

51st Annual Meeting of the Association for Computational Linguistics (ACL 2013)

Long Papers

**Sofia, Bulgaria
4 – 9 August 2013**

Volume 1 of 2

ISBN: 978-1-62748-975-1

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2013) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2013)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

| | |
|--|-----|
| <i>A Shift-Reduce Parsing Algorithm for Phrase-based String-to-Dependency Translation</i> Yang Liu | 1 |
| <i>Integrating Translation Memory into Phrase-Based Machine Translation during Decoding</i> Kun Wang, Chengqing Zong and Keh-Yih Su | 11 |
| <i>Training Nondeficient Variants of IBM-3 and IBM-4 for Word Alignment</i> Thomas Schoenemann | 22 |
| <i>Modelling Annotator Bias with Multi-task Gaussian Processes: An Application to Machine Translation Quality Estimation</i> Trevor Cohn and Lucia Specia | 32 |
| <i>Smoothed marginal distribution constraints for language modeling</i> Brian Roark, Cyril Allauzen and Michael Riley | 43 |
| <i>Grounded Language Learning from Video Described with Sentences</i> Haonan Yu and Jeffrey Mark Siskind | 53 |
| <i>Plurality, Negation, and Quantification: Towards Comprehensive Quantifier Scope Disambiguation</i> Mehdi Manshadi, Daniel Gildea and James Allen | 64 |
| <i>Joint Event Extraction via Structured Prediction with Global Features</i> Qi Li, Heng Ji and Liang Huang | 73 |
| <i>Language-Independent Discriminative Parsing of Temporal Expressions</i> Gabor Angeli and Jakob Uszkoreit | 83 |
| <i>Graph-based Local Coherence Modeling</i> Camille Guinaudeau and Michael Strube | 93 |
| <i>Recognizing Rare Social Phenomena in Conversation: Empowerment Detection in Support Group Chatrooms</i> Elijah Mayfield, David Adamson and Carolyn Penstein Rosé | 104 |
| <i>Decentralized Entity-Level Modeling for Coreference Resolution</i> Greg Durrett, David Hall and Dan Klein | 114 |
| <i>Chinese Parsing Exploiting Characters</i> Meishan Zhang, Yue Zhang, Wanxiang Che and Ting Liu | 125 |
| <i>A Transition-Based Dependency Parser Using a Dynamic Parsing Strategy</i> Francesco Sartorio, Giorgio Satta and Joakim Nivre | 135 |
| <i>General binarization for parsing and translation</i> Matthias Büchse, Alexander Koller and Heiko Vogler | 145 |
| <i>Distortion Model Considering Rich Context for Statistical Machine Translation</i> Isao Goto, Masao Utiyama, Eiichiro Sumita, Akihiro Tamura and Sadao Kurohashi | 155 |
| <i>Word Alignment Modeling with Context Dependent Deep Neural Network</i> Nan Yang, Shujie Liu, Mu Li, Ming Zhou and Nenghai Yu | 166 |

| | |
|--|-----|
| <i>Microblogs as Parallel Corpora</i> | |
| Wang Ling, Guang Xiang, Chris Dyer, Alan Black and Isabel Trancoso | 176 |
| <i>Improved Bayesian Logistic Supervised Topic Models with Data Augmentation</i> | |
| Jun Zhu, Xun Zheng and Bo Zhang | 187 |
| <i>Fast and Robust Compressive Summarization with Dual Decomposition and Multi-Task Learning</i> | |
| Miguel Almeida and Andre Martins | 196 |
| <i>Unsupervised Transcription of Historical Documents</i> | |
| Taylor Berg-Kirkpatrick, Greg Durrett and Dan Klein | 207 |
| <i>Adapting Discriminative Reranking to Grounded Language Learning</i> | |
| Joohyun Kim and Raymond Mooney | 218 |
| <i>Universal Conceptual Cognitive Annotation (UCCA)</i> | |
| Omri Abend and Ari Rappoport | 228 |
| <i>Linking Tweets to News: A Framework to Enrich Short Text Data in Social Media</i> | |
| Weiwei Guo, Hao Li, Heng Ji and Mona Diab | 239 |
| <i>A computational approach to politeness with application to social factors</i> | |
| Cristian Danescu-Niculescu-Mizil, Moritz Sudhof, Dan Jurafsky, Jure Leskovec and Christopher Potts | 250 |
| <i>Modeling Thesis Clarity in Student Essays</i> | |
| Isaac Persing and Vincent Ng | 260 |
| <i>Translating Italian connectives into Italian Sign Language</i> | |
| Camillo Lugaresi and Barbara Di Eugenio | 270 |
| <i>Stop-probability estimates computed on a large corpus improve Unsupervised Dependency Parsing</i> | |
| David Mareček and Milan Straka | 281 |
| <i>Transfer Learning for Constituency-Based Grammars</i> | |
| Yuan Zhang, Regina Barzilay and Amir Globerson | 291 |
| <i>A Context Free TAG Variant</i> | |
| Ben Swanson, Elif Yamangil, Eugene Charniak and Stuart Shieber | 302 |
| <i>Fast and Adaptive Online Training of Feature-Rich Translation Models</i> | |
| Spence Green, Sida Wang, Daniel Cer and Christopher D. Manning | 311 |
| <i>Advancements in Reordering Models for Statistical Machine Translation</i> | |
| Minwei Feng, Jan-Thorsten Peter and Hermann Ney | 322 |
| <i>A Markov Model of Machine Translation using Non-parametric Bayesian Inference</i> | |
| Yang Feng and Trevor Cohn | 333 |
| <i>Scaling Semi-supervised Naive Bayes with Feature Marginals</i> | |
| Michael Lucas and Doug Downey | 343 |
| <i>Learning Latent Personas of Film Characters</i> | |
| David Bamman, Brendan O'Connor and Noah A. Smith | 352 |

