

EMNLP 2011

**Conference on Empirical Methods in Natural Language
Processing**

Proceedings of the Conference

July 27–31, 2011

©2011 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-937284-11-4 / 1-937284-11-5

General Chair's Preface

Welcome to EMNLP 2011, Conference on Empirical Methods in Natural Language Processing, a conference organised annually by SIGDAT, the Association for Computational Linguistics' special interest group on linguistic data and corpus-based approaches to NLP. This year the conference is held from July 27th to 29th at the John McIntyre Conference Centre, Edinburgh, UK.

This year EMNLP is, for the first time, not only a stand alone conference, but also an anchor conference to several workshops, that are held on July 30th and July 31st at the Informatics Forum, Edinburgh.

This year's conference continues the successful growing trend of previous years, attracting the largest number of papers to date for EMNLP and requiring a large organisational effort. I would like to thank all the people involved.

Regina Barzilay from MIT and Mark Johnson from Macquarie University chaired a large scientific programme committee and introduced several innovations in the submission, selection and final scheduling of papers.

Marie Candito, of Paris Diderot University, in collaboration with the ACL workshop chairs, selected the workshops, seven of which have been affiliated to EMNLP. They have received a large number of submissions and generated very interesting programmes.

Special thanks go to the publication chair Wanxiang Che, from Harbin Institute of Technology, who had to deal with EMNLP's now famous just-in-time schedule.

We are very grateful to Bonnie Webber and Miles Osborne, from the University of Edinburgh, for accepting the demanding task of organising the largest and most complex EMNLP ever and allowing us to hold this conference in such a remarkable city. Thanks also go to Francesco Figari the webmaster, and to all the student volunteers that make the conference possible.

Miles Osborne is also the contact person for an interesting affiliated event sponsored by Google and the Scottish and Informatics Computer Science Alliance: an Intense Summer School on Hadoop and Natural Language Processing, that will take place in Edinburgh for two days before the conference.

Thanks also to David Yarowsky and Ken Church from SIGDAT who provided much useful information from past conferences and Graeme Hirst and Priscilla Rasmussen from ACL for their help and advice.

This year's EMNLP's sponsors are Google, Yahoo and Textkernel. We thank them for their very welcome contributions, which were obtained by the efforts of the local organisers and of Massimiliano Ciaramita (Google) and Stefan Riezler (Heidelberg University), our ACL sponsorship committee members for Europe.

Finally, and foremost, thank to all the authors and conference attendees that have made and will make this conference a success and source of inspiration.

Paola Merlo, University of Geneva
EMNLP 2011 General Chair

Program Chairs' Preface

Welcome to the EMNLP 2011 conference, hosted this year by the the extremely impressive University of Edinburgh.

EMNLP 2011 received 628 submissions, a new record for the conference. The program committee, consisted of 20 area chairs and 545 PC members from Asia, Europe, and North America, was able to accept 149 papers in total (for an acceptance rate of 23.7%). Among them, 94 (i.e., 63%) of the papers were accepted for oral presentations, and 54 (i.e., 36%) for poster presentations. Submissions from PC chairs were handled by the general chair in an off-line procedure that applied considerably stricter acceptance requirements than for ordinary submissions.

First and foremost, we would like to thank the authors who submitted their work to EMNLP 2011. The sheer number of submissions reflects how broad and active our field is. We are deeply indebted to the area chairs and the PC members for their hard work. They enabled us to make a wonderful program and to provide valuable feedback to the authors. We are extremely pleased that our invited speaker David McAllester has agreed to talk at EMNLP. Many thanks to local arrangements chairs, Bonnie Webber and Miles Osborne, who enabled the conference to be held in Edinburgh, one of the great intellectual cities of the world. We'd also like to thank the publications chair, Wanxiang Che, who put this volume together. We greatly benefited from advice from Hang Li and Lluís Màrquez who kindly shared with us their experience from EMNLP 2010. Special thanks to the general chair, Paola Merlo, who provided much valuable advice and assistance in the past months. We are grateful to David Yarowsky for assistance with a variety of aspects of the conference. Rich Gerber and the START team responded to our questions quickly, and helped us manage the large number of submissions smoothly; we would like to thank them as well.

This year's conference is innovative in several ways. The conference contains three additional plenary sessions compared to previous EMNLP conferences; these are used to highlight a diverse set of papers of interest to the entire EMNLP audience. We hope this will help counter the disciplinary fragmentation that some of us feel the standard multi-track conference structure encourages.

For the first time, submitted papers could be optionally accompanied by up to 10MB of supplementary material, which could consist of data, code, and text. Papers can reference the supplementary material in much the same way a paper might refer to software or a tech report available from the authors' web site (albeit without revealing the authors' identities). Reviewers were encouraged but not required to view the supplementary material.

A major challenge this year concerned undisclosed double submissions and plagiarism (especially self-plagiarism) involving papers accepted at other international conferences. We believe this is an issue that must be addressed by the broader research community. Our community needs clear, well-publicised, standards on double and overlapping submissions, and also needs procedures for sharing information between relevant conferences to discourage double submission and self-plagiarism.

But we don't want to blow this issue out of proportion. By far the vast majority of EMNLP submissions described creative, innovative work that taken together substantially advances the field. The success of a conference is really a result of the great efforts of everybody involved. We hope that you enjoy this year's conference in the historic city of Edinburgh!

Regina Barzilay and Mark Johnson
EMNLP 2011 Program Co-Chairs

Program Co-chairs

Regina Barzilay, Massachusetts Institute of Technology
Mark Johnson, Macquarie University

Area Chairs

Emily M. Bender, University of Washington
Phil Blunsom, University of Oxford
Eugene Charniak, Brown University
Chris Dyer, Carnegie Mellon University
Jenny Finkel, Columbia University
Radu Florian, IBM Watson Research Center
Amir Globerson, The Hebrew University of Jerusalem
Keith Hall, Google Research
Minlie Huang, Tsinghua University
Katrín Kirchoff, University of Washington
Mirella Lapata, University of Edinburgh
Chin-Yew Lin, Microsoft Research Asia
Bo Pang, Yahoo! Research
Stefan Riezler, University of Heidelberg
Ellen Riloff, University of Utah
Hinrich Schuetze, University of Stuttgart
Jian Su, Institute for Infocomm Research
Dekai Wu, Hong Kong University of Science and Technology
Luke Zettlemoyer, University of Washington

Local Arrangements Chair

Bonnie Webber, University of Edinburgh
Miles Osborne, University of Edinburgh

Publications Chair

Wanxiang Che, Harbin Institute of Technology

Reviewers

Omri Abend, Rodrigo Agerri, Eneko Agirre, Byung Gyu Ahn, Hua Ai, Enrique Alfonseca, Cecilia Ovesdotter Alm, Vamshi Ambati, Sophia Ananiadou, Shankar Ananthakrishnan, Marianna Apidianaki, Shilpa Arora, Abishek Arun, Nicholas Asher, Giuseppe Attardi, Necip Fazil Ayan, Collin Baker, Tim Baldwin, Carmen Banea, Ken Barker, Marco Baroni, Regina Barzilay, Mary Beckman, Beata Beigman Klebanov, Núria Bel, Anja Belz, Emily M. Bender, Luciana Benotti, Edward Benson, Jonathan Berant, Taylor Berg-Kirkpatrick, Sabine Bergler, Shane Bergsma,

