong-Han Tsai and Wen-Lian Hsu</td>
<td>846</td>
</tr>
<tr>
<td>Named Entity Recognition in Chinese News Comments on the Web</td>
<td>Xiaojun Wan, Liang Zong, Xiaojiang Huang, Tengfei Ma, Houping Jia, Yuqian Wu and Jianguo Xiao</td>
<td>856</td>
</tr>
<tr>
<td>Clustering Semantically Equivalent Words into Cognate Sets in Multilingual Lists</td>
<td>Bradley Hauer and Grzegorz Kondrak</td>
<td>865</td>
</tr>
<tr>
<td>Extending WordNet with Hypernyms and Siblings Acquired from Wikipedia</td>
<td>Ichiro Yamada, Jong-Hoon Oh, Chikara Hashimoto, Kentaro Torisawa, Jun’ichi Kazama, Stijn De Saeger and Takuya Kawada</td>
<td>874</td>
</tr>
<tr>
<td>From News to Comment: Resources and Benchmarks for Parsing the Language of Web 2.0</td>
<td>Jennifer Foster, Ozlem Cetinoglu, Joachim Wagner, Joseph Le Roux, Joakim Nivre, Deirdre Hogan and Josef van Genabith</td>
<td>893</td>
</tr>
<tr>
<td>Toward Finding Semantic Relations not Written in a Single Sentence: An Inference Method using Auto-Discovered Rules</td>
<td>Masaaki Tsuchida, Kentaro Torisawa, Stijn De Saeger, Jong Hoon Oh, Jun’ichi Kazama, Chikara Hashimoto and Hayato Ohwada</td>
<td>902</td>
</tr>
<tr>
<td>Title</td>
<td>Authors</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Fleshing it out: A Supervised Approach to MWE-token and MWE-type Classification</td>
<td>Richard Fothergill and Timothy Baldwin</td>
<td>911</td>
</tr>
<tr>
<td>Identification of relations between answers with global constraints for Community-based Question Answering services</td>
<td>Hikaru Yokono, Takaaki Hasegawa, Genichiro Kikui and Manabu Okumura</td>
<td>920</td>
</tr>
<tr>
<td>Automatically Generating Questions from Queries for Community-based Question Answering</td>
<td>Shiqi Zhao, Haifeng Wang, Chao Li, Ting Liu and Yi Guan</td>
<td>929</td>
</tr>
<tr>
<td>Question classification based on an extended class sequential rule model</td>
<td>Zijing Hui, Juan Liu and Lumei Ouyang</td>
<td>938</td>
</tr>
<tr>
<td>K2Q: Generating Natural Language Questions from Keywords with User Refinements</td>
<td>Zhicheng Zheng, Xiunce Si, Edward Chang and Xiaoyan Zhu</td>
<td>947</td>
</tr>
<tr>
<td>Answering Complex Questions via Exploiting Social Q&amp;A Collection</td>
<td>Youzheng Wu, Chiori Hori, Hisashi Kawai and Hideki Kashioka</td>
<td>956</td>
</tr>
<tr>
<td>Safety Information Mining — What can NLP do in a disaster—</td>
<td>Graham Neubig, Yuichiroh Matsubayashi, Masato Hagiwara and Koji Murakami</td>
<td>965</td>
</tr>
<tr>
<td>A Character-Level Machine Translation Approach for Normalization of SMS Abbreviations</td>
<td>Deana Pennell and Yang Liu</td>
<td>974</td>
</tr>
<tr>
<td>Using Text Reviews for Product Entity Completion</td>
<td>Mrinmaya Sachan, Tanveer Faruquie, L. V. Subramaniam and Mukesh Mohania</td>
<td>983</td>
</tr>
<tr>
<td>Mining bilingual topic hierarchies from unaligned text</td>
<td>Sumit Negi</td>
<td>992</td>
</tr>
<tr>
<td>Efficient Near-Duplicate Detection for Q&amp;A Forum</td>
<td>Yan Wu, Qi Zhang and Xuanjing Huang</td>
<td>1001</td>
</tr>
<tr>
<td>A Graph-based Method for Entity Linking</td>
<td>Yuhang Guo, Wanxiang Che, Ting Liu and Sheng Li</td>
<td>1010</td>
</tr>
<tr>
<td>Harvesting Related Entities with a Search Engine</td>
<td>Shuqi Sun, Shiqi Zhao, Muyun Yang, Haifeng Wang and Sheng Li</td>
<td>1019</td>
</tr>
<tr>
<td>Acquiring Strongly-related Events using Predicate-argument Co-occurring Statistics and Case Frames</td>
<td>Tomohide Shibata and Sadao Kurohashi</td>
