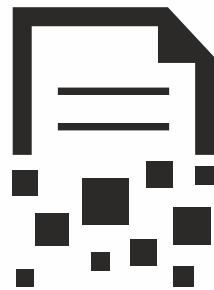


 **ACL 2013**

4 – 9 August | Sofia, Bulgaria

51st
**ANNUAL MEETING OF THE
ASSOCIATION FOR
COMPUTATIONAL LINGUISTICS**



**Proceedings of
the Conference**

VOLUME 1: Long Papers

ACL 2013

**51st Annual Meeting of the
Association for Computational Linguistics**

**Proceedings of the Conference
Volume 1: Long Papers**

August 4-9, 2013
Sofia, Bulgaria

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

PLATINUM LEVEL SPONSOR



GOLD LEVEL SPONSORS



SILVER LEVEL SPONSORS



BRONZE LEVEL SPONSORS



SUPPORTER



BEST STUDENT PAPER AWARD



STUDENT VOLUNTEER



CONFERENCE BAG SPONSOR



CONFERENCE DINNER ENTERTAINMENT SPONSOR



LOCAL ORGANIZER



©2013 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

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

ISBN 978-1-937284-50-3 (Volume 1)

Preface: General Chair

Welcome to the 51st Annual Meeting of the Association for Computational Linguistics in Sofia, Bulgaria! The first ACL meeting was held in Denver in 1963 under the name AMTCL. This makes ACL one of the longest running conferences in computer science. This year we received a record total number of 1286 submissions, which is a testament to the continued and growing importance of computational linguistics and natural language processing.

The success of an ACL conference is made possible by the dedication and hard work of many people. I thank all of them for volunteering their time and energy in service to our community.

Priscilla Rasmussen, the ACL Business Manager, and Graeme Hirst, the treasurer, did most of the groundwork in selecting Sofia as the conference site, went through several iterations of planning and shouldered a significant part of the organizational work for the conference. It was my first exposure to the logistics of organizing a large event and I was surprised at how much expertise and experience is necessary to make ACL a successful meeting.

Thanks to Svetla Koeva and her team for their work on local arrangements, including social activities (Radka Vlahova, Tsvetana Dimitrova, Svetlozara Lesseva), local sponsorship (Stoyan Mihov, Rositsa Dekova), conference handbook (Nikolay Genov, Hristina Kukova), web site (Tinko Tinchev, Emil Stoyanov, Georgi Iliev), local exhibits (Maria Todorova, Ekaterina Tarpomanova), internet, wifi and equipment (Martin Yalamov, Angel Genov, Borislav Rizov) and student volunteer management (Kalina Boncheva). Perhaps most importantly, Svetla was the liaison to the professional conference organizer AIM Group, a relationship that is crucial for the success of the conference. Doing the local arrangements is a fulltime job for an extended period of time. We are lucky that we have people in our community who are willing to provide this service without compensation.

The program co-chairs Pascale Fung and Massimo Poesio selected a strong set of papers for the main conference and invited three great keynote speakers, Harald Baayen, Chantal Prat and Lars Rasmussen. Putting together the program of the top conference in our field is a difficult job and I thank Pascale and Massimo for taking on this important responsibility.

Thanks are also due to the other key members of the ACL organizing committees: Aoife Cahill and Qun Liu (workshop co-chairs); Johan Bos and Keith Hall (tutorial co-chairs); Miriam Butt and Sarmad Hussain (demo co-chairs); Steven Bethard, Preslav Nakov and Feiyu Xu (faculty advisors to the student research workshop); Anik Dey, Eva Vecchi, Sebastian Krause and Ivelina Nikolova (co-chairs of the student research workshop); Leo Wanner (mentoring chair); and Anisava Miltenova, Ivan Derzhanski and Anna Korhonen (publicity co-chairs).

I am particularly indebted to Roberto Navigli, Jing-Shin Chang and Stefano Faralli for producing the proceedings of the conference, a bigger job than usual because of the large number of submissions and the resulting large number of acceptances.

The ACL conference and the ACL organization benefit greatly from the financial support of our sponsors. We thank the platinum level sponsor, Baidu; the three gold level sponsors; the three silver level sponsors; and six bronze level sponsors. Three other sponsors took advantage of more creative options to assist us: Facebook sponsored the Student Volunteers; IBM sponsored the Best Student Paper Award; and SDL sponsored the conference bags. We are grateful for the financial support from these organizations.

Finally, I would like to express my appreciation to the area chairs, workshop organizers, tutorial presenters and reviewers for their participation and contribution.

Of course, the ACL conference is primarily held for the people who attend the conference, including the

authors. I would like to thank all of you for your participation and wish you a productive and enjoyable meeting in Sofia!

