24th International Conference on Computational Linguistics

(COLING 2012)

Mumbai, India
8-15 December 2012

Volume 1 of 4
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi-Dimensional Feature Merger for Question Answering</td>
<td>Apoorv Agarwal, J William Murdock, Jennifer Chu-Carroll, Adam Lally and Aditya Kalyanpur</td>
<td>1</td>
</tr>
<tr>
<td>Unsupervised Discovery of Relations and Discriminative Extraction Patterns</td>
<td>Alan Akbik, Larysa Visengeriyeva, Priska Herger, Holmer Hemsen and Alexander Löser</td>
<td>17</td>
</tr>
<tr>
<td>Automatic Detection of Point of View Differences in Wikipedia</td>
<td>Khalid Al Khatib, Hinrich Schütze and Cathleen Kantner</td>
<td>33</td>
</tr>
<tr>
<td>SpeedRead: A Fast Named Entity Recognition Pipeline</td>
<td>Rami Al-Rfou' and Steven Skiena</td>
<td>51</td>
</tr>
<tr>
<td>Experiments with Term Translation</td>
<td>Mihael Arcan, Christian Federmann and Paul Buitelaar</td>
<td>67</td>
</tr>
<tr>
<td>The Floating Arabic Dictionary: An Automatic Method for Updating a Lexical Database through the Detection and Lemmatization of Unknown Words</td>
<td>Mohammed Attia, Younes Samih, Khaled Shaalan and Josef van Genabith</td>
<td>83</td>
</tr>
<tr>
<td>Contribution of Complex Lexical Information to Solve Syntactic Ambiguity in Basque</td>
<td>Aitziber Atutxa, Eneko Agirre and Kepa Sarasola</td>
<td>97</td>
</tr>
<tr>
<td>Comparative Quality Estimation: Automatic Sentence-Level Ranking of Multiple Machine Translation Outputs</td>
<td>Eleftherios Avramidis</td>
<td>115</td>
</tr>
<tr>
<td>Constructing Reference Semantic Predictions from Biomedical Knowledge Sources</td>
<td>Demeke Ayele, Jean-Pierre Chevallet, Million Meshesha and Getnet Kassie</td>
<td>133</td>
</tr>
<tr>
<td>Translation Quality-Based Supplementary Data Selection by Incremental Update of Translation Models</td>
<td>Pratyush Banerjee, Sudip Kumar Naskar, Johann Roturier, Andy Way and Josef van Genabith</td>
<td>149</td>
</tr>
<tr>
<td>Text Reuse Detection using a Composition of Text Similarity Measures</td>
<td>Daniel Bär, Torsten Zesch and Iryna Gurevych</td>
<td>167</td>
</tr>
<tr>
<td>Deriving Paraphrases for Highly Inflected Languages from Comparable Documents</td>
<td>Kfir Bar and Nachum Dershowitz</td>
<td>185</td>
</tr>
<tr>
<td>Harvesting Parallel Text in Multiple Languages with Limited Supervision</td>
<td>Luciano Barbosa, Vivek Kumar Rangarajan Sridhar, Mahsa Yarmohammadi and Srinivas Bangalore</td>
<td>201</td>
</tr>
<tr>
<td>An Evaluation of Statistical Post-Editing Systems Applied to RBMT and SMT Systems</td>
<td>Hanna Béchara, Raphaël Rubino, Yifan He, Yanjun Ma and Josef van Genabith</td>
<td>215</td>
</tr>
<tr>
<td>Prague Dependency Treebank 2.5 – a Revisited Version of PDT 2.0</td>
<td>Eduard Bejcik, Jarmila Panevova, Jan Popelka, Pavel Stranak, Magda Sevcikova, Jan Stepanekek and Zdenek Zabokrtský</td>
<td>231</td>
</tr>
</tbody>
</table>
Deriving a Lexicon for a Precision Grammar from Language Documentation Resources: A Case Study of Chintang
Emily M. Bender, Robert Schikowski and Balthasar Bickel .......................... 247

Quantifying Semantics using Complex Network Analysis
Chris Biemann, Stefanie Roos and Karsten Weihe ................................. 263

Improvements to Training an RNN parser
Richard Billingsley and James Curran.................................................. 279

Thread Specific Features are Helpful for Identifying Subjectivity Orientation of Online Forum Threads
Prakhar Biyani, Sumit Bhatia, Cornelia Caragea and Prasenjit Mitra .............. 295