| | |
|---|-----|
| <i>Scalable Decipherment for Machine Translation via Hash Sampling</i> Sujith Ravi | 362 |
| <i>Automatic Interpretation of the English Possessive</i> Stephen Tratz and Eduard Hovy | 372 |
| <i>Is a 204 cm Man Tall or Small ? Acquisition of Numerical Common Sense from the Web</i> Katsuma Narisawa, Yotaro Watanabe, Junta Mizuno, Naoaki Okazaki and Kentaro Inui | 382 |
| <i>Probabilistic Domain Modelling With Contextualized Distributional Semantic Vectors</i> Jackie Chi Kit Cheung and Gerald Penn | 392 |
| <i>Extracting bilingual terminologies from comparable corpora</i> Ahmet Aker, Monica Paramita and Rob Gaizauskas | 402 |
| <i>The Haves and the Have-Nots: Leveraging Unlabelled Corpora for Sentiment Analysis</i> Kashyap Popat, Balamurali A.R, Pushpak Bhattacharyya and Gholamreza Haffari | 412 |
| <i>Large-scale Semantic Parsing via Schema Matching and Lexicon Extension</i> Qingqing Cai and Alexander Yates | 423 |
| <i>Fast and Accurate Shift-Reduce Constituent Parsing</i> Muhua Zhu, Yue Zhang, Wenliang Chen, Min Zhang and Jingbo Zhu | 434 |
| <i>Nonconvex Global Optimization for Latent-Variable Models</i> Matthew R. Gormley and Jason Eisner | 444 |
| <i>Parsing with Compositional Vector Grammars</i> Richard Socher, John Bauer, Christopher D. Manning and Ng Andrew Y. | 455 |
| <i>Discriminative state tracking for spoken dialog systems</i> Angeliki Metallinou, Dan Bohus and Jason Williams | 466 |
| <i>Leveraging Synthetic Discourse Data via Multi-task Learning for Implicit Discourse Relation Recognition</i> Man Lan, Yu Xu and Zhengyu Niu | 476 |
| <i>Combining Intra- and Multi-sentential Rhetorical Parsing for Document-level Discourse Analysis</i> Shafiq Joty, Giuseppe Carenini, Raymond Ng and Yashar Mehdad | 486 |
| <i>Improving pairwise coreference models through feature space hierarchy learning</i> Emmanuel Lassalle and Pascal Denis | 497 |
| <i>Feature-Based Selection of Dependency Paths in Ad Hoc Information Retrieval</i> K. Tamsin Maxwell, Jon Oberlander and W. Bruce Croft | 507 |
| <i>Coordination Structures in Dependency Treebanks</i> Martin Popel, David Mareček, Jan Štěpánek, Daniel Zeman and Zdeněk Žabokrtský | 517 |
| <i>GlossBoot: Bootstrapping Multilingual Domain Glossaries from the Web</i> Flavio De Benedictis, Stefano Faralli and Roberto Navigli | 528 |
| <i>Collective Annotation of Linguistic Resources: Basic Principles and a Formal Model</i> Ulle Endriss and Raquel Fernández | 539 |

| | |
|--|-----|
| <i>ParGramBank: The ParGram Parallel Treebank</i> | |
| Sebastian Sulger, Miriam Butt, Tracy Holloway King, Paul Meurer, Tibor Laczkó, György Rákosi, Cheikh Bamba Dione, Helge Dyvik, Victoria Rosén, Koenraad De Smedt, Agnieszka Patejuk, Ozlem Cetinoglu, I Wayan Arka and Meladel Mistica | 550 |
| <i>Identifying Bad Semantic Neighbors for Improving Distributional Thesauri</i> | |
| Olivier Ferret | 561 |
| <i>Models of Semantic Representation with Visual Attributes</i> | |
| Carina Silberer, Vittorio Ferrari and Mirella Lapata | 572 |
| <i>Real-World Semi-Supervised Learning of POS-Taggers for Low-Resource Languages</i> | |
| Dan Garrette, Jason Mielens and Jason Baldridge | 583 |
| <i>Using subcategorization knowledge to improve case prediction for translation to German</i> | |
| Marion Weller, Alexander Fraser and Sabine Schulte im Walde | 593 |
| <i>Name-aware Machine Translation</i> | |
| Haibo Li, Jing Zheng, Heng Ji, Qi Li and Wen Wang | 604 |
| <i>Decipherment Complexity in 1:1 Substitution Ciphers</i> | |
| Malte Nuhn and Hermann Ney | 615 |
| <i>Non-Monotonic Sentence Alignment via Semisupervised Learning</i> | |
| Xiaojun Quan, Chunyu Kit and Yan Song | 622 |
| <i>Bootstrapping Entity Translation on Weakly Comparable Corpora</i> | |
| Taesung Lee and Seung-won Hwang | 631 |
| <i>Transfer Learning Based Cross-lingual Knowledge Extraction for Wikipedia</i> | |
| Zhigang Wang, Zhixing Li, Juanzi Li, Jie Tang and Jeff Z. Pan | 641 |
| <i>Bridging Languages through Etymology: The case of cross language text categorization</i> | |
| Vivi Nastase and Carlo Strapparava | 651 |
| <i>Creating Similarity: Lateral Thinking for Vertical Similarity Judgments</i> | |
| Tony Veale and Guofu Li | 660 |
| <i>Discovering User Interactions in Ideological Discussions</i> | |
| Arjun Mukherjee and Bing Liu | 671 |
| <i>Multilingual Affect Polarity and Valence Prediction in Metaphor-Rich Texts</i> | |
| Zornitsa Kozareva | 682 |
| <i>Large tagset labeling using Feed Forward Neural Networks. Case study on Romanian Language</i> | |
| Tiberiu Boros, Radu Ion and Dan Tufis | 692 |
| <i>Learning to lemmatise Polish noun phrases</i> | |
| Adam Radziszewski | 701 |
| <i>Using Conceptual Class Attributes to Characterize Social Media Users</i> | |
| Shane Bergsma and Benjamin Van Durme | 710 |
| <i>The Impact of Topic Bias on Quality Flaw Prediction in Wikipedia</i> | |
| Oliver Ferschke, Iryna Gurevych and Marc Rittberger | 721 |

| | |
|--|-----|
| <i>Mining Informal Language from Chinese Microtext: Joint Word Recognition and Segmentation</i> Aobo Wang and Min-Yen Kan | 731 |
| <i>Generating Synthetic Comparable Questions for News Articles</i> Oleg Rokhlenko and Idan Szpektor | 742 |
| <i>Punctuation Prediction with Transition-based Parsing</i> Dongdong Zhang, Shuangzhi Wu, Nan Yang and Mu Li | 752 |
| <i>Discriminative Learning with Natural Annotations: Word Segmentation as a Case Study</i> Wenbin Jiang, Meng Sun, Yajuan Lü, Yating Yang and Qun Liu | 761 |
| <i>Graph-based Semi-Supervised Model for Joint Chinese Word Segmentation and Part-of-Speech Tagging</i> Xiaodong Zeng, Derek F. Wong, Lidia S. Chao and Isabel Trancoso | 770 |
| <i>An Infinite Hierarchical Bayesian Model of Phrasal Translation</i> Trevor Cohn and Gholamreza Haffari | 780 |
| <i>Additive Neural Networks for Statistical Machine Translation</i> Iemao Liu, Taro Watanabe, Eiichiro Sumita and Tiejun Zhao | 791 |
| <i>Hierarchical Phrase Table Combination for Machine Translation</i> Conghui Zhu, Taro Watanabe, Eiichiro Sumita and Tiejun Zhao | 802 |
| <i>Shallow Local Multi-Bottom-up Tree Transducers in Statistical Machine Translation</i> Fabienne Braune, Nina Seemann, Daniel Quernheim and Andreas Maletti | 811 |
| <i>Enlisting the Ghost: Modeling Empty Categories for Machine Translation</i> Bing Xiang, Xiaoqiang Luo and Bowen Zhou | 822 |
| <i>A Multi-Domain Translation Model Framework for Statistical Machine Translation</i> Rico Sennrich, Holger Schwenk and Walid Aransa | 832 |
| <i>Part-of-Speech Induction in Dependency Trees for Statistical Machine Translation</i> Akihiro Tamura, Taro Watanabe, Eiichiro Sumita, Hiroya Takamura and Manabu Okumura ... | 841 |
| <i>Statistical Machine Translation Improves Question Retrieval in Community Question Answering via Matrix Factorization</i> Guangyou Zhou, Fang Liu, Yang Liu, Shizhu He and Jun Zhao | 852 |
| <i>Improved Lexical Acquisition through DPP-based Verb Clustering</i> Roi Reichart and Anna Korhonen | 862 |
| <i>Semantic Frames to Predict Stock Price Movement</i> Boyi Xie, Rebecca J. Passonneau, Leon Wu and Germán G. Creamer | 873 |
| <i>Density Maximization in Context-Sense Metric Space for All-words WSD</i> Koichi Tanigaki, Mitsuteru Shiba, Tatsuji Munaka and Yoshinori Sagisaka | 884 |
| <i>The Role of Syntax in Vector Space Models of Compositional Semantics</i> Karl Moritz Hermann and Phil Blunsom | 894 |
| <i>Margin-based Decomposed Amortized Inference</i> Gourab Kundu, Vivek Srikumar and Dan Roth | 905 |