Steven Bethard, Pushpak Bhattacharyya, Fadi Biadisy, Chris Biemann, Dan Bikel, Alexandra Birch, Steven Bird, John Blitzer, Michael Bloodgood, Phil Blunsom, Nate Bodenshteyn, Ondrej Bojar, Kalina Bontcheva, Johan Bos, Elizabeth Boschee, Jordan Boyd-Graber, S. R. K. Branavan, Thorsten Brants, Eric Breck, Chris Brew, Ted Briscoe, Chris Brockett, Sam Brody, Paul Buitelaar, Aljoscha Burchardt, Aoife Cahill, Chris Callison-Burch, Nicola Cancedda, Marie Candito, Yunbo Cao, Claire Cardie, Giuseppe Carenini, Andy Carlson, Marine Carpuat, Xavier Carreras, Francisco Casacuberta, Vittorio Castelli, Neus Català, Joyce Chai, Nate Chambers, Pi-Chuan Chang, Pi-Chuan Chang, Eugene Charniak, Wanxiang Che, Ciprian Chelba, Hsin-Hsi Chen, Colin Cherry, David Chiang, Laura Chiticariu, Yejin Choi, Christos Christodoulopoulos, Jennifer Chu-Carroll, Kenneth Church, Massimiliano Ciaramita, Philipp Cimiano, Alexander Clark, Peter Clark, Stephen Clark, James Clarke, Shay Cohen, Trevor Cohn, Nigel Collier, Michael Collins, Kevyn Collins-Thompson, Gao Cong, Marta Ruiz Costa-Jussa, Koby Crammer, Dan Cristea, Silviu Cucerzan, Aron Culotta, Walter Daelemans, Dipanjan Das, Sajib Dasgupta, Hal Daume, Dina Ddemner, Adrià de Gispert, Marie-Catherine de Marneffe, Vera Demberg, Steve DeNeefe, John DeNero, Pascal Denis, Anoop Deoras, Maarten de-Rijke, Mona Diab, Brian Dillon, Xiaowen Ding, Anna Divoli, Mark Dredze, Markus Dreyer, Rebecca Dridan, Gregory Druck, Amit Dubey, Chris Dyer, Koji Eguchi, Jacob Eisenstein, Jason Eisner, Michael Elhadad, Noemie Elhadad, Lloyd Elliott, Micha Elsner, Tanveer Faruque, Benoit Favre, Christiane Fellbaum, Donghui Feng, Yansong Feng, Elena Filatova, Denis Filimonov, Katja Filippova, Jenny Finkel, Dan Flickinger, Radu Florian, George Foster, Jennifer Foster, Anette Frank, Martin Franz, Martin Franz, Alex Fraser, Marjorie Freedman, Michel Galley, Michael Gamon, Kuzman Ganchev, Kavita Ganesan, Juri Ganitkevitch, Wei Gao, Janfeng Gao, Konstantina Garoufi, Jan Gasthaus, Eric Gaussier, Dmitriy Genzel, Kallirroi Georgila, Daniel Gildea, Kevin Gimpel, Roxana Girju, Amir Globerson, Yoav Goldberg, Yoav Goldberg, Dan Goldwasser, Sharon Goldwater, José González-Brenes, Julio González, Joao Graca, Agustin Gravano, Spence Green, Gregory Grefenstette, Ralph Grishman, Iryna Gurevych, Ben Hachey, Barry Haddow, Gholamreza Haffari, Aria Haghighi, Udo Hahn, Dilek Hakkani-Tur, Keith Hall, Ben Han, Sanda Harabagiu, Sasa Hasan, Xiadong He, Xiaodong He, John Henderson, James Henderson, James Henderson, Iris Hendrickx, Aurelie Herbelot, Hugo Hernault, Graeme Hirst, Hieu Hoang, Julia Hockenmaier, Raphael Hoffmann, Kristy Hollingshead, Yunhua Hu, Yunfeng Huang, Liang Huang, Minlie Huang, Zhongqiang Huang, Nancy Ide, Gonzalo Iglesias, Ryu Iida, Piotr Indyk, Diana Inkpen, Kentaro Inui, Abe Ittycheriah, Tommi Jaakkola, Alpa Jain, Martin Jansche, Heng Ji, Kim Jin-Dong, Richard Johansson, Howard Johnson, Mark Johnson, Rie Johnson, Kristiina Jokinen, Rosie Jones, Dan Jurafsky, Min-Yen Kan, Pallika Kanani, Hiroshi Kanayama, Rohit Kate, Simon Keizer, Frank Keller, Chloé Kiddon, Su Nam Kim, Irwin King, Tracy Holloway King, Katrin Kirchhoff, Dietrich Klakow, Dan Klein, Alex Klementiev, Kevin Knight, Philipp Koehn, Ioannis Konstas, Terry Koo, Moshe Koppel, Zornitsa Kozareva, Marco Kuhlmann, Roland Kuhn, Jonas Kuhn, Seth Kulick, Ravi Kumar, Mirella Lapata, Alex Lascarides, Alon Lavie, Claudia Leacock, Matt Lease, Lillian Lee, Yoong Keok Lee, Gregor Leusch, Abby Levenberg, Effi Levi, Gina-Anne Levow, Roger Levy, Xiao Li, Wenjie Li, Fangtao Li, Hang Li, Linlin Li, Mu Li, Shoushan Li, Yunyao Li, Zhifei Li, Percy Liang, Marc Light, Chin-Yew Lin, DeKang Lin, Christina Lioma, Qiaoling Liu, Feifan Liu, Yan Liu, Ting Liu, Yang Liu, Ed Loper, Adam Lopez, John Lowe, Yue Lu, Xiaoqiang Luo, Bill MacCartney, Klaus Macherey, Wolfgang Macherey, Nitin Madnani, Bernardo Magnini, Francois Mairesse, Andreas Maletti, Suresh Manandhar, Gideon Mann, Chris Manning, Daniel Marcu, Montserrat Marimon, Lluís Màrquez, James Martin, Andre Martins, Sameer Maskey, Fabio Massimo Zanzotto, Yuji Matsumoto, Irina Matveeva,

Arne Mauser, Michael Maxwell, Andrew McCallum, Diana McCarthy, David McClosky, Ryan McDonald, Candace McKenna, Kathy McKeown, Paul McNamee, Dan Melamed, Chris Melli-lish, Arul Menezes, Donald Metzler, Adam Meyers, Haitao Mi, Rada Mihalcea, Tim Miller, Daniel Mills, Eleni Miltsakaki, David Mimno, Chang Ming-Wei, Shachar Mirkin, Vibhu Mit-
tal, Sien Moens, Mary-Francine Moens, Saif Mohammad, Christof Monz, Taesun Moon, Ray-
mond Mooney, Robert Moore, Roser Morante, Alessandro Moschitti, Rutu Mulkar-Mehta, Dra-
gos Munteanu, Smaranda Muresan, Gabriel Murray, Tetsuji Nakagawa, Preslav Nakov, Ramesh
Nallapati, Jason Naradowsky, Tahira Naseem, Vivi Nastase, Mark-Jan Nederhof, Ani Nenkova,
Günter Neumann, Hermann Ney, Hwee Tou Ng, Vincent Ng, Malvina Nissim, Joakim Nivre,
Tadashi Nomoto, Eric Nyberg, Brendan O'Connor, Stephan Oepen, Kemal Oflazer, Manabu
Okumura, Sebastian Pado, Lluís Padró, John Paisley, Alexis Palmer, Sinno Jialin Pan, Bo Pang,
Simone Paolo Ponzetto, Carolina Parada, Chris Parisien, Marius Pasca, Siddharth Patwardhan,
Adam Pauls, Ted Pedersen, Anselmo Penas, Gerald Penn, Marco Pennacchiotti, Slav Petrov, Sasa
Petrovic, Manfred Pinkal, Emily Pitler, Massimo Poesio, Thierry Poibeau, Joseph Polifroni, Hoi-
fung Poon, Ana-Maria Popescu, Andrei Popescu-Belis, Richard Power, Sameer Pradhan, John
Prager, Rashmi Prasad, Matthew Purver, James Pustejovsky, Chris Quirk, Stephan Raaijmakers,
Dragomir Radev, Hema Raghavan, Ari Rappoport, Lev Ratinov, Sravana Reddy, Roi Reichart,
Joseph Reisinger, David Reitter, Philip Resnik, Sebastian Riedel, Sebastian Riedel, Stefan Rie-
zler, German Rigau, Ellen Riloff, Hae-Chang Rim, Laura Rimell, Eric Ringger, Alan Ritter,
Barbara Rosario, Andrew Rosenberg, Paolo Rosso, Antti-Veikko Rosti, Benjamin Roth, Salim
Roukos, Sasha Rush, Markus Saers, Kenji Sagae, Horacio Saggion, Mark Sammons, Anoop
Sarkar, Yutaka Sasaki, Christina Sauper, David Schlangen, Nathan Schneider, Hinrich Schuetze,
Karl Schultz, Lane Schwarz, Holger Schwenk, Frank Seide, Satoshi Sekine, Jean Senellart, Ab-
hinav Sethy, Hendra Setiawan, Burr Settles, Zak Shafran, Kumar Shankar, Dan Shen, Libin
Shen, Shuming Shi, Hideki Shima, Advait Siddharthan, Khalil Sima'an, Michel Simard, Sameer
Singh, Gabriel Skantze, Kevin Small, David Smith, Jason Smith, Noah Smith, Matthew Sover,
Ben Snyder, Stephen Soderland, Swapna Somasundaran, David Sontag, Valentin Spitkovsky,
Caroline Sporleder, Mark Steedman, Mark Stevenson, Suzanne Stevenson, Veselin Stoyanov,
Michael Strube, Jian Su, Venkat Subramaniam, Maosong Sun, Yun-Hsuan Sung, Mihai Sur-
deanu, Jun Suzuki, Ben Swanson, Stan Szpakowicz, Hiroya Takamura, Jie Tang, Ben Taskar,
Yee Whye Teh, Simone Teufel, Stefan Thater, Christoph Tillmann, Ivan Titov, Cigdem Toprak,
Kristina Toutonova, Roy Tromble, George Tsatsaronis, Yoshimasa Tsuruoka, Gokhan Tur, Eve-
lyne Tzoukermann, Raghavendra Udupa, Olga Uryupina, Jakob Uszkoreit, Benjamin Van Durme,
Lucy Vanderwende, Paola Verlardi, Yannick Versley, Karin Verspoor, David Vilar, Andreas Vla-
chos, Stephan Vogel, Martin Volk, Xiaojun Wan, Stephen Wan, Chong Wang, Mengqiu Wang,
Wen Wang, Wei Wang, Taro Watanabe, Bonnie Webber, David Weir, Ralph Weischedel, Ben
Wellner, Scott Wen-tau Yih, Michael Wick, Jan Wiebe, Theresa Wilson, Shuly Wintner, kamfai
wong, Kristian Woodsend, Dekai Wu, Yunqing Xia, Jinxi Xu, Peng Xu, Huichao Xue, Nian-
wen Xue, Alexander Yates, Hao Yu, Hong Yu, Yisong Yue, Francois Yvon, Richard Zens, Luke
Zettlemoyer, Chengxiang Zhai, Joy Zhang, qi zhang, Tong Zhang, Hao Zhang, Hai Zhao, Ming
Zhou, Jerry Zhu, Jun Zhu, Imed Zitouni, Andreas Zollmann, Geoff Zweig, Pierre Zweigenbaum

