

COLING 2014

**The 25th International Conference
on Computational Linguistics**

Proceedings of COLING 2014: Technical Papers

August 23-29, 2014
Dublin, Ireland

Papers marked with a Creative Commons or other specific license statement are copyright the respective authors (or their employers).

ISBN 978-1-941643-26-6

Preface

This volume contains papers from the 25th International Conference on Computational Linguistics (Coling 2014) held in Dublin, Ireland. The conference is organized by the Centre for Global Intelligent Content (CNGL) and held at the Helix Conference Centre at Dublin City University (DCU) from 25–29 August 2014, under the auspices of the International Committee on Computational Linguistics (ICCL).

COLING is almost 50 years old, its first gathering having taken place in New York in 1965. It has been organized once every two years, initially in odd years and then in even years, after COLING 1976 in Ottawa. Throughout its long history, COLING’s aspiration to provide an amicable forum for participants with broad backgrounds to present and share their ideas remains the same. We believe that the inherent complexity of language is worthy of study from diverse perspectives and that COLING provides a venue for fruitful interdisciplinary interaction.

We accepted 217 papers (138 oral presentations and 79 poster presentations) from 685 effective submissions, having received 705 submissions in total. Regardless of the format of presentation, all of the accepted papers were allocated 12 pages in the proceedings.

The review process of a large conference such as COLING is always complex and occasionally encounters difficulties. The program committee has to cope with the challenges of selecting which papers to accept among a large quantity of high quality submissions. The task of choosing 217 papers from 685 strong submissions covering the ever broadening fields of computational linguistics was not an easy one.

To cope with the anticipated difficulties, we asked six senior colleagues to join the Scientific Advisory Board (SAB) and help us through all stages of reviewing papers. They are: Ralph Grishman (New York University, USA), Yuji Matsumoto (NAIST, Japan), Joakim Nivre (Uppsala Univ., Sweden), Michael Picheny (IBM TJ Watson Research Center, USA), Donia Scott (Univ. of Sussex, UK), and Chengqing Zong (CAS, China).

We had 20 thematic areas and each area was chaired by two or more area chairs. Thanks to over 800 responsive reviewers, the review process proceeded in a very smooth manner, and each paper was read at least by three reviewers. In some cases, papers and their reviews were carefully assessed by Area Co-Chairs, one of the SAB members and by us, in our roles as Program Committee Co-Chairs. We are extremely happy with the very strong set of papers that has been accepted for presentation at the conference. It is, however, with regret that we had no choice but to reject a large number of high quality papers, due to the sheer volume of submissions received.

We would like to thank the SAB members and the Program Committee Area Chairs for their dedicated and efficient review work, and our reviewers for their professionalism in delivering high quality reviews. We also thank the authors of all the papers for submitting their fruits of labour to COLING. Although we were only able to accept a small subset of the submitted papers, we do hope that all authors and reviewers have benefited from this process of indirect dialogue.

Last but not least, we would like to thank the people who made COLING 2014 and this volume possible. We thank General Chairs, Josef van Genabith (Universität des Saarlandes/DFKI) and Andy Way (CNGL, DCU), and the chairs of the Local Organizing Committee, Cara Green (CNGL, DCU) and John Judge (CNGL/NCLT, DCU), for their tireless work. We are especially grateful to the Publications Chairs, Joachim Wagner (CNGL, DCU), Liadh Kelly (CNGL, DCU) and Lorraine Goeriot (CNGL, DCU), for their hard work in preparing the proceedings.

Prof. Jan Hajic (Charles University, Czech Republic)

Prof. Junichi Tsujii (Microsoft Research, China)

COLING 2014 Program Committee Co-Chairs

July 8, 2014

Welcome from the General Chairs

We are very pleased indeed to welcome you all to COLING 2014, the 25th International Conference on Computational Linguistics. We are particularly proud that the ICCL selected Dublin City University (DCU) as the location of COLING 2014.

DCU and its National Centre for Language Technology (NCLT) have a long track record in NLP. Unlike India, the previous COLING host country, Ireland is a very small country. A unique feature of the Irish University landscape is that universities team up with industry partners and each other to pool expertise to form large research centres. DCU is a founding member of CNGL, the Centre for Global Intelligent Content. COLING 2014 is organised by DCU in partnership with the CNGL, and as General Chairs we are proud to represent both DCU and CNGL.

The conference is taking place at the Helix Conference Centre, a stunning building added to the DCU campus in 2002. DCU is a young, dynamic and ambitious university; since admitting its first students in 1980, DCU has grown in both student numbers and size and now occupies a 72-acre site in Glasnevin, just to the north of Dublin city centre. To date almost 50,000 students have graduated from DCU and are now playing significant roles in enterprise and business globally. Today in 2014, DCU delivers more than 200 programmes to over 12,000 students across its four faculties — Humanities and Social Sciences, Science and Health, Engineering and Computing and DCU Business School. DCU's excellence is recognised internationally and it is ranked among the top-50 young Universities worldwide (QS 'Top 50 under 50' 2013). In the last eight years, DCU has twice been named Sunday Times 'University of the Year'.

At the time of writing, the total number of people registered to attend COLING has exceeded 675. With delegates from 58 countries, COLING 2014 will witness a colourful diversity of language and culture, which is appropriate given that Dublin is known as the localisation capital of the world. Some evidence for this comes from our sponsors, to whom we are extremely grateful: Baidu, eBay, Microsoft, Symantec and Google.

We are very pleased with the programme that has been assembled for you, comprising of four days for the main conference with a total of 138 oral presentations, 79 posters and a special track with 28 demo presentations, two days of workshops and tutorials before the main conference, and other satellite workshops immediately after. 18 topical workshops with a sharp focus on issues of key interest today will be attended by about 191 delegates, and the 6 high-quality tutorials are sure to attract large crowds. Social events include a welcome reception on the evening of 24th August, the conference banquet in the Guinness Storehouse on 26th, and excursions to some beautiful places of interest on 27th.

When DCU was awarded COLING two years ago, our own personal situations were quite different. One of us was away working in the translation industry in the UK, while the other was leading the Science Foundation Ireland and Industry-funded CNGL research center. Over the past few months, we have changed countries, and jobs: Andy is back as Deputy Director of the CNGL's Centre for Intelligent Content, while Josef has moved to Saarbrücken to take up a Chair and a Scientific Directorship at DFKI.

While these changes were taking place, we both had the backing of a remarkable team. The organization of an event on the scale of COLING takes enormous energy, planning and commitment from a large number of individuals. We have assembled a large, competent team of volunteers who are available to assist you while you are here in Dublin. We are sure that all of you participating at COLING — at tutorials, workshops, or the main conference — will enjoy the time you spend here in Ireland, and will look back on the event as one of the most memorable that you attend. Finally, thanks to all of you for coming. We hope you all enjoy the conference, that you benefit from the excellent programme that has been assembled, and that you go away from here having made new friends.

Prof. Josef Van Genabith (Universität des Saarlandes/DFKI, Germany)

Prof. Andy Way (CNGL, DCU, Ireland)

Organisers

General Chairs

Prof. Andy Way, CNGL, Dublin City University, Ireland

Prof. Josef van Genabith, Universität des Saarlandes/DFKI, Germany

Programme Chairs

Prof. Junichi Tsujii, Microsoft Research, China

Prof. Jan Hajic, Charles University, Czech Republic

Workshops Chairs

Dr. Jennifer Foster, CNGL, Dublin City University, Ireland

Prof. Dan Gildea, University of Rochester, USA

Prof. Tim Baldwin, University of Melbourne, Australia

Publication Chairs

Dr. Joachim Wagner, CNGL, Dublin City University, Ireland

Dr. Liadh Kelly, CNGL, Dublin City University, Ireland

Dr. Lorraine Goeriot, CNGL, Dublin City University, Ireland

Publicity Chairs

Dr. Dorothy Kenny, SALIS, Dublin City University, Ireland

Prof. Seong-Bae Park, Kyungpook National University, Korea

Local Chairs

Dr. Cara Greene, CNGL, Dublin City University, Ireland

Dr. John Judge, CNGL / NCLT, Dublin City University, Ireland

Scientific Advisory Board

Prof. Joakim Nivre, Uppsala Univ., Sweden

Prof. Yuji Matsumoto, NAIIST, Japan

Dr. Michael Picheny, IBM TJ Watson Research Center, USA

Prof. Donia Scott, Univ. of Sussex, UK

Prof. Chengqing Zong, CAS, China

Prof. Ralph Grishman, New York University, USA

Tutorial Chairs

Prof. Qun Liu, CNGL, Dublin City University, Ireland

Prof. Fei Xia, University of Washington, USA

Demo Chairs

Dr. Lamia Tounsi, CNGL, Dublin City University, Ireland

Dr. Rafal Rak, NaCTeM, University of Manchester, UK

Sponsorship Chairs

Dr. Páraic Sheridan, CNGL, Dublin City University, Ireland

Prof. Hans Uszkoreit, DFKI, Germany

Dr. Huaping Zhang, Beijing Institute of Technology, China

Dr. Hiromi Nakaiwa, NTT, Japan

Programme Committee

Programme Chairs

Junichi Tsujii, Microsoft Research, China
Jan Hajic, Charles University, Czech Republic

Linguistic Issues in CL and NLP

Emily M. Bender, University of Washington, USA
Eva Hajicova, Charles University, Czech Republic
Igor Boguslavsky, Universidad Politecnica de Madrid, Spain

Machine Learning for CL and NLP

Jason Eisner, Johns Hopkins University, USA
Yoshimasa Tsuruoka, University of Tokyo, Japan

Cognitive Issues in CL and NLP

Philippe Blache, CNRS & Université d'Aix-Marseille, France
Ted Gibson, MIT, USA

Morphology, Word Segmentation, Tagging and Chunking

Reut Tsarfaty, Weizmann Institute of Science, Israel
Yue Zhang, Singapore University of Technology and Design, Singapore

Syntax, Grammar Induction, Syntactic and Semantic Parsing

Laura Kallmeyer, Heinrich-Heine-Universität, Germany
Ryan McDonald, Google, USA

Lexical Semantics and Ontologies

Chu-Ren Huang, Hong Kong Polytechnic University, Hong Kong
Alessandro Oltramari, Carnegie Mellon University, USA

Semantic Processing, Distributional Semantics and Compositional Semantics

Stephen Clark, University of Cambridge, UK
Alessandro Lenci, University of Pisa, Italy

Modeling of Discourse and Dialogue

Nicolas Asher, CNRS & Université Paul Sabatier, France
Marilyn Walker, University of California Santa Cruz, USA

Natural Language Generation and Summarisation

Albert Gatt, University of Malta, Malta
Advait Siddharthan, University of Aberdeen, UK

Paraphrasing and Textual Entailment

Ido Dagan, Bar Ilan University, Israel
Kentaro Inui, Tohoku University, Japan

Sentiment Analysis, Opinion Mining and Social Media

Rada Mihalcea, University of Michigan, USA
Bing Liu, University of Illinois at Chicago, USA

Information Retrieval and Question Answering

Gareth Jones, CNGL, Dublin City University, Ireland
Siddharth Patwardhan, IBM Research, USA

Information Extraction and Database Linking

Danushka Bollegala, University of Liverpool, UK

Seung-won Hwang, Postec, Korea

Applications

Srinivas Bangalore, AT&T Labs-Research, USA

Heyan Huang, Beijing Institute of Technology, China

Guillaume Jacquet, Joint Research Centre, Italy

Multimodal and Natural Language Interfaces and Dialog Systems

Kristiina Jokinen, University of Helsinki, Finland

David Traum, University of Southern California, USA

Speech Recognition, Text-To-Speech, Spoken Language Understanding

Nick Campbell, Trinity College Dublin, Ireland

Alex Potamianos, National Technical University Crete, Greece

Machine Translation

Phillip Koehn, University of Edinburgh, UK / Johns Hopkins University, USA

Chris Quirk, Microsoft Research, USA

Tiejun Zhao, Harbin Institute of Technology, China

Resources

Pushpak Bhattacharyya, IIT Bombay, India

Nicoletta Calzolari, ILC-CNR, Italy

Martha Palmer, University of Colorado, USA

Languages with less resources

Steven Bird, University of Melbourne, Australia

Mark Liberman, University of Pennsylvania, USA

Rajeev Sangal, IIT Banaras Hindu University, India

Koenraad De Smedt, University of Bergen, Norway

Software and Tools

Jesús Cardeñosa, Universidad Politecnica de Madrid, Spain

Jing-Shin Chang, National Chi Nan University, Taiwan

Reviewers

Adam Meyers, Adam Pease, Adam Przepiórkowski, Adeline Nazarenko, Adrian Iftene, Afra Alishahi, Afsaneh Fazly, AiTi Aw, Alan Akbik, Alan Ritter, Aleksandrs Berdicevskis, Alessandra Zarcone, Alessandro Moschitti, Alessandro Oltramari, Alessandro Valitutti, Alexander Allauzen, Alexander Clark, Alexander M. Rush, Alexandra Balahur, Alexandra Birch, Alexandre Bouchard, Alexandre Rademaker, Alexandros Potamianos, Alexey Bogdanov, Alexis Nasr, Alexis Palmer, Alice Oh, Aline Villavicencio, Aljoscha Burchardt, Alvaro Rodrigo, Amanda Stent, Amba Kulkarni, Amin Mantrach, Amir Razavi, Amitava Das, Amit Goyal, Ananthakrishnan Ramanathan, Anastasia Krithara, Anders Sjøgaard, Andrea Gesmundo, Andrea Schalley, Andrew Kehler, Ang Sun, Ani Nenkova, Anirban Dasgupta, Anna Korhonen, Annalina Caputo, Anna Rumshisky, Anna Stavrianou, Ann Bies, Annemarie Friedrich, Annie Zaenen, Anoop Sarkar, Anselmo Peñas, Antal van den Bosch, António Branco, Antonio Pareja-Lora, Antonio Toral, Arantza Diaz de Ilarraza, Aravind Joshi, Ariadna Quattoni, Arne Mauser, Aron Culotta, Arul Menezes, Arulmozi Selvaraj, Arvind Agarwal, Arzucan Özgür, Asli Celikyilmaz, Athanasios Katsamanis, Atsushi Fujii, Atsushi Fujita, Aurélie Herbelot, Avirup Sil, Barbara Di Eugenio, Barbara Hemforth, Barbara Plank, Barry Haddow, Beatrice Alex, Beatrice Daille, Benjamin Roth, Ben King, Benoit Crabbé,