| | |
|--|------|
| <i>Semi-Supervised Semantic Tagging of Conversational Understanding using Markov Topic Regression</i> Asli Celikyilmaz, Dilek Hakkani-Tur, Gokhan Tur and Ruhi Sarikaya | 914 |
| <i>Parsing Graphs with Hyperedge Replacement Grammars</i> David Chiang, Jacob Andreas, Daniel Bauer, Karl Moritz Hermann, Bevan Jones and Kevin Knight 924 | |
| <i>Grounded Unsupervised Semantic Parsing</i> Hoifung Poon | 933 |
| <i>Automatic detection of deception in child-produced speech using syntactic complexity features</i> Maria Yancheva and Frank Rudzicz | 944 |
| <i>Sentiment Relevance</i> Christian Scheible and Hinrich Schütze | 954 |
| <i>Predicting and Eliciting Addressee’s Emotion in Online Dialogue</i> Takayuki Hasegawa, Nobuhiro Kaji, Naoki Yoshinaga and Masashi Toyoda | 964 |
| <i>Utterance-Level Multimodal Sentiment Analysis</i> Veronica Perez-Rosas, Rada Mihalcea and Louis-Philippe Morency | 973 |
| <i>Probabilistic Sense Sentiment Similarity through Hidden Emotions</i> Mitra Mohtarami, Man Lan and Chew Lim Tan | 983 |
| <i>A user-centric model of voting intention from Social Media</i> Vasileios Lampos, Daniel PreoŃiuc-Pietro and Trevor Cohn | 993 |
| <i>Using Supervised Bigram-based ILP for Extractive Summarization</i> Chen Li, Xian Qian and Yang Liu | 1004 |
| <i>Summarization Through Submodularity and Dispersion</i> Anirban Dasgupta, Ravi Kumar and Sujith Ravi | 1014 |
| <i>Subtree Extractive Summarization via Submodular Maximization</i> Hajime Morita, Ryohei Sasano, Hiroya Takamura and Manabu Okumura | 1023 |
| <i>The effect of non-tightness on Bayesian estimation of PCFGs</i> Shay B. Cohen and Mark Johnson | 1033 |
| <i>Integrating Multiple Dependency Corpora for Inducing Wide-coverage Japanese CCG Resources</i> Sumire Uematsu, Takuya Matsuzaki, Hiroki Hanaoka, Yusuke Miyao and Hideki Mima | 1042 |
| <i>Transition-based Dependency Parsing with Selectional Branching</i> Jinho D. Choi and Andrew McCallum | 1052 |
| <i>Bilingually-Guided Monolingual Dependency Grammar Induction</i> Kai Liu, Yajuan Lü, Wenbin Jiang and Qun Liu | 1063 |
| <i>Joint Word Alignment and Bilingual Named Entity Recognition Using Dual Decomposition</i> Mengqiu Wang, Wanxiang Che and Christopher D. Manning | 1073 |
| <i>Resolving Entity Morphs in Censored Data</i> Hongzhao Huang, Zhen Wen, Dian Yu, Heng Ji, Yizhou Sun, Jiawei Han and He Li | 1083 |

| | |
|--|------|
| <i>Learning to Extract International Relations from Political Context</i> Brendan O'Connor, Brandon M. Stewart and Noah A. Smith | 1094 |
| <i>Graph Propagation for Paraphrasing Out-of-Vocabulary Words in Statistical Machine Translation</i> Majid Razmara, Maryam Siahbani, Reza Haffari and Anoop Sarkar | 1105 |
| <i>Online Relative Margin Maximization for Statistical Machine Translation</i> Vladimir Eidelman, Yuval Marton and Philip Resnik | 1116 |
| <i>Handling Ambiguities of Bilingual Predicate-Argument Structures for Statistical Machine Translation</i> Feifei Zhai, Jiajun Zhang, Yu Zhou and Chengqing Zong | 1127 |
| <i>Reconstructing an Indo-European Family Tree from Non-native English Texts</i> Ryo Nagata and Edward Whittaker | 1137 |
| <i>Word Association Profiles and their Use for Automated Scoring of Essays</i> Beata Beigman Klebanov and Michael Flor | 1148 |
| <i>Adaptive Parser-Centric Text Normalization</i> Congle Zhang, Tyler Baldwin, Howard Ho, Benny Kimelfeld and Yunyao Li | 1159 |
| <i>A Random Walk Approach to Selectional Preferences Based on Preference Ranking and Propagation</i> Zhenhua Tian, Hengheng Xiang, Ziqi Liu and Qinghua Zheng | 1169 |
| <i>ImpAr: A Deterministic Algorithm for Implicit Semantic Role Labelling</i> Egoitz Laparra and German Rigau | 1180 |
| <i>Cross-lingual Transfer of Semantic Role Labeling Models</i> Mikhail Kozhevnikov and Ivan Titov | 1190 |
| <i>DERivBase: Inducing and Evaluating a Derivational Morphology Resource for German</i> Britta Zeller, Jan Šnajder and Sebastian Padó | 1201 |
| <i>Crowdsourcing Interaction Logs to Understand Text Reuse from the Web</i> Martin Potthast, Matthias Hagen, Michael Völske and Benno Stein | 1212 |
| <i>SPred: Large-scale Harvesting of Semantic Predicates</i> Tiziano Flati and Roberto Navigli | 1222 |
| <i>Towards Robust Abstractive Multi-Document Summarization: A Caseframe Analysis of Centrality and Domain</i> Jackie Chi Kit Cheung and Gerald Penn | 1233 |
| <i>HEADY: News headline abstraction through event pattern clustering</i> Enrique Alfonseca, Daniele Pighin and Guillermo Garrido | 1243 |
| <i>Conditional Random Fields for Responsive Surface Realisation using Global Features</i> Nina Dethlefs, Helen Hastie, Heriberto Cuayáhuil and Oliver Lemon | 1254 |
| <i>Two-Neighbor Orientation Model with Cross-Boundary Global Contexts</i> Hendra Setiawan, Bowen Zhou, Bing Xiang and Libin Shen | 1264 |
| <i>Cut the noise: Mutually reinforcing reordering and alignments for improved machine translation</i> Karthik Visweswariah, Mitesh M. Khapra and Ananthakrishnan Ramanathan | 1275 |