Best Reviewer Awards

Tim Baldwin, Ken Barker, Marco Baroni, Mary Beckman, Luciana Benotti, Steven Bethard,
Chris Biemann, Phil Blunsom, Chris Brockett, Sam Brody, Nate Chambers, Jennifer Chu-Carroll,

Peter Clark, Dina Ddemner, Markus Dreyer, Rebecca Dridan, Benjamin Van Durme, Chris Dyer, Michael Elhadad, Donghui Feng, Katja Filippova, Anette Frank, Kallirroi Georgila, Kevin Gimpel, Ben Hachey, John Henderson, Graeme Hirst, Dan Jurafsky, Min-Yen Kan, Chloe Kiddon, Tracy King, Kevin Knight, Terry Koo, Roger Levy, Bill MacCartney, Lluís Marquez, Andre Martins, Ryan McDonald, David Minmo, Saif Mohammad, Raymond Mooney, Ani Nenkova, Günter Neumann, Vincent Ng, Tadashi Nomoto, Alexis Palmer, Ted Pedersen, Sasa Petrovic, Emily Pitler, Hoifung Poon, Ari Rappoport, Eric Ringer, Alan Ritter, Mark Sammons, Yutaka Sasaki, Kevin Small, Noah Smith, Suzanne Stevenson, Yee Whye Teh, Simone Teufel, Raghavendra Udupa, Paola Verlardi, Andreas Vlachos, Stephen Wan, Alexander Yates, Scott Wen-tau Yih, Fabio Massimo Zanzotto, Pierre Zweigenbaum

Table of Contents

<i>Fast and Robust Joint Models for Biomedical Event Extraction</i> Sebastian Riedel and Andrew McCallum	1
<i>Predicting Thread Discourse Structure over Technical Web Forums</i> Li Wang, Marco Lui, Su Nam Kim, Joakim Nivre and Timothy Baldwin	13
<i>Exact Decoding of Phrase-Based Translation Models through Lagrangian Relaxation</i> Yin-Wen Chang and Michael Collins	26
<i>Optimal Search for Minimum Error Rate Training</i> Michel Galley and Chris Quirk	38
<i>Unsupervised Structure Prediction with Non-Parallel Multilingual Guidance</i> Shay B. Cohen, Dipanjan Das and Noah A. Smith	50
<i>Multi-Source Transfer of Delexicalized Dependency Parsers</i> Ryan McDonald, Slav Petrov and Keith Hall	62
<i>SMT Helps Bitext Dependency Parsing</i> Wenliang Chen, Jun'ichi Kazama, Min Zhang, Yoshimasa Tsuruoka, Yujie Zhang, Yiou Wang, Kentaro Torisawa and Haizhou Li	73
<i>Accurate Parsing with Compact Tree-Substitution Grammars: Double-DOP</i> Federico Sangati and Willem Zuidema	84
<i>A Generate and Rank Approach to Sentence Paraphrasing</i> Prodromos Malakasiotis and Ion Androutsopoulos	96
<i>Correcting Semantic Collocation Errors with LI-induced Paraphrases</i> Daniel Dahlmeier and Hwee Tou Ng	107
<i>Class Label Enhancement via Related Instances</i> Zornitsa Kozareva, Konstantin Voevodski and Shanghua Teng	118
<i>A Joint Model for Extended Semantic Role Labeling</i> Vivek Srikumar and Dan Roth	129
<i>Domain-Assisted Product Aspect Hierarchy Generation: Towards Hierarchical Organization of Un-structured Consumer Reviews</i> Jianxing Yu, Zheng-Jun Zha, Meng Wang, Kai Wang and Tat-Seng Chua	140
<i>Semi-Supervised Recursive Autoencoders for Predicting Sentiment Distributions</i> Richard Socher, Jeffrey Pennington, Eric H. Huang, Andrew Y. Ng and Christopher D. Manning	151

<i>Unsupervised Discovery of Discourse Relations for Eliminating Intra-sentence Polarity Ambiguities</i> Lanjun Zhou, Binyang Li, Wei Gao, Zhongyu Wei and Kam-Fai Wong	162
<i>Compositional Matrix-Space Models for Sentiment Analysis</i> Ainur Yessenalina and Claire Cardie	172
<i>Training a Parser for Machine Translation Reordering</i> Jason Katz-Brown, Slav Petrov, Ryan McDonald, Franz Och, David Talbot, Hiroshi Ichikawa, Masakazu Seno and Hideto Kazawa.....	183
<i>Inducing Sentence Structure from Parallel Corpora for Reordering</i> John DeNero and Jakob Uszkoreit	193
<i>Augmenting String-to-Tree Translation Models with Fuzzy Use of Source-side Syntax</i> Jiajun Zhang, Feifei Zhai and Chengqing Zong.....	204
<i>A novel dependency-to-string model for statistical machine translation</i> Jun Xie, Haitao Mi and Qun Liu	216
<i>Bayesian Checking for Topic Models</i> David Mimno and David Blei.....	227
<i>Dual Decomposition with Many Overlapping Components</i> Andre Martins, Noah Smith, Mario Figueiredo and Pedro Aguiar.....	238
<i>Approximate Scalable Bounded Space Sketch for Large Data NLP</i> Amit Goyal and Hal Daume III	250
<i>Optimizing Semantic Coherence in Topic Models</i> David Mimno, Hanna Wallach, Edmund Talley, Miriam Leenders and Andrew McCallum ...	262
<i>A Weakly-supervised Approach to Argumentative Zoning of Scientific Documents</i> Yufan Guo, Anna Korhonen and Thierry Poibeau	273
<i>Linear Text Segmentation Using Affinity Propagation</i> Anna Kazantseva and Stan Szpakowicz	284
<i>Minimally Supervised Event Causality Identification</i> Quang Do, Yee Seng Chan and Dan Roth.....	294
<i>A Model of Discourse Predictions in Human Sentence Processing</i> Amit Dubey, Frank Keller and Patrick Sturt	304
<i>Simple Effective Decipherment via Combinatorial Optimization</i> Taylor Berg-Kirkpatrick and Dan Klein	313
<i>Universal Morphological Analysis using Structured Nearest Neighbor Prediction</i> Young-Bum Kim, João Graça and Benjamin Snyder	322