Bernd Bohnet, Bill Byrne, Bill Keller, Bishan Yang, Bonan Min, Bonaventura Coppola, Borja Navarro, Bo Tan, Boyan Onyshkevych, Branimir Boguraev, Brian Roark, Brigitte Grau, Canasai Kruengkrai, Caner Deric, Carlos Busso, Carlos Gómez-Rodríguez, Carmen Banea, Carolina Gallardo, Caroline Brun, Caroline Hagege, Cecile Paris, Cem Akkaya, Cettolo Mauro, Changxing Wu, Chen Chen, Chengqing Zong, Chi-Ho Li, Chris Brew, Chris Callison-Burch, Chris Fox, Christian Buck, Christiane Fellbaum, Christian Scheible, Christina Lioma, Christina Sauper, Christof Monz, Christopher Mitchell, Christopher Potts, Christoph Lofi, Christoph Tillmann, Christos Christodoulopoulos, Chu-Ren Huang, Ciprian Chelba, Claire Bonial, Claire Gardent, Claudia Leacock, Colin Cherry, Conghui Zhu, Constantin Orasan, Corina Forascu, Cornelia Caragea, Costanza Navarretta, Cristina Bosco, Cyril Goutte, Daisuke Kawahara, Damianos Karakos, Dan Bikel, Dan Flickinger, Dan Garrette, Dan Goldwasser, Daniel Gildea, Daniel Gillick, Danilo Croce, Dan Lassiter, Dan Tufiş, Danushka Bollegala, Daoud Clarke, David Burkett, David Chiang, David Graff, David Hall, David Kauchak, David McClosky, David Reitter, David Schlangen, David Talbot, David Vilar, David Weir, Deepak Ramachandran, Degen Huang, Dekai Wu, Delip Rao, Derek F. Wong, Deyi Xiong, Diana Inkpen, Diana McCarthy, Diarmuid Ó Séaghdha, Diego De Cao, Diego Molla, Dietmar Zaefferer, Dipanjan Das, Dipti Sharma, Dirk Hovy, Djamel Seddah, Dominique Laurent, Dongdong Zhang, Donghong Ji, Donia Scott, Donna Byron, Dorothee Beermann, Doug Downey, Dwaipayan Roy, Edmundo Pavel Soriano Morales, Eduard Dragut, Edward Grefenstette, Egoitz Laparra, Ehud Reiter, Eiji Aramaki, Ekaterina Ovchinnikova, Ekaterina Shutova, Eleftherios Avramidis, Elena Cabrio, Elena Kozerenko, Elena Lloret, Elif Aktolga, Ellen Riloff, Ellen Voorhees, Emiel Krahmer, Emily Mower Provost, Emily Pitler, Endong Xun, Eric de la Clergerie, Erik Cambria, Eugenio Martinez-Camara, Euny-oung Ha, Eva Hajicova, Eva Maria Vecchi, Evelina Fedorenko, Fabio Celli, Fabio Massimo Zanzotto, Fabre Cécile, Fangtao Li, Farah Benamara, Federico Alberto Pozzi, Feifei Zhai, Fei Huang, Felix Hieber, Fernando Martínez-Santiago, Francis Bond, Francisco Casacuberta, Francisco Guzman, Francis Tyers, François Portet, François Yvon, Franco M. Luque, Frank Keller, Franz Och, Furu Wei, Gabriella Lapesa, Ganesh Ramakrishnan, Gary Geunbae Lee, Geli Fei, Gemma Boleda, George Foster, Georgiana Dinu, Gerald Penn, Gerard de Melo, Gerard Kempen, German Rigau, Gholamreza Haffari, Gianluca Leboni, Gina-Anne Levow, Giovanni Semeraro, Girish Jha, Giuseppe Attardi, Giuseppe Carenini, Giuseppe Di Fabrizio, Giuseppe Riccardi, Graeme Hirst, Graham Neubig, Grazyna Demenko, Gregor Leusch, Grzegorz Chrupała, Guangyou Zhou, Guenter Neumann, Guergana Savova, Guillermo Garrido, Guodong Zhou, Guo-Wei Bian, Guy Lapalme, Haifeng Wang, Hailong Cao, Haitao Mi, Haixun Wang, Hai Zhao, Hany Hassan, Hao Zhang, Harksoo Kim, Harold Somers, Helmut Horacek, Hema Raghavan, Heng Ji, Hidekazu Oiwa, Hideto Kazawa, Hikaru Yokono, Himanshu Sharma, Hinrich Schuetze, Hiroshi Kanayama, Hiroya Takamura, Hiroyuki Shindo, Hitoshi Nishikawa, Hoa Trang Dang, Holger Schwenk, Hongwei Ding, Houfeng Wang, Hristo Tanev, Hsin-Hsi Chen, Hua Wu, Hui Fang, Hwee Tou Ng, Hyuckchul Jung, Idan Szpektor, Ielka van der Sluis, Ilyas Cicekli, Ines Rehbein, Ioannis Korkontzelos, Irina Kobozeva, Iryna Gurevych, Ivan Habernal, Jaakko Vaeyrynen, Jagadeesh Jagarlamudi, James Fan, James Mayfield, Jan Niehues, Jan Šnajder, Jan van Santen, Janyce Wiebe, Jarmila Panevova, Jason D Williams, Jason Naradowsky, Jean-Gabriel Ganascia, Jean Senellart, Jeff Mitchell, Jennifer Chu-Carroll, Jennifer Culbertson, Jennifer Gillenwater, Jennifer Williams, Jens Edlund, Jeremy Reffin, Jerry Feldman, Jiajun Chen, Jiajun Zhang, Jianfeng Gao, Jian Su, Jianxing Yu, Jian Xu, Jill Burstein, Jim Blevins, Jingbo Zhu, Jing Jiang, Jingsong Su, Jinho D. Choi, Ji-Rong Wen, Joel Nothman, Joern Wuebker, Johan Bos, Johannes Hoffart, John Carroll, John Chen, John Hale, John Prager, Jonas Kuhn, Jonathan Ginzburg, Jongwuk Lee, Jordi Atserias Batalla, Jörg Tiedemann, Jose Angel Olivas, José B. Mariño, Josef Ruppenhofer, Josef van Genabith, Joseph Le Roux, Jose Pinto, Joyce Chai, Juergen Trouvain, Julien Ah-Pine, Julie Weeds, Jungi Kim, Jungyeul Park, Jun-Ping Ng, Jun Suzuki, Junyi Jessy Li,

Jun Zhao, Juri Ganitkevitch, Jyoti Pawar, Kaiqi Zhao, Kai-Wei Chang, Kai Zhao, Kalina Bontcheva, Kam-Fai Wong, Kang Liu, Kangqi Luo, Kareem Darwish, Kashif Shah, Kathy McKeown, Katja Markert, Katrin Erk, Kees van Deemter, Keh-Yih Su, Keikichi Hirose, Keisuke Sakaguchi, Keith Hall, Keith Vander Linden, Kemal Oflazer, Ken Church, Kenji Sagae, Kenny Zhu, Kentaro Torisawa, Kenton Murray, Kevin Cohen, Kevin Duh, Kevin Gimpel, Kevin Knight, Kevin Small, Key-Sun Choi, Khalil Sima'an, Khiet Truong, kim gerdes, Kiril Simov, Klaus Macherey, Koenraad De Smedt, Konstantin Zuev, Krasimir Angelov, Kristian Woodsend, Kristina Striegnitz, Kristy Boyer, Kumiko Tanaka-Ishii, Kuzman Ganchev, L. Alfonso Urena Lopez, Lane Schwartz, Laurent Besacier, Laurent Prévot, Laure Vieu, Lei Cui, Leila Kosseim, Lei Zhang, Lemaou Liu, Leonid Iomdin, Leo Wanner, Le Zhao, Liang-Chih Yu, Lidia Pivovarova, Lidia S. Chao, Li Dong, Liheng Xu, Likun Qiu, Limin Yao, Liu Zhanyi, Livia Polanyi, Lorenzo Ferrone, Louise McNally, Lucia Specia, Lucy Vanderwende, Luis Espinosa, Luke Zettlemoyer, Lun-Wei Ku, Lu Qin, Maayan Zhitomirsky-Geffet, Magnus Sahlgren, Maja Popović, Makoto Miwa, Malvina Nissim, Mamoru Komachi, Mandar Mitra, Manfred Pinkal, Manfred Stede, Marcello Federico, Marco Baroni, Marco Turchi, Margaret Mitchell, Margot Mieskes, Mária Gósy, Maria Liakata, Maria Pershina, Marie Candido, Marie-Catherine de Marneffe, Mariët Theune, Marine Carpuat, Mariona Taulé, Marjorie McShane, Marketa Lopatkova, Mark Fishel, Mark Granroth-Wilding, Mark-Jan Nederhof, Mark Liberman, Mark Sammons, Mark Steedman, Markus Dreyer, Marta R. Costa-jussà, Mary Ellen Foster, Masato Hagiwara, Massimo Poesio, Matthew Purver, Matthew Stone, Matthias Gallé, Matthias Huck, Maud Ehrmann, Md. Faisal Mahbub Chowdhury, Mengqiu Wang, Meng Zhang, Menno van Zaanen, Messina Enza, Michael Carl, Michael Elhadad, Michael Gamon, Michael Johnston, Michael Paul, Michael Picheny, Micha Elsner, Michael Strube, Michael White, Michael Wiegand, Michel Galley, Miguel Ballesteros, Mihael Arcan, Mikael Kågebäck, Mike Dillinger, Mikel Forcada, Mikel Iruskieta, Mike Thelwall, Miles Osborne, Milica Gasic, Ming Zhou, Minlie Huang, Min Zhang, Miriam Butt, Mohamed Maamouri, Mohamed Yahya, Mohit Bansal, Mona Diab, Monojit Choudhury, Mu Li, Muntsa Padró, Muyun Yang, Myungha Jang, Nadir Durrani, Nancy Ide, Naoaki Okazaki, Nathaniel Smith, Nathan Schneider, Ndapandula Nakashole, Nianwen Xue, Nikolaos Lagos, Niladri Chatterjee, Nitish Aggarwal, Nobuhiro Kaji, Noémie Elhadad, Nuria Gala, Oliver Ferschke, Oliver Niebuhr, Olivier Pietquin, Omri Abend, Ondrej Bojar, Oscar Täckström, Owen Rambow, Özlem Çetinoğlu, Paavo Alku, Pablo Duboue, Pablo Gervás, Pallika Kanani, Paola Merlo, Pascal Denis, Patrick Ehlen, Patrick Nguyen, Patrick Paroubek, Patrick Ruch, Patrik Lambert, Paul Buitelaar, Paul McNamee, Paul Meurer, Paul Piwek, Paul Soma, Paul Thompson, Pavel Braslavski, Peng Xu, Peter Ljunglöf, Peter Turney, Petya Osenova, Phil Blunsom, Philipp Cimiano, Philippe Muller, Philipp Koehn, Philip Wille, Phillippe Langlais, Piek Vossen, Pierpaolo Basile, Pierre Zweigenbaum, Prasanth Kolachina, Preslav Nakov, Pushpak Bhattacharyya, Qiang Zhou, Qin Gao, Qi Zhang, Qun Liu, Radu Florian, Raffaella Bernardi, Raghavendra Udupa, Rahul Jha, Ralph Grishman, Rami Al-Rfou, Raquel Fernandez, Raquel Hervas, Raquel Justo, Rashmi Prasad, Ravi Sinha, Reed Coke, Reid Swanson, Richard Johansson, Richard Socher, Richard Sutcliffe, Richard Tzong-Han Tsai, Richard Zens, Riyaz Bhatt, Roberto Basili, Roberto Zamparelli, Rodney Nielsen, Roi Reichart, Roland Kuhn, Roman Yangarber, Ronaldo Martins, Ron Artstein, Ronen Feldman, Roser Saurí, Roy Bar-Haim, Rui Xia, Ruli Manurung, Rune Sætre, Rutu Mulkar, Ryuichiro Higashinaka, Sabine Bergler, Sabine Schulte im Walde, Sadao Kurohashi, Saif Mohammad, Salah Ait-Mokhtar, Sameer Pradhan, Sameer Singh, Sameh Alansary, Sanda Harabagiu, Sandra Kübler, Sandra Williams, Sara Stymne, Saša Hasan, Sasha Blair-Goldensohn, Satoshi Sekine, Saurabh Kataria, Scott Martin, Sebastian Padó, Secondary Reviewer, Seong-Bae Park, Sergei Nirenburg, Seth Kulick, Seung-won Hwang, Shachar Mirkin, Shafiq Joty, Shalom Lappin, Shankar Kumar, Shan Wang, Shih-Hung Wu, Shikhar Kumar Sarma, Shou-de Lin, Shoushan Li, Shujie Liu, Shu-Kai Hsieh, Shuly Wintner, Siddharth Agrawal, Simone Teufel, Sinno J. Pan,

