

Table of Contents

Computational Linguistics Research

Lexical Resources

Invited paper:

Integrating Semantic Frames from Multiple Sources <i>Namhee Kwon, Eduard Hovy</i>	1
--	---

Invited paper:

Making Senses: Bootstrapping Sense-Tagged Lists of Semantically-Related Words <i>Nancy Ide</i>	13
--	----

Enriching Wordnets with New Relations and with Event and Argument Structures <i>Raquel Amaro, Rui Pedro Chaves, Palmira Marrafa, Sara Mendes</i> ...	28
--	----

Experiments in Cross-Language Morphological Annotation Transfer <i>Anna Feldman, Jirka Hana, Chris Brew</i>	41
--	----

Sentence Segmentation Model to Improve Tree Annotation Tool <i>So-Young Park, Dongha Shin, Ui-Sung Song</i>	51
--	----

Markov Cluster Shortest Path Founded Upon the Alibi-Breaking Algorithm <i>Jaeyoung Jung, Maki Miyake, Hiroyuki Akama</i>	55
--	----

Corpus-Based Knowledge Acquisition

Unsupervised Learning of Verb Argument Structures <i>Thiago Alexandre Salgueiro Pardo, Daniel Marcu, Maria das Graças Volpe Nunes</i>	59
--	----

A Methodology for Extracting Ontological Knowledge from Spanish Documents <i>Rafael Valencia-García, Dagoberto Castellanos-Nieves, Jesusaldo Tomás Fernández-Breis, Pedro José Vivancos-Vicente</i>	71
---	----

Automatically Determining Allowable Combinations of a Class of Flexible Multiword Expressions <i>Afsaneh Fazly, Ryan North, Suzanne Stevenson</i>	81
---	----

Web-Based Measurements of Intra-collocational Cohesion in Oxford Collocations Dictionary
Igor A. Bolshakov, Sofía Natalia Galicia-Haro 93

Probabilistic Neural Network Based English-Arabic Sentence Alignment
Mohamed Abdel Fattah, Fuji Ren, Shingo Kuroiwa 97

Morphology and Part-of-Speech Tagging

Towards the Automatic Lemmatization of 16th Century Mexican Spanish: A Stemming Scheme for the CHEM
Alfonso Medina-Urrea 101

Word Frequency Approximation for Chinese Without Using Manually-Annotated Corpus
Maosong Sun, Zhengcao Zhang, Benjamin Ka-Yin T'sou, Huaming Lu 105

Abbreviation Recognition with MaxEnt Model
Chunyu Kit, Xiaoyue Liu, Jonathan J. Webster 117

An Efficient Multi-agent System Combining POS-Taggers for Arabic Texts
Chiraz Ben Othmane Zribi, Aroua Torjmen, Mohamed Ben Ahmed ... 121

Syntax and Parsing

A Comparative Evaluation of a New Unsupervised Sentence Boundary Detection Approach on Documents in English and Portuguese
Jan Strunk, Carlos N. Silla Jr., Celso A.A. Kaestner 132

A General and Multi-lingual Phrase Chunking Model Based on Masking Method
Yu-Chieh Wu, Chia-Hui Chang, Yue-Shi Lee 144

UCSG Shallow Parser
Guntur Bharadwaja Kumar, Kavi Narayana Murthy..... 156

Evaluating the Performance of the Survey Parser with the NIST Scheme
Alex Chengyu Fang 168

Sequences of Part of Speech Tags vs. Sequences of Phrase Labels: How Do They Help in Parsing?
Gabriel Infante-Lopez, Maarten de Rijke 180

Word Sense Disambiguation and Anaphora Resolution

Verb Sense Disambiguation Using Support Vector Machines: Impact of WordNet-Extracted Features <i>Davide Buscaldi, Paolo Rosso, Ferran Pla, Encarna Segarra, Emilio Sanchis Arnal</i>	192
Preposition Senses: Generalized Disambiguation Model <i>Chutima Boonthum, Shunichi Toida, Irwin Levinstein</i>	196
An Unsupervised Language Independent Method of Name Discrimination Using Second Order Co-occurrence Features <i>Ted Pedersen, Anagha Kulkarni, Roxana Angheluta, Zornitsa Kozareva, Thamar Solorio</i>	208
Extracting Key Phrases to Disambiguate Personal Names on the Web <i>Danushka Bollegala, Yutaka Matsuo, Mitsuru Ishizuka</i>	223
Chinese Noun Phrase Metaphor Recognition with Maximum Entropy Approach <i>Zhimin Wang, Houfeng Wang, Huiming Duan, Shuang Han, Shiwen Yu</i>	235
Zero Anaphora Resolution in Chinese Discourse <i>Yuzhen Cui, Qinan Hu, Haihua Pan, Jianhua Hu</i>	245
Semantics	
<i>Invited paper:</i>	
Random Walks on Text Structures <i>Rada Mihalcea</i>	249
<i>Best paper award (first place):</i>	
Shallow Case Role Annotation Using Two-Stage Feature-Enhanced String Matching <i>Samuel W.K. Chan</i>	263
SPARTE, a Test Suite for Recognising Textual Entailment in Spanish <i>Anselmo Peñas, Álvaro Rodrigo, Felisa Verdejo</i>	275
Analysis of a Textual Entailer <i>Vasile Rus, Philip M. McCarthy, Arthur C. Graesser</i>	287