Natural Language Generation for Nature Conservation: Automating Feedback to Help Volunteers Identify Bumblebee Species
Steven Blake, Advaith Siddharthan, Hien Nguyen, Nirwan Sharma, Anne-Marie Robinson, Elaine O'Mahony, Ben Darvill, Chris Mellish and Rene van der Wal ................. 311

Studying the Effect of Input Size for Bayesian Word Segmentation on the Providence Corpus
Benjamin Börschinger, Katherine Demuth and Mark Johnson...................... 325

Bayesian Language Modelling of German Compounds
Jan A. Botha, Chris Dyer and Phil Blunsom ........................................ 341

Can Spanish Be Simpler? LexSiS: Lexical Simplification for Spanish
Stefan Bott, Luz Rello, Biljana Drndarevic and Horacio Saggion .................. 357

Identification of Social Acts in Dialogue
David Bracewell, Marc Tomlinson and Hui Wang .................................. 375

Robust, Lexicalized Native Language Identification
Julian Brooke and Graeme Hirst ..................................................... 391

Identifying Urdu Complex Predication via Bigram Extraction
Miriam Butt, Tina Bögel, Annette Hautli, Sebastian Sulger and Tafseer Ahmed ... 409

Native Language Identification using Recurring n-grams – Investigating Abstraction and Domain Dependence
Serhiy Bykh and Detmar Meurers ....................................................... 425

Analysis and Enhancement of Wikification for Microblogs with Context Expansion
Taylor Cassidy, Heng Ji, Lev-Arie Ratinov, Arkaitz Zubiaga and Hongzhao Huang .... 441

On the Effectiveness of using Sentence Compression Models for Query-Focused Multi-Document Summarization
Yllias Chali and Sadid A. Hasan ......................................................... 457

Towards Automatic Topical Question Generation
Yllias Chali and Sadid A. Hasan .......................................................... 475

Adjective Deletion for Linguistic Steganography and Secret Sharing
Ching-Yun Chang and Stephen Clark ...................................................... 493
The Secret’s in the Word Order: Text-to-Text Generation for Linguistic Steganography
Ching-Yun Chang and Stephen Clark.............................................. 511

Joint Modeling for Chinese Event Extraction with Rich Linguistic Features
Chen Chen and Vincent Ng .......................................................... 529

A Simplification-Translation-Restoration Framework for Cross-Domain SMT Applications
Han-Bin Chen, Hen-Hsen Huang, Hsin-Hsi Chen and Ching-Ting Tan........ 545

A Semi-Supervised Bayesian Network Model for Microblog Topic Classification
Yan Chen, Zhoujun Li, Liqiang Nie, Xia Hu, Xiangyu Wang, Tat-Seng Chua and Xiaoming
Zhang....................................................................................... 561

A System for Multilingual Sentiment Learning On Large Data Sets
Alex Cheng and Oles Zhulyn....................................................... 577

Extraction of Russian Sentiment Lexicon for Product Meta-Domain
Ilia Chetviorkin and Natalia Loukachevitch .................................. 593

Problems in Evaluating Grammatical Error Detection Systems
Martin Chodorow, Markus Dickinson, Ross Israel and Joel Tetreault........... 611

Unsupervised and Semi-Supervised Morphological Analysis for Information Retrieval in the Biomedical Domain
Vincent Claveau....................................................................... 629

A Hybrid Approach to Finding Phenotype Candidates in Genetic Texts
Nigel Collier, Mai-Vu Tran, Hoang-Quynh Le, Anika Oellrich, Ai Kawazoe, Martin Hall-May
and Dietrich Rebholz-Schuhmann .................................................. 647

Using Argumentative Zones for Extractive Summarization of Scientific Articles
Danish Contractor, Yufan Guo and Anna Korhonen.............................. 663

Annotation Tools and Knowledge Representation for a Text-To-Scene System
Bob Coyne, Alex Klapheke, Masoud Rouhizadeh, Richard Sproat and Daniel Bauer 679

Towards Efficient HPSG Generation for German, a Non-Configurational Language
Berthold Crysmann and Woodley Packard.................................... 695

A Corpus-Based Study of Edit Categories in Featured and Non-Featured Wikipedia Articles
Johannes Daxenberger and Iryna Gurevych...................................... 711

A Computational Cognitive Model for Semantic Sub-Network Extraction from Natural Language Queries
Suman Deb Roy and Wenjun Zeng................................................ 727

Extraction of Domain-Specific Bilingual Lexicon from Comparable Corpora: Compositional Translation and Ranking
Estelle Delpech, Béatrice Daille, Emmanuel Morin and Claire Lemaire........ 745