Slav Petrov, Smaranda Muresan, Sobha Lalitha Devi, Song Feng, Sophia Ananiadou, Srinivas Bangalore, Sriram Venkatapathy, Stefan Bott, Stefan Dumitrescu, Stefan Evert, Stefan L. Frank, Stefano Borgo, Stefano Faralli, Stefan Riezler, Stefan Scherer, Stefan Thater, Stephan Busemann, Stephanie Strassel, Stephan Oepen, Stephen Clark, Stergos Afantenos, Steve DeNeefe, Steven Bird, Steven Piantadosi, Sudeep Gandhe, Sudeshna Sarkar, Sung-Hyon Myaeng, Suresh Manandhar, Susanne Burger, Suzanne Stevenson, Sven Hartrumpf, Svetla Koeva, Svetlana Timoshenko, Swapna Somasundaran, Sylvain Pogodalla, Taesun Moon, Takenobu Tokunaga, Takuya Matsuzaki, Tamara Polajnar, Tara McIntosh, Taro Watanabe, Tatsuya Kawahara, Taylor Berg-Kirkpatrick, Tejaswini Deoskar, Teresa Herrmann, Terry Koo, Teruhisa Misu, Thierry Poibeau, Thomas Mueller, Tiejun Qian, Timothy Baldwin, Tim Van de Cruys, Ting Liu, Tom Kwiatkowski, Tommaso Caselli, Tomohide Shibata, Tomoyuki Kajiwara, Tong Xiao, Toni Marti, Tony Veale, Toshiaki Nakazawa, Toshiyuki Sadanobu, Tracy Holloway King, U Kang, Ulrich Germann, Ulrike Pado, Vahed Qazvinian, Valentin Spitzkovsky, Valeria de Paiva, Valia Kordoni, Vanni Zavarella, Vasudeva Varma, Verena Rieser, Verginica Mititelu, Verónica Pérez-Rosas, Veronique Hoste, Veselin Stoyanov, Victor Raskin, Vincent Ng, Virach Sormlertlamvanich, Vivek Kumar Rangarajan Sridhar, Viviana Mascardi, Vladimir Selegey, Wai Lam, Waleed Ammar, Wanxiang Che, Wayne Xin Zhao, Weiwei Sun, Wei Xu, Wei Zhang, Wenbin Jiang, Wen-Chih Peng, Wenjie Li, Wenliang Chen, Wen-Lian Hsu, Wenxuan Gao, William Jarrold, William Schuler, Wim Peters, Wlodek Zadrozny, Wolfgang Macherey, Wolfgang Maier, Wolf-Tilo Balke, Wray Buntine, Xavier Blanco Escoda, Xavier Carreras, Xavier Tannier, Xiaodan Zhu, Xiaodong He, Xiaodong Shi, Xiaoguang Hu, Xiaojun Wan, Xiaoqiang Luo, Xiaoyan Zhu, Xing Wang, Xin Wang, Xinyan Xiao, Xipeng Qiu, Xuanjing Huang, Xu Jian, Yael Netzer, Yajuan Lv, Yang Ding, Yanghua Xiao, Yang Liu, Yangqiu Song, Yanjun Ma, Yannick Versley, Yashar Mehdad, Yejin Choi, Yisong Yue, Yoav Artzi, Yoav Goldberg, Yoshinori Sagisaka, Young-In Song, Yue Lu, Yuhong Guo, Yujie Zhang, Yuji Matsumoto, Yukiko Nakano, Yulan He, Yunqing Xia, Yunyao Li, Yuval Marton, Yu Zhou, Zdenka Uresova, Željko Agić, Zhenghua Li, Zhiyuan Cai, Zhiyuan Chen, Zhongjun He, Zhongqiang Huang, Ziheng Lin, and Zornitsa Kozareva

Invited Speakers

Mary Harper — *Learning from 26 languages: Program Management and Science in the Babel Program*

Mary will give her invited talk on Monday August 25th.

Ted Gibson — *Language for Communication: Language as Rational Inference*

Ted will give his invited talk on Tuesday August 26th.

Qun Liu — *Annotation Adaptation and Language Adaptation in NLP*

Qun will give his invited talk on Thursday August 28th.

Martin Kay — *Does a Computational Linguist have to be a Linguist?*

Martin will give his invited talk on Friday August 29th.

After Dinner Speaker

Tony Veale — *Creative Twitterbots: Putting Words (and Wit) Into the Mouths of Bots*

Tony will give his talk after the conference dinner on Wednesday August 27th.

Table of Contents

<i>Learning from 26 Languages: Program Management and Science in the Babel Program</i> Mary Harper	1
<i>Unsupervised learning of rhetorical structure with un-topic models</i> Diarmuid Ó Séaghdha and Simone Teufel	2
<i>Cross-lingual Coreference Resolution of Pronouns</i> Michal Novak and Zdenek Zabokrtsky	14
<i>A Context-Aware NLP Approach For Noteworthiness Detection in Cellphone Conversations</i> Francesca Bonin, Jose San Pedro and Nuria Oliver	25
<i>Hierarchical Topical Segmentation with Affinity Propagation</i> Anna Kazantseva and Stan Szpakowicz	37
<i>Capturing Cultural Differences in Expressions of Intentions</i> Marc Tomlinson, David Bracewell and Wayne Krug	48
<i>Identification of Implicit Topics in Twitter Data Not Containing Explicit Search Queries</i> Suzi Park and Hyopil Shin	58
<i>Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts</i> Cicero dos Santos and Maira Gatti	69
<i>Joint Inference and Disambiguation of Implicit Sentiments via Implicature Constraints</i> Lingjia Deng, Janyce Wiebe and Yoonjung Choi	79
<i>Group Non-negative Matrix Factorization with Natural Categories for Question Retrieval in Community Question Answer Archives</i> Guangyou Zhou, Yubo Chen, Daojian Zeng and Jun Zhao	89
<i>Multi-Objective Search Results Clustering</i> Sudipta Acharya, Sriparna Saha, Jose G. Moreno and Gaël Dias	99
<i>Query-by-Example Image Retrieval using Visual Dependency Representations</i> Desmond Elliott, Victor Lavrenko and Frank Keller	109
<i>Augmenting Business Entities with Salient Terms from Twitter</i> Riham Mansour, Nesma Refaei and Vanessa Murdock	121
<i>A PAC-Bayesian Approach to Minimum Perplexity Language Modeling</i> Sujeeth Bharadwaj and Mark Hasegawa-Johnson	130
<i>Co-learning of Word Representations and Morpheme Representations</i> Siyu Qiu, Qing Cui, Jiang Bian, Bin Gao and Tie-Yan Liu	141
<i>A Probabilistic Model for Learning Multi-Prototype Word Embeddings</i> Fei Tian, Hanjun Dai, Jiang Bian, Bin Gao, Rui Zhang, Enhong Chen and Tie-Yan Liu	151
<i>Learning Task-specific Bilexical Embeddings</i> Pranava Swaroop Madhyastha, Xavier Carreras and Ariadna Quattoni	161

<i>Building Large-Scale Twitter-Specific Sentiment Lexicon : A Representation Learning Approach</i> Duyu Tang, Furu Wei, Bing Qin, Ming Zhou and Ting Liu	172
<i>Political Tendency Identification in Twitter using Sentiment Analysis Techniques</i> Ferran Pla and Lluís-F. Hurtado	183
<i>A Study of using Syntactic and Semantic Structures for Concept Segmentation and Labeling</i> Iman Saleh, Scott Cyphers, Jim Glass, Shafiq Joty, Lluís Màrquez, Alessandro Moschitti and Preslav Nakov	193
<i>Time-aware Personalized Hashtag Recommendation on Social Media</i> Qi Zhang, Yeyun Gong, Xuyang Sun and Xuanjing Huang	203
<i>Sarcasm Detection on Czech and English Twitter</i> Tomáš Ptáček, Ivan Habernal and Jun Hong	213
<i>A Three-Step Transition-Based System for Non-Projective Dependency Parsing</i> Ophélie Lacroix and Denis Béchet	224
<i>Collaborative Topic Regression with Multiple Graphs Factorization for Recommendation in Social Me- dia</i> Qing Zhang and Houfeng Wang	233
<i>High Performance Word Sense Alignment by Joint Modeling of Sense Distance and Gloss Similarity</i> Michael Matuschek and Iryna Gurevych	245
<i>Multi-view Chinese Treebanking</i> Likun Qiu, Yue Zhang, Peng Jin and Houfeng Wang	257
<i>Rapid Development of a Corpus with Discourse Annotations using Two-stage Crowdsourcing</i> Daisuke Kawahara, Yuichiro Machida, Tomohide Shibata, Sadao Kurohashi, Hayato Kobayashi and Manabu Sassano	269
<i>Chinese Word Ordering Errors Detection and Correction for Non-Native Chinese Language Learners</i> Shuk-Man Cheng, Chi-Hsin Yu and Hsin-Hsi Chen	279
<i>Unsupervised Training Set Generation for Automatic Acquisition of Technical Terminology in Patents</i> Alex Judea, Hinrich Schütze and Soeren Bruegmann	290
<i>A Data Driven Approach for Person Name Disambiguation in Web Search Results</i> Agustín D. Delgado, Raquel Martínez, Víctor Fresno and Soto Montalvo	301
<i>Picking the Amateur’s Mind - Predicting Chess Player Strength from Game Annotations</i> Christian Scheible and Hinrich Schütze	311
<i>Zipf’s Law and Statistical Data on Modern Tibetan</i> Huidan Liu, Minghua Nuo and Jian Wu	322
<i>Simple or Complex? Assessing the readability of Basque Texts</i> Itziar Gonzalez-Dios, María Jesús Aranzabe, Arantza Díaz de Ilarraza and Haritz Salaberri ...	334
<i>Influence of Target Reader Background and Text Features on Text Readability in Bangla: A Computa- tional Approach</i> Manjira Sinha, Tirthankar Dasgupta and Anupam Basu	345

<i>Inducing Word Sense with Automatically Learned Hidden Concepts</i> Baobao Chang, Wenzhe Pei and Miaohong Chen	355
<i>Inferring Knowledge with Word Refinements in a Crowdsourced Lexical-Semantic Network</i> Manel Zarrouk and Mathieu Lafourcade	365
<i>A Supervised Learning Approach Towards Profiling the Preservation of Authorial Style in Literary Translations</i> Gerard Lynch	376
<i>Author Verification Using Common N-Gram Profiles of Text Documents</i> Magdalena Jankowska, Evangelos Milios and Vlado Keselj	387
<i>Dynamically Integrating Cross-Domain Translation Memory into Phrase-Based Machine Translation during Decoding</i> Kun Wang, Chengqing Zong and Keh-Yih Su	398
<i>Machine Translation Quality Estimation Across Domains</i> José G. C. de Souza, Marco Turchi and Matteo Negri	409
<i>Investigating the Usefulness of Generalized Word Representations in SMT</i> Nadir Durrani, Philipp Koehn, Helmut Schmid and Alexander Fraser	421
<i>Confusion Network for Arabic Name Disambiguation and Transliteration in Statistical Machine Translation</i> Young-Suk Lee	433
<i>Fourteen Light Tasks for comparing Analogical and Phrase-based Machine Translation</i> Rafik Rhouma and Phillippe Langlais	444
<i>Finding Zelig in Text: A Measure for Normalising Linguistic Accommodation</i> Simon Jones, Rachel Cotterill, Nigel Dewdney, Kate Muir and Adam Joinson	455
<i>The annotation of the Central Unit in Rhetorical Structure Trees: A Key Step in Annotating Rhetorical Relations</i> Mikel Iruskietia, Arantza Díaz de Ilarraza and Mikel Lersundi	466
<i>Measuring Lexical Cohesion: Beyond Word Repetition</i> Anna Kazantseva and Stan Szpakowicz	476
<i>Fast Tweet Retrieval with Compact Binary Codes</i> Weiwei Guo, Wei Liu and Mona Diab	486
<i>Learning Sense-specific Word Embeddings By Exploiting Bilingual Resources</i> Jiang Guo, Wanxiang Che, Haifeng Wang and Ting Liu	497
<i>Using unmarked contexts in nominal lexical semantic classification</i> Lauren Romeo, Sara Mendes and Núria Bel	508
<i>Skill Inference with Personal and Skill Connections</i> Zhongqing Wang, Shoushan Li, Hanxiao Shi and Guodong Zhou	520
<i>Jointly or Separately: Which is Better for Parsing Heterogeneous Dependencies?</i> Meishan Zhang, Wanxiang Che, Yanqiu Shao and Ting Liu	530