| | |
|--|------|
| <i>Vector Space Model for Adaptation in Statistical Machine Translation</i> | |
| Boxing Chen, Roland Kuhn and George Foster | 1285 |
| <i>From Natural Language Specifications to Program Input Parsers</i> | |
| Tao Lei, Fan Long, Regina Barzilay and Martin Rinard | 1294 |
| <i>Entity Linking for Tweets</i> | |
| Xiaohua Liu, Yitong Li, Haocheng Wu, Ming Zhou, Furu Wei and Yi Lu | 1304 |
| <i>Identification of Speakers in Novels</i> | |
| Hua He, Denilson Barbosa and Grzegorz Kondrak | 1312 |
| <i>Language Acquisition and Probabilistic Models: keeping it simple</i> | |
| Aline Villavicencio, Marco Idiart, Robert Berwick and Igor Malioutov | 1321 |
| <i>A Two Level Model for Context Sensitive Inference Rules</i> | |
| Oren Melamud, Jonathan Berant, Ido Dagan, Jacob Goldberger and Idan Szpektor | 1331 |
| <i>Align, Disambiguate and Walk: A Unified Approach for Measuring Semantic Similarity</i> | |
| Mohammad Taher Pilehvar, David Jurgens and Roberto Navigli | 1341 |
| <i>Linking and Extending an Open Multilingual Wordnet</i> | |
| Francis Bond and Ryan Foster | 1352 |
| <i>FrameNet on the Way to Babel: Creating a Bilingual FrameNet Using Wiktionary as Interlingual Connection</i> | |
| Silvana Hartmann and Iryna Gurevych | 1363 |
| <i>Dirt Cheap Web-Scale Parallel Text from the Common Crawl</i> | |
| Jason R. Smith, Herve Saint-Amand, Magdalena Plamada, Philipp Koehn, Chris Callison-Burch and Adam Lopez | 1374 |
| <i>A Sentence Compression Based Framework to Query-Focused Multi-Document Summarization</i> | |
| Lu Wang, Hema Raghavan, Vittorio Castelli, Radu Florian and Claire Cardie | 1384 |
| <i>Domain-Independent Abstract Generation for Focused Meeting Summarization</i> | |
| Lu Wang and Claire Cardie | 1395 |
| <i>A Statistical NLG Framework for Aggregated Planning and Realization</i> | |
| Ravi Kondadadi, Blake Howald and Frank Schilder | 1406 |
| <i>Models of Translation Competitions</i> | |
| Mark Hopkins and Jonathan May | 1416 |
| <i>Learning a Phrase-based Translation Model from Monolingual Data with Application to Domain Adaptation</i> | |
| Jiajun Zhang and Chengqing Zong | 1425 |
| <i>SenseSpotting: Never let your parallel data tie you to an old domain</i> | |
| Marine Carpuat, Hal Daume III, Katharine Henry, Ann Irvine, Jagadeesh Jagarlamudi and Rachel Rudinger | 1435 |
| <i>BRAINSUP: Brainstorming Support for Creative Sentence Generation</i> | |
| Gözde Özbal, Daniele Pighin and Carlo Strapparava | 1446 |

| | |
|--|------|
| <i>Grammatical Error Correction Using Integer Linear Programming</i> Yuanbin Wu and Hwee Tou Ng | 1456 |
| <i>Text-Driven Toponym Resolution using Indirect Supervision</i> Michael Speriosu and Jason Baldridge | 1466 |
| <i>Argument Inference from Relevant Event Mentions in Chinese Argument Extraction</i> Peifeng Li, Qiaoming Zhu and Guodong Zhou | 1477 |
| <i>Fine-grained Semantic Typing of Emerging Entities</i> Ndapandula Nakashole, Tomasz Tylenda and Gerhard Weikum | 1488 |
| <i>Embedding Semantic Similarity in Tree Kernels for Domain Adaptation of Relation Extraction</i> Barbara Plank and Alessandro Moschitti | 1498 |
| <i>A joint model of word segmentation and phonological variation for English word-final /t/-deletion</i> Benjamin Börschinger, Mark Johnson and Katherine Demuth | 1508 |
| <i>Compositional-ly Derived Representations of Morphologically Complex Words in Distributional Semantics</i> Angeliki Lazaridou, Marco Marelli, Roberto Zamparelli and Marco Baroni | 1517 |
| <i>Unsupervised Consonant-Vowel Prediction over Hundreds of Languages</i> Young-Bum Kim and Benjamin Snyder | 1527 |
| <i>Improving Text Simplification Language Modeling Using Unsimplified Text Data</i> David Kauchak | 1537 |
| <i>Combining Referring Expression Generation and Surface Realization: A Corpus-Based Investigation of Architectures</i> Sina Zarrieß and Jonas Kuhn | 1547 |
| <i>Named Entity Recognition using Cross-lingual Resources: Arabic as an Example</i> Kareem Darwish | 1558 |
| <i>Beam Search for Solving Substitution Ciphers</i> Malte Nuhn, Julian Schamper and Hermann Ney | 1568 |
| <i>Social Text Normalization using Contextual Graph Random Walks</i> Hany Hassan and Arul Menezes | 1577 |
| <i>Integrating Phrase-based Reordering Features into a Chart-based Decoder for Machine Translation</i> ThuyLinh Nguyen and Stephan Vogel | 1587 |
| <i>Machine Translation Detection from Monolingual Web-Text</i> Yuki Arase and Ming Zhou | 1597 |
| <i>Paraphrase-Driven Learning for Open Question Answering</i> Anthony Fader, Luke Zettlemoyer and Oren Etzioni | 1608 |
| <i>Aid is Out There: Looking for Help from Tweets during a Large Scale Disaster</i> István Varga, Motoki Sano, Kentaro Torisawa, Chikara Hashimoto, Kiyonori Ohtake, Takao Kawai, Jong-Hoon Oh and Stijn De Saeger | 1619 |
| <i>A Bayesian Model for Joint Unsupervised Induction of Sentiment, Aspect and Discourse Representations</i> Angeliki Lazaridou, Ivan Titov and Caroline Sporleder | 1630 |

