Table of Contents

Testing SDRT’s Right Frontier  
Stergos Afantenos and Nicholas Asher ........................................... 1

Identifying Multi-word Expressions by Leveraging Morphological and Syntactic Idiosyncrasy 
Hassan Al-Haj and Shuly Wintner .................................................. 10

Robust Measurement and Comparison of Context Similarity for Finding Translation Pairs 
Daniel Andrade, Tetsuya Nasukawa and Junichi Tsujii ........................ 19

Multilingual Subjectivity: Are More Languages Better? 
Carmen Banea, Rada Mihalcea and Janyce Wiebe ................................ 28

Plagiarism Detection across Distant Language Pairs 
Alberto Barrón-Cedeño, Paolo Rosso, Eneko Agirre and Gorka Labaka ....... 37

Automatic Detection of Non-deverbal Event Nouns for Quick Lexicon Production 
Nuria Bel, Maria Coll and Gabriela Resnik ....................................... 46

Fast and Accurate Arc Filtering for Dependency Parsing 
Shane Bergsma and Colin Cherry ..................................................... 53

A Hierarchical Classifier Applied to Multi-way Sentiment Detection 
Adrian Bickerstaffe and Ingrid Zukerman ......................................... 62

Fluency Constraints for Minimum Bayes-Risk Decoding of Statistical Machine Translation Lattices 
Graeme Blackwood, Adrià de Gispert and William Byrne ...................... 71

Self-Annotation for fine-grained geospatial relation extraction 
Andre Blessing and Hinrich Schütze ................................................. 80

Top Accuracy and Fast Dependency Parsing is not a Contradiction 
Bernd Bohnet .................................................................................. 89

Broad Coverage Multilingual Deep Sentence Generation with a Stochastic Multi-Level Realizer 
Bernd Bohnet, Leo Wanner, Simon Mill and Alicia Burga ...................... 98

Towards an optimal weighting of context words based on distance 
Bernard Brosseau-Villeneuve, Jian-Yun Nie and Noriko Kando ................. 107

Measuring the Non-compositionality of Multiword Expressions 
Fan Bu, Xiaoyan Zhu and Ming Li ...................................................... 116

A Utility-Driven Approach to Question Ranking in Social QA 
Razvan Bunescu and Yunfeng Huang ................................................. 125

Simultaneous Ranking and Clustering of Sentences: A Reinforcement Approach to Multi-Document Summarization 
Xiaoyan Cai, Wenjie Li, You Ouyang and Hong Yan ............................... 134
End-to-End Coreference Resolution via Hypergraph Partitioning
Jie Cai and Michael Strube ......................................................... 143

Exploiting Background Knowledge for Relation Extraction
Yee Seng Chan and Dan Roth......................................................... 152

Jointly Modeling WSD and SRL with Markov Logic
Wanxiang Che and Ting Liu............................................................ 161

Bipolar Person Name Identification of Topic Documents Using Principal Component Analysis
Chien Chin Chen and Chen-Yuan Wu.............................................. 170

Emotion Cause Detection with Linguistic Constructions
Ying Chen, Sophia Yat Mei Lee, Shoushan Li and Chu-Ren Huang .......... 179

A Twin-Candidate Based Approach for Event Pronoun Resolution using Composite Kernel
Bin Chen, Jian Su and Chew Lim Tan.............................................. 188

Unsupervised Synthesis of Multilingual Wikipedia Articles
Yuncong Chen and Pascale Fung..................................................... 197

Simplicity is Better: Revisiting Single Kernel PPI Extraction
Sung-Pil Choi and Sung-Hyon Myaeng ........................................... 206

An ontology-driven system for detecting global health events
Nigel Collier, Reiko Matsuda Goodwin, John McCrae, Son Doan, Ai Kawazoe, Mike Conway, Asanee Kawtrakul, Koichi Takeuchi and Dinh Dien ......................................................... 215

Constraining robust constructions for broad-coverage parsing with precision grammars
Bart Cramer and Yi Zhang ............................................................. 223

Local lexical adaptation in Machine Translation through triangulation: SMT helping SMT
Josep Maria Crego, Aurélien Max and François Yvon .................................. 232

Automated Translation of Semantic Relationships
Dmitry Davidov and Ari Rappoport ................................................. 241

