



ACL-HLT 2011

**Proceedings of the 49th Annual Meeting of the
Association for Computational Linguistics:
Human Language Technologies
Volume 2: Short Papers**



We wish to thank our sponsors

PLATINUM SPONSORS



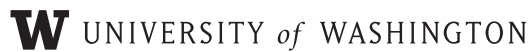
GOLD SPONSOR



SILVER SPONSORS



BRONZE SPONSORS



SUPPORTERS



ACL HLT 2011

**The 49th Annual Meeting of the
Association for Computational Linguistics:
Human Language Technologies**

Proceedings of the Conference

June 19-24, 2011
Portland, Oregon, USA

Production and Manufacturing by
Omnipress, Inc.
2600 Anderson Street
Madison, WI 53704 USA

©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 978-1-932432-88-6

Organizing Committee

General Chair

Dekang Lin, Google

Local Arrangements Chair

Brian Roark, Oregon Health & Science University

Program Co-Chairs

Yuji Matsumoto, Nara Institute of Science and Technology

Rada Mihalcea, University of North Texas

Local Arrangements Committee

Nate Bodenstab, Oregon Health & Science University

Aaron Dunlop, Oregon Health & Science University

Peter Heeman, Oregon Health & Science University

Meg Mitchell, Oregon Health & Science University

Christian Monson, Nuance

Zak Shafran, Oregon Health & Science University

Richard Sproat, Oregon Health & Science University

Masoud Rouhizadeh, Oregon Health & Science University

Mahsa Yarmohammadi, Oregon Health & Science University

Publications Chair

Guodong Zhou, Suzhou University

Sponsorship Chairs

Haifeng Wang, Baidu

Kevin Duh, National Inst. of Information and Communications Technology

Massimiliano Ciaramita, Google

Michael Gamon, Microsoft

Priscilla Rasmussen, Association for Computational Linguistics

Srinivas Bangalore, AT&T

Stephen Pulman, Oxford University

Tutorial Co-chairs

Patrick Pantel, Microsoft Research

Andy Way, Dublin City University

Workshop Co-chairs

Hal Daume III, University of Maryland
John Carroll, University of Sussex

Demo Chair

Sadao Kurohashi, Kyoto University

Mentoring

Chair

Tim Baldwin, University of Melbourne

Committee

Chris Biemann, TU Darmstadt

Mark Dras, Macquarie University

Jeremy Nicholson, University of Melbourne

Student Research Workshop

Student Co-chairs

Sasa Petrovic, University of Edinburgh

Emily Pitler, University of Pennsylvania

Ethan Selfridge, Oregon Health & Science University

Faculty Advisors

Miles Osborne, University of Edinburgh

Thamar Solorio, University of Alabama at Birmingham

ACL Conference Coordination Committee

Ido Dagan, Bar Ilan University (chair)

Chris Brew, Ohio State University

Graeme Hirst, University of Toronto

Lori Levin, Carnegie Mellon University

Christopher Manning, Stanford University

Dragomir Radev, University of Michigan

Owen Rambow, Columbia University

Priscilla Rasmussen, Association for Computational Linguistics

Suzanne Stevenson, University of Toronto

ACL Business Manager

Priscilla Rasmussen, Association for Computational Linguistics

Program Committee

Program Co-chairs

Yuji Matsumoto, Nara Institute of Science and Technology
Rada Mihalcea, University of North Texas

Area Chairs

Razvan Bunescu, Ohio University
Xavier Carreras, Technical University of Catalonia
Anna Feldman, Montclair University
Pascale Fung, Hong Kong University of Science and Technology
Chu-Ren Huang, Hong Kong Polytechnic University
Kentaro Inui, Tohoku University
Greg Kondrak, University of Alberta
Shankar Kumar, Google
Yang Liu, University of Texas at Dallas
Bernardo Magnini, Fondazione Bruno Kessler
Elliott Macklovitch, Marque d'Or
Katja Markert, University of Leeds
Lluís Marquez, Technical University of Catalonia
Diana McCarthy, Lexical Computing Ltd
Ryan McDonald, Google
Alessandro Moschitti, University of Trento
Vivi Nastase, Heidelberg Institute for Theoretical Studies
Manabu Okumura, Tokyo Institute of Technology
Vasile Rus, University of Memphis
Fabrizio Sebastiani, National Research Council of Italy
Michel Simard, National Research Council of Canada
Thamar Solorio, University of Alabama at Birmingham
Svetlana Stoyanchev, Open University
Carlo Strapparava, Fondazione Bruno Kessler
Dan Tufis, Romanian Academy of Artificial Intelligence
Xiaojun Wan, Peking University
Taro Watanabe, National Inst. of Information and Communications Technology
Alexander Yates, Temple University
Deniz Yuret, Koc University

Program Committee

Ahmed Abbasi, Eugene Agichtein, Eneko Agirre, Lars Ahrenberg, Gregory Aist, Enrique Alfonso, Laura Alonso i Alemany, Gianni Amati, Alina Andreevskaia, Ion Androutsopoulos, Abhishek Arun, Masayuki Asahara, Nicholas Asher, Giuseppe Attardi, Necip Fazil Ayan

Collin Baker, Jason Baldridge, Tim Baldwin, Krisztian Balog, Carmen Banea, Verginica Barbu

Mititelu, Marco Baroni, Regina Barzilay, Roberto Basili, John Bateman, Tilman Becker, Lee Becker, Beata Beigman-Klebanov, Cosmin Bejan, Ron Bekkerman, Daisuke Bekki, Kedar Bel-lare, Anja Belz, Sabine Bergler, Shane Bergsma, Raffaella Bernardi, Nicola Bertoldi, Pushpak Bhattacharyya, Archana Bhattarai, Tim Bickmore, Chris Biemann, Dan Bikel, Alexandra Birch, Maria Biryukov, Alan Black, Roi Blanco, John Blitzer, Phil Blunsom, Gemma Boleda, Francis Bond, Kalina Bontcheva, Johan Bos, Gosse Bouma, Kristy Boyer, S.R.K. Branavan, Thorsten Brants, Eric Breck, Ulf Brefeld, Chris Brew, Ted Briscoe, Samuel Brody

Michael Cafarella, Aoife Cahill, Chris Callison-Burch, Rafael Calvo, Nicoletta Calzolari, Nicola Cancedda, Claire Cardie, Giuseppe Carenini, Claudio Carpineto, Marine Carpuat, Xavier Car-reras, John Carroll, Ben Carterette, Francisco Casacuberta, Helena Caseli, Julio Castillo, Mauro Cettolo, Hakan Ceylan, Joyce Chai, Pi-Chuan Chang, Vinay Chaudhri, Berlin Chen, Ying Chen, Hsin-Hsi Chen, John Chen, Colin Cherry, David Chiang, Yejin Choi, Jennifer Chu-Carroll, Grace Chung, Kenneth Church, Massimiliano Ciaramita, Philipp Cimiano, Stephen Clark, Shay Co-hen, Trevor Cohn, Nigel Collier, Michael Collins, John Conroy, Paul Cook, Ann Copestake, Bonaventura Coppola, Fabrizio Costa, Koby Crammer, Dan Cristea, Montse Cuadros, Silviu-Petru Cucerzan, Aron Culotta, James Curran

