

Table of Contents

<i>Optimization in Multimodal Interpretation</i> Joyce Y. Chai, Pengyu Hong, Michelle X. Xhou and Zahar Prasov	1
<i>Constructivist Development of Grounded Construction Grammar</i> Luc Steels	9
<i>A Distributional Model of Semantic Context Effects in Lexical Processing</i> Scott McDonald and Chris Brew	17
<i>Analysis of Mixed Natural and Symbolic Input in Mathematical Dialogs</i> Magdalena Wolska and Ivana Kruijff-KorBayová	25
<i>A TAG-based noisy-channel model of speech repairs</i> Mark Johnson and Eugene Charniak	33
<i>Attention Shifting for Parsing Speech</i> Keith B. Hall and Mark Johnson	40
<i>Discriminative Language Modeling with Conditional Random Fields and the Perceptron Algorithm</i> Brian Roark, Murat Saraclar, Michael Collins and Mark Johnson	47
<i>Statistical Modeling for Unit Selection in Speech Synthesis</i> Mehryar Mohri, Cyril Allauzen and Michael Riley	55
<i>Developing a Flexible Spoken Dialog System Using Simulation</i> Grace Chung	63
<i>Data-Driven Strategies for an Automated Dialogue System</i> Hilda Hardy, Tomek Strzalkowski, Min Wu, Cristian Ursu, Nick Webb, Alan Biermann, R. Bryce Inouye and Ashley McKenzie	71
<i>Trainable Sentence Planning for Complex Information Presentations in Spoken Dialog Systems</i> Amanda Stent, Rashmi Prasad and Marilyn Walker	79
<i>User Expertise Modeling and Adaptivity in a Speech-Based E-Mail System</i> Kristiina Jokinen and Kari Kanto	87
<i>Discriminative Training of a Neural Network Statistical Parser</i> James Henderson	95
<i>Parsing the WSJ Using CCG and Log-Linear Models</i> Stephen Clark and James R. Curran	103
<i>Incremental Parsing with the Perceptron Algorithm</i> Michael Collins and Brian Roark	111
<i>Convolution Kernels with Feature Selection for Natural Language Processing Tasks</i> Jun Suzuki, Hideki Isozaki and Eisaku Maeda	119
<i>Improving Pronoun Resolution by Incorporating Coreferential Information of Candidates</i> Xiaofeng Yang, Jian Su, Guodong Zhou and Chew-Lim Tan	127
<i>A Mention-Synchronous Coreference Resolution Algorithm Based On the Bell Tree</i> Xiaoqiang Luo, Abe Ittycheriah, Hongyan Jing, Nanda Kambhatla and Salim Roukos	135
<i>Learning to Resolve Bridging References</i> Massimo Poesio, Rahul Mehta, Axel Maroudas and Janet Hitzeman	143

<i>Learning Noun Phrase Anaphoricity to Improve Conference Resolution: Issues in Representation and Optimization</i> Vincent Ng	151
<i>A Joint Source-Channel Model for Machine Transliteration</i> Haizhou Li, Min Zhang and Jian Su	159
<i>Collocation Translation Acquisition Using Monolingual Corpora</i> Yajuan Lü and Ming Zhou	167
<i>Statistical Machine Translation with Word- and Sentence-Aligned Parallel Corpora</i> Chris Callison-Burch, David Talbot and Miles Osborne	175
<i>Finding Ideographic Representations of Japanese Names Written in Latin Script via Language Identification and Corpus Validation</i> Yan Qu and Gregory Grefenstette	183
<i>Extracting Regulatory Gene Expression Networks From Pubmed</i> Jasmin Saric, Lars J. Jensen, Peer Bork, Rossitza Ouzounova and Isabel Rojas	191
<i>Linguistic Profiling for Authorship Recognition and Verification</i> Hans van Halteren	199
<i>An Empirical Study of Information Synthesis Task</i> Enrique Amigo, Julio Gonzalo, Victor Peinado, Anselmo Peñas and Felisa Verdejo	207
<i>Mining Metalinguistic Activity in Corpora to Create Lexical Resources Using Information Extraction Techniques: the MOP System</i> Carlos Rodriguez Penagos	215
<i>Optimizing Typed Feature Structure Grammar Parsing through Non-Statistical Indexing</i> Cosmin Munteanu and Gerald Penn	223
<i>Head-Driven Parsing for Word Lattices</i> Christopher Collins, Bob Carpenter and Gerald Penn	231
<i>Balancing Clarity and Efficiency in Typed Feature Logic Through Delaying</i> Gerald Penn	239
<i>Minimal Recursion Semantics as Dominance Constraints: Translation, Evaluation, and Analysis</i> Ruth Fuchss, Alexander Koller, Joachim Niehren and Stefan Thater	247
<i>Learning with Unlabeled Data for Text Categorization Using a Bootstrapping and a Feature Projection Technique</i> Youngjoong Ko and Jungyun Seo	255
<i>The Sentimental Factor: Improving Review Classification Via Human-Provided Information</i> Philip Beineke, Trevor Hastie and Shivakumar Vaithyanathan	263
<i>A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts</i> Bo Pang and Lillian Lee	271
<i>Finding Predominant Word Senses in Untagged Text</i> Diana McCarthy, Rob Koeling, Julie Weeds and John Carroll	279
<i>Unsupervised Sense Disambiguation Using Bilingual Probabilistic Models</i> Indrajit Bhattacharya, Lise Getoor and Yoshua Bengio	287
<i>Chinese Verb Sense Discrimination Using an EM Clustering Model with Rich Linguistic Features</i> Jinying Chen and Martha Palmer	295
<i>Relieving the data Acquisition Bottleneck in Word Sense Disambiguation</i>	