| | |
|--|------|
| <i>Joint Inference for Fine-grained Opinion Extraction</i> | |
| Bishan Yang and Claire Cardie | 1640 |
| <i>Linguistic Models for Analyzing and Detecting Biased Language</i> | |
| Marta Recasens, Cristian Danescu-Niculescu-Mizil and Dan Jurafsky | 1650 |
| <i>Evaluating a City Exploration Dialogue System with Integrated Question-Answering and Pedestrian Navigation</i> | |
| Srinivasan Janarthanam, Oliver Lemon, Phil Bartie, Tiphaine Dalmás, Anna Dickinson, Xingkun Liu, William Mackaness and Bonnie Webber | 1660 |
| <i>Lightly Supervised Learning of Procedural Dialog Systems</i> | |
| Svitlana Volkova, Pallavi Choudhury, Chris Quirk, Bill Dolan and Luke Zettlemoyer | 1669 |
| <i>Public Dialogue: Analysis of Tolerance in Online Discussions</i> | |
| Arjun Mukherjee, Vivek Venkataraman, Bing Liu and Sharon Meraz | 1680 |
| <i>Offspring from Reproduction Problems: What Replication Failure Teaches Us</i> | |
| Antske Fokkens, Marieke van Erp, Marten Postma, Ted Pedersen, Piek Vossen and Nuno Freire | 1691 |
| <i>Evaluating Text Segmentation using Boundary Edit Distance</i> | |
| Chris Fournier | 1702 |
| <i>Crowd Prefers the Middle Path: A New IAA Metric for Crowdsourcing Reveals Turker Biases in Query Segmentation</i> | |
| Rohan Ramanath, Monojit Choudhury, Kalika Bali and Rishiraj Saha Roy | 1713 |
| <i>Deceptive Answer Prediction with User Preference Graph</i> | |
| Fangtao Li, Yang Gao, Shuchang Zhou, Xiance Si and Decheng Dai | 1723 |
| <i>Why-Question Answering using Intra- and Inter-Sentential Causal Relations</i> | |
| Jong-Hoon Oh, Kentaro Torisawa, Chikara Hashimoto, Motoki Sano, Stijn De Saeger and Kiyonori Ohtake | 1733 |
| <i>Question Answering Using Enhanced Lexical Semantic Models</i> | |
| Wen-tau Yih, Ming-Wei Chang, Christopher Meek and Andrzej Pastusiak | 1744 |
| <i>Syntactic Patterns versus Word Alignment: Extracting Opinion Targets from Online Reviews</i> | |
| Kang Liu, Liheng Xu and Jun Zhao | 1754 |
| <i>Mining Opinion Words and Opinion Targets in a Two-Stage Framework</i> | |
| Liheng Xu, Kang Liu, Siwei Lai, Yubo Chen and Jun Zhao | 1764 |
| <i>Connotation Lexicon: A Dash of Sentiment Beneath the Surface Meaning</i> | |
| Song Feng, Jun Seok Kang, Polina Kuznetsova and Yejin Choi | 1774 |

51st Annual Meeting of the Association for Computational Linguistics (ACL 2013)

Short Papers

**Sofia, Bulgaria
4 – 9 August 2013**

Volume 2 of 2

ISBN: 978-1-62748-975-1

Table of Contents

| | |
|--|----|
| <i>Translating Dialectal Arabic to English</i> Hassan Sajjad, Kareem Darwish and Yonatan Belinkov | 1 |
| <i>Exact Maximum Inference for the Fertility Hidden Markov Model</i> Chris Quirk | 7 |
| <i>A Tale about PRO and Monsters</i> Preslav Nakov, Francisco Guzmán and Stephan Vogel | 12 |
| <i>Supervised Model Learning with Feature Grouping based on a Discrete Constraint</i> Jun Suzuki and Masaaki Nagata | 18 |
| <i>Exploiting Topic based Twitter Sentiment for Stock Prediction</i> Jianfeng Si, Arjun Mukherjee, Bing Liu, Qing Li, Huayi Li and Xiaotie Deng | 24 |
| <i>Learning Entity Representation for Entity Disambiguation</i> Zhengyan He, Shujie Liu, Mu Li, Ming Zhou, Longkai Zhang and Houfeng Wang | 30 |
| <i>Natural Language Models for Predicting Programming Comments</i> Dana Movshovitz-Attias and William W. Cohen | 35 |
| <i>Paraphrasing Adaptation for Web Search Ranking</i> Chenguang Wang, Nan Duan, Ming Zhou and Ming Zhang | 41 |
| <i>Semantic Parsing as Machine Translation</i> Jacob Andreas, Andreas Vlachos and Stephen Clark | 47 |
| <i>A relatedness benchmark to test the role of determiners in compositional distributional semantics</i> Raffaella Bernardi, Georgiana Dinu, Marco Marelli and Marco Baroni | 53 |
| <i>An Empirical Study on Uncertainty Identification in Social Media Context</i> Zhongyu Wei, Junwen Chen, Wei Gao, Binyang Li, Lanjun Zhou, Yulan He and Kam-Fai Wong | 58 |
| <i>PARMA: A Predicate Argument Aligner</i> Travis Wolfe, Benjamin Van Durme, Mark Dredze, Nicholas Andrews, Charley Beller, Chris Callison-Burch, Jay DeYoung, Justin Snyder, Jonathan Weese, Tan Xu and Xuchen Yao | 63 |
| <i>Aggregated Word Pair Features for Implicit Discourse Relation Disambiguation</i> Or Biran and Kathleen McKeown | 69 |
| <i>Implicatures and Nested Beliefs in Approximate Decentralized-POMDPs</i> Adam Vogel, Christopher Potts and Dan Jurafsky | 74 |
| <i>Domain-Specific Coreference Resolution with Lexicalized Features</i> Nathan Gilbert and Ellen Riloff | 81 |
| <i>Learning to Order Natural Language Texts</i> Jiwei Tan, Xiaojun Wan and Jianguo Xiao | 87 |
| <i>Universal Dependency Annotation for Multilingual Parsing</i> Ryan McDonald, Joakim Nivre, Yvonne Quirnbach-Brundage, Yoav Goldberg, Dipanjan Das, Kuzman Ganchev, Keith Hall, Slav Petrov, Hao Zhang, Oscar Täckström, Claudia Bedini, Núria Bertomeu Castelló and Jungmee Lee | 92 |

| | |
|--|-----|
| <i>An Empirical Examination of Challenges in Chinese Parsing</i> | |
| Jonathan K. Kummerfeld, Daniel Tse, James R. Curran and Dan Klein | 98 |
| <i>Joint Inference for Heterogeneous Dependency Parsing</i> | |
| Guangyou Zhou and Jun Zhao | 104 |
| <i>Easy-First POS Tagging and Dependency Parsing with Beam Search</i> | |
| Ji Ma, Jingbo Zhu, Tong Xiao and Nan Yang | 110 |
| <i>Arguments and Modifiers from the Learner’s Perspective</i> | |
| Leon Bergen, Edward Gibson and Timothy J. O’Donnell | 115 |
| <i>Benefactive/Malefactive Event and Writer Attitude Annotation</i> | |
| Lingjia Deng, Yoonjung Choi and Janyce Wiebe | 120 |
| <i>GuiTAR-based Pronominal Anaphora Resolution in Bengali</i> | |
| Apurbalal Senapati and Utpal Garain | 126 |
| <i>A Decade of Automatic Content Evaluation of News Summaries: Reassessing the State of the Art</i> | |
| Peter A. Rankel, John M. Conroy, Hoa Trang Dang and Ani Nenkova | 131 |
| <i>On the Predictability of Human Assessment: when Matrix Completion Meets NLP Evaluation</i> | |
| Guillaume Wisniewski | 137 |
| <i>Automated Pyramid Scoring of Summaries using Distributional Semantics</i> | |
| Rebecca J. Passonneau, Emily Chen, Weiwei Guo and Dolores Perin | 143 |
| <i>Are Semantically Coherent Topic Models Useful for Ad Hoc Information Retrieval?</i> | |
| Romain Deveaud, Eric SanJuan and Patrice Bellot | 148 |
| <i>Post-Retrieval Clustering Using Third-Order Similarity Measures</i> | |
| Jose G. Moreno, Gaël Dias and Guillaume Cleuziou | 153 |
| <i>Automatic Coupling of Answer Extraction and Information Retrieval</i> | |
| Xuchen Yao, Benjamin Van Durme and Peter Clark | 159 |
| <i>An improved MDL-based compression algorithm for unsupervised word segmentation</i> | |
| Ruey-Cheng Chen | 166 |
| <i>Co-regularizing character-based and word-based models for semi-supervised Chinese word segmentation</i> | |
| Xiaodong Zeng, Derek F. Wong, Lidia S. Chao and Isabel Trancoso | 171 |
| <i>Improving Chinese Word Segmentation on Micro-blog Using Rich Punctuations</i> | |
| Longkai Zhang, Li Li, Zhengyan He, Houfeng Wang and Ni Sun | 177 |
| <i>Accurate Word Segmentation using Transliteration and Language Model Projection</i> | |
| Masato Hagiwara and Satoshi Sekine | 183 |
| <i>Broadcast News Story Segmentation Using Manifold Learning on Latent Topic Distributions</i> | |
| Xiaoming Lu, Lei Xie, Cheung-Chi Leung, Bin Ma and Haizhou Li | 190 |
| <i>Is word-to-phone mapping better than phone-phone mapping for handling English words?</i> | |
| Naresh Kumar Elluru, Anandaswarup Vadapalli, Raghavendra Elluru, Hema Murthy and Kishore Prahallad | 196 |