Comparison of different algebras for inducing the temporal structure of texts
Pascal Denis and Philippe Muller .................................................... 250

Generating Learner-Like Morphological Errors in Russian
Markus Dickinson ........................................................................... 259

Resolving Object and Attribute Coreference in Opinion Mining
Xiaowen Ding and Bing Liu ............................................................ 268

Entity Disambiguation for Knowledge Base Population
Mark Dredze, Paul McNamee, Delip Rao, Adam Gerber and Tim Finin ........ 277

A Discriminative Latent Variable-Based “DE” Classifier for Chinese-English SMT
Jinhua Du and Andy Way ............................................................... 286
An Empirical Study on Learning to Rank of Tweets
Yajuan Duan, Long Jiang, Tao Qin, Ming Zhou and Heung-Yeung Shum .............................. 295

Translation Model Generalization using Probability Averaging for Machine Translation
Nan Duan, Hong Sun and Ming Zhou .................................................................................. 304

Mixture Model-based Minimum Bayes Risk Decoding using Multiple Machine Translation Systems
Nan Duan, Mu Li, Dongdong Zhang and Ming Zhou ............................................................. 313

Multi-Sentence Compression: Finding Shortest Paths in Word Graphs
Katja Filippova.................................................................................................................. 322

Enriching Dictionaries with Images from the Internet - Targeting Wikipedia and a Japanese Semantic Lexicon: Lexeed
Sanae Fujita and Masaaki Nagata ...................................................................................... 331

Opinion: A Graph Based Approach to Abstractive Summarization of Highly Redundant Opinions
Kavita Ganesan, ChengXiang Zhai and Jiawei Han ............................................................ 340

EMDC: A Semi-supervised Approach for Word Alignment
Qin Gao, Francisco Guzman and Stephan Vogel ................................................................. 349

A Large Scale Ranker-Based System for Search Query Spelling Correction
Jianfeng Gao, Xiaolong Li, Daniel Micol, Chris Quirk and Xu Sun ...................................... 358

RTG based surface realisation for TAG
Claire Gardent and Laura Perez-Beltrachini ...................................................................... 367

Automatically Learning Source-side Reordering Rules for Large Scale Machine Translation
Dmitriy Genzel .................................................................................................................. 376

Comparing Language Similarity across Genetic and Typologically-Based Groupings
Ryan Georgi, Fei Xia and William Lewis ............................................................................... 385

Better Arabic Parsing: Baselines, Evaluations, and Analysis
Spence Green and Christopher D. Manning ....................................................................... 394

Paraphrase Alignment for Synonym Evidence Discovery
Gintarė Grigonytė, João Paulo Cordeiro, Gaël Dias, Rumen Moraliyski and Pavel Brazdil ... 403

Finding the Storyteller: Automatic Spoiler Tagging using Linguistic Cues
Sheng Guo and Naren Ramakrishnan ............................................................................... 412

Detection of Simple Plagiarism in Computer Science Papers
Yaakov HaCohen-Kerner, Aharon Tayeb and Natan Ben-Dror ........................................... 421

A Structured Vector Space Model for Hidden Attribute Meaning in Adjective-Noun Phrases
Matthias Hartung and Anette Frank .................................................................................. 430

Hierarchical Phrase-based Machine Translation with Word-based Reordering Model
Katsuhiko Hayashi, Hajime Tsukada, Katsuhito Sudoh, Kevin Duh and Seiichi Yamamoto ... 439

xiv
A Novel Reordering Model Based on Multi-layer Phrase for Statistical Machine Translation
Yanqing He, Yu Zhou, Chengqing Zong and Huilin Wang ........................................... 447

Standardizing Wordnets in the ISO Standard LMF: Wordnet-LMF for GermaNet
Verena Henrich and Erhard Hinrichs ............................................................................. 456

Normal-form parsing for Combinatory Categorial Grammars with generalized composition and type-raising
Julia Hockenmaier and Yonatan Bisk .......................................................................... 465

An Empirical Study on Web Mining of Parallel Data
Gumwon Hong, Chi-Ho Li, Ming Zhou and Hae-Chang Rim ...................................... 474

Enhancing Cross Document Coreference of Web Documents with Context Similarity and Very Large Scale Text Categorization
Jian Huang, Pucktada Treeratpituk, Sarah Taylor and C. Lee Giles ............................. 483