<i>Training a Log-Linear Parser with Loss Functions via Softmax-Margin</i>	
Michael Auli and Adam Lopez	333
<i>Large-Scale Cognate Recovery</i>	
David Hall and Dan Klein	344
<i>Domain Adaptation via Pseudo In-Domain Data Selection</i>	
Amittai Axelrod, Xiaodong He and Jianfeng Gao	355
<i>Language Models for Machine Translation: Original vs. Translated Texts</i>	
Gennadi Lembersky, Noam Ordan and Shuly Wintner	363
<i>Better Evaluation Metrics Lead to Better Machine Translation</i>	
Chang Liu, Daniel Dahlmeier and Hwee Tou Ng	375
<i>Evaluating Dependency Parsing: Robust and Heuristics-Free Cross-Annotation Evaluation</i>	
Reut Tsarfaty, Joakim Nivre and Evelina Andersson	385
<i>Parser Evaluation over Local and Non-Local Deep Dependencies in a Large Corpus</i>	
Emily M. Bender, Dan Flickinger, Stephan Oepen and Yi Zhang	397
<i>Learning to Simplify Sentences with Quasi-Synchronous Grammar and Integer Programming</i>	
Kristian Woodsend and Mirella Lapata	409
<i>Bootstrapping Semantic Parsers from Conversations</i>	
Yoav Artzi and Luke Zettlemoyer	421
<i>Timeline Generation through Evolutionary Trans-Temporal Summarization</i>	
Rui Yan, Liang Kong, Congrui Huang, Xiaojun Wan, Xiaoming Li and Yan Zhang	433
<i>Corpus-Guided Sentence Generation of Natural Images</i>	
Yezhou Yang, Ching Teo, Hal Daume III and Yiannis Aloimonos	444
<i>Corroborating Text Evaluation Results with Heterogeneous Measures</i>	
Enrique Amigó, Julio Gonzalo, Jesus Gimenez and Felisa Verdejo	455
<i>Ranking Human and Machine Summarization Systems</i>	
Peter Rankel, John Conroy, Eric Slud and Dianne O’Leary	467
<i>Quasi-Synchronous Phrase Dependency Grammars for Machine Translation</i>	
Kevin Gimpel and Noah A. Smith	474
<i>A Word Reordering Model for Improved Machine Translation</i>	
Karthik Visweswariah, Rajakrishnan Rajkumar, Ankur Gandhe, Ananthkrishnan Ramanathan and Jiri Navratil	486
<i>Feature-Rich Language-Independent Syntax-Based Alignment for Statistical Machine Translation</i>	
Jason Riesa, Ann Irvine and Daniel Marcu	497

<i>Literal and Metaphorical Sense Identification through Concrete and Abstract Context</i> Peter Turney, Yair Neuman, Dan Assaf and Yohai Cohen	680
<i>Syntactic Decision Tree LMs: Random Selection or Intelligent Design?</i> Denis Filimonov and Mary Harper	691
<i>The Imagination of Crowds: Conversational AAC Language Modeling using Crowdsourcing and Large Data Sources</i> Keith Vertanen and Per Ola Kristensson	700
<i>Using Syntactic and Semantic Structural Kernels for Classifying Definition Questions in Jeopardy!</i> Alessandro Moschitti, Jennifer Chu-carroll, Siddharth Patwardhan, James Fan and Giuseppe Ricciardi	712
<i>Multiword Expression Identification with Tree Substitution Grammars: A Parsing tour de force with French</i> Spence Green, Marie-Catherine de Marneffe, John Bauer and Christopher D. Manning	725
<i>Modelling Discourse Relations for Arabic</i> Amal Al-Saif and Katja Markert	736
<i>Classifying Sentences as Speech Acts in Message Board Posts</i> Ashequl Qadir and Ellen Riloff	748
<i>Learning Local Content Shift Detectors from Document-level Information</i> Richard Farkas	759
<i>Collaborative Ranking: A Case Study on Entity Linking</i> Zheng Chen and Heng Ji	771
<i>Robust Disambiguation of Named Entities in Text</i> Johannes Hoffart, Mohamed Amir Yosef, Iaria Bordino, Hagen Fürstenau, Manfred Pinkal, Marc Spaniol, Bilyana Taneva, Stefan Thater and Gerhard Weikum	782
<i>A Cascaded Classification Approach to Semantic Head Recognition</i> Lukas Michelbacher, Alok Kothari, Martin Forst, Christina Lioma and Hinrich Schütze	793
<i>Linking Entities to a Knowledge Base with Query Expansion</i> Swapna Gottipati and Jing Jiang	804
<i>Unsupervised Information Extraction with Distributional Prior Knowledge</i> Cane Wing-ki Leung, Jing Jiang, Kian Ming A. Chai, Hai Leong Chieu and Loo-Nin Teow ..	814
<i>Relation Acquisition using Word Classes and Partial Patterns</i> Stijn De Saeger, Kentaro Torisawa, Masaaki Tsuchida, Jun'ichi Kazama, Chikara Hashimoto, Ichiro Yamada, Jong Hoon Oh, Istvan Varga and Yulan Yan	825
<i>Identification of Multi-word Expressions by Combining Multiple Linguistic Information Sources</i> Yulia Tsvetkov and Shuly Wintner	836

<i>Analyzing Methods for Improving Precision of Pivot Based Bilingual Dictionaries</i>	
Xabier Saralegi, Iker Manterola and Iñaki San Vicente	846
<i>Soft Dependency Constraints for Reordering in Hierarchical Phrase-Based Translation</i>	
Yang Gao, Philipp Koehn and Alexandra Birch	857
<i>Statistical Machine Translation with Local Language Models</i>	
Christof Monz	869
<i>Fast Generation of Translation Forest for Large-Scale SMT Discriminative Training</i>	
Xinyan Xiao, Yang Liu, Qun Liu and Shouxun Lin	880
<i>A Correction Model for Word Alignments</i>	
J. Scott McCarley, Abraham Ittycheriah, Salim Roukos, Bing Xiang and Jian-ming Xu	889
<i>Heuristic Search for Non-Bottom-Up Tree Structure Prediction</i>	
Andrea Gesmundo and James Henderson	899
<i>Cache-based Document-level Statistical Machine Translation</i>	
Zhengxian Gong, Min Zhang and Guodong Zhou	909
<i>Minimum Imputed-Risk: Unsupervised Discriminative Training for Machine Translation</i>	
Zhifei Li, Ziyuan Wang, Jason Eisner, Sanjeev Khudanpur and Brian Roark	920
<i>Improving Bilingual Projections via Sparse Covariance Matrices</i>	
Jagadeesh Jagarlamudi, Raghavendra Udupa, Hal Daume III and Abhijit Bhole	930
<i>Entire Relaxation Path for Maximum Entropy Problems</i>	
Moshe Dubiner and Yoram Singer	941
<i>A Non-negative Matrix Factorization Based Approach for Active Dual Supervision from Document and Word Labels</i>	
Chao Shen and Tao Li	949
<i>Splitting Noun Compounds via Monolingual and Bilingual Paraphrasing: A Study on Japanese Katakana Words</i>	
Nobuhiro Kaji and Masaru Kitsuregawa	959
<i>Enhancing Chinese Word Segmentation Using Unlabeled Data</i>	
Weiwei Sun and Jia Xu	970
<i>Unsupervised Learning of Selectional Restrictions and Detection of Argument Coercions</i>	
Kirk Roberts and Sanda Harabagiu	980
<i>Harnessing different knowledge sources to measure semantic relatedness under a uniform model</i>	
Ziqi Zhang, Anna Lisa Gentile and Fabio Ciravegna	991
<i>Refining the Notions of Depth and Density in WordNet-based Semantic Similarity Measures</i>	
Tong Wang and Graeme Hirst	1003