<i>An LR-inspired generalized lexicalized phrase structure parser</i>	
Benoit Crabbé	541
<i>Modeling Review Argumentation for Robust Sentiment Analysis</i>	
Henning Wachsmuth, Martin Trenkmann, Benno Stein and Gregor Engels	553
<i>Biber Redux: Reconsidering Dimensions of Variation in American English</i>	
Rebecca J. Passonneau, Nancy Ide, Songqiao Su and Jesse Stuart	565
<i>Cross-lingual Discourse Relation Analysis: A corpus study and a semi-supervised classification system</i>	
Junyi Jessy Li, Marine Carpuat and Ani Nenkova	577
<i>Enforcing Topic Diversity in a Document Recommender for Conversations</i>	
Maryam Habibi and Andrei Popescu-Belis	588
<i>Identifying Important Features for Graph Retrieval</i>	
Zhuo Li, Sandra Carberry, Hui Fang and Kathleen McCoy	600
<i>Inducing Discourse Connectives from Parallel Texts</i>	
Majid Laali and Leila Kosseim	610
<i>Lyrics-based Analysis and Classification of Music</i>	
Michael Fell and Caroline Sporleder	620
<i>Interpretation of Chinese Discourse Connectives for Explicit Discourse Relation Recognition</i>	
Hen-Hsen Huang, Tai-Wei Chang, Huan-Yuan Chen and Hsin-Hsi Chen	632
<i>Unsupervised Coreference Resolution by Utilizing the Most Informative Relations</i>	
Nafise Sadat Moosavi and Michael Strube	644
<i>Knowledge Sharing via Social Login: Exploiting Microblogging Service for Warming up Social Question Answering Websites</i>	
Yang Xiao, Wayne Xin Zhao, Kun Wang and Zhen Xiao	656
<i>Review Topic Discovery with Phrases using the Pólya Urn Model</i>	
Geli Fei, Zhiyuan Chen and Bing Liu	667
<i>Joint Opinion Relation Detection Using One-Class Deep Neural Network</i>	
Liheng Xu, Kang Liu and Jun Zhao	677
<i>A Generative Model for Identifying Target Companies of Microblogs</i>	
Yeyun Gong, Yaqian Zhou, Ya Guo, Qi Zhang and Xuanjing Huang	688
<i>Inducing Latent Semantic Relations for Structured Distributional Semantics</i>	
Sujay Kumar Jauhar and Eduard Hovy	698
<i>Improving distributional thesauri by exploring the graph of neighbors</i>	
Vincent Claveau, Ewa Kijak and Olivier Ferret	709
<i>Towards Syntax-aware Compositional Distributional Semantic Models</i>	
Lorenzo Ferrone and Fabio Massimo Zanzotto	721
<i>Low-Dimensional Manifold Distributional Semantic Models</i>	
Georgia Athanasopoulou, Elias Iosif and Alexandros Potamianos	731

<i>An Entity-Centric Coreference Resolution System for Person Entities with Rich Linguistic Information</i> Marcos Garcia and Pablo Gamallo	741
<i>Unsupervised Multiword Segmentation of Large Corpora using Prediction-Driven Decomposition of n-grams</i> Julian Brooke, Vivian Tsang, Graeme Hirst and Fraser Shein	753
<i>docrep: A lightweight and efficient document representation framework</i> Tim Dawborn and James R. Curran	762
<i>Why Implementation Matters: Evaluation of an Open-source Constraint Grammar Parser</i> Dávid Márk Nemeskey, Francis Tyers and Mans Hulden	772
<i>Language for Communication: Language as Rational Inference</i> Edward Gibson	781
<i>Soft Cross-lingual Syntax Projection for Dependency Parsing</i> Zhenghua Li, Min Zhang and Wenliang Chen	783
<i>Automatic Feature Selection for Agenda-Based Dependency Parsing</i> Miguel Ballesteros and Bernd Bohnet	794
<i>Predicate-Argument Structure Analysis with Zero-Anaphora Resolution for Dialogue Systems</i> Kenji Imamura, Ryuichiro Higashinaka and Tomoko Izumi	806
<i>Feature Embedding for Dependency Parsing</i> Wenliang Chen, Yue Zhang and Min Zhang	816
<i>Identifying Emotional and Informational Support in Online Health Communities</i> Prakhar Biyani, Cornelia Caragea, Prasenjit Mitra and John Yen	827
<i>Identifying Emotion Labels from Psychiatric Social Texts Using Independent Component Analysis</i> Liang-Chih Yu and Chun-Yuan Ho	837
<i>Modeling Mutual Influence Between Social Actions and Social Ties</i> Xiaofeng Yu and Junqing Xie	848
<i>Discovering Topical Aspects in Microblogs</i> Abhimanyu Das and Anitha Kannan	860
<i>Utilizing Microblogs for Automatic News Highlights Extraction</i> Zhongyu Wei and Wei Gao	872
<i>A Step Towards Usable Privacy Policy: Automatic Alignment of Privacy Statements</i> Fei Liu, Rohan Ramanath, Norman Sadeh and Noah A. Smith	884
<i>An Off-the-shelf Approach to Authorship Attribution</i> Jamal A. Nasir, Nico Görnitz and Ulf Brefeld	895
<i>Automatic Prediction of Aesthetics and Interestingness of Text Passages</i> Debasis Ganguly, Johannes Leveling and Gareth Jones	905
<i>Triple based Background Knowledge Ranking for Document Enrichment</i> Muyu Zhang, Bing Qin, Ting Liu and Mao Zheng	917

<i>Towards an open-domain conversational system fully based on natural language processing</i> Ryuichiro Higashinaka, Kenji Imamura, Toyomi Meguro, Chiaki Miyazaki, Nozomi Kobayashi, Hiroaki Sugiyama, Toru Hirano, Toshiro Makino and Yoshihiro Matsuo	928
<i>The Impact of Deep Hierarchical Discourse Structures in the Evaluation of Text Coherence</i> Vanessa Wei Feng, Ziheng Lin and Graeme Hirst	940
<i>Lexical Chaining for Measuring Discourse Coherence Quality in Test-taker Essays</i> Swapna Somasundaran, Jill Burstein and Martin Chodorow	950
<i>Improving Cloze Test Performance of Language Learners Using Web N-Grams</i> Martin Potthast, Matthias Hagen, Anna Beyer and Benno Stein	962
<i>A Framework for Translating SMS Messages</i> Vivek Kumar Rangarajan Sridhar, John Chen, Srinivas Bangalore and Ron Shacham	974
<i>A Hybrid Approach to Features Representation for Fine-grained Arabic Named Entity Recognition</i> Fahd Alotaibi and Mark Lee	984
<i>Prior-informed Distant Supervision for Temporal Evidence Classification</i> Ridho Reinanda and Maarten de Rijke	996
<i>Identification of Basic Phrases for Kazakh Language using Maximum Entropy Model</i> Gulila Altenbek, Xiaolong Wang and Gulizhada Haisha	1007
<i>Collecting Bilingual Audio in Remote Indigenous Communities</i> Steven Bird, Lauren Gawne, Katie Gelbart and Isaac McAlister	1015
<i>Inclusive yet Selective: Supervised Distributional Hypernymy Detection</i> Stephen Roller, Katrin Erk and Gemma Boleda	1025
<i>Automatic Discovery of Adposition Typology</i> Rishiraj Saha Roy, Rahul Katare, Niloy Ganguly and Monojit Choudhury	1037
<i>What good are 'Nominalkomposita' for 'noun compounds': Multilingual Extraction and Structure Analysis of Nominal Compositions using Linguistic Restrictors</i> Patrick Ziering and Lonneke van der Plas	1047
<i>Automatic Classification of Communicative Functions of Definiteness</i> Archna Bhatia, Chu-Cheng Lin, Nathan Schneider, Yulia Tsvetkov, Fatima Talib Al-Raisi, Laleh Roostapour, Jordan Bender, Abhimanu Kumar, Lori Levin, Mandy Simons and Chris Dyer	1059
<i>Argument structure of adverbial derivatives in Russian</i> Igor Boguslavsky	1071
<i>Active Learning in Noisy Conditions for Spoken Language Understanding</i> Hossein Hadian and Hossein Sameti	1081
<i>A Self-adaptive Classifier for Efficient Text-stream Processing</i> Naoki Yoshinaga and Masaru Kitsuregawa	1091
<i>A Dependency Edge-based Transfer Model for Statistical Machine Translation</i> Hongshen Chen, Jun Xie, Fandong Meng, Wenbin Jiang and Qun Liu	1103
<i>Fast Domain Adaptation of SMT models without in-Domain Parallel Data</i> Prashant Mathur, Sriram Venkatapathy and Nicola Cancedda	1114

<i>Discriminative Language Models as a Tool for Machine Translation Error Analysis</i> Koichi Akabe, Graham Neubig, Sakriani Sakti, Tomoki Toda and Satoshi Nakamura	1124
<i>A Structured Language Model for Incremental Tree-to-String Translation</i> Heng Yu, Haitao Mi, Liang Huang and Qun Liu	1133
<i>A Lexicalized Reordering Model for Hierarchical Phrase-based Translation</i> Hailong Cao, Dongdong Zhang, Mu Li, Ming Zhou and Tiejun Zhao	1144
<i>Automatic Corpus Expansion for Chinese Word Segmentation by Exploiting the Redundancy of Web Information</i> Xipeng Qiu, ChaoChao Huang and Xuanjing Huang	1154
<i>Fast High-Accuracy Part-of-Speech Tagging by Independent Classifiers</i> Robert Moore	1165
<i>Morfessor FlatCat: An HMM-Based Method for Unsupervised and Semi-Supervised Learning of Morphology</i> Stig-Arne Grönroos, Sami Virpioja, Peter Smit and Mikko Kurimo	1177
<i>Japanese Word Reordering Integrated with Dependency Parsing</i> Kazushi Yoshida, Tomohiro Ohno, Yoshihide Kato and Shigeki Matsubara	1186
<i>Query-focused Multi-Document Summarization: Combining a Topic Model with Graph-based Semi-supervised Learning</i> Yanran Li and Sujian Li	1197
<i>Ranking Multidocument Event Descriptions for Building Thematic Timelines</i> Kiem-Hieu Nguyen, Xavier Tannier and Véronique Moriceau	1208
<i>Integrating Language and Vision to Generate Natural Language Descriptions of Videos in the Wild</i> Jesse Thomason, Subhashini Venugopalan, Sergio Guadarrama, Kate Saenko and Raymond Mooney	1218
<i>Cross-Topic Authorship Attribution: Will Out-Of-Topic Data Help?</i> Upendra Sapkota, Thamar Solorio, Manuel Montes, Steven Bethard and Paolo Rosso	1228
<i>Online Gaming for Crowd-sourcing Phrase-equivalents</i> A Kumaran, Melissa Densmore and Shaishav Kumar	1238
<i>Unsupervised Verb Inference from Nouns Crossing Root Boundary</i> Soon Gill Hong, Sin-hee Cho and Mun Yong Yi	1248
<i>Enriching Wikipedia's Intra-language Links by their Cross-language Transfer</i> Takashi Tsunakawa, Makoto Araya and Hiroyuki Kaji	1260
<i>Chinese Irony Corpus Construction and Ironic Structure Analysis</i> Yi-jie Tang and Hsin-Hsi Chen	1269
<i>Global Methods for Cross-lingual Semantic Role and Predicate Labelling</i> Lonneke van der Plas, Marianna Apidianaki and Chenhua Chen	1279
<i>Multilingual Semantic Parsing : Parsing Multiple Languages into Semantic Representations</i> Zhanming Jie and Wei Lu	1291

<i>Unsupervised Word Sense Induction using Distributional Statistics</i> Kartik Goyal and Eduard Hovy	1302
<i>Group based Self Training for E-Commerce Product Record Linkage</i> Xin Zhao, Yuexin Wu, Hongfei Yan and Xiaoming Li	1311
<i>Reducing Over-Weighting in Supervised Term Weighting for Sentiment Analysis</i> Haibing Wu and Xiaodong Gu	1322
<i>Sentiment Classification with Graph Co-Regularization</i> Guangyou Zhou, Jun Zhao and Daojian Zeng	1331
<i>Hybrid Deep Belief Networks for Semi-supervised Sentiment Classification</i> Shusen Zhou, Qingcai Chen, Xiaolong Wang and Xiaoling Li	1341
<i>Latent Dynamic Model with Category Transition Constraint for Opinion Classification</i> Takeshi Kobayakawa	1350
<i>Sentence Compression for Target-Polarity Word Collocation Extraction</i> Yanyan Zhao, Wanxiang Che, Honglei Guo, Bing Qin, Zhong Su and Ting Liu	1360
<i>Hybrid Grammars for Discontinuous Parsing</i> Mark-Jan Nederhof and Heiko Vogler	1370
<i>From neighborhood to parenthood: the advantages of dependency representation over bigrams in Brown clustering</i> Simon Suster and Gertjan van Noord	1382
<i>An Empirical Evaluation of Automatic Conversion from Constituency to Dependency in Hungarian</i> Katalin Ilona Simkó, Veronika Vincze, Zsolt Szántó and Richárd Farkas	1392
<i>Deep-Syntactic Parsing</i> Miguel Ballesteros, Bernd Bohnet, Simon Mille and Leo Wanner	1402
<i>Modeling Newswire Events using Neural Networks for Anomaly Detection</i> Pradeep Dasigi and Eduard Hovy	1414
<i>Million-scale Derivation of Semantic Relations from a Manually Constructed Predicate Taxonomy</i> Motoki Sano, Kentaro Torisawa, Julien Kloetzer, Chikara Hashimoto, István Varga and Jong-Hoon Oh	1423
<i>Combining Supervised and Unsupervised Parsing for Distributional Similarity</i> Martin Riedl, Irina Alles and Chris Biemann	1435
<i>A Markovian approach to distributional semantics with application to semantic compositionality</i> Edouard Grave, Guillaume Obozinski and Francis Bach	1447
<i>A Beam-Search Decoder for Disfluency Detection</i> Xuancong Wang, Hwee Tou Ng and Khe Chai Sim	1457
<i>Single Document Keyphrase Extraction Using Label Information</i> Sumit Negi	1468
<i>Predicting Interesting Things in Text</i> Michael Gamon, Arjun Mukherjee and Patrick Pantel	1477