Walter Daelemans, Robert Damper, Hoa Dang, Dipanjan Das, Hal Daume, Adria de Gispert, Marie-Catherine de Marneffe, Gerard de Melo, Maarten de Rijke, Vera Demberg, Steve DeNeefe, John DeNero, Pascal Denis, Ann Devitt, Giuseppe Di Fabrizio, Mona Diab, Markus Dickinson, Mike Dillinger, Bill Dolan, Doug Downey, Markus Dreyer, Greg Druck, Kevin Duh, Chris Dyer, Marc Dymetman

Markus Egg, Koji Eguchi, Andreas Eisele, Jacob Eisenstein, Jason Eisner, Michael Elhadad, Tomaz Erjavec, Katrin Erk, Hugo Escalante, Andrea Esuli

Hui Fang, Alex Chengyu Fang, Benoit Favre, Anna Feldman, Christiane Fellbaum, Donghui Feng, Raquel Fernandez, Nicola Ferro, Katja Filippova, Jenny Finkel, Seeger Fisher, Margaret Fleck, Dan Flickinger, Corina Forascu, Kate Forbes-Riley, Mikel L. Forcada, Eric Fosler-Lussier, Jennifer Foster, George Foster, Anette Frank, Alex Fraser, Dayne Freitag, Guohong Fu, Hagen Fuerstenau, Pascale Fung, Sadaoki Furui

Evgeniy Gabrilovich, Robert Gaizauskas, Michel Galley, Michael Gamon, Kuzman Ganchev, Jianfeng Gao, Claire Gardent, Thomas Gärtner, Albert Gatt, Dmitriy Genzel, Kallirroi Georgila, Carlo Geraci, Pablo Gervas, Shlomo Geva, Daniel Gildea, Alastair Gill, Dan Gillick, Jesus Gimenez, Kevin Gimpel, Roxana Girju, Claudio Giuliano, Amir Globerson, Yoav Goldberg, Sharon Goldwater, Carlos Gomez Rodriguez, Julio Gonzalo, Brigitte Grau, Stephan Greene, Ralph Grishman, Tunga Gungor, Zhou GuoDong, Iryna Gurevych, David Guthrie

Nizar Habash, Ben Hachey, Barry Haddow, Gholamreza Haffari, Aria Haghighi, Udo Hahn, Jan Hajic, Dilek Hakkani-Tür, Keith Hall, Jirka Hana, John Hansen, Sanda Harabagiu, Mark Hasegawa-Johnson, Koiti Hasida, Ahmed Hassan, Katsuhiko Hayashi, Ben He, Xiaodong He, Ulrich Heid, Michael Heilman, Ilana Heintz, Jeff Heinz, John Henderson, James Henderson, Iris

Hendrickx, Aurelie Herbelot, Erhard Hinrichs, Tsutomu Hirao, Julia Hirschberg, Graeme Hirst, Julia Hockenmaier, Tracy Holloway King, Bo-June (Paul) Hsu, Xuanjing Huang, Liang Huang, Jimmy Huang, Jian Huang, Chu-Ren Huang, Juan Huerta, Rebecca Hwa

Nancy Ide, Gonzalo Iglesias, Gabriel Infante-López, Diana Inkpen, Radu Ion, Elena Irimia, Pierre Isabelle, Mitsuru Ishizuka, Aminul Islam, Abe Ittycheriah, Tomoharu Iwata

Martin Jansche, Sittichai Jiampojarn, Jing Jiang, Valentin Jijkoun, Richard Johansson, Mark Johnson, Aravind Joshi

Nanda Kambhatla, Min-Yen Kan, Kyoko Kanzaki, Rohit Kate, Junichi Kazama, Bill Keller, Andre Kempe, Philipp Keohn, Fazel Keshtkar, Adam Kilgarriff, Jin-Dong Kim, Su Nam Kim, Brian Kingsbury, Katrin Kirchhoff, Ioannis Klapaftis, Dan Klein, Alexandre Klementiev, Kevin Knight, Rob Koeling, Oskar Kohonen, Alexander Kolcz, Alexander Koller, Kazunori Komatani, Terry Koo, Moshe Koppel, Valia Kordoni, Anna Korhonen, Andras Kornai, Zornitsa Kozareva, Lun-Wei Ku, Sandra Kuebler, Marco Kuhlmann, Roland Kuhn, Mikko Kurimo, Oren Kurland, Olivia Kwong

Krista Lagus, Philippe Langlais, Guy Lapalme, Mirella Lapata, Dominique Laurent, Alberto Lavelli, Matthew Lease, Gary Lee, Kiyong Lee, Els Lefever, Alessandro Lenci, James Lester, Gina-Anne Levow, Tao Li, Shoushan LI, Fangtao Li, Zhifei Li, Haizhou Li, Hang Li, Wenjie Li, Percy Liang, Chin-Yew Lin, Frank Lin, Mihai Lintean, Ken Litkowski, Diane Litman, Marina Litvak, Yang Liu, Bing Liu, Qun Liu, Jingjing Liu, Elena Lloret, Birte Loenneker-Rodman, Adam Lopez, Annie Louis, Xiaofei Lu, Yue Lu

Tengfei Ma, Wolfgang Macherey, Klaus Macherey, Elliott Macklovitch, Nitin Madnani, Bernardo Magnini, Suresh Manandhar, Gideon Mann, Chris Manning, Daniel Marcu, David Martínez, Andre Martins, Yuval Marton, Sameer Maskey, Spyros Matsoukas, Mausam, Arne Mauser, Jon May, David McAllester, Andrew McCallum, David McClosky, Ryan McDonald, Bridget McInnes, Tara McIntosh, Kathleen McKeown, Paul McNamee, Yashar Mehdad, Qiaozhu Mei, Arul Menezes, Paola Merlo, Donald Metzler, Adam Meyers, Haitao Mi, Jeff Mielke, Einat Minkov, Yusuke Miyao, Dunja Mladenic, Marie-Francine Moens, Saif Mohammad, Dan Moldovan, Diego Molla, Christian Monson, Manuel Montes y Gomez, Raymond Mooney, Robert Moore, Tatsunori Mori, Glyn Morrill, Sara Morrissey, Alessandro Moschitti, Jack Mostow, Smaranda Muresan, Gabriel Murray, Gabriele Musillo, Sung-Hyon Myaeng

Tetsuji Nakagawa, Mikio Nakano, Preslav Nakov, Ramesh Nallapati, Vivi Nastase, Borja Navarro-Colorado, Roberto Navigli, Mark-Jan Nederhof, Matteo Negri, Ani Nenkova, Graham Neubig, Guenter Neumann, Vincent Ng, Hwee Tou Ng, Patrick Nguyen, Jian-Yun Nie, Rodney Nielsen, Joakim Nivre, Tadashi Nomoto, Scott Nowson

Diarmuid Ó Séaghdha, Sharon O'Brien, Franz Och, Stephan Oepen, Kemal Oflazer, Jong-Hoon Oh, Constantin Orasan, Miles Osborne, Gozde Ozbek

Sebastian Pado, Tim Paek, Bo Pang, Patrick Pantel, Soo-Min Pantel, Ivandre Paraboni, Cecile Paris, Marius Pasca, Gabriella Pasi, Andrea Passerini, Rebecca J. Passonneau, Siddharth Patwardhan, Adam Pauls, Adam Pease, Ted Pedersen, Anselmo Penas, Anselmo Peñas, Jing Peng, Fuchun Peng, Gerald Penn, Marco Pennacchiotti, Wim Peters, Slav Petrov, Emanuele Pianta, Michael Picheny, Daniele Pighin, Manfred Pinkal, David Pinto, Stelios Piperidis, Paul Piwek, Benjamin Piwowarski, Massimo Poesio, Livia Polanyi, Simone Paolo Ponzetto, Hoi-fung Poon, Ana-Maria Popescu, Andrei Popescu-Belis, Maja Popovic, Martin Potthast, Richard Power, Sameer Pradhan, John Prager, Rashmi Prasad, Partha Pratim Talukdar, Adam Przepiórkowski, Vasin Punyakanok, Matthew Purver, Sampo Pyysalo