<td>1028</td>
</tr>
<tr>
<td>Relevance Feedback using Latent Information</td>
<td>Jun Harashima and Sadao Kurohashi</td>
<td>1037</td>
</tr>
<tr>
<td>Passage Retrieval for Information Extraction using Distant Supervision</td>
<td>Wei Xu, Ralph Grishman and Le Zhao</td>
<td>1046</td>
</tr>
<tr>
<td>Using Context Inference to Improve Sentence Ordering for Multi-document Summarization</td>
<td>Peifeng Li, Guangxi Deng and Qiaoming Zhu</td>
<td>1055</td>
</tr>
<tr>
<td>Enhancing extraction based summarization with outside word space</td>
<td>Christian Smith and Arne Jönsson</td>
<td>1062</td>
</tr>
<tr>
<td>Title</td>
<td>Authors</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Shallow Discourse Parsing with Conditional Random Fields</td>
<td>Sucheta Ghosh, Richard Johansson, Giuseppe Riccardi and Sara Tonelli</td>
<td>1071</td>
</tr>
<tr>
<td>Relational Lasso —An Improved Method Using the Relations Among Features—</td>
<td>Kotaro Kitagawa and Kumiko Tanaka-Ishii</td>
<td>1080</td>
</tr>
<tr>
<td>Enhance Top-down method with Meta-Classification for Very Large-scale Hierarchical Classification</td>
<td>Xiao-lin Wang, Hai Zhao and Bao-Liang Lu</td>
<td>1089</td>
</tr>
<tr>
<td>Automatic Determination of a Domain Adaptation Method for Word Sense Disambiguation Using Decision Tree Learning</td>
<td>Kanako Komiya and Manabu Okumura</td>
<td>1107</td>
</tr>
<tr>
<td>Learning from Chinese-English Parallel Data for Chinese Tense Prediction</td>
<td>Feifan Liu, Fei Liu and Yang Liu</td>
<td>1116</td>
</tr>
<tr>
<td>Jointly Extracting Japanese Predicate-Argument Relation with Markov Logic</td>
<td>Katsumasa Yoshikawa, Masayuki Asahara and Yuji Matsumoto</td>
<td>1125</td>
</tr>
<tr>
<td>Word Meaning in Context: A Simple and Effective Vector Model</td>
<td>Stefan Thater, Hagen Fürstenau and Manfred Pinkal</td>
<td>1134</td>
</tr>
<tr>
<td>Automatic Analysis of Semantic Coherence in Academic Abstracts Written in Portuguese</td>
<td>Vinícius Mourão Alves de Souza and Valéria Delisandra Feltrim</td>
<td>1144</td>
</tr>
<tr>
<td>Sentence Subjectivity Detection with Weakly-Supervised Learning</td>
<td>Chenghua Lin, Yulan He and Richard Everson</td>
<td>1153</td>
</tr>
<tr>
<td>Opinion Expression Mining by Exploiting Keyphrase Extraction</td>
<td>Gábor Berend</td>
<td>1162</td>
</tr>
<tr>
<td>Extracting Resource Terms for Sentiment Analysis</td>
<td>Lei Zhang and Bing Liu</td>
<td>1171</td>
</tr>
<tr>
<td>Towards Context-Based Subjectivity Analysis</td>
<td>Farah Benamara, Baptiste Chardon, Yannick Mathieu and Vladimir Popescu</td>
<td>1180</td>
</tr>
<tr>
<td>Compression Methods by Code Mapping and Code Dividing for Chinese Dictionary Stored in a Double-Array Trie</td>
<td>Huidan Liu, Minghua Nuo, Longlong Ma, Jian Wu and Yeping He</td>
<td>1189</td>
</tr>
<tr>
<td>Functional Elements and POS Categories</td>
<td>Qiuye Zhao and Mitch Marcus</td>
<td>1198</td>
</tr>
<tr>
<td>Joint Alignment and Artificial Data Generation: An Empirical Study of Pivot-based Machine Transliteration</td>
<td>Min Zhang, Xiangyu Duan, Ming Liu, Yunqing Xia and Haizhou Li</td>
<td>1207</td>
</tr>
<tr>
<td>Incremental Joint POS Tagging and Dependency Parsing in Chinese</td>
<td>Jun Hatori, Takuya Matsuzaki, Yusuke Miyao and Jun’ichi Tsujii</td>
<td>1216</td>
</tr>
</tbody>
</table>
Extending the adverbial coverage of a NLP oriented resource for French
Elsa Tolone and Stavroula Voyatzi ............................................. 1225