| | |
|--|-----|
| <i>Enriching Entity Translation Discovery using Selective Temporality</i> Gae-won You, Young-rok Cha, Jinhan Kim and Seung-won Hwang | 201 |
| <i>Combination of Recurrent Neural Networks and Factored Language Models for Code-Switching Language Modeling</i> Heike Adel, Ngoc Thang Vu and Tanja Schultz | 206 |
| <i>Latent Semantic Matching: Application to Cross-language Text Categorization without Alignment Information</i> Tsutomu Hirao, Tomoharu Iwata and Masaaki Nagata | 212 |
| <i>TopicSpam: a Topic-Model based approach for spam detection</i> Jiwei Li, Claire Cardie and Sujian Li | 217 |
| <i>Semantic Neighborhoods as Hypergraphs</i> Chris Quirk and Pallavi Choudhury | 222 |
| <i>Unsupervised joke generation from big data</i> Saša Petrović and David Matthews | 228 |
| <i>Modeling of term-distance and term-occurrence information for improving n-gram language model performance</i> Tze Yuang Chong, Rafael E. Banchs, Eng Siong Chng and Haizhou Li | 233 |
| <i>Discriminative Approach to Fill-in-the-Blank Quiz Generation for Language Learners</i> Keisuke Sakaguchi, Yuki Arase and Mamoru Komachi | 238 |
| <i>"Let Everything Turn Well in Your Wife": Generation of Adult Humor Using Lexical Constraints</i> Alessandro Valitutti, Hannu Toivonen, Antoine Doucet and Jukka M. Toivanen | 243 |
| <i>Random Walk Factoid Annotation for Collective Discourse</i> Ben King, Rahul Jha, Dragomir Radev and Robert Mankoff | 249 |
| <i>Identifying English and Hungarian Light Verb Constructions: A Contrastive Approach</i> Veronika Vincze, István Nagy T. and Richárd Farkas | 255 |
| <i>English-to-Russian MT evaluation campaign</i> Pavel Braslavski, Alexander Beloborodov, Maxim Khalilov and Serge Sharoff | 262 |
| <i>IndoNet: A Multilingual Lexical Knowledge Network for Indian Languages</i> Brijesh Bhatt, Lahari Poddar and Pushpak Bhattacharyya | 268 |
| <i>Building Japanese Textual Entailment Specialized Data Sets for Inference of Basic Sentence Relations</i> Kimi Kaneko, Yusuke Miyao and Daisuke Bekki | 273 |
| <i>Building Comparable Corpora Based on Bilingual LDA Model</i> Zede Zhu, Miao Li, Lei Chen and Zhenxin Yang | 278 |
| <i>Using Lexical Expansion to Learn Inference Rules from Sparse Data</i> Oren Melamud, Ido Dagan, Jacob Goldberger and Idan Szpektor | 283 |
| <i>Mining Equivalent Relations from Linked Data</i> Ziqi Zhang, Anna Lisa Gentile, Isabelle Augenstein, Eva Blomqvist and Fabio Ciravegna | 289 |
| <i>Context-Dependent Multilingual Lexical Lookup for Under-Resourced Languages</i> Lian Tze Lim, Lay-Ki Soon, Tek Yong Lim, Enya Kong Tang and Bali Ranaivo-Malançon | 294 |

| | |
|--|-----|
| <i>Sorani Kurdish versus Kurmanji Kurdish: An Empirical Comparison</i> Kyumars Sheykh Esmaili and Shahin Salavati | 300 |
| <i>Enhanced and Portable Dependency Projection Algorithms Using Interlinear Glossed Text</i> Ryan Georgi, Fei Xia and William D. Lewis | 306 |
| <i>Cross-lingual Projections between Languages from Different Families</i> Mo Yu, Tiejun Zhao, Yalong Bai, Hao Tian and Dianhai Yu | 312 |
| <i>Using Context Vectors in Improving a Machine Translation System with Bridge Language</i> Samira Tofighi Zahabi, Somayeh Bakhshaei and Shahram Khadivi | 318 |
| <i>Task Alternation in Parallel Sentence Retrieval for Twitter Translation</i> Felix Hieber, Laura Jehl and Stefan Riezler | 323 |
| <i>Sign Language Lexical Recognition With Propositional Dynamic Logic</i> Arturo Curiel and Christophe Collet | 328 |
| <i>Stacking for Statistical Machine Translation</i> Majid Razmara and Anoop Sarkar | 334 |
| <i>Bilingual Data Cleaning for SMT using Graph-based Random Walk</i> Lei Cui, Dongdong Zhang, Shujie Liu, Mu Li and Ming Zhou | 340 |
| <i>Automatically Predicting Sentence Translation Difficulty</i> Abhijit Mishra, Pushpak Bhattacharyya and Michael Carl | 346 |
| <i>Learning to Prune: Context-Sensitive Pruning for Syntactic MT</i> Wenduan Xu, Yue Zhang, Philip Williams and Philipp Koehn | 352 |
| <i>A Novel Graph-based Compact Representation of Word Alignment</i> Qun Liu, Zhaopeng Tu and Shouxun Lin | 358 |
| <i>Stem Translation with Affix-Based Rule Selection for Agglutinative Languages</i> Zhiyang Wang, Yajuan Lü, Meng Sun and Qun Liu | 364 |
| <i>A Novel Translation Framework Based on Rhetorical Structure Theory</i> Mei Tu, Yu Zhou and Chengqing Zong | 370 |
| <i>Improving machine translation by training against an automatic semantic frame based evaluation metric</i> Chi-kiu Lo, Karteek Addanki, Markus Saers and Dekai Wu | 375 |
| <i>Bilingual Lexical Cohesion Trigger Model for Document-Level Machine Translation</i> Guosheng Ben, Deyi Xiong, Zhiyang Teng, Yajuan Lü and Qun Liu | 382 |
| <i>Generalized Reordering Rules for Improved SMT</i> Fei Huang and Cezar Pendus | 387 |
| <i>A Tightly-coupled Unsupervised Clustering and Bilingual Alignment Model for Transliteration</i> Tingting Li, Tiejun Zhao, Andrew Finch and Chunyue Zhang | 393 |
| <i>Can Markov Models Over Minimal Translation Units Help Phrase-Based SMT?</i> Nadir Durrani, Alexander Fraser, Helmut Schmid, Hieu Hoang and Philipp Koehn | 399 |
| <i>Learning Non-linear Features for Machine Translation Using Gradient Boosting Machines</i> Kristina Toutanova and Byung-Gyu Ahn | 406 |