Silvia Quarteroni, Ariadna Quattoni, Chris Quirk

Stephan Raaijmakers, Dragomir Radev, Filip Radlinski, Bhuvana Ramabhadran, Ganesh Ramakrishnan, Owen Rambow, Aarne Ranta, Delip Rao, Ari Rappoport, Lev Ratinov, Antoine Raux, Emmanuel Rayner, Roi Reichart, Ehud Reiter, Steve Renals, Philip Resnik, Giuseppe Riccardi, Sebastian Riedel, Stefan Riezler, German Rigau, Ellen Riloff, Laura Rimell, Eric Ringger, Horacio Rodríguez, Paolo Rosso, Antti-Veikko Rosti, Rachel Edita Roxas, Alex Rudnicky, Marta Ruiz Costa-Jussa, Vasile Rus, Graham Russell, Anton Rytting

Rune Sætre, Kenji Sagae, Horacio Saggion, Tapio Salakoski, Agnes Sandor, Sudeshna Sarkar, Anoop Sarkar, Giorgio Satta, Hassan Sawaf, Frank Schilder, Anne Schiller, David Schlangen, Sabine Schulte im Walde, Tanja Schultz, Holger Schwenk, Donia Scott, Yohei Seki, Satoshi Sekine, Stephanie Seneff, Jean Senellart, Violeta Seretan, Burr Settles, Serge Sharoff, Dou Shen, Wade Shen, Libin Shen, Kiyooki Shirai, Luo Si, Grigori Sidorov, Mário Silva, Fabrizio Silvestri, Khalil Simaan, Michel Simard, Gabriel Skantze, Noah Smith, Matthew Snover, Rion Snow, Benjamin Snyder, Stephen Soderland, Marina Sokolova, Tamar Solorio, Swapna Somasundaran, Lucia Specia, Valentin Spitkovsky, Richard Sproat, Manfred Stede, Mark Steedman, Amanda Stent, Mark Stevenson, Svetlana Stoyanchev, Veselin Stoyanov, Michael Strube, Sara Stymne, Keh-Yih Su, Fangzhong Su, Jian Su, L Venkata Subramaniam, David Suendermann, Maosong Sun, Mihai Surdeanu, Richard Sutcliffe, Charles Sutton, Jun Suzuki, Stan Szpakowicz, Idan Szpektor

Hiroya Takamura, David Talbot, Irina Temnikova, Michael Tepper, Simone Teufel, Stefan Thater, Allan Third, Jörg Tiedemann, Christoph Tillmann, Ivan Titov, Takenobu Tokunaga, Kentaro Torisawa, Kristina Toutanova, Isabel Trancoso, Richard Tsai, Vivian Tsang, Dan Tufis

Takehito Utsuro

Shivakumar Vaithyanathan, Alessandro Valitutti, Antal van den Bosch, Hans van Halteren, Gertjan van Noord, Lucy Vanderwende, Vasudeva Varma, Tony Veale, Olga Vechtomova, Paola Veldardi, Rene Venegas, Ashish Venugopal, Jose Luis Vicedo, Evelyne Viegas, David Vilar, Begona Villada Moiron, Sami Virpioja, Andreas Vlachos, Stephan Vogel, Piek Vossen

Michael Walsh, Xiaojun Wan, Xinglong Wang, Wei Wang, Haifeng Wang, Justin Washtell, Andy

Way, David Weir, Ben Wellner, Ji-Rong Wen, Chris Wendt, Michael White, Ryen White, Richard Wicentowski, Jan Wiebe, Sandra Williams, Jason Williams, Theresa Wilson, Shuly Wintner, Kam-Fai Wong, Fei Wu

Deyi Xiong, Peng Xu, Jinxi Xu, Nianwen Xue

Scott Wen-tau Yih, Emine Yilmaz

David Zajic, Fabio Zanzotto, Richard Zens, Torsten Zesch, Hao Zhang, Bing Zhang, Min Zhang, Huarui Zhang, Jun Zhao, Bing Zhao, Jing Zheng, Li Hai Zhou, Michael Zock, Andreas Zollmann, Geoffrey Zweig, Pierre Zweigenbaum

Secondary Reviewers

Omri Abend, Rodrigo Agerri, Paolo Annesi, Wilker Aziz, Tyler Baldwin, Verginica Barbu Mititelu, David Batista, Delphine Bernhard, Stephen Boxwell, Janez Brank, Chris Brockett, Tim Buckwalter, Wang Bukang, Alicia Burga, Steven Burrows, Silvia Calegari, Marie Candito, Marina Cardenas, Bob Carpenter, Paula Carvalho, Diego Ceccarelli, Asli Celikyilmaz, Soumaya Chaffar, Bin Chen, Danilo Croce, Daniel Dahlmeier, Hong-Jie Dai, Mariam Daoud, Steven DeNeefe, Leon Derczynski, Elina Desypri, Sobha Lalitha Devi, Gideon Dror, Loic Dugast, Eraldo Fernandes, Jody Foo, Kotaro Funakoshi, Jing Gao, Wei Gao, Diman Ghazi, Julius Goth, Joseph Grafsgaard, Eun Young Ha, Robbie Haertel, Matthias Hagen, Enrique Henestroza, Hieu Hoang, Maria Holmqvist, Dennis Hoppe, Yunhua Hu, Yun Huang, Radu Ion, Elena Irimia, Jagadeesh Jagarlamudi, Antonio Juárez-González, Sun Jun, Evangelos Kanoulas, Aaron Kaplan, Caroline Lavecchia, Lianhau Lee, Michael Levit, Ping Li, Thomas Lin, Wang Ling, Ying Liu, José David Lopes, Bin Lu, Jia Lu, Saab Mansour, Raquel Martinez-Unanue, Haitao Mi, Simon Mille, Teruhisa Misu, Behrang Mohit, Sílvio Moreira, Rutu Mulkar-Mehta, Jason Naradowsky, Sudip Naskar, Heung-Seon Oh, You Ouyang, Lluís Padró, Sujith Ravi, Marta Recasens, Luz Rello, Stefan Rigo, Alan Ritter, Alvaro Rodrigo, Hasim Sak, Kevin Seppi, Aliaksei Severyn, Chao Shen, Shuming Shi, Laurianne Sitbon, Jun Sun, György Szarvas, Eric Tang, Alberto Téllez-Valero, Luong Minh Thang, Gabriele Tolomei, David Tomás, Diana Trandabat, Zhaopeng Tu, Gokhan Tur, Kateryna Tymoshenko, Fabienne Venant, Esaú Villatoro-Tello, Joachim Wagner, Dan Walker, Wei Wei, Xinyan Xiao, Jun Xie, Hao Xiong, Gu Xu, Jun Xu, Huichao Xue, Taras Zagibalov, Beñat Zapirain, Kalliopi Zervanou, Renxian Zhang, Daqi Zheng, Arkaitz Zubiaga