Linguistic Phenomena, Analyses, and Representations: Understanding Conversion between Treebanks
Rajesh Bhatt, Owen Rambow and Fei Xia ........................................ 1234

Automatic Transformation of the Thai Categorial Grammar Treebank to Dependency Trees
Christian Rishøj, Taneth Ruangrajitpakorn, Prachya Boonkwan and Thepchai Supnithi .... 1243

Parse Reranking Based on Higher-Order Lexical Dependencies
Zhiguo Wang and Chengqing Zong .............................................. 1251

Improving Part-of-speech Tagging for Context-free Parsing
Xiao Chen and Chunyu Kit ......................................................... 1260

Models Cascade for Tree-Structured Named Entity Detection
Marco Dinarelli and Sophie Rosset .............................................. 1269

Clausal parsing helps data-driven dependency parsing: Experiments with Hindi
Samar Husain, Phani Gadde, Joakim Nivre and Rajeev Sangal ......................... 1279

Word-reordering for Statistical Machine Translation Using Trigram Language Model
Jing He and Hongyu Liang ........................................................ 1288

Extracting Hierarchical Rules from a Weighted Alignment Matrix
Zhaopeng Tu, Yang Liu, Qun Liu and Shouxun Lin .................................. 1294

Integration of Reduplicated Multiword Expressions and Named Entities in a Phrase Based Statistical Machine Translation System
Thoudam Doren Singh and Sivaji Bandyopadhyay .................................. 1304

Regularizing Mono- and Bi-Word Models for Word Alignment
Thomas Schoenemann ............................................................. 1313

Parametric Weighting of Parallel Data for Statistical Machine Translation
Kashif Shah, Loïc Barrault and Holger Schwenk ................................ 1323

An Effective and Robust Framework for Transliteration Exploration
EA-EE JAN, Niyu Ge, Shih-Hsiang Lin and Berlin Chen .......................... 1332
**Part B: Short Papers**