<i>Latent Vector Weighting for Word Meaning in Context</i>	
Tim Van de Cruys, Thierry Poibeau and Anna Korhonen	1012
<i>Hierarchical Verb Clustering Using Graph Factorization</i>	
Lin Sun and Anna Korhonen	1023
<i>Structured Lexical Similarity via Convolution Kernels on Dependency Trees</i>	
Danilo Croce, Alessandro Moschitti and Roberto Basili	1034
<i>Probabilistic models of similarity in syntactic context</i>	
Diarmuid Ó Séaghdha and Anna Korhonen	1047
<i>Lexical Co-occurrence, Statistical Significance, and Word Association</i>	
Dipak L. Chaudhari, Om P. Damani and Srivatsan Laxman	1058
<i>Learning the Information Status of Noun Phrases in Spoken Dialogues</i>	
Altaf Rahman and Vincent Ng	1069
<i>Harnessing WordNet Senses for Supervised Sentiment Classification</i>	
Balamurali AR, Aditya Joshi and Pushpak Bhattacharyya	1081
<i>Learning General Connotation of Words using Graph-based Algorithms</i>	
Song Feng, Ritwik Bose and Yejin Choi	1092
<i>Hypotheses Selection Criteria in a Reranking Framework for Spoken Language Understanding</i>	
Marco Dinarelli and Sophie Rosset	1104
<i>A Fast Re-scoring Strategy to Capture Long-Distance Dependencies</i>	
Anoop Deoras, Tomas Mikolov and Kenneth Church	1116
<i>Efficient Subsampling for Training Complex Language Models</i>	
Puyang Xu, Asela Gunawardana and Sanjeev Khudanpur	1128
<i>Generating Aspect-oriented Multi-Document Summarization with Event-aspect model</i>	
Peng Li, Yinglin Wang, Wei Gao and Jing Jiang	1137
<i>Syntax-Based Grammaticality Improvement using CCG and Guided Search</i>	
Yue Zhang and Stephen Clark	1147
<i>Generating Subsequent Reference in Shared Visual Scenes: Computation vs Re-Use</i>	
Jette Viethen, Robert Dale and Markus Guhe	1158
<i>Learning Sentential Paraphrases from Bilingual Parallel Corpora for Text-to-Text Generation</i>	
Juri Ganitkevitch, Chris Callison-Burch, Courtney Napoles and Benjamin Van Durme	1168
<i>Joint Models for Chinese POS Tagging and Dependency Parsing</i>	
Zhenghua Li, Min Zhang, Wanxiang Che, Ting Liu, Wenliang Chen and Haizhou Li	1180
<i>Relaxed Cross-lingual Projection of Constituent Syntax</i>	
Wenbin Jiang, Qun Liu and Yajuan Lv	1192

<i>Computing Logical Form on Regulatory Texts</i>	
Nikhil Dinesh, Aravind Joshi and Insup Lee	1202
<i>Computation of Infix Probabilities for Probabilistic Context-Free Grammars</i>	
Mark-Jan Nederhof and Giorgio Satta	1213
<i>Parse Correction with Specialized Models for Difficult Attachment Types</i>	
Enrique Henestroza Anguiano and Marie Candito	1222
<i>Exact Inference for Generative Probabilistic Non-Projective Dependency Parsing</i>	
Shay B. Cohen, Carlos Gómez-Rodríguez and Giorgio Satta	1234
<i>Semi-supervised CCG Lexicon Extension</i>	
Emily Thomforde and Mark Steedman	1246
<i>A Fast, Accurate, Non-Projective, Semantically-Enriched Parser</i>	
Stephen Tratz and Eduard Hovy	1257
<i>Lateen EM: Unsupervised Training with Multiple Objectives, Applied to Dependency Grammar Induction</i>	
Valentin I. Spitzkovsky, Hiyan Alshawi and Daniel Jurafsky	1269
<i>Unsupervised Dependency Parsing without Gold Part-of-Speech Tags</i>	
Valentin I. Spitzkovsky, Hiyan Alshawi, Angel X. Chang and Daniel Jurafsky	1281
<i>Exploiting Syntactic and Distributional Information for Spelling Correction with Web-Scale N-gram Models</i>	
Wei Xu, Joel Tetreault, Martin Chodorow, Ralph Grishman and Le Zhao	1291
<i>Discriminating Gender on Twitter</i>	
John D. Burger, John Henderson, George Kim and Guido Zarrella	1301
<i>Identifying and Following Expert Investors in Stock Microblogs</i>	
Roy Bar-Haim, Elad Dinur, Ronen Feldman, Moshe Fresko and Guy Goldstein	1310
<i>Unsupervised Semantic Role Induction with Graph Partitioning</i>	
Joel Lang and Mirella Lapata	1320
<i>Structural Opinion Mining for Graph-based Sentiment Representation</i>	
Yuanbin Wu, Qi Zhang, Xuanjing Huang and Lide Wu	1332
<i>Summarize What You Are Interested In: An Optimization Framework for Interactive Personalized Summarization</i>	
Rui Yan, Jian-Yun Nie and Xiaoming Li	1342
<i>Tuning as Ranking</i>	
Mark Hopkins and Jonathan May	1352

<i>Watermarking the Outputs of Structured Prediction with an application in Statistical Machine Translation.</i>	
Ashish Venugopal, Jakob Uszkoreit, David Talbot, Franz Och and Juri Ganitkevitch	1363
<i>Hierarchical Phrase-based Translation Representations</i>	
Gonzalo Iglesias, Cyril Allauzen, William Byrne, Adrià de Gispert and Michael Riley	1373
<i>Improved Transliteration Mining Using Graph Reinforcement</i>	
Ali El Kahki, Kareem Darwish, Ahmed Saad El Din, Mohamed Abd El-Wahab, Ahmed Hefny and Waleed Ammar	1384
<i>Experimental Support for a Categorical Compositional Distributional Model of Meaning</i>	
Edward Grefenstette and Mehrnoosh Sadrzadeh	1394
<i>Cross-Cutting Models of Lexical Semantics</i>	
Joseph Reisinger and Raymond Mooney	1405
<i>Reducing Grounded Learning Tasks To Grammatical Inference</i>	
Benjamin Börschinger, Bevan K. Jones and Mark Johnson	1416
<i>Relation Extraction with Relation Topics</i>	
Chang Wang, James Fan, Aditya Kalyanpur and David Gondek	1426
<i>Extreme Extraction – Machine Reading in a Week</i>	
Marjorie Freedman, Lance Ramshaw, Elizabeth Boschee, Ryan Gabbard, Gary Kratkiewicz, Nicolas Ward and Ralph Weischedel	1437
<i>Discovering Relations between Noun Categories</i>	
Thahir Mohamed, Estevam Hruschka and Tom Mitchell	1447
<i>Structured Relation Discovery using Generative Models</i>	
Limin Yao, Aria Haghighi, Sebastian Riedel and Andrew McCallum	1456
<i>Closing the Loop: Fast, Interactive Semi-Supervised Annotation With Queries on Features and Instances</i>	
Burr Settles	1467
<i>Third-order Variational Reranking on Packed-Shared Dependency Forests</i>	
Katsuhiko Hayashi, Taro Watanabe, Masayuki Asahara and Yuji Matsumoto	1479
<i>Training dependency parsers by jointly optimizing multiple objectives</i>	
Keith Hall, Ryan McDonald, Jason Katz-Brown and Michael Ringgaard	1489
<i>Structured Sparsity in Structured Prediction</i>	
Andre Martins, Noah Smith, Mario Figueiredo and Pedro Aguiar	1500
<i>Lexical Generalization in CCG Grammar Induction for Semantic Parsing</i>	
Tom Kwiatkowski, Luke Zettlemoyer, Sharon Goldwater and Mark Steedman	1512

<i>Named Entity Recognition in Tweets: An Experimental Study</i>	
Alan Ritter, Sam Clark, Mausam and Oren Etzioni	1524
<i>Identifying Relations for Open Information Extraction</i>	
Anthony Fader, Stephen Soderland and Oren Etzioni	1535
<i>Active Learning with Amazon Mechanical Turk</i>	
Florian Laws, Christian Scheible and Hinrich Schütze	1546
<i>Bootstrapped Named Entity Recognition for Product Attribute Extraction</i>	
Duangmanee Putthividhya and Junling Hu	1557
<i>Twitter Catches The Flu: Detecting Influenza Epidemics using Twitter</i>	
Eiji ARAMAKI, Sachiko MASKAWA and Mizuki MORITA	1568
<i>A Simple Word Trigger Method for Social Tag Suggestion</i>	
Zhiyuan Liu, Xinxiong Chen and Maosong Sun	1577
<i>Rumor has it: Identifying Misinformation in Microblogs</i>	
Vahed Qazvinian, Emily Rosengren, Dragomir R. Radev and Qiaozhu Mei	1589
<i>Exploiting Parse Structures for Native Language Identification</i>	
Sze-Meng Jojo Wong and Mark Dras	1600
<i>A Probabilistic Forest-to-String Model for Language Generation from Typed Lambda Calculus Expressions</i>	
Wei Lu and Hwee Tou Ng	1611