Table of Contents

<i>Lexicographic Semirings for Exact Automata Encoding of Sequence Models</i> Brian Roark, Richard Sproat and Izhak Shafran	1
<i>Good Seed Makes a Good Crop: Accelerating Active Learning Using Language Modeling</i> Dmitriy Dligach and Martha Palmer	6
<i>Temporal Restricted Boltzmann Machines for Dependency Parsing</i> Nikhil Garg and James Henderson	11
<i>Efficient Online Locality Sensitive Hashing via Reservoir Counting</i> Benjamin Van Durme and Ashwin Lall	18
<i>An Empirical Investigation of Discounting in Cross-Domain Language Models</i> Greg Durrett and Dan Klein	24
<i>HITS-based Seed Selection and Stop List Construction for Bootstrapping</i> Tetsuo Kiso, Masashi Shimbo, Mamoru Komachi and Yuji Matsumoto	30
<i>The Arabic Online Commentary Dataset: an Annotated Dataset of Informal Arabic with High Dialectal Content</i> Omar F. Zaidan and Chris Callison-Burch	37
<i>Part-of-Speech Tagging for Twitter: Annotation, Features, and Experiments</i> Kevin Gimpel, Nathan Schneider, Brendan O’Connor, Dipanjan Das, Daniel Mills, Jacob Eisenstein, Michael Heilman, Dani Yogatama, Jeffrey Flanigan and Noah A. Smith	42
<i>Semi-supervised condensed nearest neighbor for part-of-speech tagging</i> Anders Søgaard	48
<i>Latent Class Transliteration based on Source Language Origin</i> Masato Hagiwara and Satoshi Sekine	53
<i>Tier-based Strictly Local Constraints for Phonology</i> Jeffrey Heinz, Chetan Rawal and Herbert G. Tanner	58
<i>Lost in Translation: Authorship Attribution using Frame Semantics</i> Steffen Hedegaard and Jakob Grue Simonsen	65
<i>Insertion, Deletion, or Substitution? Normalizing Text Messages without Pre-categorization nor Supervision</i> Fei Liu, Fuliang Weng, Bingqing Wang and Yang Liu	71
<i>Unsupervised Discovery of Rhyme Schemes</i> Sravana Reddy and Kevin Knight	77

<i>Language of Vandalism: Improving Wikipedia Vandalism Detection via Stylometric Analysis</i> Manoj Harpalani, Michael Hart, Sandesh Singh, Rob Johnson and Yejin Choi	83
<i>That’s What She Said: Double Entendre Identification</i> Chloe Kiddon and Yuriy Brun	89
<i>Joint Identification and Segmentation of Domain-Specific Dialogue Acts for Conversational Dialogue Systems</i> Fabrizio Morbini and Kenji Sagae	95
<i>Extracting Opinion Expressions and Their Polarities – Exploration of Pipelines and Joint Models</i> Richard Johansson and Alessandro Moschitti	101
<i>Subjective Natural Language Problems: Motivations, Applications, Characterizations, and Implications</i> Cecilia Ovesdotter Alm	107
<i>Entrainment in Speech Preceding Backchannels.</i> Rivka Levitan, Agustin Gravano and Julia Hirschberg	113
<i>Question Detection in Spoken Conversations Using Textual Conversations</i> Anna Margolis and Mari Ostendorf	118
<i>Extending the Entity Grid with Entity-Specific Features</i> Micha Elsner and Eugene Charniak	125
<i>French TimeBank: An ISO-TimeML Annotated Reference Corpus</i> André Bittar, Pascal Amsili, Pascal Denis and Laurence Danlos	130
<i>Search in the Lost Sense of “Query”: Question Formulation in Web Search Queries and its Temporal Changes</i> Bo Pang and Ravi Kumar	135
<i>A Corpus of Scope-disambiguated English Text</i> Mehdi Manshadi, James Allen and Mary Swift	141
<i>From Bilingual Dictionaries to Interlingual Document Representations</i> Jagadeesh Jagarlamudi, Hal Daume III and Raghavendra Udupa	147
<i>AM-FM: A Semantic Framework for Translation Quality Assessment</i> Rafael E. Banchs and Haizhou Li	153
<i>Automatic Evaluation of Chinese Translation Output: Word-Level or Character-Level?</i> Maoxi Li, Chengqing Zong and Hwee Tou Ng	159
<i>How Much Can We Gain from Supervised Word Alignment?</i> Jinxi Xu and Jinying Chen	165
<i>Word Alignment via Submodular Maximization over Matroids</i> Hui Lin and Jeff Bilmes	170

<i>Better Hypothesis Testing for Statistical Machine Translation: Controlling for Optimizer Instability</i> Jonathan H. Clark, Chris Dyer, Alon Lavie and Noah A. Smith	176
<i>Bayesian Word Alignment for Statistical Machine Translation</i> Coskun Mermer and Murat Saraclar	182
<i>Transition-based Dependency Parsing with Rich Non-local Features</i> Yue Zhang and Joakim Nivre	188
<i>Reversible Stochastic Attribute-Value Grammars</i> Daniël de Kok, Barbara Plank and Gertjan van Noord	194
<i>Joint Training of Dependency Parsing Filters through Latent Support Vector Machines</i> Colin Cherry and Shane Bergsma	200
<i>Insertion Operator for Bayesian Tree Substitution Grammars</i> Hiroyuki Shindo, Akinori Fujino and Masaaki Nagata	206
<i>Language-Independent Parsing with Empty Elements</i> Shu Cai, David Chiang and Yoav Goldberg	212
<i>Judging Grammaticality with Tree Substitution Grammar Derivations</i> Matt Post	217
<i>Query Snowball: A Co-occurrence-based Approach to Multi-document Summarization for Question Answering</i> Hajime Morita, Tetsuya Sakai and Manabu Okumura	223
<i>Discrete vs. Continuous Rating Scales for Language Evaluation in NLP</i> Anja Belz and Eric Kow	230
<i>Semi-Supervised Modeling for Prenominal Modifier Ordering</i> Margaret Mitchell, Aaron Dunlop and Brian Roark	236
<i>Data-oriented Monologue-to-Dialogue Generation</i> Paul Piwek and Svetlana Stoyanchev	242
<i>Towards Style Transformation from Written-Style to Audio-Style</i> Amjad Abu-Jbara, Barbara Rosario and Kent Lyons	248
<i>Optimal and Syntactically-Informed Decoding for Monolingual Phrase-Based Alignment</i> Kapil Thadani and Kathleen McKeown	254
<i>Can Document Selection Help Semi-supervised Learning? A Case Study On Event Extraction</i> Shasha Liao and Ralph Grishman	260
<i>Relation Guided Bootstrapping of Semantic Lexicons</i> Tara McIntosh, Lars Yencken, James R. Curran and Timothy Baldwin	266

