

Contents

Editors' Foreword	ix
--------------------------	----

I. INVITED TALKS

<i>Chris Fox & Shalom Lappin</i> A type-theoretic approach to anaphora and ellipsis resolution	1
<i>Yorick Wilks, Nick Webb, Andrea Setzer, Mark Hepple & Roberta Catizone</i> Human dialogue modelling using machine learning	17
<i>Stephen G. Pulman & Maria Liakata</i> Learning domain theories	29
<i>Inderjeet Mani</i> Recent developments in temporal information extraction	45
<i>Branimir K. Boguraev</i> Annotation-based finite state processing in a large-scale NLP architecture	61

II. LEXICAL SEMANTICS AND LEXICAL KNOWLEDGE ACQUISITION

<i>Oren Glickman & Ido Dagan</i> Acquiring lexical paraphrases from a single corpus	81
<i>Violeta Seretan, Luka Nerima & Eric Wehrli</i> Multi-word collocation extraction by syntactic composition of collocation bigrams	91
<i>Peter D. Turney, Michael L. Littman, Jeffrey Bigham & Victor Shnayder</i> Combining independent modules in lexical multiple-choice problems	101
<i>Mario Jarmasz & Stan Szpakowicz</i> Roget's thesaurus and semantic similarity	111
<i>Eneko Agirre & Oier Lopez de Lacalle</i> Clustering WordNet word senses	121

- Roberto Basili, Maria Teresa Pazienza & Fabio Massimo Zanzotto*
Inducing hyperlinking rules in text collections 131
- Diana Zaiu Inkpen & Graeme Hirst*
Near-synonym choice in natural language generation 141

III. TAGGING, PARSING AND SYNTAX

- Jesús Giménez & Lluís Màrquez*
Fast and accurate part-of-speech tagging: The SVM approach revisited 153
- Virginia Savova & Leon Peshkin*
Part-of-speech tagging with minimal lexicalization 163
- António Branco & João Silva*
Accurate annotation: An efficiency metric 173
- Mary Hearne & Khalil Sima'an*
Structured parameter estimation for LFG-DOP 183
- Sandra Kübler*
Parsing without grammar — Using complete trees instead 193
- Xavier Carreras & Lluís Màrquez*
Phrase recognition by filtering and ranking with perceptrons 205
- Sebastian van Delden & Fernando Gomez*
Cascaded finite-state partial parsing: A larger-first approach 217
- Henning Christiansen*
A constraint-based bottom-up counterpart to definite clause grammars 227

IV. INFORMATION EXTRACTION

- Mary McGee Wood, Susannah J. Lydon, Valentin Tablan, Diana Maynard & Hamish Cunningham*
Using parallel texts to improve recall in botany 237
- Elena Filatova & Vasileios Hatzivassiloglou*
Marking atomic events in sets of related texts 247

<i>Svetla Boytcheva, Milena Yankova & Albena Strupchanska</i> Semantically driven approach for scenario recognition in the IE system FRET	257
<i>Richard J. Evans</i> A framework for named entity recognition in the open domain	267
V. TEXT SUMMARISATION AND DOCUMENT PROCESSING	
<i>Tristan Miller</i> Latent semantic analysis and the construction of coherent extracts	277
<i>Ani Nenkova & Amit Bagga</i> Facilitating email thread access by extractive summary generation	287
<i>Preslav Nakov, Elena Valchanova & Galia Angelova</i> Towards deeper understanding of the latent semantic analysis performance	297
<i>Bruno Pouliquen, Ralf Steinberger & Camelia Ignat</i> Automatic linking of similar texts across languages	307
VI. OTHER NLP TOPICS	
<i>Leif Arda Nielsen</i> Verb phrase ellipsis detection using machine learning techniques	317
<i>Kiril Iv. Simov</i> HPSG-based annotation scheme for corpora development and parsing evaluation	327
<i>Allan Ramsay & Hanady Mansour</i> Arabic Morpho-syntax for Text-to-Speech	337
<i>Preslav Nakov, Yury Bonev, Galia Angelova, Evelyn Gius & Walther von Hahn</i> Guessing morphological classes of unknown German nouns	347

<i>Rada Mihalcea & Timothy Chklovski</i> Building sense tagged corpora with volunteer contributions over the Web	357
<i>Anette Hulth</i> Reducing false positives by expert combination in automatic keyword indexing	367
<i>Hristo T. Tanev</i> Socrates: A question answering prototype for Bulgarian	377
<i>Rada Mihalcea</i> Unsupervised natural language disambiguation using non-ambiguous words	387
List of Contributors	397
Index of Subjects and Terms	399