Conference Program

Wednesday, July 27, 2011

9:00–10:20 Plenary session: Opening and Invited Talk
Chair: Paola Merlo
Invited Talk: Object Detection Grammars (David McAllester)

10:20–11:00 Coffee break

Session 1: Plenary session

Chair: Jason Eisner

11:00–11:25 *Fast and Robust Joint Models for Biomedical Event Extraction*
Sebastian Riedel and Andrew McCallum

11:25–11:50 *Predicting Thread Discourse Structure over Technical Web Forums*
Li Wang, Marco Lui, Su Nam Kim, Joakim Nivre and Timothy Baldwin

11:50–12:15 *Exact Decoding of Phrase-Based Translation Models through Lagrangian Relaxation*
Yin-Wen Chang and Michael Collins

12:15–12:40 *Optimal Search for Minimum Error Rate Training*
Michel Galley and Chris Quirk

12:40–14:10 Lunch

Session 2A: Syntax and Parsing

Chair: Stephen Clark

14:10–14:35 *Unsupervised Structure Prediction with Non-Parallel Multilingual Guidance*
Shay B. Cohen, Dipanjan Das and Noah A. Smith

14:35–15:00 *Multi-Source Transfer of Delexicalized Dependency Parsers*
Ryan McDonald, Slav Petrov and Keith Hall

15:00–15:25 *SMT Helps Bitext Dependency Parsing*
Wenliang Chen, Jun'ichi Kazama, Min Zhang, Yoshimasa Tsuruoka, Yujie Zhang, You Wang, Kentaro Torisawa and Haizhou Li

15:25–15:50 *Accurate Parsing with Compact Tree-Substitution Grammars: Double-DOP*
Federico Sangati and Willem Zuidema

Wednesday, July 27, 2011 (continued)

Session 2B: Semantics

Chair: Mirella Lapata

- 14:10–14:35 *A Generate and Rank Approach to Sentence Paraphrasing*
Prodromos Malakasiotis and Ion Androutsopoulos
- 14:35–15:00 *Correcting Semantic Collocation Errors with LI-induced Paraphrases*
Daniel Dahlmeier and Hwee Tou Ng
- 15:00–15:25 *Class Label Enhancement via Related Instances*
Zornitsa Kozareva, Konstantin Voevodski and Shanghua Teng
- 15:25–15:50 *A Joint Model for Extended Semantic Role Labeling*
Vivek Srikumar and Dan Roth

Session 2C: Sentiment Analysis and Opinion Mining

Chair: Bo Pang

- 14:10–14:35 *Domain-Assisted Product Aspect Hierarchy Generation: Towards Hierarchical Organization of Unstructured Consumer Reviews*
Jianxing Yu, Zheng-Jun Zha, Meng Wang, Kai Wang and Tat-Seng Chua
- 14:35–15:00 *Semi-Supervised Recursive Autoencoders for Predicting Sentiment Distributions*
Richard Socher, Jeffrey Pennington, Eric H. Huang, Andrew Y. Ng and Christopher D. Manning
- 15:00–15:25 *Unsupervised Discovery of Discourse Relations for Eliminating Intra-sentence Polarity Ambiguities*
Lanjuan Zhou, Binyang Li, Wei Gao, Zhongyu Wei and Kam-Fai Wong
- 15:25–15:50 *Compositional Matrix-Space Models for Sentiment Analysis*
Ainur Yessenalina and Claire Cardie

Wednesday, July 27, 2011 (continued)

Session 3A: Machine Translation

Chair: Phil Blunsom

- 16:20–16:45 *Training a Parser for Machine Translation Reordering*
Jason Katz-Brown, Slav Petrov, Ryan McDonald, Franz Och, David Talbot, Hiroshi Ichikawa, Masakazu Seno and Hideto Kazawa
- 16:45–17:10 *Inducing Sentence Structure from Parallel Corpora for Reordering*
John DeNero and Jakob Uszkoreit
- 17:10–17:35 *Augmenting String-to-Tree Translation Models with Fuzzy Use of Source-side Syntax*
Jiajun Zhang, Feifei Zhai and Chengqing Zong
- 17:35–18:00 *A novel dependency-to-string model for statistical machine translation*
Jun Xie, Haitao Mi and Qun Liu

Session 3B: NLP related Machine Learning

Chair: David Smith

- 16:20–16:45 *Bayesian Checking for Topic Models*
David Mimno and David Blei
- 16:45–17:10 *Dual Decomposition with Many Overlapping Components*
Andre Martins, Noah Smith, Mario Figueiredo and Pedro Aguiar
- 17:10–17:35 *Approximate Scalable Bounded Space Sketch for Large Data NLP*
Amit Goyal and Hal Daume III
- 17:35–18:00 *Optimizing Semantic Coherence in Topic Models*
David Mimno, Hanna Wallach, Edmund Talley, Miriam Leenders and Andrew McCallum

Wednesday, July 27, 2011 (continued)

Session 3C: Discourse Dialogue and Pragmatics

Chair: Oliver Lemon

- 16:20–16:45 *A Weakly-supervised Approach to Argumentative Zoning of Scientific Documents*
Yufan Guo, Anna Korhonen and Thierry Poibeau
- 16:45–17:10 *Linear Text Segmentation Using Affinity Propagation*
Anna Kazantseva and Stan Szpakowicz
- 17:10–17:35 *Minimally Supervised Event Causality Identification*
Quang Do, Yee Seng Chan and Dan Roth
- 17:35–18:00 *A Model of Discourse Predictions in Human Sentence Processing*
Amit Dubey, Frank Keller and Patrick Sturt

Thursday, July 28, 2011

Session 4: Plenary session

Chair: Michael Collins

- 9:05–9:30 *Simple Effective Decipherment via Combinatorial Optimization*
Taylor Berg-Kirkpatrick and Dan Klein
- 9:30–9:55 *Universal Morphological Analysis using Structured Nearest Neighbor Prediction*
Young-Bum Kim, João Graça and Benjamin Snyder
- 9:55–10:20 *Training a Log-Linear Parser with Loss Functions via Softmax-Margin*
Michael Auli and Adam Lopez
- 10:20–11:00 Coffee break

Thursday, July 28, 2011 (continued)

Session 5A: Machine Translation

Chair: Philipp Koehn

11:00–11:25 *Large-Scale Cognate Recovery*

David Hall and Dan Klein

11:25–11:50 *Domain Adaptation via Pseudo In-Domain Data Selection*

Amittai Axelrod, Xiaodong He and Jianfeng Gao

11:50–12:15 *Language Models for Machine Translation: Original vs. Translated Texts*

Gennadi Lembersky, Noam Ordan and Shuly Wintner

12:15–12:40 *Better Evaluation Metrics Lead to Better Machine Translation*

Chang Liu, Daniel Dahlmeier and Hwee Tou Ng

Session 5B: Syntax and Parsing

Chair: Ryan McDonald

11:00–11:25 *Evaluating Dependency Parsing: Robust and Heuristics-Free Cross-Annotation Evaluation*

Reut Tsarfaty, Joakim Nivre and Evelina Andersson

11:25–11:50 *Parser Evaluation over Local and Non-Local Deep Dependencies in a Large Corpus*

Emily M. Bender, Dan Flickinger, Stephan Oepen and Yi Zhang

11:50–12:15 *Learning to Simplify Sentences with Quasi-Synchronous Grammar and Integer Programming*

Kristian Woodsend and Mirella Lapata

12:15–12:40 *Bootstrapping Semantic Parsers from Conversations*

Yoav Artzi and Luke Zettlemoyer

Thursday, July 28, 2011 (continued)

Session 5C: Summarization and Generation

Chair: Johanna Moore

- 11:00–11:25 *Timeline Generation through Evolutionary Trans-Temporal Summarization*
Rui Yan, Liang Kong, Congrui Huang, Xiaojun Wan, Xiaoming Li and Yan Zhang
- 11:25–11:50 *Corpus-Guided Sentence Generation of Natural Images*
Yezhou Yang, Ching Teo, Hal Daume III and Yiannis Aloimonos
- 11:50–12:15 *Corroborating Text Evaluation Results with Heterogeneous Measures*
Enrique Amigó, Julio Gonzalo, Jesus Gimenez and Felisa Verdejo
- 12:15–12:40 *Ranking Human and Machine Summarization Systems*
Peter Rinkel, John Conroy, Eric Slud and Dianne O’Leary
- 12:40–14:10 Lunch