<i>Model-Portability Experiments for Textual Temporal Analysis</i>	
Oleksandr Kolomiyets, Steven Bethard and Marie-Francine Moens	271
<i>End-to-End Relation Extraction Using Distant Supervision from External Semantic Repositories</i>	
Truc Vien T. Nguyen and Alessandro Moschitti	277
<i>Automatic Extraction of Lexico-Syntactic Patterns for Detection of Negation and Speculation Scopes</i>	
Emilia Apostolova, Noriko Tomuro and Dina Demner-Fushman	283
<i>Coreference for Learning to Extract Relations: Yes Virginia, Coreference Matters</i>	
Ryan Gabbard, Marjorie Freedman and Ralph Weischedel	288
<i>Corpus Expansion for Statistical Machine Translation with Semantic Role Label Substitution Rules</i>	
Qin Gao and Stephan Vogel	294
<i>Scaling up Automatic Cross-Lingual Semantic Role Annotation</i>	
Lonneke van der Plas, Paola Merlo and James Henderson	299
<i>Towards Tracking Semantic Change by Visual Analytics</i>	
Christian Rohrdantz, Annette Hautli, Thomas Mayer, Miriam Butt, Daniel A. Keim and Frans Plank	305
<i>Improving Classification of Medical Assertions in Clinical Notes</i>	
Youngjun Kim, Ellen Riloff and Stéphane Meystre	311
<i>ParaSense or How to Use Parallel Corpora for Word Sense Disambiguation</i>	
Els Lefever, Véronique Hoste and Martine De Cock	317
<i>Models and Training for Unsupervised Preposition Sense Disambiguation</i>	
Dirk Hovy, Ashish Vaswani, Stephen Tratz, David Chiang and Eduard Hovy	323
<i>Types of Common-Sense Knowledge Needed for Recognizing Textual Entailment</i>	
Peter LoBue and Alexander Yates	329
<i>Modeling Wisdom of Crowds Using Latent Mixture of Discriminative Experts</i>	
Derya Ozkan and Louis-Philippe Morency	335
<i>Language Use: What can it tell us?</i>	
Marjorie Freedman, Alex Baron, Vasin Punyakanok and Ralph Weischedel	341
<i>Automatic Detection and Correction of Errors in Dependency Treebanks</i>	
Alexander Volokh and Günter Neumann	346
<i>Temporal Evaluation</i>	
Naushad UzZaman and James Allen	351
<i>A Corpus for Modeling Morpho-Syntactic Agreement in Arabic: Gender, Number and Rationality</i>	
Sarah Alkuhlani and Nizar Habash	357

<i>NULEX: An Open-License Broad Coverage Lexicon</i>	
Clifton McFate and Kenneth Forbus	363
<i>Even the Abstract have Color: Consensus in Word-Colour Associations</i>	
Saif Mohammad	368
<i>Detection of Agreement and Disagreement in Broadcast Conversations</i>	
Wen Wang, Sibel Yaman, Kristin Precoda, Colleen Richey and Geoffrey Raymond.....	374
<i>Dealing with Spurious Ambiguity in Learning ITG-based Word Alignment</i>	
Shujian Huang, Stephan Vogel and Jiajun Chen.....	379
<i>Clause Restructuring For SMT Not Absolutely Helpful</i>	
Susan Howlett and Mark Dras	384
<i>Improving On-line Handwritten Recognition using Translation Models in Multimodal Interactive Machine Translation</i>	
Vicent Alabau, Alberto Sanchis and Francisco Casacuberta	389
<i>Monolingual Alignment by Edit Rate Computation on Sentential Paraphrase Pairs</i>	
Houda Bouamor, Aurélien Max and Anne Vilnat	395
<i>Terminal-Aware Synchronous Binarization</i>	
Licheng Fang, Tagyoung Chung and Daniel Gildea	401
<i>Domain Adaptation for Machine Translation by Mining Unseen Words</i>	
Hal Daume III and Jagadeesh Jagarlamudi	407
<i>Issues Concerning Decoding with Synchronous Context-free Grammar</i>	
Tagyoung Chung, Licheng Fang and Daniel Gildea	413
<i>Improving Decoding Generalization for Tree-to-String Translation</i>	
Jingbo Zhu and Tong Xiao	418
<i>Discriminative Feature-Tied Mixture Modeling for Statistical Machine Translation</i>	
Bing Xiang and Abraham Ittycheriah.....	424
<i>Is Machine Translation Ripe for Cross-Lingual Sentiment Classification?</i>	
Kevin Duh, Akinori Fujino and Masaaki Nagata.....	429
<i>Reordering Constraint Based on Document-Level Context</i>	
Takashi Onishi, Masao Utiyama and Eiichiro Sumita.....	434
<i>Confidence-Weighted Learning of Factored Discriminative Language Models</i>	
Viet Ha Thuc and Nicola Cancedda	439
<i>On-line Language Model Biasing for Statistical Machine Translation</i>	
Sankaranarayanan Ananthakrishnan, Rohit Prasad and Prem Natarajan.....	445

<i>Reordering Modeling using Weighted Alignment Matrices</i>	
Wang Ling, Tiago Luís, João Graça, Isabel Trancoso and Luísa Coheur	450
<i>Two Easy Improvements to Lexical Weighting</i>	
David Chiang, Steve DeNeeffe and Michael Pust	455
<i>Why Initialization Matters for IBM Model 1: Multiple Optima and Non-Strict Convexity</i>	
Kristina Toutanova and Michel Galley	461
<i>“I Thou Thee, Thou Traitor”: Predicting Formal vs. Informal Address in English Literature</i>	
Manaal Faruqi and Sebastian Padó	467
<i>Clustering Comparable Corpora For Bilingual Lexicon Extraction</i>	
Bo Li, Eric Gaussier and Akiko Aizawa	473
<i>Identifying Word Translations from Comparable Corpora Using Latent Topic Models</i>	
Ivan Vulić, Wim De Smet and Marie-Francine Moens	479
<i>Why Press Backspace? Understanding User Input Behaviors in Chinese Pinyin Input Method</i>	
Yabin Zheng, Lixing Xie, Zhiyuan Liu, Maosong Sun, Yang Zhang and Liyun Ru	485
<i>Automatic Assessment of Coverage Quality in Intelligence Reports</i>	
Samuel Brody and Paul Kantor	491
<i>Putting it Simply: a Context-Aware Approach to Lexical Simplification</i>	
Or Biran, Samuel Brody and Noemie Elhadad	496
<i>Automatically Predicting Peer-Review Helpfulness</i>	
Wenting Xiong and Diane Litman	502
<i>They Can Help: Using Crowdsourcing to Improve the Evaluation of Grammatical Error Detection Systems</i>	
Nitin Madnani, Martin Chodorow, Joel Tetreault and Alla Rozovskaya	508
<i>Typed Graph Models for Learning Latent Attributes from Names</i>	
Delip Rao and David Yarowsky	514
<i>Interactive Group Suggesting for Twitter</i>	
Zhonghua Qu and Yang Liu	519
<i>Improved Modeling of Out-Of-Vocabulary Words Using Morphological Classes</i>	
Thomas Mueller and Hinrich Schuetze	524
<i>Pointwise Prediction for Robust, Adaptable Japanese Morphological Analysis</i>	
Graham Neubig, Yosuke Nakata and Shinsuke Mori	529
<i>Nonparametric Bayesian Machine Transliteration with Synchronous Adaptor Grammars</i>	
Yun Huang, Min Zhang and Chew Lim Tan	534