Twitter Topic Summarization by Ranking Tweets using Social Influence and Content Quality
Yajuan Duan, Zhumin Chen, Furu Wei, Ming Zhou and Heung-Yeung Shum... 763

S-Restricted Monotone Alignments: Algorithm, Search Space, and Applications
Steffen Eger........................................................................... 781
Mining Words in the Minds of Second Language Learners: Learner-Specific Word Difficulty  
Yo Ehara, Issei Sato, HidekazuOiwa and Hiroshi Nakagawa ........................ 799

Jointly Disambiguating and Clustering Concepts and Entities with Markov Logic  
Angela Fahrni and Michael Strube ................................................. 815

Flexible Structural Analysis of Near-Meet-Semilattices for Typed Unification-Based Grammar Design  
Rouzbeh Farahmand and Gerald Penn ............................................. 833

Stacking of Dependency and Phrase Structure Parsers  
Richárd Farkas and Bernd Bohnet ................................................. 849

Semantic Cohesion Model for Phrase-Based SMT  
Minwei Feng, Weiwei Sun and Hermann Ney ...................................... 867

Comparing Taxonomies for Organising Collections of Documents  
Samuel Fernando, Mark Hall, Eneko Agirre, Aitor Soroa, Paul Clough and Mark Stevenson 879

Modeling the Complexity of Manual Annotation Tasks: a Grid of Analysis  
Karën Fort, Adeline Nazarenko and Sophie Rosset ............................ 895

Extractive Multi-Document Summarization with Integer Linear Programming and Support Vector Regression  
Dimitrios Galanis, Gerasimos Lampouras and Ion Androutsopoulos .............. 911

Cross-Lingual Topical Relevance Models  
Debasis Ganguly, Johannes Leveling and Gareth Jones .......................... 927

Structured Term Recognition in Medical Text  
Michael Glass and Alfio Gliozzo ................................................... 943

A Dynamic Oracle for Arc-Eager Dependency Parsing  
Yoav Goldberg and Joakim Nivre .................................................. 959

Statistical Mechanical Analysis of Semantic Orientations on Lexical Network  
Takuma Goto, Yoshiyuki Kabashima and Hiroya Takamura ......................... 977

Finding Thoughtful Comments from Social Media  
Swapna Gottipati and Jing Jiang .................................................. 995

A Distributed Platform for Sanskrit Processing  
Pawan Goyal, Gérard Huet, Amba Kulkarni, Peter Scharf and Ralph Bunker ...... 1011

Understanding the Performance of Statistical MT Systems: A Linear Regression Framework  
Francisco Guzman and Stephan Vogel ............................................. 1029

Geolocation Prediction in Social Media Data by Finding Location Indicative Words  
Bo Han, Paul Cook and Timothy Baldwin ........................................ 1045