Session 6A: Machine Translation

Chair: Stefan Riezler

- 14:10–14:35 *Quasi-Synchronous Phrase Dependency Grammars for Machine Translation*
Kevin Gimpel and Noah A. Smith
- 14:35–15:00 *A Word Reordering Model for Improved Machine Translation*
Karthik Visweswariah, Rajakrishnan Rajkumar, Ankur Gandhe, Ananthkrishnan Ramanathan and Jiri Navratil
- 15:00–15:25 *Feature-Rich Language-Independent Syntax-Based Alignment for Statistical Machine Translation*
Jason Riesa, Ann Irvine and Daniel Marcu
- 15:25–15:50 *Efficient retrieval of tree translation examples for Syntax-Based Machine Translation*
Fabien Cromieres and Sadao Kurohashi

Thursday, July 28, 2011 (continued)

Session 7A: Phonology Morphology Tagging Chunking and Segmentation

Chair: Noah Smith

- 16:20–16:45 *Non-parametric Bayesian Segmentation of Japanese Noun Phrases*
Yugo Murawaki and Sadao Kurohashi
- 16:45–17:10 *Discovering Morphological Paradigms from Plain Text Using a Dirichlet Process Mixture Model*
Markus Dreyer and Jason Eisner
- 17:10–17:35 *Multilayer Sequence Labeling*
Ai Azuma and Yuji Matsumoto
- 17:35–18:00 *A Bayesian Mixture Model for PoS Induction Using Multiple Features*
Christos Christodoulopoulos, Sharon Goldwater and Mark Steedman

Session 7B: Semantics

Chair: Mark Stevenson

- 16:20–16:45 *Large-Scale Noun Compound Interpretation Using Bootstrapping and the Web as a Corpus*
Su Nam Kim and Preslav Nakov
- 16:45–17:10 *Linguistic Redundancy in Twitter*
Fabio Massimo Zanzotto, Marco Pennacchiotti and Kostas Tsioutsoulis
- 17:10–17:35 *Divide and Conquer: Crowdsourcing the Creation of Cross-Lingual Textual Entailment Corpora*
Matteo Negri, Luisa Bentivogli, Yashar Mehdad, Danilo Giampiccolo and Alessandro Marchetti
- 17:35–18:00 *Literal and Metaphorical Sense Identification through Concrete and Abstract Context*
Peter Turney, Yair Neuman, Dan Assaf and Yohai Cohen

Thursday, July 28, 2011 (continued)

Session 7C: Spoken Language + IR

Chair: Steve Renals

- 16:20–16:45 *Syntactic Decision Tree LMs: Random Selection or Intelligent Design?*
Denis Filimonov and Mary Harper
- 16:45–17:10 *The Imagination of Crowds: Conversational AAC Language Modeling using Crowdsourcing and Large Data Sources*
Keith Vertanen and Per Ola Kristensson
- 17:10–17:35 *Using Syntactic and Semantic Structural Kernels for Classifying Definition Questions in Jeopardy!*
Alessandro Moschitti, Jennifer Chu-carroll, Siddharth Patwardhan, James Fan and Giuseppe Riccardi
- 17:35–18:00 *Multiword Expression Identification with Tree Substitution Grammars: A Parsing tour de force with French*
Spence Green, Marie-Catherine de Marneffe, John Bauer and Christopher D. Manning
- 18:30–21:30 Poster session and Reception
- Modelling Discourse Relations for Arabic*
Amal Al-Saif and Katja Markert
- Classifying Sentences as Speech Acts in Message Board Posts*
Ashequl Qadir and Ellen Riloff
- Learning Local Content Shift Detectors from Document-level Information*
Richard Farkas
- Collaborative Ranking: A Case Study on Entity Linking*
Zheng Chen and Heng Ji
- Robust Disambiguation of Named Entities in Text*
Johannes Hoffart, Mohamed Amir Yosef, Ilaria Bordino, Hagen Fürstenau, Manfred Pinkal, Marc Spaniol, Bilyana Taneva, Stefan Thater and Gerhard Weikum
- A Cascaded Classification Approach to Semantic Head Recognition*
Lukas Michelbacher, Alok Kothari, Martin Forst, Christina Lioma and Hinrich Schütze
- Linking Entities to a Knowledge Base with Query Expansion*
Swapna Gottipati and Jing Jiang

Thursday, July 28, 2011 (continued)

Unsupervised Information Extraction with Distributional Prior Knowledge

Cane Wing-ki Leung, Jing Jiang, Kian Ming A. Chai, Hai Leong Chieu and Loo-Nin Teow

Relation Acquisition using Word Classes and Partial Patterns

Stijn De Saeger, Kentaro Torisawa, Masaaki Tsuchida, Jun'ichi Kazama, Chikara Hashimoto, Ichiro Yamada, Jong Hoon Oh, Istvan Varga and Yulan Yan

Identification of Multi-word Expressions by Combining Multiple Linguistic Information Sources

Yulia Tsvetkov and Shuly Wintner

Analyzing Methods for Improving Precision of Pivot Based Bilingual Dictionaries

Xabier Saralegi, Iker Manterola and Iñaki San Vicente

Soft Dependency Constraints for Reordering in Hierarchical Phrase-Based Translation

Yang Gao, Philipp Koehn and Alexandra Birch

Statistical Machine Translation with Local Language Models

Christof Monz

Fast Generation of Translation Forest for Large-Scale SMT Discriminative Training

Xinyan Xiao, Yang Liu, Qun Liu and Shouxun Lin

A Correction Model for Word Alignments

J. Scott McCarley, Abraham Ittycheriah, Salim Roukos, Bing Xiang and Jian-ming Xu

Heuristic Search for Non-Bottom-Up Tree Structure Prediction

Andrea Gesmundo and James Henderson

Cache-based Document-level Statistical Machine Translation

Zhengxian Gong, Min Zhang and Guodong Zhou

Minimum Imputed-Risk: Unsupervised Discriminative Training for Machine Translation

Zhifei Li, Ziyuan Wang, Jason Eisner, Sanjeev Khudanpur and Brian Roark

Improving Bilingual Projections via Sparse Covariance Matrices

Jagadeesh Jagarlamudi, Raghavendra Udupa, Hal Daume III and Abhijit Bhole

Thursday, July 28, 2011 (continued)

Entire Relaxation Path for Maximum Entropy Problems

Moshe Dubiner and Yoram Singer

A Non-negative Matrix Factorization Based Approach for Active Dual Supervision from Document and Word Labels

Chao Shen and Tao Li

Splitting Noun Compounds via Monolingual and Bilingual Paraphrasing: A Study on Japanese Katakana Words

Nobuhiro Kaji and Masaru Kitsuregawa

Enhancing Chinese Word Segmentation Using Unlabeled Data

Weiwei Sun and Jia Xu

Unsupervised Learning of Selectional Restrictions and Detection of Argument Coercions

Kirk Roberts and Sanda Harabagiu

Harnessing different knowledge sources to measure semantic relatedness under a uniform model

Ziqi Zhang, Anna Lisa Gentile and Fabio Ciravegna

Refining the Notions of Depth and Density in WordNet-based Semantic Similarity Measures

Tong Wang and Graeme Hirst

Latent Vector Weighting for Word Meaning in Context

Tim Van de Cruys, Thierry Poibeau and Anna Korhonen

Hierarchical Verb Clustering Using Graph Factorization

Lin Sun and Anna Korhonen

Structured Lexical Similarity via Convolution Kernels on Dependency Trees

Danilo Croce, Alessandro Moschitti and Roberto Basili

Probabilistic models of similarity in syntactic context

Diarmuid Ó Séaghdha and Anna Korhonen

Lexical Co-occurrence, Statistical Significance, and Word Association

Dipak L. Chaudhari, Om P. Damani and Srivatsan Laxman

Thursday, July 28, 2011 (continued)

Learning the Information Status of Noun Phrases in Spoken Dialogues
Altaf Rahman and Vincent Ng

Harnessing WordNet Senses for Supervised Sentiment Classification
Balamurali AR, Aditya Joshi and Pushpak Bhattacharyya

Learning General Connotation of Words using Graph-based Algorithms
Song Feng, Ritwik Bose and Yejin Choi