Mona Diab	303
<i>Enriching the Output of a Parser Using Memory-based Learning</i> Valentin Jijkoun and Maarten de Rijke	311
<i>Long-Distance Dependency Resolution in Automatically Acquired Wide-Coverage PCFG-Based LFG Approximations</i> Aoife Cahill, Michael Burke, Ruth O'Donovan, Josef Van Genabith and Andy Way	319
<i>Deep Dependencies from Context-Free Statistical Parsers: Correcting the Surface Dependency Approximation</i> Roger Levy and Christopher Manning	327
<i>A Study on Convolution Kernels for Shallow Statistic Parsing</i> Alessandro Moschitti	335
<i>Combining Acoustic and Pragmatic Features to Predict Recognition Performance in Spoken Dialogue Systems</i> Malte Gabsdil and Oliver Lemon	343
<i>Predicting Student Emotions in Computer-Human Tutoring Dialogues</i> Diane J. Litman and Kate Forbes-Riley	351
<i>Building Verb Predicates: A Computational View</i> Fernando Gomez	359
<i>Large-Scale Induction and Evaluation of Lexical Resources from the Penn-II Treebank</i> Ruth O'Donovan, Michael Burke, Aoife Cahill, Josef Van Genabith and Andy Way	367
<i>Inducing Frame Semantic Verb Classes from WordNet and LDOCE</i> Rebecca Green, Bonnie J. Dorr and Philip Resnik	375
<i>Paragraph-, Word-, and Coherence-based Approaches to Sentence Ranking: A Comparison of Algorithm and Human Performance</i> Florian Wolf and Edward Gibson	383
<i>Evaluating Centering-Based Metrics of Coherence</i> Nikiforos Karamanis, Massimo Poesio, Chris Mellish and Jon Oberlander	391
<i>Computing Locally Coherent Discourses</i> Ernst Althaus, Nikiforos Karamanis and Alexander Koller	399
<i>Generating Referring Expressions in Open Domains</i> Advaith Siddharthan and Ann Copestake	407
<i>Discovering Relations among Named Entities from Large Corpora</i> Takaaki Hasegawa, Satoshi Sekine and Ralph Grishman	415
<i>Dependency Tree Kernels for Relation Extraction</i> Aron Culotta and Jeffrey Sorensen	423
<i>Classifying Semantic Relations in Bioscience Texts</i> Barbara Rosario and Marti Hearst	430
<i>Collective Information Extraction with Relational Markov Networks</i> Razvan Bunescu and Raymond Mooney	438
<i>Error Mining for Wide-Coverage Grammar Engineering</i> Gertjan van Noord	446
<i>Alternative approaches for Generating Bodies of Grammar Rules</i> Gabriel Infante-Lopez and Maarten de Rijke	454