Readability Classification for German using Lexical, Syntactic, and Morphological Features  
Julia Hancke, Sowmya Vajjala and Detmar Meurers ................................ 1063
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walk-based Computation of Contextual Word Similarity</td>
<td>Kazuo Hara, Ikumi Suzuki, Masashi Shimbo and Yuji Matsumoto</td>
<td>1081</td>
</tr>
<tr>
<td>Flexible Japanese Sentence Compression by Relaxing Unit Constraints</td>
<td>Jun Harashima and Sadao Kurohashi</td>
<td>1097</td>
</tr>
<tr>
<td>Approximating Theoretical Linguistics Classification in Real Data:</td>
<td>Boris Haselbach, Kerstin Eckart, Wolfgang Seeker, Kurt Eberle and Ulrich</td>
<td>1113</td>
</tr>
<tr>
<td>the Case of German “nach” Particle Verbs</td>
<td>Heid</td>
<td></td>
</tr>
<tr>
<td>Bridging the Gap between Intrinsic and Perceived Relevance in Snippet</td>
<td>Jing He, Pablo Duboue and Jian-Yun Nie</td>
<td>1129</td>
</tr>
<tr>
<td>Generation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A Comparison and Improvement of Online Learning Algorithms for</td>
<td>Zhengyan He and Houfeng Wang</td>
<td>1147</td>
</tr>
<tr>
<td>Sequence Labeling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creating an Extended Named Entity Dictionary from Wikipedia</td>
<td>Ryuichiro Higashinaka, Kugatsu Sadamitsu, Kuniko Saito, Toshiro Makino</td>
<td>1163</td>
</tr>
<tr>
<td>and Yoshihiro Matsuo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistical Method of Building Dialect Language Models for ASR Systems</td>
<td>Naoki Hirayama, Shinsuke Mori and Hiroshi G. Okuno</td>
<td>1179</td>
</tr>
<tr>
<td>Tailored Feature Extraction for Lexical Disambiguation of English</td>
<td>Martin Holub, Vincent Kríž, Silvie Cinková and Eckhard Bick</td>
<td>1195</td>
</tr>
<tr>
<td>Verbs Based on Corpus Pattern Analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Method Mention Extraction from Scientific Research Papers</td>
<td>Hospice Houngbo and Robert E. Mercer</td>
<td>1211</td>
</tr>
<tr>
<td>Context-Enhanced Personalized Social Summarization</td>
<td>Po Hu, Donghong Ji, Chong Teng and Yujing Gu</td>
<td>1223</td>
</tr>
<tr>
<td>Tweet Ranking Based on Heterogeneous Networks</td>
<td>Hongzhao Huang, Arkaitz Zubiaga, Heng Ji, Hongbo Deng, Dong Wang,</td>
<td>1239</td>
</tr>
<tr>
<td></td>
<td>Hieu Le, Tarek Abdelzaher, Jiawei Han, Alice Leung, John Hancock and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Clare Voss</td>
<td></td>
</tr>
<tr>
<td>Improved Combinatory Categorial Grammar Induction with Boundary</td>
<td>Yun Huang, Min Zhang and Chew-Lim Tan</td>
<td>1257</td>
</tr>
<tr>
<td>Words and Bayesian Inference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mining Rules for Rewriting States in a Transition-based Dependency</td>
<td>Akihiro Inokuchi and Ayumu Yamaoka</td>
<td>1275</td>
</tr>
<tr>
<td>Parser for English</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coreference Resolution with ILP-based Weighted Abduction</td>
<td>Naoya Inoue, Ekaterina Ovchinnikova, Kentaro Inui and Jerry Hobbs</td>
<td>1291</td>
</tr>
<tr>
<td>N-gram Fragment Sequence Based Unsupervised Domain-Specific Document</td>
<td>Shoaib Jameel, Xiaojun Qian and Wai Lam</td>
<td>1309</td>
</tr>
<tr>
<td>Readability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Using Knowledge and Constraints To Find the Best Antecedent</td>
<td>Prateek Jindal and Dan Roth</td>
<td>1327</td>
</tr>
<tr>
<td>Towards a Generic and Flexible Citation Classifier Based on a</td>
<td>Charles Jochim and Hinrich Schütze</td>
<td>1343</td>
</tr>
<tr>
<td>Faceted Classification Scheme</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Semantics-Based Machine Translation with Hyperedge Replacement Grammars
Bevan Jones, Jacob Andreas, Daniel Bauer, Karl Moritz Hermann and Kevin Knight 1359

Answering Yes/No Questions via Question Inversion
Hiroshi Kanayama, Yusuke Miyao and John Prager 1377

Improving Topic Classification for Highly Inflective Languages
Jurgita Kapociute-Dzikiene, Frederik Vaassen, Walter Daelemans and Algis Krupavičius 1393

Generating “A for Alpha” When There Are Thousands of Characters
Hiroaki Kawasaki, Ryohei Sasano, Hiroya Takamura and Manabu Okumura 1411

A Machine Learning Approach for Phenotype Name Recognition
Maryam Khordad, Robert E Mercer and Peter Rogan 1425

Improving Combinatory Categorial Grammar Parse Reranking with Dependency Grammar Features
Sunghwan Mac Kim, Dominick Ng, Mark Johnson and James Curran 1441

Inducing Crosslingual Distributed Representations of Words
Alexandre Klementiev, Ivan Titov and Binod Bhattarai 1459

Exploring Local and Global Semantic Information for Event Pronoun Resolution
Fang Kong and Guodong Zhou 1475

Semantic Processing of Compounds in Indian Languages
Amba Kulkarni, Soma Paul, Malhar Kulkarni, Anil Kumar and Nitesh Surtani 1489

Unsupervised Japanese-Chinese Opinion Word Translation using Dependency Distance and Feature-Opinion Association Weight
Guo-Hau Lai, Ying-Mei Guo and Richard Tzong-Han Tsai 1503

On-line Trend Analysis with Topic Models: #twitter Trends Detection Topic Model Online
Jey Han Lau, Nigel Collier and Timothy Baldwin 1519

Learning Compositional Semantics for Open Domain Semantic Parsing
Phong Le and Willem Zuidema 1535

Evaluating Different Methods for Automatically Collecting Large General Corpora for Basque from the Web
Igor Leturia 1553