<i>Context Dependent Claim Detection</i>	
Ran Levy, Yonatan Bilu, Daniel Hershcovich, Ehud Aharoni and Noam Slonim	1489
<i>Annotating Argument Components and Relations in Persuasive Essays</i>	
Christian Stab and Iryna Gurevych	1501
<i>Building a Hierarchically Aligned Chinese-English Parallel Treebank</i>	
Dun Deng and Nianwen Xue	1511
<i>3arif: A Corpus of Modern Standard and Egyptian Arabic Tweets Annotated for Epistemic Modality Using Interactive Crowdsourcing</i>	
Rania Al-Sabbagh, Roxana Girju and Jana Diesner	1521
<i>Empirical Analysis of Aggregation Methods for Collective Annotation</i>	
Ciyang Qing, Ulle Endriss, Raquel Fernandez and Justin Kruger	1533
<i>Annotation Adaptation and Language Adaptation in NLP</i>	
Qun Liu	1543
<i>Collective Named Entity Disambiguation using Graph Ranking and Clique Partitioning Approaches</i>	
Ayman Alhelbawy and Robert Gaizauskas	1544
<i>Analysis and Refinement of Temporal Relation Aggregation</i>	
Taylor Cassidy and Heng Ji	1556
<i>The Wisdom of Minority: Unsupervised Slot Filling Validation based on Multi-dimensional Truth-Finding</i>	
Dian Yu, Hongzhao Huang, Taylor Cassidy, Heng Ji, Chi Wang, Shi Zhi, Jiawei Han, Clare Voss and Malik Magdon-Ismael	1567
<i>Common Space Embedding of Primal-Dual Relation Semantic Spaces</i>	
Hidekazu Oiwa and Jun'ichi Tsujii	1579
<i>An Enhanced Lesk Word Sense Disambiguation Algorithm through a Distributional Semantic Model</i>	
Pierpaolo Basile, Annalina Caputo and Giovanni Semeraro	1591
<i>Word Sense Induction Using Lexical Chain based Hypergraph Model</i>	
Tao Qian, Donghong Ji, Mingyao Zhang, Chong Teng and Congling Xia	1601
<i>Minimally Supervised Classification to Semantic Categories using Automatically Acquired Symmetric Patterns</i>	
Roy Schwartz, Roi Reichart and Ari Rappoport	1612
<i>Novel Word-sense Identification</i>	
Paul Cook, Jey Han Lau, Diana McCarthy and Timothy Baldwin	1624
<i>Learning to Summarise Related Sentences</i>	
Emmanouil Tzouridis, Jamal Nasir and Ulf Brefeld	1636
<i>Learning to Generate Coherent Summary with Discriminative Hidden Semi-Markov Model</i>	
Hitoshi Nishikawa, Kazuho Arita, Katsumi Tanaka, Tsutomu Hirao, Toshiro Makino and Yoshihiro Matsuo	1648
<i>Query-Focused Opinion Summarization for User-Generated Content</i>	
Lu Wang, Hema Raghavan, Claire Cardie and Vittorio Castelli	1660

<i>Generating Supplementary Travel Guides from Social Media</i>	
Liu Yang, Jing Jiang, Lifu Huang, Minghui Qiu and Lizi Liao	1670
<i>Ensemble-Based Medical Relation Classification</i>	
Jennifer D’Souza and Vincent Ng	1682
<i>Combining Natural and Artificial Examples to Improve Implicit Discourse Relation Identification</i>	
Chloé Braud and Pascal Denis	1694
<i>Reinforcement Learning of Cooperative Persuasive Dialogue Policies using Framing</i>	
Takuya Hiraoka, Graham Neubig, Sakriani Sakti, Tomoki Toda and Satoshi Nakamura	1706
<i>Towards multimodal modeling of physicians’ diagnostic confidence and self-awareness using medical narratives</i>	
Joseph Bullard, Cecilia Ovesdotter Alm, Qi Yu, Pengcheng Shi and Anne Haake	1718
<i>Towards Semantic Validation of a Derivational Lexicon</i>	
Britta Zeller, Sebastian Padó and Jan Šnajder	1728
<i>Detecting Learner Errors in the Choice of Content Words Using Compositional Distributional Semantics</i>	
Ekaterina Kochmar and Ted Briscoe	1740
<i>A Novel Distributional Approach to Multilingual Conceptual Metaphor Recognition</i>	
Michael Mohler, Bryan Rink, David Bracewell and Marc Tomlinson	1752
<i>Part of Speech Tagging for French Social Media Data</i>	
Farhad Nooralahzadeh, Caroline Brun and Claude Roux	1764
<i>Morphological Analysis for Japanese Noisy Text based on Character-level and Word-level Normalization</i>	
Itsumi Saito, Kugatsu Sadamitsu, Hisako Asano and Yoshihiro Matsuo	1773
<i>Adapting taggers to Twitter with not-so-distant supervision</i>	
Barbara Plank, Dirk Hovy, Ryan McDonald and Anders Søgaard	1783
<i>Interpolated Dirichlet Class Language Model for Speech Recognition Incorporating Long-distance N-grams</i>	
Md. Akmal Haidar and Douglas O’Shaughnessy	1793
<i>Situated Incremental Natural Language Understanding using a Multimodal, Linguistically-driven Update Model</i>	
Casey Kennington, Spyros Kousidis and David Schlangen	1803
<i>Quality Estimation for Automatic Speech Recognition</i>	
Matteo Negri, Marco Turchi, José G. C. de Souza and Falavigna Daniele	1813
<i>A Generic Anaphora Resolution Engine for Indian Languages</i>	
Sobha Lalitha Devi, Vijay Sundar Ram and Pattabhi RK Rao	1824
<i>Converting Phrase Structures to Dependency Structures in Sanskrit</i>	
Pawan Goyal and Amba Kulkarni	1834
<i>Uncertainty Detection in Hungarian Texts</i>	
Veronika Vincze	1844
<i>Rediscovering Annotation Projection for Cross-Lingual Parser Induction</i>	
Jörg Tiedemann	1854

<i>Synchronous Constituent Context Model for Inducing Bilingual Synchronous Structures</i> Xiangyu Duan, Min Zhang and Qiaoming Zhu	1865
<i>Syntactic Parsing and Compound Recognition via Dual Decomposition: Application to French</i> Joseph Le Roux, Antoine Rozenknop and Matthieu Constant	1875
<i>Learning the Taxonomy of Function Words for Parsing</i> Dongchen Li, Xiantao Zhang, Dingsheng Luo and Xihong Wu	1886
<i>A Neural Reordering Model for Phrase-based Translation</i> Peng Li, Yang Liu, Maosong Sun, Tatsuya Izuha and Dakun Zhang	1897
<i>Recurrent Neural Network-based Tuple Sequence Model for Machine Translation</i> Youzheng Wu, Taro Watanabe and Chiori Hori	1908
<i>Class-Based Language Modeling for Translating into Morphologically Rich Languages</i> Arianna Bisazza and Christof Monz	1918
<i>Latent Domain Translation Models in Mix-of-Domains Haystack</i> Cuong Hoang and Khalil Sima'an	1928
<i>Language Family Relationship Preserved in Non-native English</i> Ryo Nagata	1940
<i>Why Gender and Age Prediction from Tweets is Hard: Lessons from a Crowdsourcing Experiment</i> Dong Nguyen, Dolf Trieschnigg, A. Seza Dođruöz, Rilana Gravel, Mariet Theune, Theo Meder and Franciska De Jong	1950
<i>Exploring Syntactic Features for Native Language Identification: A Variationist Perspective on Feature Encoding and Ensemble Optimization</i> Serhiy Bykh and Detmar Meurers	1962
<i>Applying automatically parsed corpora to the study of language variation</i> Jelke Bloem, Arjen Versloot and Fred Weerman	1974
<i>Empirical analysis of exploiting review helpfulness for extractive summarization of online reviews</i> Wenting Xiong and Diane Litman	1985
<i>Lexico-syntactic text simplification and compression with typed dependencies</i> Mandya Angrosh, Tadashi Nomoto and Advait Siddharthan	1996
<i>Learning when to point: A data-driven approach</i> Albert Gatt and Patrizia Paggio	2007
<i>Generating Acrostics via Paraphrasing and Heuristic Search</i> Benno Stein, Matthias Hagen and Christof Bräutigam	2018
<i>Does a Computational Linguist have to be a Linguist?</i> Martin Kay	2030
<i>Query Lattice for Translation Retrieval</i> Meiping Dong, Yong Cheng, Yang Liu, Jia Xu, Maosong Sun, Tatsuya Izuha and Jie Hao	2031
<i>RED: A Reference Dependency Based MT Evaluation Metric</i> Hui Yu, Xiaofeng Wu, Jun Xie, Wenbin Jiang, Qun Liu and Shouxun Lin	2042

<i>Quality Estimation of English-French Machine Translation: A Detailed Study of the Role of Syntax</i> Rasoul Kaljahi, Jennifer Foster, Johann Roturier and Raphael Rubino	2052
<i>Effective Incorporation of Source Syntax into Hierarchical Phrase-based Translation</i> Tong Xiao, Adrià de Gispert, Jingbo Zhu and Bill Byrne	2064
<i>BEL: Bagging for Entity Linking</i> Zhe Zuo, Gjergji Kasneci, Toni Gruetze and Felix Naumann	2075
<i>Exploratory Relation Extraction in Large Text Corpora</i> Alan Akbik, Thilo Michael and Christoph Boden	2087
<i>An Analysis of Causality between Events and its Relation to Temporal Information</i> Paramita Mirza and Sara Tonelli	2097
<i>Exploring Fine-grained Entity Type Constraints for Distantly Supervised Relation Extraction</i> Yang Liu, Kang Liu, Liheng Xu and Jun Zhao	2107
<i>Using Collections of Human Language Intuitions to Measure Corpus Representativeness</i> Reinhard Rapp	2117
<i>Limited memory incremental coreference resolution</i> Kellie Webster and James R. Curran	2129
<i>Left-corner Transitions on Dependency Parsing</i> Hiroshi Noji and Yusuke Miyao	2140
<i>Data-driven Measurement of Child Language Development with Simple Syntactic Templates</i> Shannon Lubetich and Kenji Sagae	2151
<i>Employing Event Inference to Improve Semi-Supervised Chinese Event Extraction</i> Peifeng Li, Qiaoming Zhu and Guodong Zhou	2161
<i>Supervised Ranking of Co-occurrence Profiles for Acquisition of Continuous Lexical Attributes</i> Julian Brooke and Graeme Hirst	2172
<i>Unsupervised extraction of semantic relations using discourse cues</i> Juliette Conrath, Stergos Afantenos, Nicholas Asher and Philippe Muller	2184
<i>HARPY: Hypernyms and Alignment of Relational Paraphrases</i> Adam Grycner and Gerhard Weikum	2195
<i>Limitations of MT Quality Estimation Supervised Systems: The Tails Prediction Problem</i> Erwan Moreau and Carl Vogel	2205
<i>Augment Dependency-to-String Translation with Fixed and Floating Structures</i> Jun Xie, Jinan Xu and Qun Liu	2217
<i>Soft Dependency Matching for Hierarchical Phrase-based Machine Translation</i> Hailong Cao, Dongdong Zhang, Ming Zhou and Tiejun Zhao	2227
<i>Using Spreading Activation to Evaluate and Improve Ontologies</i> Ronan Mac an tSaoir	2237
<i>Learning to Distinguish Hypernyms and Co-Hyponyms</i> Julie Weeds, Daoud Clarke, Jeremy Reffin, David Weir and Bill Keller	2249

<i>"One Entity per Discourse" and "One Entity per Collocation" Improve Named-Entity Disambiguation</i> Ander Barrena, Eneko Agirre, Bernardo Cabaleiro, Anselmo Peñas and Aitor Soroa	2260
<i>Comparable Study of Event Extraction in Newswire and Biomedical Domains</i> Makoto Miwa, Paul Thompson, Ioannis Korkontzelos and Sophia Ananiadou	2270
<i>A Probabilistic Co-Bootstrapping Method for Entity Set Expansion</i> Bei Shi, Zhenzhong Zhang, Le Sun and Xianpei Han	2280
<i>Separating Brands from Types: an Investigation of Different Features for the Food Domain</i> Michael Wiegand and Dietrich Klakow	2291
<i>Unsupervised Instance-Based Part of Speech Induction Using Probable Substitutes</i> Deniz Yuret, Mehmet Ali Yatbaz and Enis Sert	2303
<i>Solving Substitution Ciphers with Combined Language Models</i> Bradley Hauer, Ryan Hayward and Grzegorz Kondrak	2314
<i>Unsupervised Word Segmentation in Context</i> Gabriel Synnaeve, Isabelle Dautriche, Benjamin Börschinger, Mark Johnson and Emmanuel Dupoux 2326	
<i>Relation Classification via Convolutional Deep Neural Network</i> Daojian Zeng, Kang Liu, Siwei Lai, Guangyou Zhou and Jun Zhao	2335
<i>A context-based model for Sentiment Analysis in Twitter</i> Andrea Vanzo, Danilo Croce and Roberto Basili	2345