| | |
|---|-----|
| <i>Language Independent Connectivity Strength Features for Phrase Pivot Statistical Machine Translation</i> Ahmed El Kholy, Nizar Habash, Gregor Leusch, Evgeny Matusov and Hassan Sawaf | 412 |
| <i>Semantic Roles for String to Tree Machine Translation</i> Marzieh Bazrafshan and Daniel Gildea | 419 |
| <i>Minimum Bayes Risk based Answer Re-ranking for Question Answering</i> Nan Duan | 424 |
| <i>Question Classification Transfer</i> Anne-Laure Ligozat | 429 |
| <i>Latent Semantic Tensor Indexing for Community-based Question Answering</i> Xipeng Qiu, Le Tian and Xuanjing Huang | 434 |
| <i>Measuring semantic content in distributional vectors</i> Aur lie Herbelot and Mohan Ganesalingam | 440 |
| <i>Modeling Human Inference Process for Textual Entailment Recognition</i> Hen-Hsen Huang, Kai-Chun Chang and Hsin-Hsi Chen | 446 |
| <i>Recognizing Partial Textual Entailment</i> Omer Levy, Torsten Zesch, Ido Dagan and Iryna Gurevych | 451 |
| <i>Sentence Level Dialect Identification in Arabic</i> Heba Elfardy and Mona Diab | 456 |
| <i>Leveraging Domain-Independent Information in Semantic Parsing</i> Dan Goldwasser and Dan Roth | 462 |
| <i>A Structured Distributional Semantic Model for Event Co-reference</i> Kartik Goyal, Sujay Kumar Jauhar, Huiying Li, Mrinmaya Sachan, Shashank Srivastava and Eduard Hovy | 467 |
| <i>Text Classification from Positive and Unlabeled Data using Misclassified Data Correction</i> Fumiyo Fukumoto, Yoshimi Suzuki and Suguru Matsuyoshi | 474 |
| <i>Character-to-Character Sentiment Analysis in Shakespeare’s Plays</i> Eric T. Nalisnick and Henry S. Baird | 479 |
| <i>A Novel Classifier Based on Quantum Computation</i> Ding Liu, Xiaofang Yang and Minghu Jiang | 484 |
| <i>Re-embedding words</i> Igor Labutov and Hod Lipson | 489 |
| <i>LABR: A Large Scale Arabic Book Reviews Dataset</i> Mohamed Aly and Amir Atiya | 494 |
| <i>Generating Recommendation Dialogs by Extracting Information from User Reviews</i> Kevin Reschke, Adam Vogel and Dan Jurafsky | 499 |
| <i>Exploring Sentiment in Social Media: Bootstrapping Subjectivity Clues from Multilingual Twitter Streams</i> Svitlana Volkova, Theresa Wilson and David Yarowsky | 505 |

| | |
|---|-----|
| <i>Joint Modeling of News Reader’s and Comment Writer’s Emotions</i> Huanhuan Liu, Shoushan Li, Guodong Zhou, Chu-ren Huang and Peifeng Li | 511 |
| <i>An annotated corpus of quoted opinions in news articles</i> Tim O’Keefe, James R. Curran, Peter Ashwell and Irena Koprinska | 516 |
| <i>Dual Training and Dual Prediction for Polarity Classification</i> Rui Xia, Tao Wang, Xuelei Hu, Shoushan Li and Chengqing Zong | 521 |
| <i>Co-Regression for Cross-Language Review Rating Prediction</i> Xiaojun Wan | 526 |
| <i>Extracting Definitions and Hypernym Relations relying on Syntactic Dependencies and Support Vector Machines</i> Guido Boella and Luigi Di Caro | 532 |
| <i>Neighbors Help: Bilingual Unsupervised WSD Using Context</i> Sudha Bhingardive, Samiulla Shaikh and Pushpak Bhattacharyya | 538 |
| <i>Reducing Annotation Effort for Quality Estimation via Active Learning</i> Daniel Beck, Lucia Specia and Trevor Cohn | 543 |
| <i>Reranking with Linguistic and Semantic Features for Arabic Optical Character Recognition</i> Nadi Tomeh, Nizar Habash, Ryan Roth, Noura Farra, Pradeep Dasigi and Mona Diab | 549 |
| <i>Evolutionary Hierarchical Dirichlet Process for Timeline Summarization</i> Jiwei Li and Sujian Li | 556 |
| <i>Using Integer Linear Programming in Concept-to-Text Generation to Produce More Compact Texts</i> Gerasimos Lampouras and Ion Androutsopoulos | 561 |
| <i>Sequential Summarization: A New Application for Timely Updated Twitter Trending Topics</i> Dehong Gao, Wenjie Li and Renxian Zhang | 567 |
| <i>A System for Summarizing Scientific Topics Starting from Keywords</i> Rahul Jha, Amjad Abu-Jbara and Dragomir Radev | 572 |
| <i>A Unified Morpho-Syntactic Scheme of Stanford Dependencies</i> Reut Tsarfaty | 578 |
| <i>Dependency Parser Adaptation with Subtrees from Auto-Parsed Target Domain Data</i> Xuezhe Ma and Fei Xia | 585 |
| <i>Iterative Transformation of Annotation Guidelines for Constituency Parsing</i> Xiang Li, Wenbin Jiang, Yajuan Lü and Qun Liu | 591 |
| <i>Nonparametric Bayesian Inference and Efficient Parsing for Tree-adjointing Grammars</i> Elif Yamangil and Stuart M. Shieber | 597 |
| <i>Using CCG categories to improve Hindi dependency parsing</i> Bharat Ram Ambati, Tejaswini Deoskar and Mark Steedman | 604 |
| <i>The Effect of Higher-Order Dependency Features in Discriminative Phrase-Structure Parsing</i> Greg Coppola and Mark Steedman | 610 |