Feature-Rich Discriminative Phrase Rescoring for SMT
Fei Huang and Bing Xiang ......................................................................................... 492

FactRank: Random Walks on a Web of Facts
Alpa Jain and Patrick Pantel .......................................................................................... 501

Open Entity Extraction from Web Search Query Logs
Alpa Jain and Marco Pennacchiotti .............................................................................. 510

Reranking Models in Fine-grained Opinion Analysis
Richard Johansson and Alessandro Moschitti ............................................................. 519

Unsupervised phonemic Chinese word segmentation using Adaptor Grammars
Mark Johnson and Katherine Demuth ........................................................................... 528

Data-Driven Parsing with Probabilistic Linear Context-Free Rewriting Systems
Laura Kallmeyer and Wolfgang Maier ......................................................................... 537

Learning to Predict Readability using Diverse Linguistic Features
Rohit Kate, Xiaoliang Luo, Siddharth Patwardhan, Martin Franz, Radu Florian, Raymond Mooney, Salim Roukos and Chris Welty ................................................................. 546

Value for Money: Balancing Annotation Effort, Lexicon Building and Accuracy for Multilingual WSD
Mitesh Khapra, Saurabh Sohoney, Anup Kulkarni and Pushpak Bhattacharyya ............. 555

A Cross-lingual Annotation Projection Approach for Relation Detection
Seokhwan Kim, Minwoo Jeong, Jonghoon Lee and Gary Geunbae Lee ...................... 564

Evaluating N-gram based Evaluation Metrics for Automatic Keyphrase Extraction
Su Nam Kim, Timothy Baldwin and Min-Yen Kan ......................................................... 572

Improving the Quality of Text Understanding by Delaying Ambiguity Resolution
Doo Soon Kim, Ken Barker and Bruce Porter ............................................................... 581
Automatic generation of inter-passage links based on semantic similarity
Petr Knoth, Jakub Novotny and Zdenek Zdrahal ................................. 590

Dependency-driven Anaphoricity Determination for Coreference Resolution
Fang Kong, Guodong Zhou, Longhua Qian and Qiaoming Zhu .................. 599

Phrase Clustering for Smoothing TM Probabilities - or, How to Extract Paraphrases from Phrase Tables
Roland Kuhn, Boxing Chen, George Foster and Evan Stratford .................. 608

Revisiting Context-based Projection Methods for Term-Translation Spotting in Comparable Corpora
Audrey Laroche and Philippe Langlais ............................................. 617

Constituent Reordering and Syntax Models for English-to-Japanese Statistical Machine Translation
Young-Suk Lee, Bing Zhao and Xiaoqian Luo ...................................... 626

Sentiment Classification and Polarity Shifting
Shoushan Li, Sophia Y. M. Lee, Ying Chen, Chu-Ren Huang and Guodong Zhou 635

Improving Corpus Comparability for Bilingual Lexicon Extraction from Comparable Corpora
Bo Li and Eric Gaussier ............................................................... 644

Structure-Aware Review Mining and Summarization
Fangtao Li, Chao Han, Minlie Huang, Xiaoyan Zhu, Ying-Ju Xia, Shu Zhang and Hao Yu 653

Adaptive Development Data Selection for Log-linear Model in Statistical Machine Translation
Mu Li, Yinggong Zhao, Dongdong Zhang and Ming Zhou ......................... 662

Learning the Scope of Negation via Shallow Semantic Parsing
Junhui Li, Guodong Zhou, Hongling Wang and Qiaoming Zhu .................... 671

Filtered Ranking for Bootstrapping in Event Extraction
Shasha Liao and Ralph Grishman ................................................... 680

Exploring variation across biomedical subdomains
Tom Lippincott, Diarmuid Ó Séaghdha, Lin Sun and Anna Korhonen ........... 689

Semantic Role Labeling for News Tweets
Xiaohua Liu, Kuan Li, Bo Han, Ming Zhou, Long Jiang, Zhongyang Xiong and Changning Huang 698

Joint Parsing and Translation
Yang Liu and Qun Liu .............................................................. 707

Semantic Role Features for Machine Translation
Ding Liu and Daniel Gildea .......................................................... 716