<i>Fully Unsupervised Word Segmentation with BVE and MDL</i>	
Daniel Hewlett and Paul Cohen	540
<i>An Empirical Evaluation of Data-Driven Paraphrase Generation Techniques</i>	
Donald Metzler, Eduard Hovy and Chunliang Zhang	546
<i>Identification of Domain-Specific Senses in a Machine-Readable Dictionary</i>	
Fumiyo Fukumoto and Yoshimi Suzuki	552
<i>A Probabilistic Modeling Framework for Lexical Entailment</i>	
Eyal Shnarch, Jacob Goldberger and Ido Dagan	558
<i>Liars and Saviors in a Sentiment Annotated Corpus of Comments to Political Debates</i>	
Paula Carvalho, Luís Sarmiento, Jorge Teixeira and Mário J. Silva	564
<i>Semi-supervised latent variable models for sentence-level sentiment analysis</i>	
Oscar Täckström and Ryan McDonald	569
<i>Identifying Noun Product Features that Imply Opinions</i>	
Lei Zhang and Bing Liu	575
<i>Identifying Sarcasm in Twitter: A Closer Look</i>	
Roberto González-Ibáñez, Smaranda Muresan and Nina Wacholder	581
<i>Subjectivity and Sentiment Analysis of Modern Standard Arabic</i>	
Muhammad Abdul-Mageed, Mona Diab and Mohammed Korayem	587
<i>Identifying the Semantic Orientation of Foreign Words</i>	
Ahmed Hassan, Amjad AbuJbara, Rahul Jha and Dragomir Radev	592
<i>Hierarchical Text Classification with Latent Concepts</i>	
Xipeng Qiu, Xuanjing Huang, Zhao Liu and Jinlong Zhou	598
<i>Semantic Information and Derivation Rules for Robust Dialogue Act Detection in a Spoken Dialogue System</i>	
Wei-Bin Liang, Chung-Hsien Wu and Chia-Ping Chen	603
<i>Predicting Relative Prominence in Noun-Noun Compounds</i>	
Taniya Mishra and Srinivas Bangalore	609
<i>Contrasting Multi-Lingual Prosodic Cues to Predict Verbal Feedback for Rapport</i>	
Siwei Wang and Gina-Anne Levow	614
<i>Generalized Interpolation in Decision Tree LM</i>	
Denis Filimonov and Mary Harper	620
<i>A Scalable Probabilistic Classifier for Language Modeling</i>	
Joel Lang	625

<i>Chinese sentence segmentation as comma classification</i>	
Nianwen Xue and Yaqin Yang	631
<i>Learning Condensed Feature Representations from Large Unsupervised Data Sets for Supervised Learning</i>	
Jun Suzuki, Hideki Isozaki and Masaaki Nagata	636
<i>Probabilistic Document Modeling for Syntax Removal in Text Summarization</i>	
William M. Darling and Fei Song	642
<i>Comparative News Summarization Using Linear Programming</i>	
Xiaojiang Huang, Xiaojun Wan and Jianguo Xiao	648
<i>Hierarchical Reinforcement Learning and Hidden Markov Models for Task-Oriented Natural Language Generation</i>	
Nina Dethlefs and Heriberto Cuayahuitl	654
<i>Does Size Matter – How Much Data is Required to Train a REG Algorithm?</i>	
Mariët Theune, Ruud Koolen, Emiel Krahmer and Sander Wubben	660
<i>Simple English Wikipedia: A New Text Simplification Task</i>	
William Coster and David Kauchak	665
<i>A Hierarchical Model of Web Summaries</i>	
Yves Petinot, Kathleen McKeown and Kapil Thadani	670
<i>Unary Constraints for Efficient Context-Free Parsing</i>	
Nathan Bodenstab, Kristy Hollingshead and Brian Roark	676
<i>Data point selection for cross-language adaptation of dependency parsers</i>	
Anders Søgaard	682
<i>Getting the Most out of Transition-based Dependency Parsing</i>	
Jinho D. Choi and Martha Palmer	687
<i>Using Derivation Trees for Treebank Error Detection</i>	
Seth Kulick, Ann Bies and Justin Mott	693
<i>Improving Dependency Parsing with Semantic Classes</i>	
Eneko Agirre, Kepa Bengoetxea, Koldo Gojenola and Joakim Nivre	699
<i>Joint Hebrew Segmentation and Parsing using a PCFGLA Lattice Parser</i>	
Yoav Goldberg and Michael Elhadad	704
<i>An Ensemble Model that Combines Syntactic and Semantic Clustering for Discriminative Dependency Parsing</i>	
Gholamreza Haffari, Marzieh Razavi and Anoop Sarkar	710
<i>Better Automatic Treebank Conversion Using A Feature-Based Approach</i>	
Muhua Zhu, Jingbo Zhu and Minghan Hu	715

<i>The Surprising Variance in Shortest-Derivation Parsing</i>	
Mohit Bansal and Dan Klein	720
<i>Entity Set Expansion using Topic information</i>	
Kugatsu Sadamitsu, Kuniko Saito, Kenji Imamura and Genichiro Kikui	726

Conference Program

Tuesday, June 21, 2011

Session 4-A: (9:00-10:30) Best Paper Session

Lexicographic Semirings for Exact Automata Encoding of Sequence Models

Brian Roark, Richard Sproat and Izhak Shafran

Session 5-A: (11:00-12:15) Machine Learning Methods

Good Seed Makes a Good Crop: Accelerating Active Learning Using Language Modeling

Dmitriy Dligach and Martha Palmer

Temporal Restricted Boltzmann Machines for Dependency Parsing

Nikhil Garg and James Henderson

Efficient Online Locality Sensitive Hashing via Reservoir Counting

Benjamin Van Durme and Ashwin Lall

An Empirical Investigation of Discounting in Cross-Domain Language Models

Greg Durrett and Dan Klein

HITS-based Seed Selection and Stop List Construction for Bootstrapping

Tetsuo Kiso, Masashi Shimbo, Mamoru Komachi and Yuji Matsumoto

Session 5-B: (11:00-12:15) Phonology/Morphology & POSTagging

The Arabic Online Commentary Dataset: an Annotated Dataset of Informal Arabic with High Dialectal Content

Omar F. Zaidan and Chris Callison-Burch

Part-of-Speech Tagging for Twitter: Annotation, Features, and Experiments

Kevin Gimpel, Nathan Schneider, Brendan O'Connor, Dipanjan Das, Daniel Mills, Jacob Eisenstein, Michael Heilman, Dani Yogatama, Jeffrey Flanigan and Noah A. Smith

Semi-supervised condensed nearest neighbor for part-of-speech tagging

Anders Søgaard

Latent Class Transliteration based on Source Language Origin

Masato Hagiwara and Satoshi Sekine

Tuesday, June 21, 2011 (continued)

Tier-based Strictly Local Constraints for Phonology

Jeffrey Heinz, Chetan Rawal and Herbert G. Tanner

Session 5-C: (11:00-12:15) Linguistic Creativity

Lost in Translation: Authorship Attribution using Frame Semantics

Steffen Hedegaard and Jakob Grue Simonsen

Insertion, Deletion, or Substitution? Normalizing Text Messages without Pre-categorization nor Supervision

Fei Liu, Fuliang Weng, Bingqing Wang and Yang Liu

Unsupervised Discovery of Rhyme Schemes

Sravana Reddy and Kevin Knight

Language of Vandalism: Improving Wikipedia Vandalism Detection via Stylometric Analysis

Manoj Harpalani, Michael Hart, Sandesh Signh, Rob Johnson and Yejin Choi

That's What She Said: Double Entendre Identification

Chloe Kiddon and Yuriy Brun

Session 5-D: (11:00-12:15) Opinion Analysis and Textual and Spoken Conversations

Joint Identification and Segmentation of Domain-Specific Dialogue Acts for Conversational Dialogue Systems

Fabrizio Morbini and Kenji Sagae

Extracting Opinion Expressions and Their Polarities – Exploration of Pipelines and Joint Models

Richard Johansson and Alessandro Moschitti

Subjective Natural Language Problems: Motivations, Applications, Characterizations, and Implications

Cecilia Ovesdotter Alm

Entrainment in Speech Preceding Backchannels.