Conference Program

Monday, August 25, 2014

08:45-09:00 Welcome and Opening by General, Program and Local Chairs

Session Mo11: (09:00-10:15) Invited Talk 1

09:00–10:15 *Learning from 26 Languages: Program Management and Science in the Babel Program*

Mary Harper

10:15-10:45 Coffee Break

Session Mo21: (10:45-12:25) Modeling of Discourse and Dialogue I

10:45–11:10 *Unsupervised learning of rhetorical structure with un-topic models*

Diarmuid Ó Séaghdha and Simone Teufel

11:10–11:35 *Cross-lingual Coreference Resolution of Pronouns*

Michal Novak and Zdenek Zabokrtsky

11:35–12:00 *A Context-Aware NLP Approach For Noteworthiness Detection in Cellphone Conversations*

Francesca Bonin, Jose San Pedro and Nuria Oliver

12:00–12:25 *Hierarchical Topical Segmentation with Affinity Propagation*

Anna Kazantseva and Stan Szpakowicz

Session Mo22: (10:45-12:25) Sentiment Analysis, Opinion Mining and Social Media I

10:45–11:10 *Capturing Cultural Differences in Expressions of Intentions*

Marc Tomlinson, David Bracewell and Wayne Krug

11:10–11:35 *Identification of Implicit Topics in Twitter Data Not Containing Explicit Search Queries*

Suzi Park and Hyopil Shin

11:35–12:00 *Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts*

Cicero dos Santos and Maira Gatti

Monday, August 25, 2014 (continued)

12:00–12:25 *Joint Inference and Disambiguation of Implicit Sentiments via Implicature Constraints*
Lingjia Deng, Janyce Wiebe and Yoonjung Choi

Session Mo23: (10:45-12:25) Information Retrieval and Question Answering

10:45–11:10 *Group Non-negative Matrix Factorization with Natural Categories for Question Retrieval in Community Question Answer Archives*
Guangyou Zhou, Yubo Chen, Daojian Zeng and Jun Zhao

11:10–11:35 *Multi-Objective Search Results Clustering*
Sudipta Acharya, Sriparna Saha, Jose G. Moreno and Gaël Dias

11:35–12:00 *Query-by-Example Image Retrieval using Visual Dependency Representations*
Desmond Elliott, Victor Lavrenko and Frank Keller

12:00–12:25 *Augmenting Business Entities with Salient Terms from Twitter*
Riham Mansour, Nesma Refaei and Vanessa Murdock

Session Mo24: (10:45-12:25) Machine Learning for CL and NLP

10:45–11:10 *A PAC-Bayesian Approach to Minimum Perplexity Language Modeling*
Sujeeth Bharadwaj and Mark Hasegawa-Johnson

11:10–11:35 *Co-learning of Word Representations and Morpheme Representations*
Siyu Qiu, Qing Cui, Jiang Bian, Bin Gao and Tie-Yan Liu

11:35–12:00 *A Probabilistic Model for Learning Multi-Prototype Word Embeddings*
Fei Tian, Hanjun Dai, Jiang Bian, Bin Gao, Rui Zhang, Enhong Chen and Tie-Yan Liu

12:00–12:25 *Learning Task-specific Bilexical Embeddings*
Pranava Swaroop Madhyastha, Xavier Carreras and Ariadna Quattoni

12:25-14:00 Lunch Break

Monday, August 25, 2014 (continued)

Session Mo3P: (14:00-15:15) Posters I

Building Large-Scale Twitter-Specific Sentiment Lexicon : A Representation Learning Approach

Duyu Tang, Furu Wei, Bing Qin, Ming Zhou and Ting Liu

Political Tendency Identification in Twitter using Sentiment Analysis Techniques

Ferran Pla and Lluís-F. Hurtado

A Study of using Syntactic and Semantic Structures for Concept Segmentation and Labeling

Iman Saleh, Scott Cyphers, Jim Glass, Shafiq Joty, Lluís Màrquez, Alessandro Moschitti and Preslav Nakov

Time-aware Personalized Hashtag Recommendation on Social Media

Qi Zhang, Yeyun Gong, Xuyang Sun and Xuanjing Huang

Sarcasm Detection on Czech and English Twitter

Tomáš Ptáček, Ivan Habernal and Jun Hong

A Three-Step Transition-Based System for Non-Projective Dependency Parsing

Ophélie Lacroix and Denis Béchet

Collaborative Topic Regression with Multiple Graphs Factorization for Recommendation in Social Media

Qing Zhang and Houfeng Wang

High Performance Word Sense Alignment by Joint Modeling of Sense Distance and Gloss Similarity

Michael Matuschek and Iryna Gurevych

Multi-view Chinese Treebanking

Likun Qiu, Yue Zhang, Peng Jin and Houfeng Wang

Rapid Development of a Corpus with Discourse Annotations using Two-stage Crowdsourcing

Daisuke Kawahara, Yuichiro Machida, Tomohide Shibata, Sadao Kurohashi, Hayato Kobayashi and Manabu Sassano

Chinese Word Ordering Errors Detection and Correction for Non-Native Chinese Language Learners

Shuk-Man Cheng, Chi-Hsin Yu and Hsin-Hsi Chen

Unsupervised Training Set Generation for Automatic Acquisition of Technical Terminology in Patents

Alex Judea, Hinrich Schütze and Soeren Bruegmann

Monday, August 25, 2014 (continued)

A Data Driven Approach for Person Name Disambiguation in Web Search Results

Agustín D. Delgado, Raquel Martínez, Víctor Fresno and Soto Montalvo

Picking the Amateur's Mind - Predicting Chess Player Strength from Game Annotations

Christian Scheible and Hinrich Schütze

Zipf's Law and Statistical Data on Modern Tibetan

Huidan Liu, Minghua Nuo and Jian Wu

Simple or Complex? Assessing the readability of Basque Texts

Itziar Gonzalez-Dios, María Jesús Aranzabe, Arantza Díaz de Ilarraza and Haritz Salaberri

Influence of Target Reader Background and Text Features on Text Readability in Bangla: A Computational Approach

Manjira Sinha, Tirthankar Dasgupta and Anupam Basu

Inducing Word Sense with Automatically Learned Hidden Concepts

Baobao Chang, Wenzhe Pei and Miaohong Chen

Inferring Knowledge with Word Refinements in a Crowdsourced Lexical-Semantic Network

Manel Zarrouk and Mathieu Lafourcade

A Supervised Learning Approach Towards Profiling the Preservation of Authorial Style in Literary Translations

Gerard Lynch

Author Verification Using Common N-Gram Profiles of Text Documents

Magdalena Jankowska, Evangelos Milios and Vlado Keselj

Dynamically Integrating Cross-Domain Translation Memory into Phrase-Based Machine Translation during Decoding

Kun Wang, Chengqing Zong and Keh-Yih Su

Machine Translation Quality Estimation Across Domains

José G. C. de Souza, Marco Turchi and Matteo Negri

Investigating the Usefulness of Generalized Word Representations in SMT

Nadir Durrani, Philipp Koehn, Helmut Schmid and Alexander Fraser

Monday, August 25, 2014 (continued)

Confusion Network for Arabic Name Disambiguation and Transliteration in Statistical Machine Translation

Young-Suk Lee

Fourteen Light Tasks for comparing Analogical and Phrase-based Machine Translation

Rafik Rhouma and Phillippe Langlais

Finding Zelig in Text: A Measure for Normalising Linguistic Accommodation

Simon Jones, Rachel Cotterill, Nigel Dewdney, Kate Muir and Adam Joinson

The annotation of the Central Unit in Rhetorical Structure Trees: A Key Step in Annotating Rhetorical Relations

Mikel Iruskieta, Arantza Díaz de Ilarraza and Mikel Lersundi

Measuring Lexical Cohesion: Beyond Word Repetition

Anna Kazantseva and Stan Szpakowicz

Fast Tweet Retrieval with Compact Binary Codes

Weiwei Guo, Wei Liu and Mona Diab

Learning Sense-specific Word Embeddings By Exploiting Bilingual Resources

Jiang Guo, Wanxiang Che, Haifeng Wang and Ting Liu

Using unmarked contexts in nominal lexical semantic classification

Lauren Romeo, Sara Mendes and Núria Bel

Skill Inference with Personal and Skill Connections

Zhongqing Wang, Shoushan Li, Hanxiao Shi and Guodong Zhou

Jointly or Separately: Which is Better for Parsing Heterogeneous Dependencies?

Meishan Zhang, Wanxiang Che, Yanqiu Shao and Ting Liu

An LR-inspired generalized lexicalized phrase structure parser

Benoit Crabbé

Modeling Review Argumentation for Robust Sentiment Analysis

Henning Wachsmuth, Martin Trenkmann, Benno Stein and Gregor Engels

Monday, August 25, 2014 (continued)

Biber Redux: Reconsidering Dimensions of Variation in American English

Rebecca J. Passonneau, Nancy Ide, Songqiao Su and Jesse Stuart

Cross-lingual Discourse Relation Analysis: A corpus study and a semi-supervised classification system

Junyi Jessy Li, Marine Carpuat and Ani Nenkova

Enforcing Topic Diversity in a Document Recommender for Conversations

Maryam Habibi and Andrei Popescu-Belis

Identifying Important Features for Graph Retrieval

Zhuo Li, Sandra Carberry, Hui Fang and Kathleen McCoy

15:15-15:45 Coffee Break

Session Mo41: (15:45-17:25) Modeling of Discourse and Dialogue II

15:45–16:10 *Inducing Discourse Connectives from Parallel Texts*

Majid Laali and Leila Kosseim

16:10–16:35 *Lyrics-based Analysis and Classification of Music*

Michael Fell and Caroline Sporleder

16:35–17:00 *Interpretation of Chinese Discourse Connectives for Explicit Discourse Relation Recognition*

Hen-Hsen Huang, Tai-Wei Chang, Huan-Yuan Chen and Hsin-Hsi Chen

17:00–17:25 *Unsupervised Coreference Resolution by Utilizing the Most Informative Relations*

Nafise Sadat Moosavi and Michael Strube

Monday, August 25, 2014 (continued)

Session Mo42: (15:45-17:25) Sentiment Analysis, Opinion Mining and Social Media II

- 15:45–16:10 *Knowledge Sharing via Social Login: Exploiting Microblogging Service for Warming up Social Question Answering Websites*
Yang Xiao, Wayne Xin Zhao, Kun Wang and Zhen Xiao
- 16:10–16:35 *Review Topic Discovery with Phrases using the Pólya Urn Model*
Geli Fei, Zhiyuan Chen and Bing Liu
- 16:35–17:00 *Joint Opinion Relation Detection Using One-Class Deep Neural Network*
Liheng Xu, Kang Liu and Jun Zhao
- 17:00–17:25 *A Generative Model for Identifying Target Companies of Microblogs*
Yeyun Gong, Yaqian Zhou, Ya Guo, Qi Zhang and Xuanjing Huang

Session Mo43: (15:45-17:25) Semantic Processing, Distributional Semantics and Compositional Semantics I

- 15:45–16:10 *Inducing Latent Semantic Relations for Structured Distributional Semantics*
Sujay Kumar Jauhar and Eduard Hovy
- 16:10–16:35 *Improving distributional thesauri by exploring the graph of neighbors*
Vincent Claveau, Ewa Kijak and Olivier Ferret
- 16:35–17:00 *Towards Syntax-aware Compositional Distributional Semantic Models*
Lorenzo Ferrone and Fabio Massimo Zanzotto
- 17:00–17:25 *Low-Dimensional Manifold Distributional Semantic Models*
Georgia Athanasopoulou, Elias Iosif and Alexandros Potamianos

Monday, August 25, 2014 (continued)

Session Mo44: (15:45-17:25) Software, Tools

- 15:45–16:10 *An Entity-Centric Coreference Resolution System for Person Entities with Rich Linguistic Information*
Marcos Garcia and Pablo Gamallo
- 16:10–16:35 *Unsupervised Multiword Segmentation of Large Corpora using Prediction-Driven Decomposition of n-grams*
Julian Brooke, Vivian Tsang, Graeme Hirst and Fraser Shein
- 16:35–17:00 *docrep: A lightweight and efficient document representation framework*
Tim Dawborn and James R. Curran
- 17:00–17:25 *Why Implementation Matters: Evaluation of an Open-source Constraint Grammar Parser*
Dávid Márk Nemeskey, Francis Tyers and Mans Hulden

Tuesday, August 26, 2014

Session Tu11: (09:00-10:15) Invited Talk 2

- 09:00–10:15 *Language for Communication: Language as Rational Inference*
Edward Gibson
- 10:15-10:45 Coffee Break

Session Tu21: (10:45-12:25) Syntax, Grammar Induction, Syntactic and Semantic Parsing I

- 10:45–11:10 *Soft Cross-lingual Syntax Projection for Dependency Parsing*
Zhenghua Li, Min Zhang and Wenliang Chen
- 11:10–11:35 *Automatic Feature Selection for Agenda-Based Dependency Parsing*
Miguel Ballesteros and Bernd Bohnet
- 11:35–12:00 *Predicate-Argument Structure Analysis with Zero-Anaphora Resolution for Dialogue Systems*
Kenji Imamura, Ryuichiro Higashinaka and Tomoko Izumi
- 12:00–12:25 *Feature Embedding for Dependency Parsing*
Wenliang Chen, Yue Zhang and Min Zhang

Tuesday, August 26, 2014 (continued)