<i>Adaptive Chinese Word Segmentation</i> Jianfeng Gao, Andi Wu, Cheng-Ning Huang, Hong qiao Li, Xinsong Xia and Hauwei Qin	462
<i>Experiments in parallel-text based grammar induction</i> Jonas Kuhn	470
<i>Corpus-Based Induction of Syntactic Structure: Models of Dependency and Constituency</i> Dan Klein and Christopher Manning	478
<i>Annealing Techniques For Unsupervised Statistical Language Learning</i> Noah A. Smith and Jason Eisner	486
<i>Multi-Engine Machine Translation with Voted Language Model</i> Tadashi Nomoto	494
<i>Aligning words using matrix factorisation</i> Cyril Goutte, Kenji Yamada and Eric Gaussier	502
<i>FSA: An Efficient and Flexible C++ Toolkit for Finite State Automata Using On-Demand Computation</i> Stephan Kanthak and Hermann Ney	510
<i>Improving IBM Word Alignment Model 1</i> Robert C. Moore	518
<i>A Geometric View on Bilingual Lexicon Extraction from Comparable Corpora</i> Eric Gaussier, J.M. Renders, I. Matveeva, C. Goutte and H. Dejean	526
<i>Creating Multilingual Translation Lexicons with Regional Variations Using Web Corpora</i> Pu-Jen Cheng, Wen-Hsiang Lu, Jei-Wen Teng and Lee-Feng Chien	534
<i>Probabilistic Parsing Strategies</i> Mark-Jan Nederhof and Giorgio Satta	542
<i>An Alternative Method of Training Probabilistic LR Parsers</i> Mark-Jan Nederhof and Giorgio Satta	550
<i>Wrapping of Trees</i> James Rogers	558
<i>Splitting Complex Temporal Questions for Question Answering Systems</i> E. Saquete, P. Martínez-Barco, R. Muñoz and J.L. Vicedo	566
<i>Question Answering Using Constraint Satisfaction: QA-By-Dossier-With-Constraints</i> John Prager, Jennifer Chu-Carroll and Krzysztof Czuba	574
<i>Applying Machine Learning to Chinese Temporal Relation Resolution</i> Wenjie Li, Kam-Fai Wong, Guihong Cao and Chunfa Yuan	582
<i>Multi-Criteria-based Active Learning for Named Entity Recognition</i> Dan Shen, Jie Zhang, Jian Su, Guodong Zhou and Chew-Lim Tan	589
<i>Weakly Supervised Learning for Cross-document Person Name Disambiguation Supported by Information Ex- traction</i> Cheng Niu, Wei Li and Rohini K. Srihari	597
<i>Automatic Evaluation of Machine Translation Quality Using Longest Common Subsequence and Skip-Bigram Statistics</i> Chin-Yew Lin and Franz Josef Och	605
<i>A Unified Framework For Automatic Evaluation Using 4-Gram Co-occurrence Statistics</i> Radu Soricut and Eric Brill	613

<i>Extending the BLEU MT Evaluation Method with Frequency Weightings</i> Bogdan Babych and Tony Hartley	621
<i>Learning Word Sense With Feature Selection and Order Identification Capabilities</i> Zheng-Yu Niu, Dong-Hong Ji and Chew-Lim Tan	629
<i>A Kernel PCA Method for Superior Word Sense Disambiguation</i> Dekai Wu, Weifeng Su and Marine Carpuat	637
<i>Using Linguistic Principles to Recover Empty Categories</i> Richard Campbell	645
<i>Statistical Machine Translation by Parsing</i> I. Dan Melamed	653
<i>Generalized Multitext Grammars</i> I. Dan Melamed, Giorgio Satta and Benjamin Wellington	661
<i>Identifying Agreement and Disagreement in Conversational Speech: Use of Bayesian Networks to Model Pragmatic Dependencies</i> Michel Galley, Kathleen McKeown, Julia Hirschberg and Elizabeth Shriberg	669
<i>Using Conditional Random Fields to Predict Pitch Accents in Conversational Speech</i> Michelle Gregory and Yasemin Altun	677
<i>Acquiring the Meaning of Discourse Markers</i> Ben Hutchinson	684
<i>FLSA: Extending Latent Semantic Analysis with Features for Dialogue Act Classification</i> Riccardo Serafin and Barbara Di Eugenio	692