| | |
|--|-----|
| <i>Turning on the Turbo: Fast Third-Order Non-Projective Turbo Parsers</i> Andre Martins, Miguel Almeida and Noah A. Smith | 617 |
| <i>A Lattice-based Framework for Joint Chinese Word Segmentation, POS Tagging and Parsing</i> Zhiguo Wang, Chengqing Zong and Nianwen Xue | 623 |
| <i>Efficient Implementation of Beam-Search Incremental Parsers</i> Yoav Goldberg, Kai Zhao and Liang Huang | 628 |
| <i>Simpler unsupervised POS tagging with bilingual projections</i> Long Duong, Paul Cook, Steven Bird and Pavel Pecina | 634 |
| <i>Part-of-speech tagging with antagonistic adversaries</i> Anders Søgaard | 640 |
| <i>Temporal Signals Help Label Temporal Relations</i> Leon Derczynski and Robert Gaizauskas | 645 |
| <i>Diverse Keyword Extraction from Conversations</i> Maryam Habibi and Andrei Popescu-Belis | 651 |
| <i>Understanding Tables in Context Using Standard NLP Toolkits</i> Vidhya Govindaraju, Ce Zhang and Christopher Ré | 658 |
| <i>Filling Knowledge Base Gaps for Distant Supervision of Relation Extraction</i> Wei Xu, Raphael Hoffmann, Le Zhao and Ralph Grishman | 665 |
| <i>Joint Apposition Extraction with Syntactic and Semantic Constraints</i> Will Radford and James R. Curran | 671 |
| <i>Adaptation Data Selection using Neural Language Models: Experiments in Machine Translation</i> Kevin Duh, Graham Neubig, Katsuhito Sudoh and Hajime Tsukada | 678 |
| <i>Mapping Source to Target Strings without Alignment by Analogical Learning: A Case Study with Transliteration</i> Phillippe Langlais | 684 |
| <i>Scalable Modified Kneser-Ney Language Model Estimation</i> Kenneth Heafield, Ivan Pouzyrevsky, Jonathan H. Clark and Philipp Koehn | 690 |
| <i>Incremental Topic-Based Translation Model Adaptation for Conversational Spoken Language Translation</i> Sanjika Hewavitharana, Dennis Mehay, Sankaranarayanan Ananthakrishnan and Prem Natarajan | 697 |
| <i>A Lightweight and High Performance Monolingual Word Aligner</i> Xuchen Yao, Benjamin Van Durme, Chris Callison-Burch and Peter Clark | 702 |
| <i>A Learner Corpus-based Approach to Verb Suggestion for ESL</i> Yu Sawai, Mamoru Komachi and Yuji Matsumoto | 708 |
| <i>Learning Semantic Textual Similarity with Structural Representations</i> Aliaksei Severyn, Massimo Nicosia and Alessandro Moschitti | 714 |
| <i>Typesetting for Improved Readability using Lexical and Syntactic Information</i> Ahmed Salama, Kemal Oflazer and Susan Hagan | 719 |

| | |
|--|-----|
| <i>Annotation of regular polysemy and underspecification</i> | |
| Héctor Martínez Alonso, Bolette Sandford Pedersen and Núria Bel | 725 |
| <i>Derivational Smoothing for Syntactic Distributional Semantics</i> | |
| Sebastian Padó, Jan Šnajder and Britta Zeller | 731 |
| <i>Diathesis alternation approximation for verb clustering</i> | |
| Lin Sun, Diana McCarthy and Anna Korhonen | 736 |
| <i>Outsourcing FrameNet to the Crowd</i> | |
| Marco Fossati, Claudio Giuliano and Sara Tonelli | 742 |
| <i>Smatch: an Evaluation Metric for Semantic Feature Structures</i> | |
| Shu Cai and Kevin Knight | 748 |
| <i>Variable Bit Quantisation for LSH</i> | |
| Sean Moran, Victor Lavrenko and Miles Osborne | 753 |
| <i>Context Vector Disambiguation for Bilingual Lexicon Extraction from Comparable Corpora</i> | |
| Dhouha Bouamor, Nasredine Semmar and Pierre Zweigenbaum | 759 |
| <i>The Effects of Lexical Resource Quality on Preference Violation Detection</i> | |
| Jesse Dunietz, Lori Levin and Jaime Carbonell | 765 |
| <i>Exploiting Qualitative Information from Automatic Word Alignment for Cross-lingual NLP Tasks</i> | |
| José G.C. de Souza, Miquel Esplà-Gomis, Marco Turchi and Matteo Negri | 771 |
| <i>An Information Theoretic Approach to Bilingual Word Clustering</i> | |
| Manaal Faruqui and Chris Dyer | 777 |
| <i>Building and Evaluating a Distributional Memory for Croatian</i> | |
| Jan Šnajder, Sebastian Padó and Željko Agić | 784 |
| <i>Generalizing Image Captions for Image-Text Parallel Corpus</i> | |
| Polina Kuznetsova, Vicente Ordonez, Alexander Berg, Tamara Berg and Yejin Choi | 790 |
| <i>Recognizing Identical Events with Graph Kernels</i> | |
| Goran Glavaš and Jan Šnajder | 797 |
| <i>Automatic Term Ambiguity Detection</i> | |
| Tyler Baldwin, Yunyao Li, Bogdan Alexe and Ioana R. Stanoi | 804 |
| <i>Towards Accurate Distant Supervision for Relational Facts Extraction</i> | |
| Xingxing Zhang, Jianwen Zhang, Junyu Zeng, Jun Yan, Zheng Chen and Zhifang Sui | 810 |
| <i>Extra-Linguistic Constraints on Stance Recognition in Ideological Debates</i> | |
| Kazi Saidul Hasan and Vincent Ng | 816 |
| <i>Are School-of-thought Words Characterizable?</i> | |
| Xiaorui Jiang, Xiaoping Sun and Hai Zhuge | 822 |
| <i>Identifying Opinion Subgroups in Arabic Online Discussions</i> | |
| Amjad Abu-Jbara, Ben King, Mona Diab and Dragomir Radev | 829 |
| <i>Extracting Events with Informal Temporal References in Personal Histories in Online Communities</i> | |
| Miaomiao Wen, Zeyu Zheng, Hyeju Jang, Guang Xiang and Carolyn Penstein Rosé | 836 |

| | |
|---|-----|
| <i>Multimodal DBN for Predicting High-Quality Answers in cQA portals</i> Haifeng Hu, Bingquan Liu, Baoxun Wang, Ming Liu and Xiaolong Wang | 843 |
| <i>Bi-directional Inter-dependencies of Subjective Expressions and Targets and their Value for a Joint Model</i> Roman Klinger and Philipp Cimiano | 848 |
| <i>Identifying Sentiment Words Using an Optimization-based Model without Seed Words</i> Hongliang Yu, Zhi-Hong Deng and Shiyinxue Li | 855 |
| <i>Detecting Turnarounds in Sentiment Analysis: Thwarting</i> Ankit Ramteke, Akshat Malu, Pushpak Bhattacharyya and J. Saketha Nath | 860 |
| <i>Explicit and Implicit Syntactic Features for Text Classification</i> Matt Post and Shane Bergsma | 866 |
| <i>Does Korean defeat phonotactic word segmentation?</i> Robert Daland and Kie Zuraw | 873 |
| <i>Word surprisal predicts N400 amplitude during reading</i> Stefan L. Frank, Leun J. Otten, Giulia Galli and Gabriella Vigliocco | 878 |
| <i>Computerized Analysis of a Verbal Fluency Test</i> James O. Ryan, Serguei Pakhomov, Susan Marino, Charles Bernick and Sarah Banks | 884 |
| <i>A New Set of Norms for Semantic Relatedness Measures</i> Sean Szumlanski, Fernando Gomez and Valerie K. Sims | 890 |