Session Tu22: (10:45-12:25) Sentiment Analysis, Opinion Mining and Social Media III

- 10:45–11:10 *Identifying Emotional and Informational Support in Online Health Communities*
Prakhar Biyani, Cornelia Caragea, Prasenjit Mitra and John Yen
- 11:10–11:35 *Identifying Emotion Labels from Psychiatric Social Texts Using Independent Component Analysis*
Liang-Chih Yu and Chun-Yuan Ho
- 11:35–12:00 *Modeling Mutual Influence Between Social Actions and Social Ties*
Xiaofeng Yu and Junqing Xie
- 12:00–12:25 *Discovering Topical Aspects in Microblogs*
Abhimanyu Das and Anitha Kannan

Session Tu23: (10:45-12:25) Applications I

- 10:45–11:10 *Utilizing Microblogs for Automatic News Highlights Extraction*
Zhongyu Wei and Wei Gao
- 11:10–11:35 *A Step Towards Usable Privacy Policy: Automatic Alignment of Privacy Statements*
Fei Liu, Rohan Ramanath, Norman Sadeh and Noah A. Smith
- 11:35–12:00 *An Off-the-shelf Approach to Authorship Attribution*
Jamal A. Nasir, Nico Görnitz and Ulf Brefeld
- 12:00–12:25 *Automatic Prediction of Aesthetics and Interestingness of Text Passages*
Debasis Ganguly, Johannes Leveling and Gareth Jones

Tuesday, August 26, 2014 (continued)

Session Tu24: (10:45-12:25) Modeling of Discourse and Dialogue III

- 10:45–11:10 *Triple based Background Knowledge Ranking for Document Enrichment*
Muyu Zhang, Bing Qin, Ting Liu and Mao Zheng
- 11:10–11:35 *Towards an open-domain conversational system fully based on natural language processing*
Ryuichiro Higashinaka, Kenji Imamura, Toyomi Meguro, Chiaki Miyazaki, Nozomi Kobayashi, Hiroaki Sugiyama, Toru Hirano, Toshiro Makino and Yoshihiro Matsuo
- 11:35–12:00 *The Impact of Deep Hierarchical Discourse Structures in the Evaluation of Text Coherence*
Vanessa Wei Feng, Ziheng Lin and Graeme Hirst
- 12:00–12:25 *Lexical Chaining for Measuring Discourse Coherence Quality in Test-taker Essays*
Swapna Somasundaran, Jill Burstein and Martin Chodorow
- 12:25-14:00 Lunch Break

Session Tu3P: (14:00-15:15) Posters II

- Improving Cloze Test Performance of Language Learners Using Web N-Grams*
Martin Potthast, Matthias Hagen, Anna Beyer and Benno Stein
- A Framework for Translating SMS Messages*
Vivek Kumar Rangarajan Sridhar, John Chen, Srinivas Bangalore and Ron Shacham
- A Hybrid Approach to Features Representation for Fine-grained Arabic Named Entity Recognition*
Fahd Alotaibi and Mark Lee
- Prior-informed Distant Supervision for Temporal Evidence Classification*
Ridho Reinanda and Maarten de Rijke
- Identification of Basic Phrases for Kazakh Language using Maximum Entropy Model*
Gulila Altenbek, Xiaolong Wang and Gulizhada Haisha
- Collecting Bilingual Audio in Remote Indigenous Communities*
Steven Bird, Lauren Gawne, Katie Gelbart and Isaac McAlister

Tuesday, August 26, 2014 (continued)

Inclusive yet Selective: Supervised Distributional Hypernymy Detection

Stephen Roller, Katrin Erk and Gemma Boleda

Automatic Discovery of Adposition Typology

Rishiraj Saha Roy, Rahul Katare, Niloy Ganguly and Monojit Choudhury

What good are 'Nominalkomposita' for 'noun compounds': Multilingual Extraction and Structure Analysis of Nominal Compositions using Linguistic Restrictors

Patrick Ziering and Lonneke van der Plas

Automatic Classification of Communicative Functions of Definiteness

Archana Bhatia, Chu-Cheng Lin, Nathan Schneider, Yulia Tsvetkov, Fatima Talib Al-Raisi, Laleh Roostapour, Jordan Bender, Abhimanu Kumar, Lori Levin, Mandy Simons and Chris Dyer

Argument structure of adverbial derivatives in Russian

Igor Boguslavsky

Active Learning in Noisy Conditions for Spoken Language Understanding

Hossein Hadian and Hossein Sameti

A Self-adaptive Classifier for Efficient Text-stream Processing

Naoki Yoshinaga and Masaru Kitsuregawa

A Dependency Edge-based Transfer Model for Statistical Machine Translation

Hongshen Chen, Jun Xie, Fandong Meng, Wenbin Jiang and Qun Liu

Fast Domain Adaptation of SMT models without in-Domain Parallel Data

Prashant Mathur, Sriram Venkatapathy and Nicola Cancedda

Discriminative Language Models as a Tool for Machine Translation Error Analysis

Koichi Akabe, Graham Neubig, Sakriani Sakti, Tomoki Toda and Satoshi Nakamura

A Structured Language Model for Incremental Tree-to-String Translation

Heng Yu, Haitao Mi, Liang Huang and Qun Liu

A Lexicalized Reordering Model for Hierarchical Phrase-based Translation

Hailong Cao, Dongdong Zhang, Mu Li, Ming Zhou and Tiejun Zhao

Tuesday, August 26, 2014 (continued)

Automatic Corpus Expansion for Chinese Word Segmentation by Exploiting the Redundancy of Web Information

Xipeng Qiu, ChaoChao Huang and Xuanjing Huang

Fast High-Accuracy Part-of-Speech Tagging by Independent Classifiers

Robert Moore

Morfessor FlatCat: An HMM-Based Method for Unsupervised and Semi-Supervised Learning of Morphology

Stig-Arne Grönroos, Sami Virpioja, Peter Smit and Mikko Kurimo

Japanese Word Reordering Integrated with Dependency Parsing

Kazushi Yoshida, Tomohiro Ohno, Yoshihide Kato and Shigeki Matsubara

Query-focused Multi-Document Summarization: Combining a Topic Model with Graph-based Semi-supervised Learning

Yanran Li and Sujian Li

Ranking Multidocument Event Descriptions for Building Thematic Timelines

Kiem-Hieu Nguyen, Xavier Tannier and Véronique Moriceau

Integrating Language and Vision to Generate Natural Language Descriptions of Videos in the Wild

Jesse Thomason, Subhashini Venugopalan, Sergio Guadarrama, Kate Saenko and Raymond Mooney

Cross-Topic Authorship Attribution: Will Out-Of-Topic Data Help?

Upendra Sapkota, Thamar Solorio, Manuel Montes, Steven Bethard and Paolo Rosso

Online Gaming for Crowd-sourcing Phrase-equivalents

A Kumaran, Melissa Densmore and Shaishav Kumar

Unsupervised Verb Inference from Nouns Crossing Root Boundary

Soon Gill Hong, Sin-hee Cho and Mun Yong Yi

Enriching Wikipedia's Intra-language Links by their Cross-language Transfer

Takashi Tsunakawa, Makoto Araya and Hiroyuki Kaji

Chinese Irony Corpus Construction and Ironic Structure Analysis

Yi-jie Tang and Hsin-Hsi Chen

Tuesday, August 26, 2014 (continued)

Global Methods for Cross-lingual Semantic Role and Predicate Labelling

Lonneke van der Plas, Marianna Apidianaki and Chenhua Chen

Multilingual Semantic Parsing : Parsing Multiple Languages into Semantic Representations

Zhanming Jie and Wei Lu

Unsupervised Word Sense Induction using Distributional Statistics

Kartik Goyal and Eduard Hovy

Group based Self Training for E-Commerce Product Record Linkage

Xin Zhao, Yuexin Wu, Hongfei Yan and Xiaoming Li

Reducing Over-Weighting in Supervised Term Weighting for Sentiment Analysis

Haibing Wu and Xiaodong Gu

Sentiment Classification with Graph Co-Regularization

Guangyou Zhou, Jun Zhao and Daojian Zeng

Hybrid Deep Belief Networks for Semi-supervised Sentiment Classification

Shusen Zhou, Qingcai Chen, Xiaolong Wang and Xiaoling Li

Latent Dynamic Model with Category Transition Constraint for Opinion Classification

Takeshi Kobayakawa

Sentence Compression for Target-Polarity Word Collocation Extraction

Yanyan Zhao, Wanxiang Che, Honglei Guo, Bing Qin, Zhong Su and Ting Liu

15:15-15:45 Coffee Break

Tuesday, August 26, 2014 (continued)

Session Tu41: (15:45-17:25) Syntax, Grammar Induction, Syntactic and Semantic Parsing II

- 15:45–16:10 *Hybrid Grammars for Discontinuous Parsing*
Mark-Jan Nederhof and Heiko Vogler
- 16:10–16:35 *From neighborhood to parenthood: the advantages of dependency representation over bigrams in Brown clustering*
Simon Suster and Gertjan van Noord
- 16:35–17:00 *An Empirical Evaluation of Automatic Conversion from Constituency to Dependency in Hungarian*
Katalin Ilona Simkó, Veronika Vincze, Zsolt Szántó and Richárd Farkas
- 17:00–17:25 *Deep-Syntactic Parsing*
Miguel Ballesteros, Bernd Bohnet, Simon Mille and Leo Wanner

Session Tu42: (15:45-17:25) Semantic Processing, Distributional Semantics and Compositional Semantics II

- 15:45–16:10 *Modeling Newswire Events using Neural Networks for Anomaly Detection*
Pradeep Dasigi and Eduard Hovy
- 16:10–16:35 *Million-scale Derivation of Semantic Relations from a Manually Constructed Predicate Taxonomy*
Motoki Sano, Kentaro Torisawa, Julien Kloetzer, Chikara Hashimoto, István Varga and Jong-Hoon Oh
- 16:35–17:00 *Combining Supervised and Unsupervised Parsing for Distributional Similarity*
Martin Riedl, Irina Alles and Chris Biemann
- 17:00–17:25 *A Markovian approach to distributional semantics with application to semantic compositionality*
Edouard Grave, Guillaume Obozinski and Francis Bach

Tuesday, August 26, 2014 (continued)

Session Tu43: (15:45-17:25) Applications II

- 15:45–16:10 *A Beam-Search Decoder for Disfluency Detection*
Xuancong Wang, Hwee Tou Ng and Khe Chai Sim
- 16:10–16:35 *Single Document Keyphrase Extraction Using Label Information*
Sumit Negi
- 16:35–17:00 *Predicting Interesting Things in Text*
Michael Gamon, Arjun Mukherjee and Patrick Pantel
- 17:00–17:25 *Context Dependent Claim Detection*
Ran Levy, Yonatan Bilu, Daniel Hershcovich, Ehud Aharoni and Noam Slonim

Session Tu44: (15:45-17:25) Language Resources

- 15:45–16:10 *Annotating Argument Components and Relations in Persuasive Essays*
Christian Stab and Iryna Gurevych
- 16:10–16:35 *Building a Hierarchically Aligned Chinese-English Parallel Treebank*
Dun Deng and Nianwen Xue
- 16:35–17:00 *3arif: A Corpus of Modern Standard and Egyptian Arabic Tweets Annotated for Epistemic Modality Using Interactive Crowdsourcing*
Rania Al-Sabbagh, Roxana Girju and Jana Diesner
- 17:00–17:25 *Empirical Analysis of Aggregation Methods for Collective Annotation*
Ciyang Qing, Ulle Endriss, Raquel Fernandez and Justin Kruger

Wednesday, August 27, 2014

Full Day Excursions

Thursday, August 28, 2014

Session Th11: (09:00-10:15) Invited Talk 3

09:00–10:15 *Annotation Adaptation and Language Adaptation in NLP*
Qun Liu

10:15-10:45 Coffee Break

Session Th21: (10:45-12:25) IE/Database Linking I

10:45–11:10 *Collective Named Entity Disambiguation using Graph Ranking and Clique Partitioning Approaches*
Ayman Alhelbawy and Robert Gaizauskas

11:10–11:35 *Analysis and Refinement of Temporal Relation Aggregation*
Taylor Cassidy and Heng Ji

11:35–12:00 *The Wisdom of Minority: Unsupervised Slot Filling Validation based on Multi-dimensional Truth-Finding*
Dian Yu, Hongzhao Huang, Taylor Cassidy, Heng Ji, Chi Wang, Shi Zhi, Jiawei Han, Clare Voss and Malik Magdon-Ismail

12:00–12:25 *Common Space Embedding of Primal-Dual Relation Semantic Spaces*
Hidekazu Oiwa and Jun'ichi Tsujii

Thursday, August 28, 2014 (continued)

Session Th22: (10:45-12:25) Lexical Semantics and Ontologies I

- 10:45–11:10 *An Enhanced Lesk Word Sense Disambiguation Algorithm through a Distributional Semantic Model*
Pierpaolo Basile, Annalina Caputo and Giovanni Semeraro
- 11:10–11:35 *Word Sense Induction Using Lexical Chain based Hypergraph Model*
Tao Qian, Donghong Ji, Mingyao Zhang, Chong Teng and Congling Xia
- 11:35–12:00 *Minimally Supervised Classification to Semantic Categories using Automatically Acquired Symmetric Patterns*
Roy Schwartz, Roi Reichart and Ari Rappoport
- 12:00–12:25 *Novel Word-sense Identification*
Paul Cook, Jey Han Lau, Diana McCarthy and Timothy Baldwin