Approximate Sentence Retrieval for Scalable and Efficient Example-Based Machine Translation
Johannes Leveling, Debasis Ganguly, Sandipan Dandapat and Gareth Jones 1571

Improving Text Normalization using Character-Blocks Based Models and System Combination
Chen Li and Yang Liu 1587

Update Summarization using a Multi-level Hierarchical Dirichlet Process Model
Jiwei Li, Sujian Li, Xun Wang, Ye Tian and Baobao Chang 1603

Employing Morphological Structures and Sememes for Chinese Event Extraction
Peifeng Li and Guodong Zhou 1619
Joint Modeling of Trigger Identification and Event Type Determination in Chinese Event Extraction
Peifeng Li, Qiaoming Zhu, Hongjun Diao and Guodong Zhou .............................. 1635

Integrating Surface and Abstract Features for Robust Cross-Domain Chinese Word Segmentation
Xiaoqing Li, Kun Wang, Chengqing Zong and Keh-Yih Su ............................ 1653

Code-Switch Language Model with Inversion Constraints for Mixed Language Speech Recognition
Ying Li and Pascale Fung .................................................................................. 1671

A Separately Passive-Aggressive Training Algorithm for Joint POS Tagging and Dependency Parsing
Zhenghua Li, Min Zhang, Wanxiang Che and Ting Liu .................................. 1681

Graph-Based Multi-Tweet Summarization using Social Signals
Xiaohua Liu, Yitong Li, Furu Wei and Ming Zhou ............................................ 1699

Topical Word Trigger Model for Keyphrase Extraction
Zhiyuan Liu, Chen Liang and Maosong Sun .................................................... 1715

Easy-First Chinese POS Tagging and Dependency Parsing
Ji Ma, Tong Xiao, Jingbo Zhu and Feiliang Ren .............................................. 1731

Recognizing Personal Characteristics of Readers using Eye-Movements and Text Features
Pascual Martínez-Gómez, Tadayoshi Hara and Akiko Aizawa .......................... 1747

To Exhibit is not to Loiter: A Multilingual, Sense-Disambiguated Wiktionary for Measuring Verb Similarity
Christian M. Meyer and Iryna Gurevych ...................................................... 1763

Using Distributional Similarity for Lexical Expansion in Knowledge-based Word Sense Disambiguation
Tristan Miller, Chris Biemann, Torsten Zesch and Iryna Gurevych .................. 1781

Revising the Compositional Method for Terminology Acquisition from Comparable Corpora
Emmanuel Morin and Béatrice Daille ............................................................ 1797

Is Bad Structure Better Than No Structure?: Unsupervised Parsing for Realisation Ranking
Yasaman Motazedi, Mark Dras and François Larue ........................................ 1811

Analysis of Linguistic Style Accommodation in Online Debates
Arjun Mukherjee and Bing Liu ......................................................................... 1831

Sentiment Analysis in Twitter with Lightweight Discourse Analysis
Subhabrata Mukherjee and Pushpak Bhattacharyya ......................................... 1847

Subhabrata Mukherjee and Pushpak Bhattacharyya ......................................... 1865

Constrained Decoding for Text-Level Discourse Parsing
Philippe Muller, Stergos Afantenos, Pascal Denis and Nicholas Asher ............ 1883

Incremental Learning of Affix Segmentation
Wondwossen Mulugeta, Michael Gasser and Baye Yimam .............................. 1901
Semi-Supervised Noun Compound Analysis with Edge and Span Features
Yugo Murawaki and Sadao Kurohashi ............................................ 1915

Learning Effective and Interpretable Semantic Models using Non-Negative Sparse Embedding
Brian Murphy, Partha Talukdar and Tom Mitchell ............................................ 1933

Combining Wordnet and Morphosyntactic Information in Terminology Clustering
Agnieszka Mykowiecka and Malgorzata Marciniak ............................................ 1951

Alignment by Bilingual Generation and Monolingual Derivation
Toshiaki Nakazawa and Sadao Kurohashi ............................................ 1963

Optimizing for Sentence-Level BLEU+1 Yields Short Translations
Preslav Nakov, Francisco Guzman and Stephan Vogel ............................................ 1979

Grammarless Parsing for Joint Inference
Jason Naradowsky, Tim Vieira and David Smith ............................................ 1995

Error Mining with Susicion Trees: Seeing the Forest for the Trees
Shashi Narayan and Claire Gardent ............................................ 2011

Structure-Driven Lexicalist Generation
Shashi Narayan and Claire Gardent ............................................ 2027

A Comparison of Syntactic Reordering Methods for English-German Machine Translation
Jiri Navratil, Karthik Visweswariah and Ananthakrishnan Ramanathan ............................................ 2043