Rivka Levitan, Agustin Gravano and Julia Hirschberg

Question Detection in Spoken Conversations Using Textual Conversations

Anna Margolis and Mari Ostendorf

Tuesday, June 21, 2011 (continued)

Session 5-E: (11:00-12:15) Corpus & Document Analysis

Extending the Entity Grid with Entity-Specific Features

Micha Elsner and Eugene Charniak

French TimeBank: An ISO-TimeML Annotated Reference Corpus

André Bittar, Pascal Amsili, Pascal Denis and Laurence Danlos

Search in the Lost Sense of “Query”: Question Formulation in Web Search Queries and its Temporal Changes

Bo Pang and Ravi Kumar

A Corpus of Scope-disambiguated English Text

Mehdi Manshadi, James Allen and Mary Swift

From Bilingual Dictionaries to Interlingual Document Representations

Jagadeesh Jagarlamudi, Hal Daume III and Raghavendra Udupa

(12:15 - 2:00) Lunch

Session 6-A: (2:00 - 3:30) Machine Translation

AM-FM: A Semantic Framework for Translation Quality Assessment

Rafael E. Banchs and Haizhou Li

Automatic Evaluation of Chinese Translation Output: Word-Level or Character-Level?

Maoxi Li, Chengqing Zong and Hwee Tou Ng

How Much Can We Gain from Supervised Word Alignment?

Jinxi Xu and Jinying Chen

Word Alignment via Submodular Maximization over Matroids

Hui Lin and Jeff Bilmes

Better Hypothesis Testing for Statistical Machine Translation: Controlling for Optimizer Instability

Jonathan H. Clark, Chris Dyer, Alon Lavie and Noah A. Smith

Tuesday, June 21, 2011 (continued)

Bayesian Word Alignment for Statistical Machine Translation

Coskun Mermer and Murat Saraclar

Session 6-B: (2:00 - 3:30) Syntax & Parsing

Transition-based Dependency Parsing with Rich Non-local Features

Yue Zhang and Joakim Nivre

Reversible Stochastic Attribute-Value Grammars

Daniël de Kok, Barbara Plank and Gertjan van Noord

Joint Training of Dependency Parsing Filters through Latent Support Vector Machines

Colin Cherry and Shane Bergsma

Insertion Operator for Bayesian Tree Substitution Grammars

Hiroyuki Shindo, Akinori Fujino and Masaaki Nagata

Language-Independent Parsing with Empty Elements

Shu Cai, David Chiang and Yoav Goldberg

Judging Grammaticality with Tree Substitution Grammar Derivations

Matt Post

Session 6-C: (2:00 - 3:30) Summarization & Generation

Query Snowball: A Co-occurrence-based Approach to Multi-document Summarization for Question Answering

Hajime Morita, Tetsuya Sakai and Manabu Okumura

Discrete vs. Continuous Rating Scales for Language Evaluation in NLP

Anja Belz and Eric Kow

Semi-Supervised Modeling for Prenominal Modifier Ordering

Margaret Mitchell, Aaron Dunlop and Brian Roark

Data-oriented Monologue-to-Dialogue Generation

Paul Piwek and Svetlana Stoyanchev

Tuesday, June 21, 2011 (continued)

Towards Style Transformation from Written-Style to Audio-Style

Amjad Abu-Jbara, Barbara Rosario and Kent Lyons

Optimal and Syntactically-Informed Decoding for Monolingual Phrase-Based Alignment

Kapil Thadani and Kathleen McKeown

Session 6-D: (2:00 - 3:30) Information Extraction

Can Document Selection Help Semi-supervised Learning? A Case Study On Event Extraction

Shasha Liao and Ralph Grishman

Relation Guided Bootstrapping of Semantic Lexicons

Tara McIntosh, Lars Yencken, James R. Curran and Timothy Baldwin

Model-Portability Experiments for Textual Temporal Analysis

Oleksandr Kolomiyets, Steven Bethard and Marie-Francine Moens

End-to-End Relation Extraction Using Distant Supervision from External Semantic Repositories

Truc Vien T. Nguyen and Alessandro Moschitti

Automatic Extraction of Lexico-Syntactic Patterns for Detection of Negation and Speculation Scopes

Emilia Apostolova, Noriko Tomuro and Dina Demner-Fushman

Coreference for Learning to Extract Relations: Yes Virginia, Coreference Matters

Ryan Gabbard, Marjorie Freedman and Ralph Weischedel

Tuesday, June 21, 2011 (continued)

Session 6-E: (2:00 - 3:30) Semantics

Corpus Expansion for Statistical Machine Translation with Semantic Role Label Substitution Rules

Qin Gao and Stephan Vogel

Scaling up Automatic Cross-Lingual Semantic Role Annotation

Lonneke van der Plas, Paola Merlo and James Henderson

Towards Tracking Semantic Change by Visual Analytics

Christian Rohrdantz, Annette Hautli, Thomas Mayer, Miriam Butt, Daniel A. Keim and Frans Plank

Improving Classification of Medical Assertions in Clinical Notes

Youngjun Kim, Ellen Riloff and Stéphane Meystre

ParaSense or How to Use Parallel Corpora for Word Sense Disambiguation

Els Lefever, Véronique Hoste and Martine De Cock

Models and Training for Unsupervised Preposition Sense Disambiguation

Dirk Hovy, Ashish Vaswani, Stephen Tratz, David Chiang and Eduard Hovy

Monday, June 20, 2011

(6:00-8:30) Poster Session (Short papers)

Types of Common-Sense Knowledge Needed for Recognizing Textual Entailment

Peter LoBue and Alexander Yates

Modeling Wisdom of Crowds Using Latent Mixture of Discriminative Experts

Derya Ozkan and Louis-Philippe Morency

Language Use: What can it tell us?

Marjorie Freedman, Alex Baron, Vasin Punyakanok and Ralph Weischedel

Automatic Detection and Correction of Errors in Dependency Treebanks

Alexander Volokh and Günter Neumann

Monday, June 20, 2011 (continued)

Temporal Evaluation

Naushad UzZaman and James Allen

A Corpus for Modeling Morpho-Syntactic Agreement in Arabic: Gender, Number and Rationality

Sarah Alkuhlani and Nizar Habash

NULEX: An Open-License Broad Coverage Lexicon

Clifton McFate and Kenneth Forbus

Even the Abstract have Color: Consensus in Word-Colour Associations

Saif Mohammad

Detection of Agreement and Disagreement in Broadcast Conversations

Wen Wang, Sibel Yaman, Kristin Precoda, Colleen Richey and Geoffrey Raymond

Dealing with Spurious Ambiguity in Learning ITG-based Word Alignment

Shujian Huang, Stephan Vogel and Jiajun Chen

Clause Restructuring For SMT Not Absolutely Helpful

Susan Howlett and Mark Dras

Improving On-line Handwritten Recognition using Translation Models in Multimodal Interactive Machine Translation

Vicent Alabau, Alberto Sanchis and Francisco Casacuberta

Monolingual Alignment by Edit Rate Computation on Sentential Paraphrase Pairs

Houda Bouamor, Aurélien Max and Anne Vilnat

Terminal-Aware Synchronous Binarization

Licheng Fang, Tagyoung Chung and Daniel Gildea

Domain Adaptation for Machine Translation by Mining Unseen Words

Hal Daume III and Jagadeesh Jagarlamudi

Issues Concerning Decoding with Synchronous Context-free Grammar

Tagyoung Chung, Licheng Fang and Daniel Gildea

Monday, June 20, 2011 (continued)

Improving Decoding Generalization for Tree-to-String Translation

Jingbo Zhu and Tong Xiao

Discriminative Feature-Tied Mixture Modeling for Statistical Machine Translation

Bing Xiang and Abraham Ittycheriah

Is Machine Translation Ripe for Cross-Lingual Sentiment Classification?