Session Th23: (10:45-12:25) Natural Language Generation and Summarization I

- 10:45–11:10 *Learning to Summarise Related Sentences*
Emmanouil Tzouridis, Jamal Nasir and Ulf Brefeld
- 11:10–11:35 *Learning to Generate Coherent Summary with Discriminative Hidden Semi-Markov Model*
Hitoshi Nishikawa, Kazuho Arita, Katsumi Tanaka, Tsutomu Hirao, Toshiro Makino and Yoshihiro Matsuo
- 11:35–12:00 *Query-Focused Opinion Summarization for User-Generated Content*
Lu Wang, Hema Raghavan, Claire Cardie and Vittorio Castelli
- 12:00–12:25 *Generating Supplementary Travel Guides from Social Media*
Liu Yang, Jing Jiang, Lifu Huang, Minghui Qiu and Lizi Liao

Thursday, August 28, 2014 (continued)

Session Th24: (10:45-12:25) Modeling of Discourse and Dialogue IV and Multimodal Processing

- 10:45–11:10 *Ensemble-Based Medical Relation Classification*
Jennifer D’Souza and Vincent Ng
- 11:10–11:35 *Combining Natural and Artificial Examples to Improve Implicit Discourse Relation Identification*
Chloé Braud and Pascal Denis
- 11:35–12:00 *Reinforcement Learning of Cooperative Persuasive Dialogue Policies using Framing*
Takuya Hiraoka, Graham Neubig, Sakriani Sakti, Tomoki Toda and Satoshi Nakamura
- 12:00–12:25 *Towards multimodal modeling of physicians’ diagnostic confidence and self-awareness using medical narratives*
Joseph Bullard, Cecilia Ovesdotter Alm, Qi Yu, Pengcheng Shi and Anne Haake
- 12:25-14:00 Lunch Break

Session Th31: (14:00-15:15) Semantic Processing, Distributional Semantics and Compositional Semantics III

- 14:00–14:25 *Towards Semantic Validation of a Derivational Lexicon*
Britta Zeller, Sebastian Padó and Jan Šnajder
- 14:25–14:50 *Detecting Learner Errors in the Choice of Content Words Using Compositional Distributional Semantics*
Ekaterina Kochmar and Ted Briscoe
- 14:50–15:15 *A Novel Distributional Approach to Multilingual Conceptual Metaphor Recognition*
Michael Mohler, Bryan Rink, David Bracewell and Marc Tomlinson

Thursday, August 28, 2014 (continued)

Session Th32: (14:00-15:15) Morphology, Word Segmentation, Tagging and Chunking I

- 14:00–14:25 *Part of Speech Tagging for French Social Media Data*
Farhad Nooralahzadeh, Caroline Brun and Claude Roux
- 14:25–14:50 *Morphological Analysis for Japanese Noisy Text based on Character-level and Word-level Normalization*
Itsumi Saito, Kugatsu Sadamitsu, Hisako Asano and Yoshihiro Matsuo
- 14:50–15:15 *Adapting taggers to Twitter with not-so-distant supervision*
Barbara Plank, Dirk Hovy, Ryan McDonald and Anders Søgaard

Session Th33: (14:00-15:15) Speech Recognition, Text-To-Speech, Spoken Language Understanding

- 14:00–14:25 *Interpolated Dirichlet Class Language Model for Speech Recognition Incorporating Long-distance N-grams*
Md. Akmal Haidar and Douglas O’Shaughnessy
- 14:25–14:50 *Situated Incremental Natural Language Understanding using a Multimodal, Linguistically-driven Update Model*
Casey Kennington, Spyros Kousidis and David Schlangen
- 14:50–15:15 *Quality Estimation for Automatic Speech Recognition*
Matteo Negri, Marco Turchi, José G. C. de Souza and Falavigna Daniele

Session Th34: (14:00-15:15) Lesser Resourced Languages

- 14:00–14:25 *A Generic Anaphora Resolution Engine for Indian Languages*
Sobha Lalitha Devi, Vijay Sundar Ram and Pattabhi RK Rao
- 14:25–14:50 *Converting Phrase Structures to Dependency Structures in Sanskrit*
Pawan Goyal and Amba Kulkarni
- 14:50–15:15 *Uncertainty Detection in Hungarian Texts*
Veronika Vincze
- 15:15-15:45 Coffee Break

Thursday, August 28, 2014 (continued)

Session Th41: (15:45-17:25) Syntax, Grammar Induction, Syntactic and Semantic Parsing III

- 15:45–16:10 *Rediscovering Annotation Projection for Cross-Lingual Parser Induction*
Jörg Tiedemann
- 16:10–16:35 *Synchronous Constituent Context Model for Inducing Bilingual Synchronous Structures*
Xiangyu Duan, Min Zhang and Qiaoming Zhu
- 16:35–17:00 *Syntactic Parsing and Compound Recognition via Dual Decomposition: Application to French*
Joseph Le Roux, Antoine Rozenknop and Matthieu Constant
- 17:00–17:25 *Learning the Taxonomy of Function Words for Parsing*
Dongchen Li, Xiantao Zhang, Dingsheng Luo and Xihong Wu

Session Th42: (15:45-17:25) Machine Translation I

- 15:45–16:10 *A Neural Reordering Model for Phrase-based Translation*
Peng Li, Yang Liu, Maosong Sun, Tatsuya Izuha and Dakun Zhang
- 16:10–16:35 *Recurrent Neural Network-based Tuple Sequence Model for Machine Translation*
Youzheng Wu, Taro Watanabe and Chiori Hori
- 16:35–17:00 *Class-Based Language Modeling for Translating into Morphologically Rich Languages*
Arianna Bisazza and Christof Monz
- 17:00–17:25 *Latent Domain Translation Models in Mix-of-Domains Haystack*
Cuong Hoang and Khalil Sima'an

Thursday, August 28, 2014 (continued)

Session Th43: (15:45-17:25) Linguistic and Cognitive Issues in CL and NLP I

- 15:45–16:10 *Language Family Relationship Preserved in Non-native English*
Ryo Nagata
- 16:10–16:35 *Why Gender and Age Prediction from Tweets is Hard: Lessons from a Crowdsourcing Experiment*
Dong Nguyen, Dolf Trieschnigg, A. Seza Dođruöz, Rilana Gravel, Mariet Theune, Theo Meder and Franciska De Jong
- 16:35–17:00 *Exploring Syntactic Features for Native Language Identification: A Variationist Perspective on Feature Encoding and Ensemble Optimization*
Serhiy Bykh and Detmar Meurers
- 17:00–17:25 *Applying automatically parsed corpora to the study of language variation*
Jelke Bloem, Arjen Versloot and Fred Weerman

Session Th44: (15:45-17:25) Natural Language Generation and Summarization II and Paraphrasing

- 15:45–16:10 *Empirical analysis of exploiting review helpfulness for extractive summarization of online reviews*
Wenting Xiong and Diane Litman
- 16:10–16:35 *Lexico-syntactic text simplification and compression with typed dependencies*
Mandya Angrosh, Tadashi Nomoto and Advait Siddharthan
- 16:35–17:00 *Learning when to point: A data-driven approach*
Albert Gatt and Patrizia Paggio
- 17:00–17:25 *Generating Acrostics via Paraphrasing and Heuristic Search*
Benno Stein, Matthias Hagen and Christof Bräutigam

Friday, August 29, 2014

Session Fr11: (09:00-10:15) Invited Talk 4

09:00–10:15 *Does a Computational Linguist have to be a Linguist?*
Martin Kay

10:15-10:45 Coffee Break

Session Fr21: (10:45-12:25) Machine Translation II

10:45–11:10 *Query Lattice for Translation Retrieval*
Meiping Dong, Yong Cheng, Yang Liu, Jia Xu, Maosong Sun, Tatsuya Izuha and Jie Hao

11:10–11:35 *RED: A Reference Dependency Based MT Evaluation Metric*
Hui Yu, Xiaofeng Wu, Jun Xie, Wenbin Jiang, Qun Liu and Shouxun Lin

11:35–12:00 *Quality Estimation of English-French Machine Translation: A Detailed Study of the Role of Syntax*
Rasoul Kaljahi, Jennifer Foster, Johann Roturier and Raphael Rubino

12:00–12:25 *Effective Incorporation of Source Syntax into Hierarchical Phrase-based Translation*
Tong Xiao, Adrià de Gispert, Jingbo Zhu and Bill Byrne

Session Fr22: (10:45-12:25) IE/Database Linking II

10:45–11:10 *BEL: Bagging for Entity Linking*
Zhe Zuo, Gjergji Kasneci, Toni Gruetze and Felix Naumann

11:10–11:35 *Exploratory Relation Extraction in Large Text Corpora*
Alan Akbik, Thilo Michael and Christoph Boden

11:35–12:00 *An Analysis of Causality between Events and its Relation to Temporal Information*
Paramita Mirza and Sara Tonelli

12:00–12:25 *Exploring Fine-grained Entity Type Constraints for Distantly Supervised Relation Extraction*
Yang Liu, Kang Liu, Liheng Xu and Jun Zhao

Friday, August 29, 2014 (continued)

Session Fr23: (10:45-12:25) Linguistic and Cognitive Issues in CL and NLP II

- 10:45–11:10 *Using Collections of Human Language Intuitions to Measure Corpus Representativeness*
Reinhard Rapp
- 11:10–11:35 *Limited memory incremental coreference resolution*
Kellie Webster and James R. Curran
- 11:35–12:00 *Left-corner Transitions on Dependency Parsing*
Hiroshi Noji and Yusuke Miyao
- 12:00–12:25 *Data-driven Measurement of Child Language Development with Simple Syntactic Templates*
Shannon Lubetich and Kenji Sagae

Session Fr24: (10:45-12:25) Lexical Semantics and Ontologies II

- 10:45–11:10 *Employing Event Inference to Improve Semi-Supervised Chinese Event Extraction*
Peifeng Li, Qiaoming Zhu and Guodong Zhou
- 11:10–11:35 *Supervised Ranking of Co-occurrence Profiles for Acquisition of Continuous Lexical Attributes*
Julian Brooke and Graeme Hirst
- 11:35–12:00 *Unsupervised extraction of semantic relations using discourse cues*
Juliette Conrath, Stergos Afantenos, Nicholas Asher and Philippe Muller
- 12:00–12:25 *HARPY: Hypernyms and Alignment of Relational Paraphrases*
Adam Grycner and Gerhard Weikum
- 12:25-14:00 Lunch Break

Friday, August 29, 2014 (continued)

Session Fr31: (14:00-15:15) Machine Translation III

- 14:00–14:25 *Limitations of MT Quality Estimation Supervised Systems: The Tails Prediction Problem*
Erwan Moreau and Carl Vogel
- 14:25–14:50 *Augment Dependency-to-String Translation with Fixed and Floating Structures*
Jun Xie, Jinan Xu and Qun Liu
- 14:50–15:15 *Soft Dependency Matching for Hierarchical Phrase-based Machine Translation*
Hailong Cao, Dongdong Zhang, Ming Zhou and Tiejun Zhao

Session Fr32: (14:00-15:15) Lexical Semantics and Ontologies III

- 14:00–14:25 *Using Spreading Activation to Evaluate and Improve Ontologies*
Ronan Mac an tSaoir
- 14:25–14:50 *Learning to Distinguish Hypernyms and Co-Hyponyms*
Julie Weeds, Daoud Clarke, Jeremy Reffin, David Weir and Bill Keller
- 14:50–15:15 *"One Entity per Discourse" and "One Entity per Collocation" Improve Named-Entity Disambiguation*
Ander Barrena, Eneko Agirre, Bernardo Cabaleiro, Anselmo Peñas and Aitor Soroa

Session Fr33: (14:00-15:15) IE/Database Linking III

- 14:00–14:25 *Comparable Study of Event Extraction in Newswire and Biomedical Domains*
Makoto Miwa, Paul Thompson, Ioannis Korkontzelos and Sophia Ananiadou
- 14:25–14:50 *A Probabilistic Co-Bootstrapping Method for Entity Set Expansion*
Bei Shi, Zhenzhong Zhang, Le Sun and Xianpei Han
- 14:50–15:15 *Separating Brands from Types: an Investigation of Different Features for the Food Domain*
Michael Wiegand and Dietrich Klakow

Friday, August 29, 2014 (continued)

Session Fr34: (14:00-15:15) Morphology, Word Segmentation, Tagging and Chunking II

14:00–14:25 *Unsupervised Instance-Based Part of Speech Induction Using Probable Substitutes*

Deniz Yuret, Mehmet Ali Yatbaz and Enis Sert

14:25–14:50 *Solving Substitution Ciphers with Combined Language Models*

Bradley Hauer, Ryan Hayward and Grzegorz Kondrak

14:50–15:15 *Unsupervised Word Segmentation in Context*

Gabriel Synnaeve, Isabelle Dautriche, Benjamin Börschinger, Mark Johnson and Emmanuel Dupoux

15:15-15:45 Coffee Break

Session Fr41: (15:45-17:25) Best Paper Talks and Closing

15:45–16:15 *Relation Classification via Convolutional Deep Neural Network*

Daojian Zeng, Kang Liu, Siwei Lai, Guangyou Zhou and Jun Zhao

16:15–16:45 *A context-based model for Sentiment Analysis in Twitter*

Andrea Vanzo, Danilo Croce and Roberto Basili

16:45-17:25 Closing Ceremony