Grounded Language Acquisition: A Minimal Commitment Approach
Sushobhan Nayak and Amitabha Mukerjee ............................................ 2059

Bayesian Text Segmentation for Index Term Identification and Keyphrase Extraction
David Newman, Nagendra Koilada, Jey Han Lau and Timothy Baldwin ............................................ 2077

Exploiting Category-Specific Information for Multi-Document Summarization
Jun-Ping Ng, Praveen Bysani, Ziheng Lin, Min-Yen Kan and Chew-Lim Tan ............................................ 2093

Improved Temporal Relation Classification using Dependency Parses and Selective Crowdsourced Annotations
Jun-Ping Ng and Min-Yen Kan ............................................ 2109

Accurate Unbounded Dependency Recovery using Generalized Categorial Grammars
Luan Nguyen, Marten Van Schijndel and William Schuler ............................................ 2125

Tibetan Base Noun Phrase Identification Framework Based on Chinese-Tibetan Sentence Aligned Corpus
Ming Hua Nuo, Hui Dan Liu, Wei Na Zhao, Long Long Ma, Jian Wu and Zhi Ming Ding ............................................ 2141

A Pipeline Arabic Named Entity Recognition using a Hybrid Approach
Mai Oudah and Khaled Shaalan ............................................ 2159

Attribute Extraction from Conjectural Queries
Marius Pasca ............................................ 2177
A Comprehensive Analysis of Constituent Coordination for Grammar Engineering
Agnieszka Patejuk and Adam Przepiórkowski ........................................ 2191

Simple and Effective Parameter Tuning for Domain Adaptation of Statistical Machine Translation
Pavel Pecina, Antonio Toral and Josef van Genabith ............................... 2209

A Supervised Aggregation Framework for Multi-Document Summarization
Yulong Pei, Wenpeng Yin, Qifeng Fan and Lian’en Huang ........................ 2225

Collective Search for Concept Disambiguation
Anja Pilz and Gerhard Paaß ...................................................... 2243

Who’s (Really) the Boss? Perception of Situational Power in Written Interactions
Vinodkumar Prabhakaran, Owen Rambow and Mona Diab ....................... 2259

Bilingual Lexicon Construction from Comparable Corpora via Dependency Mapping
Longhua Qian, Hongling Wang, Guodong Zhou and Qiaoming Zhu .......... 2275

A MWE Acquisition and Lexicon Builder Web Service
Valeria Quochi, Francesca Frontini and Francesco Rubino ....................... 2291

A Diverse Dirichlet Process Ensemble for Unsupervised Induction of Syntactic Categories
Roi Reichart, Gal Elidan and Ari Rappoport ....................................... 2307

From Finite-State to Inversion Transductions: Toward Unsupervised Bilingual Grammar Induction
Markus Saers, Karteek Addanki and Dekai Wu .................................... 2325

Underspecified Query Refinement via Natural Language Question Generation
Hassan Sajjad, Patrick Pantel and Michael Gamon ................................ 2341

Joint English Spelling Error Correction and POS Tagging for Language Learners Writing
Keisuke Sakaguchi, Tomoya Mizumoto, Mamoru Komachi and Yuji Matsumoto ... 2357

Automatic Detection of Psychological Distress Indicators and Severity Assessment from Online Forum Posts
Shirin Saleem, Rohit Prasad, Shiv Vitaladevuni, Maciej Pacula, Michael Crystal, Brian Marx,
Denise Sloan, Jennifer Vasterling and Theodore Speroff ................................... 2375

Ant Colony Algorithm for the Unsupervised Word Sense Disambiguation of Texts: Comparison and Evaluation
Didier Schwab, Jérôme Goulian, Andon Tchechmedjiev and Hervé Blanchon .... 2389

Learnability-Based Syntactic Annotation Design
Roy Schwartz, Omri Abend and Ari Rappoport .................................... 2405

Improving Supervised Sense Disambiguation with Web-Scale Selectors
H. Andrew Schwartz, Fernando Gomez and Lyle Ungar ........................... 2423

The French Social Media Bank: a Treebank of Noisy User Generated Content
Djamé Seddah, Benoit Sagot, Marie Candito, Virginie Mouilleron and Vanessa Combet ... 2441