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reducing Asymmetry between language-pairs to Improve Alignment and Translation Quality</td>
<td>Rashmi Gangadharaiah</td>
<td>1346</td>
</tr>
<tr>
<td>Clause-Based Reordering Constraints to Improve Statistical Machine Translation</td>
<td>Ananthakrishnan Ramanathan, Pushpak Bhattacharyya, Karthik Visweswariah, Kushal Ladha and Ankur Gandhe</td>
<td>1351</td>
</tr>
<tr>
<td>Generalized Minimum Bayes Risk System Combination</td>
<td>Kevin Duh, Katsuhito Sudoh, Xianchao Wu, Hajime Tsukada and Masaaki Nagata</td>
<td>1356</td>
</tr>
<tr>
<td>Enhancing scarce-resource language translation through pivot combinations</td>
<td>Marta R. Costa-jussà, Carlos Henríquez and Rafael E. Banchs</td>
<td>1361</td>
</tr>
<tr>
<td>A Baseline System for Chinese Near-Synonym Choice</td>
<td>Liang-Chih Yu, Wei-Nan Chien and Shih-Ting Chen</td>
<td>1366</td>
</tr>
<tr>
<td>Cluster Labelling based on Concepts in a Machine-Readable Dictionary</td>
<td>Fumiyo Fukumoto and Yoshimi Suzuki</td>
<td>1371</td>
</tr>
<tr>
<td>Text Patterns and Compression Models for Semantic Class Learning</td>
<td>Chung-Yao Chuang, Yi-Hsun Lee and Wen-Lian Hsu</td>
<td>1376</td>
</tr>
<tr>
<td>Potts Model on the Case Fillers for Word Sense Disambiguation</td>
<td>Hiroya Takamura and Manabu Okumura</td>
<td>1382</td>
</tr>
<tr>
<td>Improving Word Sense Induction by Exploiting Semantic Relevance</td>
<td>Zhenzhong Zhang and Le Sun</td>
<td>1387</td>
</tr>
<tr>
<td>Predicting Word Clipping with Latent Semantic Analysis</td>
<td>Julian Brooke, Tong Wang and Graeme Hirst</td>
<td>1392</td>
</tr>
<tr>
<td>A Semantic Relatedness Measure Based on Combined Encyclopedic, Ontological and Collocational Knowledge</td>
<td>Yannis Haralambous and Vitaly Klyuev</td>
<td>1397</td>
</tr>
<tr>
<td>Going Beyond Text: A Hybrid Image-Text Approach for Measuring Word Relatedness</td>
<td>Chee Wee Leong and Rada Mihalcea</td>
<td>1403</td>
</tr>
<tr>
<td>Domain Independent Model for Product Attribute Extraction from User Reviews using Wikipedia</td>
<td>Sudheer Kovelamudi, Sethu Ramalingam, Arpit Sood and Vasudeva Varma</td>
<td>1408</td>
</tr>
<tr>
<td>Finding Problem Solving Threads in Online Forum</td>
<td>Zhonghua Qu and Yang Liu</td>
<td>1413</td>
</tr>
</tbody>
</table>
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 .......................... 1418

Mining the Sentiment Expectation of Nouns Using Bootstrapping Method
Miaomiao Wen and Yunfang Wu ............................................................... 1423

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 ................................. 1428

Diversifying Information Needs in Results of Question Retrieval
Yaoyun Zhang, Xiaolong Wang, Xuan Wang, Ruifeng Xu, Jun Xu and ShiXi Fan .............. 1432

Beyond Normalization: Pragmatics of Word Form in Text Messages
Tyler Baldwin and Joyce Chai .............................................................. 1437

Chinese Discourse Relation Recognition
Hen-Hsen Huang and Hsin-Hsi Chen .......................................................... 1442

Improving Chinese POS Tagging with Dependency Parsing
Zhenghua Li, Wanxiang Che and Ting Liu ................................................... 1447

Exploring self training for Hindi dependency parsing
Rahul Goutam and Bharat Ram Ambati ...................................................... 1452

Reduction of Search Space to Annotate Monolingual Corpora
Prajol Shrestha, Christine Jacquin and Beatrice Daille ........................................... 1457

Toward a Parallel Corpus of Spoken Cantonese and Written Chinese
John Lee .......................................................... 1462

Query Expansion for IR using Knowledge-Based Relatedness
Arantxa Otegi, Xabier Arregi and Eneko Agirre .............................................. 1467

Word Sense Disambiguation Corpora Acquisition via Confirmation Code
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
(10:50-12:05) Machine Learning I

Active Learning Strategies for Support Vector Machines, Application to Temporal Relation Classification
Seyed Abolghasem Miroshandel, 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 Ananthakrishnan 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

xxx
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

$S^3$ - 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


*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

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, Yusuке Miyao and Jun’ichi Tsujii

*A Discriminative Approach toJapanese 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, Yusuке Miayo, 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
(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 Ananthakrishnan, 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

xliii
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 Faruquie, 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
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
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 Fürstenuau 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
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 Categorial Grammar Treebank to Dependency Trees
Christian Rishøj, Taneth Ruangraitpakorn, 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