Hypotheses Selection Criteria in a Reranking Framework for Spoken Language Understanding
Marco Dinarelli and Sophie Rosset

A Fast Re-scoring Strategy to Capture Long-Distance Dependencies
Anoop Deoras, Tomas Mikolov and Kenneth Church

Efficient Subsampling for Training Complex Language Models
Puyang Xu, Asela Gunawardana and Sanjeev Khudanpur

Generating Aspect-oriented Multi-Document Summarization with Event-aspect model
Peng Li, Yinglin Wang, Wei Gao and Jing Jiang

Syntax-Based Grammaticality Improvement using CCG and Guided Search
Yue Zhang and Stephen Clark

Generating Subsequent Reference in Shared Visual Scenes: Computation vs Re-Use
Jette Viethen, Robert Dale and Markus Guhe

Learning Sentential Paraphrases from Bilingual Parallel Corpora for Text-to-Text Generation
Juri Ganitkevitch, Chris Callison-Burch, Courtney Napoles and Benjamin Van Durme

Joint Models for Chinese POS Tagging and Dependency Parsing
Zhenghua Li, Min Zhang, Wanxiang Che, Ting Liu, Wenliang Chen and Haizhou Li

Relaxed Cross-lingual Projection of Constituent Syntax
Wenbin Jiang, Qun Liu and Yajuan Lv

Thursday, July 28, 2011 (continued)

Computing Logical Form on Regulatory Texts

Nikhil Dinesh, Aravind Joshi and Insup Lee

Computation of Infix Probabilities for Probabilistic Context-Free Grammars

Mark-Jan Nederhof and Giorgio Satta

Parse Correction with Specialized Models for Difficult Attachment Types

Enrique Henestroza Anguiano and Marie Candito

Exact Inference for Generative Probabilistic Non-Projective Dependency Parsing

Shay B. Cohen, Carlos Gómez-Rodríguez and Giorgio Satta

Semi-supervised CCG Lexicon Extension

Emily Thomforde and Mark Steedman

A Fast, Accurate, Non-Projective, Semantically-Enriched Parser

Stephen Tratz and Eduard Hovy

Lateen EM: Unsupervised Training with Multiple Objectives, Applied to Dependency Grammar Induction

Valentin I. Spitkovsky, Hiyan Alshawi and Daniel Jurafsky

Unsupervised Dependency Parsing without Gold Part-of-Speech Tags

Valentin I. Spitkovsky, Hiyan Alshawi, Angel X. Chang and Daniel Jurafsky

Exploiting Syntactic and Distributional Information for Spelling Correction with Web-Scale N-gram Models

Wei Xu, Joel Tetreault, Martin Chodorow, Ralph Grishman and Le Zhao

Discriminating Gender on Twitter

John D. Burger, John Henderson, George Kim and Guido Zarrella

Identifying and Following Expert Investors in Stock Microblogs

Roy Bar-Haim, Elad Dinur, Ronen Feldman, Moshe Fresko and Guy Goldstein

Friday, July 29, 2011

Session 8: Plenary session

Chair: Shuly Wintner

- 9:05–9:30 *Unsupervised Semantic Role Induction with Graph Partitioning*
Joel Lang and Mirella Lapata
- 9:30–9:55 *Structural Opinion Mining for Graph-based Sentiment Representation*
Yuanbin Wu, Qi Zhang, Xuanjing Huang and Lide Wu
- 9:55–10:20 *Summarize What You Are Interested In: An Optimization Framework for Interactive Personalized Summarization*
Rui Yan, Jian-Yun Nie and Xiaoming Li

10:20–10:50 Coffee break

Session 9A: Machine Translation

Chair: John DeNero

- 10:50–11:15 *Tuning as Ranking*
Mark Hopkins and Jonathan May
- 11:15–11:40 *Watermarking the Outputs of Structured Prediction with an application in Statistical Machine Translation.*
Ashish Venugopal, Jakob Uszkoreit, David Talbot, Franz Och and Juri Ganitkevitch
- 11:40–12:05 *Hierarchical Phrase-based Translation Representations*
Gonzalo Iglesias, Cyril Allauzen, William Byrne, Adrià de Gispert and Michael Riley
- 12:05–12:30 *Improved Transliteration Mining Using Graph Reinforcement*
Ali El Kahki, Kareem Darwish, Ahmed Saad El Din, Mohamed Abd El-Wahab, Ahmed Hefny and Waleed Ammar

Friday, July 29, 2011 (continued)

Session 9B: Semantics

Chair: Peter Turney

- 10:50–11:15 *Experimental Support for a Categorical Compositional Distributional Model of Meaning*
Edward Grefenstette and Mehrnoosh Sadrzadeh
- 11:15–11:40 *Cross-Cutting Models of Lexical Semantics*
Joseph Reisinger and Raymond Mooney
- 11:40–12:05 *Reducing Grounded Learning Tasks To Grammatical Inference*
Benjamin Börschinger, Bevan K. Jones and Mark Johnson
- 12:05–12:30 *Relation Extraction with Relation Topics*
Chang Wang, James Fan, Aditya Kalyanpur and David Gondek

Session 9C: Information Extraction

Chair: Alessandro Moschitti

- 10:50–11:15 *Extreme Extraction – Machine Reading in a Week*
Marjorie Freedman, Lance Ramshaw, Elizabeth Boschee, Ryan Gabbard, Gary Kratkiewicz, Nicolas Ward and Ralph Weischedel
- 11:15–11:40 *Discovering Relations between Noun Categories*
Thahir Mohamed, Estevam Hruschka and Tom Mitchell
- 11:40–12:05 *Structured Relation Discovery using Generative Models*
Limin Yao, Aria Haghighi, Sebastian Riedel and Andrew McCallum
- 12:05–12:30 *Closing the Loop: Fast, Interactive Semi-Supervised Annotation With Queries on Features and Instances*
Burr Settles
- 12:30–13:00 SIGDAT Business Meeting
- 13:00–14:10 Lunch

Friday, July 29, 2011 (continued)

Session 10A: Syntax and Parsing

Chair: Mark Steedman

14:10–14:35 *Third-order Variational Reranking on Packed-Shared Dependency Forests*
Katsuhiko Hayashi, Taro Watanabe, Masayuki Asahara and Yuji Matsumoto

14:35–15:00 *Training dependency parsers by jointly optimizing multiple objectives*
Keith Hall, Ryan McDonald, Jason Katz-Brown and Michael Ringgaard

15:00–15:25 *Structured Sparsity in Structured Prediction*
Andre Martins, Noah Smith, Mario Figueiredo and Pedro Aguiar

15:25–15:50 *Lexical Generalization in CCG Grammar Induction for Semantic Parsing*
Tom Kwiatkowski, Luke Zettlemoyer, Sharon Goldwater and Mark Steedman

Session 10B: Information Extraction

Chair: Sebastian Riedel

14:10–14:35 *Named Entity Recognition in Tweets: An Experimental Study*
Alan Ritter, Sam Clark, Mausam and Oren Etzioni

14:35–15:00 *Identifying Relations for Open Information Extraction*
Anthony Fader, Stephen Soderland and Oren Etzioni

15:00–15:25 *Active Learning with Amazon Mechanical Turk*
Florian Laws, Christian Scheible and Hinrich Schütze

15:25–15:50 *Bootstrapped Named Entity Recognition for Product Attribute Extraction*
Duangmanee Putthividhya and Junling Hu

Friday, July 29, 2011 (continued)

Session 10C: Text Mining and NLP Applications

Chair: Alexandre Klementiev

14:10–14:35 *Twitter Catches The Flu: Detecting Influenza Epidemics using Twitter*

Eiji ARAMAKI, Sachiko MASKAWA and Mizuki MORITA

14:35–15:00 *A Simple Word Trigger Method for Social Tag Suggestion*

Zhiyuan Liu, Xinxiong Chen and Maosong Sun

15:00–15:25 *Rumor has it: Identifying Misinformation in Microblogs*

Vahed Qazvinian, Emily Rosengren, Dragomir R. Radev and Qiaozhu Mei

15:25–15:50 *Exploiting Parse Structures for Native Language Identification*

Sze-Meng Jojo Wong and Mark Dras

15:50–16:20 Coffee break

Plenary session – Best Paper Award + Closing

Chair: Mark Johnson

16:20–17:10 *A Probabilistic Forest-to-String Model for Language Generation from Typed Lambda Calculus Expressions*

Wei Lu and Hwee Tou Ng