Initial Explorations on using CRFs for Turkish Named Entity Recognition
Gökhan Akın Şeker and Gülşen Eryiğit ........................................... 2459

xxxiii
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Differential Evolution Based Feature Selection and Classifier Ensemble for Named Entity Recognition</td>
<td>Utpal Kumar Sikdar, Asif Ekbal and Sriparna Saha</td>
<td>2475</td>
</tr>
<tr>
<td>Noun Group and Verb Group Identification for Hindi</td>
<td>Smriti Singh, Om P Damani and Vaijayanti M. Sarma</td>
<td>2491</td>
</tr>
<tr>
<td>Named Entity Recognition System for Urdu</td>
<td>UmrinderPal Singh, Vishal Goyal and Gurpreet Singh Lehal</td>
<td>2507</td>
</tr>
<tr>
<td>Easy-first Coreference Resolution</td>
<td>Veselin Stoyanov and Jason Eisner</td>
<td>2519</td>
</tr>
<tr>
<td>Modeling Leadership and Influence in Multi-party Online Discourse</td>
<td>Tomek Strzalkowski, Samira Shaikh, Ting Liu, George Aaron Broadwell, Jenny Stromer-Galley, Sarah Taylor, Umit Boz, Veena Ravishankar and Xiaoai Ren</td>
<td>2535</td>
</tr>
<tr>
<td>Evaluating the Translation Accuracy of a Novel Language-Independent MT Methodology</td>
<td>George Tambouratzis, Sokratis Sofianopoulos and Marina Vassiliou</td>
<td>2569</td>
</tr>
<tr>
<td>Native Tongues, Lost and Found: Resources and Empirical Evaluations in Native Language Identification</td>
<td>Joel Tetreault, Daniel Blanchard, Aoife Cahill and Martin Chodorow</td>
<td>2585</td>
</tr>
<tr>
<td>Inverse Document Density: A Smooth Measure for Location-Dependent Term Irregularities</td>
<td>Dennis Thom, Harald Bosch and Thomas Ertl</td>
<td>2603</td>
</tr>
<tr>
<td>Efficient Discrimination Between Closely Related Languages</td>
<td>Jörg Tiedemann and Nikola Ljubešić</td>
<td>2619</td>
</tr>
<tr>
<td>Semi-Supervised Semantic Role Labeling: Approaching from an Unsupervised Perspective</td>
<td>Ivan Titov and Alexandre Klementiev</td>
<td>2635</td>
</tr>
<tr>
<td>Hunting for Entailing Pairs in the Penn Discourse Treebank</td>
<td>Sara Tonelli and Elena Cabrio</td>
<td>2653</td>
</tr>
<tr>
<td>Implicitness of Discourse Relations</td>
<td>Fatemeh Torabi Asr and Vera Demberg</td>
<td>2669</td>
</tr>
<tr>
<td>Combining Statistical Translation Techniques for Cross-Language Information Retrieval</td>
<td>Ferhan Ture, Jimmy Lin and Douglas Oard</td>
<td>2685</td>
</tr>
<tr>
<td>Multi-way Tensor Factorization for Unsupervised Lexical Acquisition</td>
<td>Tim van de Cruys, Laura Rimell, Thierry Poibeau and Anna Korhonen</td>
<td>2703</td>
</tr>
<tr>
<td>Sub-corpora Sampling with an Application to Bilingual Lexicon Extraction</td>
<td>Ivan Vulić and Marie-Francine Moens</td>
<td>2721</td>
</tr>
<tr>
<td>The Utility of Discourse Structure in Identifying Resolved Threads in Technical User Forums</td>
<td>Li Wang, Su Nam Kim and Timothy Baldwin</td>
<td>2739</td>
</tr>
</tbody>
</table>
Implicit Discourse Relation Recognition by Selecting Typical Training Examples
   Xun Wang, Sujian Li, Jiwei Li and Wenjie Li ........................................ 2757

Chinese Evaluative Information Analysis
   Yiou Wang, Jun’ichi Kazama, Takuya Kawada and Kentaro Torisawa ........... 2773

Harnessing the CRF Complexity with Domain-Specific Constraints. The Case of Morphosyntactic Tagging of a Highly Inflected Language
   Jakub Waszczuk ................................................................. 2789

A Latent Discriminative Model for Compositional Entailment Relation Recognition using Natural Logic
   Yotaro Watanabe, Junta Mizuno, Eric Nichols, Naoaki Okazaki and Kentaro Inui ... 2805

Strategies for Mixed-Initiative Conversation Management using Question-Answer Pairs
   Wilson Wong, Lawrence Cavedon, John Thangarajah and Lin Padgham ............ 2821

Factored Language Model based on Recurrent Neural Network
   Youzheng Wu, Xugang Lu, Hitoshi Yamamoto, Shigeki Matsuda, Chiori Hori and Hideki Kashioka................................................................. 2835

