

Table of Contents – Part II

Natural Language Generation

Exploring Extensive Linguistic Feature Sets in Near-Synonym Lexical Choice	1
<i>Mari-Sanna Paukkeri, Jaakko Väyrynen, and Antti Arppe</i>	
Abduction in Games for a Flexible Approach to Discourse Planning	13
<i>Ralf Klabunde, Sebastian Reuße, and Björn Schlünder</i>	

Machine Translation and Multilingualism

Document-Specific Statistical Machine Translation for Improving Human Translation Productivity (Invited Paper)	25
<i>Salim Roukos, Abraham Ittycheriah, and Jian-Ming Xu</i>	
Minimum Bayes Risk Decoding with Enlarged Hypothesis Space in System Combination	40
<i>Tsuyoshi Okita and Josef van Genabith</i>	
Phrasal Syntactic Category Sequence Model for Phrase-Based MT	52
<i>Hailong Cao, Eiichiro Sumita, Tiejun Zhao, and Sheng Li</i>	
Integration of a Noun Compound Translator Tool with Moses for English-Hindi Machine Translation and Evaluation	60
<i>Prashant Mathur and Soma Paul</i>	
Neoclassical Compound Alignments from Comparable Corpora	72
<i>Rima Harastani, Béatrice Daille, and Emmanuel Morin</i>	
QAlign: A New Method for Bilingual Lexicon Extraction from Comparable Corpora	83
<i>Amir Hazem and Emmanuel Morin</i>	
Aligning the Un-Alignable — A Pilot Study Using a Noisy Corpus of Nonstandardized, Semi-parallel Texts	97
<i>Florian Petran</i>	
Parallel Corpora for WordNet Construction: Machine Translation vs. Automatic Sense Tagging	110
<i>Antoni Oliver and Salvador Climent</i>	
Method to Build a Bilingual Lexicon for Speech-to-Speech Translation Systems	122
<i>Keiji Yasuda, Andrew Finch, and Eiichiro Sumita</i>	

Text Categorization and Clustering

A Fast Subspace Text Categorization Method Using Parallel Classifiers	132
<i>Nandita Tripathi, Michael Oakes, and Stefan Wermter</i>	
Research on Text Categorization Based on a Weakly-Supervised Transfer Learning Method	144
<i>Dequan Zheng, Chenghe Zhang, Geli Fei, and Tiejun Zhao</i>	
Fuzzy Combinations of Criteria: An Application to Web Page Representation for Clustering	157
<i>Alberto Pérez García-Plaza, Víctor Fresno, and Raquel Martínez</i>	
Clustering Short Text and Its Evaluation	169
<i>Prajol Shrestha, Christine Jacquin, and Béatrice Daille</i>	

Information Extraction and Text Mining

Information Extraction from Webpages Based on DOM Distances	181
<i>Carlos Castillo, Héctor Valero, José Guadalupe Ramos, and Josep Silva</i>	
Combining Flat and Structured Approaches for Temporal Slot Filling or: How Much to Compress?	194
<i>Qi Li, Javier Artiles, Taylor Cassidy, and Heng Ji</i>	
Event Annotation Schemes and Event Recognition in Spanish Texts	206
<i>Dina Wonsever, Aiala Rosá, Marisa Malcuori, Guillermo Moncecchi, and Alan Descoins</i>	
Automatically Generated Noun Lexicons for Event Extraction	219
<i>Béatrice Arnulphy, Xavier Tannier, and Anne Vilnat</i>	
Lexical Acquisition for Clinical Text Mining Using Distributional Similarity (Invited Paper)	232
<i>John Carroll, Rob Koeling, and Shivani Puri</i>	
Developing an Algorithm for Mining Semantics in Texts	247
<i>Minhua Huang and Robert M. Haralick</i>	
Mining Market Trend from Blog Titles Based on Lexical Semantic Similarity	261
<i>Fei Wang and Yunfang Wu</i>	

Information Retrieval and Question Answering

Ensemble Approach for Cross Language Information Retrieval	274
<i>Dinesh Mavaluru, R. Shriram, and W. Aisha Banu</i>	

Web Image Annotation Using an Effective Term Weighting	286
<i>Vundavalli Srinivasarao and Vasudeva Varma</i>	
Metaphone-pt_BR: The Phonetic Importance on Search and Correction of Textual Information	297
<i>Carlos C. Jordão and João Luís G. Rosa</i>	
Robust and Fast Two-Pass Search Method for Lyric Search Covering Erroneous Queries Due to Mishearing	306
<i>Xin Xu and Tsuneo Kato</i>	
Bootstrap-Based Equivalent Pattern Learning for Collaborative Question Answering	318
<i>Tianyong Hao and Eugene Agichtein</i>	
How to Answer Yes/No Spatial Questions Using Qualitative Reasoning?	330
<i>Marcin Walas</i>	
Question Answering and Multi-search Engines in Geo-Temporal Information Retrieval	342
<i>Fernando S. Peregrino, David Tomás, and Fernando Llopis Pascual</i>	

Document Summarization

Using Graph Based Mapping of Co-occurring Words and Closeness Centrality Score for Summarization Evaluation	353
<i>Niraj Kumar, Kannan Srinathan, and Vasudeva Varma</i>	
Combining Syntax and Semantics for Automatic Extractive Single-Document Summarization	366
<i>Araly Barrera and Rakesh Verma</i>	
Combining Summaries Using Unsupervised Rank Aggregation	378
<i>Girish Keshav Palshikar, Shailesh Deshpande, and G. Athiappan</i>	
Using Wikipedia Anchor Text and Weighted Clustering Coefficient to Enhance the Traditional Multi-document Summarization	390
<i>Niraj Kumar, Kannan Srinathan, and Vasudeva Varma</i>	
Extraction of Relevant Figures and Tables for Multi-document Summarization (Verifiability Award)	402
<i>Ashish Sadh, Amit Sahu, Devesh Srivastava, Ratna Sanyal, and Sudip Sanyal</i>	
Towards Automatic Generation of Catchphrases for Legal Case Reports (Best Paper Award, Third Place)	414
<i>Filippo Galgani, Paul Compton, and Achim Hoffmann</i>	

Applications

A Dataset for the Evaluation of Lexical Simplification
(Invited Paper) 426
Jan De Belder and Marie-Francine Moens

Text Content Reliability Estimation in Web Documents: A New
Proposal 438
Luis Sanz, Héctor Allende, and Marcelo Mendoza

Fine-Grained Certainty Level Annotations Used for Coarser-Grained
E-Health Scenarios: Certainty Classification of Diagnostic Statements
in Swedish Clinical Text 450
Sumithra Vechupillai and Maria Kvist

Combining Confidence Score and Mal-rule Filters for Automatic
Creation of Bangla Error Corpus: Grammar Checker Perspective 462
*Bibekananda Kundu, Sutanu Chakraborti, and
Sanjay Kumar Choudhury*

Predictive Text Entry for Agglutinative Languages Using Unsupervised
Morphological Segmentation (Best Student Paper Award) 478
Mikka Silfverberg, Krister Lindén, and Mirka Hyvärinen

Comment Spam Classification in Blogs through Comment Analysis and
Comment-Blog Post Relationships 490
Ashwin Rajadesingan and Anand Mahendran

Detecting Players Personality Behavior with Any Effort
of Concealment 502
Fazel Keshtkar, Candice Burkett, Arthur Graesser, and Haiying Li

Author Index 515