Kevin Duh, Akinori Fujino and Masaaki Nagata

Reordering Constraint Based on Document-Level Context

Takashi Onishi, Masao Utiyama and Eiichiro Sumita

Confidence-Weighted Learning of Factored Discriminative Language Models

Viet Ha Thuc and Nicola Cancedda

On-line Language Model Biasing for Statistical Machine Translation

Sankaranarayanan Ananthakrishnan, Rohit Prasad and Prem Natarajan

Reordering Modeling using Weighted Alignment Matrices

Wang Ling, Tiago Luís, João Graça, Isabel Trancoso and Luísa Coheur

Two Easy Improvements to Lexical Weighting

David Chiang, Steve DeNeefe and Michael Pust

Why Initialization Matters for IBM Model 1: Multiple Optima and Non-Strict Convexity

Kristina Toutanova and Michel Galley

“I Thou Thee, Thou Traitor” : Predicting Formal vs. Informal Address in English Literature

Manaal Faruqui and Sebastian Padó

Clustering Comparable Corpora For Bilingual Lexicon Extraction

Bo Li, Eric Gaussier and Akiko Aizawa

Identifying Word Translations from Comparable Corpora Using Latent Topic Models

Ivan Vulić, Wim De Smet and Marie-Francine Moens

Monday, June 20, 2011 (continued)

Why Press Backspace? Understanding User Input Behaviors in Chinese Pinyin Input Method

Yabin Zheng, Lixing Xie, Zhiyuan Liu, Maosong Sun, Yang Zhang and Liyun Ru

Automatic Assessment of Coverage Quality in Intelligence Reports

Samuel Brody and Paul Kantor

Putting it Simply: a Context-Aware Approach to Lexical Simplification

Or Biran, Samuel Brody and Noemie Elhadad

Automatically Predicting Peer-Review Helpfulness

Wenting Xiong and Diane Litman

They Can Help: Using Crowdsourcing to Improve the Evaluation of Grammatical Error Detection Systems

Nitin Madnani, Martin Chodorow, Joel Tetreault and Alla Rozovskaya

Typed Graph Models for Learning Latent Attributes from Names

Delip Rao and David Yarowsky

Interactive Group Suggesting for Twitter

Zhonghua Qu and Yang Liu

Improved Modeling of Out-Of-Vocabulary Words Using Morphological Classes

Thomas Mueller and Hinrich Schuetze

Pointwise Prediction for Robust, Adaptable Japanese Morphological Analysis

Graham Neubig, Yosuke Nakata and Shinsuke Mori

Nonparametric Bayesian Machine Transliteration with Synchronous Adaptor Grammars

Yun Huang, Min Zhang and Chew Lim Tan

Fully Unsupervised Word Segmentation with BVE and MDL

Daniel Hewlett and Paul Cohen

An Empirical Evaluation of Data-Driven Paraphrase Generation Techniques

Donald Metzler, Eduard Hovy and Chunliang Zhang

Monday, June 20, 2011 (continued)

Identification of Domain-Specific Senses in a Machine-Readable Dictionary

Fumiyo Fukumoto and Yoshimi Suzuki

A Probabilistic Modeling Framework for Lexical Entailment

Eyal Shnarch, Jacob Goldberger and Ido Dagan

Liars and Saviors in a Sentiment Annotated Corpus of Comments to Political Debates

Paula Carvalho, Luís Sarmiento, Jorge Teixeira and Mário J. Silva

Semi-supervised latent variable models for sentence-level sentiment analysis

Oscar Täckström and Ryan McDonald

Identifying Noun Product Features that Imply Opinions

Lei Zhang and Bing Liu

Identifying Sarcasm in Twitter: A Closer Look

Roberto González-Ibáñez, Smaranda Muresan and Nina Wacholder

Subjectivity and Sentiment Analysis of Modern Standard Arabic

Muhammad Abdul-Mageed, Mona Diab and Mohammed Korayem

Identifying the Semantic Orientation of Foreign Words

Ahmed Hassan, Amjad AbuJbara, Rahul Jha and Dragomir Radev

Hierarchical Text Classification with Latent Concepts

Xipeng Qiu, Xuanjing Huang, Zhao Liu and Jinlong Zhou

Semantic Information and Derivation Rules for Robust Dialogue Act Detection in a Spoken Dialogue System

Wei-Bin Liang, Chung-Hsien Wu and Chia-Ping Chen

Predicting Relative Prominence in Noun-Noun Compounds

Taniya Mishra and Srinivas Bangalore

Contrasting Multi-Lingual Prosodic Cues to Predict Verbal Feedback for Rapport

Siwei Wang and Gina-Anne Levow

Monday, June 20, 2011 (continued)

Generalized Interpolation in Decision Tree LM

Denis Filimonov and Mary Harper

A Scalable Probabilistic Classifier for Language Modeling

Joel Lang

Chinese sentence segmentation as comma classification

Nianwen Xue and Yaqin Yang

Learning Condensed Feature Representations from Large Unsupervised Data Sets for Supervised Learning

Jun Suzuki, Hideki Isozaki and Masaaki Nagata

Probabilistic Document Modeling for Syntax Removal in Text Summarization

William M. Darling and Fei Song

Comparative News Summarization Using Linear Programming

Xiaojiang Huang, Xiaojun Wan and Jianguo Xiao

Hierarchical Reinforcement Learning and Hidden Markov Models for Task-Oriented Natural Language Generation

Nina Dethlefs and Heriberto Cuayahuitl

Does Size Matter – How Much Data is Required to Train a REG Algorithm?

Mariët Theune, Ruud Koolen, Emiel Krahmer and Sander Wubben

Simple English Wikipedia: A New Text Simplification Task

William Coster and David Kauchak

A Hierarchical Model of Web Summaries

Yves Petinot, Kathleen McKeown and Kapil Thadani

Unary Constraints for Efficient Context-Free Parsing

Nathan Bodenshtab, Kristy Hollingshead and Brian Roark

Data point selection for cross-language adaptation of dependency parsers

Anders Søgaard

Monday, June 20, 2011 (continued)

Getting the Most out of Transition-based Dependency Parsing

Jinho D. Choi and Martha Palmer

Using Derivation Trees for Treebank Error Detection

Seth Kulick, Ann Bies and Justin Mott

Improving Dependency Parsing with Semantic Classes

Eneko Agirre, Kepa Bengoetxea, Koldo Gojenola and Joakim Nivre

Joint Hebrew Segmentation and Parsing using a PCFGLA Lattice Parser

Yoav Goldberg and Michael Elhadad

An Ensemble Model that Combines Syntactic and Semantic Clustering for Discriminative Dependency Parsing

Gholamreza Haffari, Marzieh Razavi and Anoop Sarkar

Better Automatic Treebank Conversion Using A Feature-Based Approach

Muhua Zhu, Jingbo Zhu and Minghan Hu

The Surprising Variance in Shortest-Derivation Parsing

Mohit Bansal and Dan Klein

Entity Set Expansion using Topic information

Kugatsu Sadamitsu, Kuniko Saito, Kenji Imamura and Genichiro Kikui