Multi-View AdaBoost for Multilingual Subjectivity Analysis
   Min Xiao and Yuhong Guo ....................................................... 2851

Semi-supervised Representation Learning for Domain Adaptation using Dynamic Dependency Networks
   Min Xiao, Yuhong Guo and Alexander Yates ........................................ 2867

Unsupervised Discriminative Induction of Synchronous Grammar for Machine Translation
   Xinyan Xiao, Deyi Xiong, Yang Liu, Qun Liu and Shouxun Lin ...................... 2883

Paraphrasing for Style
   Wei Xu, Alan Ritter, Bill Dolan, Ralph Grishman and Colin Cherry ............... 2899

Modeling ESL Word Choice Similarities By Representing Word Intensions and Extensions
   Huichao Xue and Rebecca Hwa .................................................. 2915

ISO-TimeML Event Extraction in Persian Text
   Yadollah Yaghoobzadeh, Gholamreza Ghassem-Sani, Seyed Abolghassem Mirroshandel and Mahbaneh Eshaghzadeh ................................................. 2931

Measuring the Similarity between TV Programs using Semantic Relations
   Ichiro Yamada, Masaru Miyazaki, Hideki Sumiyoshi, Atsushi Matsu, Hironori Furumiya and Hideki Tanaka ................................................................. 2945

RelationListwise for Query-Focused Multi-Document Summarization
   Wepeng Yin, Lifu Huang, Yulong Pei and Lian’en Huang ........................... 2961

SentTopic-MultiRank: a Novel Ranking Model for Multi-Document Summarization
   Wepeng Yin, Yulong Pei, Fan Zhang and Lian’en Huang ............................. 2977

Language Modeling for Spoken Dialogue System based on Filtering using Predicate-Argument Structures
   Koichiro Yoshino, Shinsuke Mori and Tatsuya Kawahara ........................... 2993
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detecting Word Ordering Errors in Chinese Sentences for Learning Chinese as a Foreign Language</td>
<td>Chi-Hsin Yu and Hsin-Hsi Chen</td>
<td>3003</td>
</tr>
<tr>
<td>Machine Translation by Modeling Predicate-Argument Structure Transformation</td>
<td>Feifei Zhai, Jiajun Zhang, Yu Zhou and Chengqing Zong</td>
<td>3019</td>
</tr>
<tr>
<td>Tree-based Translation without using Parse Trees</td>
<td>Feifei Zhai, Jiajun Zhang, Yu Zhou and Chengqing Zong</td>
<td>3037</td>
</tr>
<tr>
<td>Constructing Chinese Abbreviation Dictionary: A Stacked Approach</td>
<td>Longkai Zhang, Sujian Li, Houfeng Wang, Ni Sun and Xinfan Meng</td>
<td>3055</td>
</tr>
<tr>
<td>Stacking Heterogeneous Joint Models of Chinese POS Tagging and Dependency Parsing</td>
<td>Meishan Zhang, Wanxiang Che, Ting Liu and Zhenghua Li</td>
<td>3071</td>
</tr>
<tr>
<td>A Lazy Learning Model for Entity Linking using Query-Specific Information</td>
<td>Wei Zhang, Jian Su, Chew-Lim Tan, Yunbo Cao and Chin-Yew Lin</td>
<td>3089</td>
</tr>
<tr>
<td>The Use of Dependency Relation Graph to Enhance the Term Weighting in Question Retrieval</td>
<td>Weinan Zhang, Zhaoyan Ming, Yu Zhang, Liqiang Nie, Ting Liu and Tat-Seng Chua</td>
<td>3105</td>
</tr>
<tr>
<td>Long-Tail Distributions and Unsupervised Learning of Morphology</td>
<td>Qiuye Zhao and Mitch Marcus</td>
<td>3121</td>
</tr>
<tr>
<td>User Behaviors Lend a Helping Hand: Learning Paraphrase Query Patterns from Search Log Sessions</td>
<td>Shiqi Zhao, Haifeng Wang and Ting Liu</td>
<td>3137</td>
</tr>
<tr>
<td>Exploiting Bilingual Translation for Question Retrieval in Community-Based Question Answering</td>
<td>Guangyou Zhou, Kang Liu and Jun Zhao</td>
<td>3153</td>
</tr>
<tr>
<td>Exploiting Lexical Dependencies from Large-Scale Data for Better Shift-Reduce Constituency Parsing</td>
<td>Muhua Zhu, Jingbo Zhu and Huizhen Wang</td>
<td>3171</td>
</tr>
</tbody>
</table>