Text Generation

Referring Via Document Parts
Ivandr  Paraboni, Kees van Deemter 299

Generation of Natural Language Explanations of Rules in an Expert System
Mar a de los  ngeles Alonso-Lavernia, Argelio V ctor De-la-Cruz-Rivera, Grigori Sidorov 311

Generating a Set of Rules to Determine Honorific Expression Using Decision Tree Learning
Kanako Komiya, Yasuhiro Tajima, Nobuo Inui, Yoshiyuki Kotani 315

Natural Language Interfaces and Speech Processing

NLP (Natural Language Processing) for NLP (Natural Language Programming)
Rada Mihalcea, Hugo Liu, Henry Lieberman 319

Balancing Transactions in Practical Dialogues
Luis Albreto Pineda Cort s, Hayde Castellanos, Sergio Rafael Coria Olguin, Varinia Estrada, Fernanda L pez, Isabel L pez, Ivan Meza, Iv n Moreno, Patricia P rez, Carlos Rodr guez 331

Best paper award (second place):

Finite State Grammar Transduction from Distributed Collected Knowledge
Rakesh Gupta, Ken Hennacy 343

Predicting Dialogue Acts from Prosodic Information
Sergio Rafael Coria Olguin, Luis Albreto Pineda Cort s 355

Disambiguation Based on Wordnet for Transliteration of Arabic Numerals for Korean TTS
Youngim Jung, Aesun Yoon, Hyuk-Chul Kwon 366

Intelligent Text Processing Applications

Information Retrieval

MFCRank: A Web Ranking Algorithm Based on Correlation of Multiple Features
Yunming Ye, Yan Li, Xiaofei Xu, Joshua Huang, Xiaojun Chen 378

On Text Ranking for Information Retrieval Based on Degree of Preference <i>Bo-Yeong Kang, Dae-Won Kim</i>	389
Lexical Normalization and Relationship Alternatives for a Term Dependence Model in Information Retrieval <i>Marco Gonzalez, Vera Lúcia Strube de Lima, José Valdeni de Lima</i> ..	394
Web Search Model for Dynamic and Fuzzy Directory Search <i>Bumghi Choi, Ju-Hong Lee, Sun Park, Tae-Su Park</i>	406
Information Retrieval from Spoken Documents <i>Michal Fapšo, Pavel Smrž, Petr Schwarz, Igor Szöke, Milan Schwarz, Jan Černocký, Martin Karafiát, Lukáš Burget</i>	410
Automatic Image Annotation Based on WordNet and Hierarchical Ensembles <i>Wei Li, Maosong Sun</i>	417
Creating a Testbed for the Evaluation of Automatically Generated Back-of-the-Book Indexes <i>Andras Csomai, Rada F. Mihalcea</i>	429
Question Answering	
<i>Best paper award (third place):</i>	
Automatic Acquisition of Semantic-Based Question Reformulations for Question Answering <i>Jamileh Yousefi, Leila Kosseim</i>	441
Using N-gram Models to Combine Query Translations in Cross-Language Question Answering <i>Rita M. Aceves-Pérez, Luis Villaseñor-Pineda, Manuel Montes-y-Gómez</i>	453
A Question Answering System on Special Domain and the Implementation of Speech Interface <i>Haiqing Hu, Fuji Ren, Shingo Kuroiwa, Shuwu Zhang</i>	458
Text Summarization	
Multi-document Summarization Based on BE-Vector Clustering <i>Dexi Liu, Yanxiang He, Donghong Ji, Hua Yang</i>	470

Deriving Event Relevance from the Ontology Constructed with Formal Concept Analysis <i>Wei Xu, Wenjie Li, Mingli Wu, Wei Li, Chunfa Yuan</i>	480
A Sentence Compression Module for Machine-Assisted Subtitling <i>Nadjet Bouayad-Agha, Angel Gil, Oriol Valentin, Victor Pascual</i>	490

Information Extraction and Text Mining

Application of Semi-supervised Learning to Evaluative Expression Classification <i>Yasuhiro Suzuki, Hiroya Takamura, Manabu Okumura</i>	502
A New Algorithm for Fast Discovery of Maximal Sequential Patterns in a Document Collection <i>René Arnulfo García-Hernández, José Francisco Martínez-Trinidad, Jesús Ariel Carrasco-Ochoa</i>	514
A Machine Learning Based Approach for Separating Head from Body in Web-Tables <i>Sung-Won Jung, Hyuk-Chul Kwon</i>	524

Text Classification

Best student paper award:

Clustering Abstracts of Scientific Texts Using the Transition Point Technique <i>David Pinto, Héctor Jiménez-Salazar, Paolo Rosso</i>	536
Sense Cluster Based Categorization and Clustering of Abstracts <i>Davide Buscaldi, Paolo Rosso, Mikhail Alexandrov, Alfons Juan Ciscar</i>	547
Analysing Part-of-Speech for Portuguese Text Classification <i>Teresa Gonçalves, Cassiana Silva, Paulo Quaresma, Renata Vieira</i> ...	551
Improving kNN Text Categorization by Removing Outliers from Training Set <i>Kwangcheol Shin, Ajith Abraham, Sang Yong Han</i>	563

Authoring Tools and Spelling Correction

Writing for Language-Impaired Readers <i>Aurélien Max</i>	567
--	-----

Document Copy Detection System Based on Plagiarism Patterns <i>NamOh Kang, SangYong Han</i>	571
Regional vs. Global Robust Spelling Correction <i>Manuel Vilares Ferro, Juan Otero Pombo,</i> <i>Víctor Manuel Darriba Bilbao</i>	575
Author Index	587