ACL 2013 General Chair
Hinrich Schuetze, University of Munich

Preface: Programme Committee Co-Chairs

Welcome to the 2013 Conference of the Association for Computational Linguistics! Our community continues to grow, and this year's conference has set a new record for paper submissions. We received 1286 submissions, which is 12% more than the previous record; we are particularly pleased to see a striking increase in the number of short papers submitted - 624, which is 21.8% higher than the previous record set in 2011.

Another encouraging trend in recent years is the increasing number of aspects of language processing, and forms of language, of interest to our community. In order to reflect this greater diversity, this year's conference has a much larger number of tracks than previous conferences, 26. Consequently, many more area chairs and reviewers were recruited than in the past, thus involving an even greater subset of the community in the selection of the program. We feel this, too, is a very positive development. We thank the area chairs and reviewers for their hard work.

A key innovation introduced this year is the presentation at the conference of sixteen papers accepted by the new ACL journal, Transactions of the Association for Computational Linguistics (TACL). We have otherwise maintained most of the innovations introduced in recent years, including accepting papers accompanied by supplemental materials such as corpora or software.

Another new practice this year is the presence of an industrial keynote speaker in addition to the two traditional keynote speakers. We are delighted to have as invited speakers two scholars as distinguished as Prof. Harald Baayen of Tuebingen and Alberta and Prof. Chantel Prat from the University of Wisconsin. Prof. Baayen will talk about using eye-tracking to study the semantics of compounds, an issue of great interest for work on distributional semantics. Prof. Prat will talk about research studying language in bilinguals using methods from neuroscience. The industrial keynote speaker, Dr. Lars Rasmussen from Facebook, will talk about the new graph search algorithm recently announced by the company. Last, but not least, the recipient of this year's ACL Lifetime Achievement Award will give a plenary lecture during the final day of the conference.

The list of people to thank for their contribution to this year's program is very long. First of all we wish to thank the authors who submitted top quality work to the conference; we would not have such a strong program without them, nor without the hard work of area chairs and reviewers, who enabled us to make often very difficult choices and to provide valuable feedback to the authors. As usual, Rich Gerber and the START team gave us crucial help with an amazing speed. The general conference chair Hinrich Schuetze provided valuable guidance and kept the timetable ticking along. We thank the local arrangements committee headed by Svetla Koeva, who played a key role in finalizing the program. We also thank the publication chairs, Jing-Shin Chang and Roberto Navigli, and their collaborator Stefano Faralli, who together produced this volume; and Priscilla Rasmussen, Drago Radev and Graeme Hirst, who provided enormously useful guidance and support. Finally, we wish to thank previous program chairs, and in particular John Carroll, Stephen Clark, and Jian Su, for their insight on the process.

We hope you will be as pleased as we are with the result and that you'll enjoy the conference in Sofia this Summer.

ACL 2013 Program Co-Chairs

Pascale Fung, Hong-Kong University of Science and Technology

Massimo Poesio, University of Essex

Organizing Committee

General Chair:

Hinrich Schuetze, University of Munich

Program Co-Chairs:

Pascale Fung, The Hong Kong University of Science and Technology
Massimo Poesio, University of Essex

Local Chair:

Svetla Koeva, Bulgarian Academy of Sciences

Workshop Co-Chairs:

Aoife Cahill, Educational Testing Service
Qun Liu, Dublin City University & Chinese Academy of Sciences

Tutorial Co-Chairs:

Johan Bos, University of Groningen
Keith Hall, Google

Demo Co-Chairs:

Miriam Butt, University of Konstanz
Sarmad Hussain, Al-Khawarizmi Institute of Computer Science

Publication Chairs:

Roberto Navigli, Sapienza University of Rome (Chair)
Jing-Shin Chang, National Chi Nan University (Co-Chair)
Stefano Faralli, Sapienza University of Rome

Faculty Advisors (Student Research Workshop):

Steven Bethard, University of Colorado Boulder & KU Leuven
Preslav I. Nakov, Qatar Computing Research Institute
Feiyu Xu, DFKI, German Research Center for Artificial Intelligence

Student Chairs (Student Research Workshop):

Anik Dey, The Hong Kong University of Science & Technology
Eva Vecchi, Università di Trento

Sebastian Krause, DFKI, German Research Center for Artificial Intelligence
Ivelina Nikolova, Bulgarian Academy of Sciences

Mentoring Chair:

Leo Wanner, Universitat Pompeu Fabra

Publicity Co-Chairs:

Anisava Miltenova, Bulgarian Academy of Sciences
Ivan Derzhanski, Bulgarian Academy of Sciences
Anna Korhonen, University of Cambridge

Business Manager:

Priscilla Rasmussen, ACL

Area Chairs:

Frank Keller, University of Edinburgh
Roger Levy, UC San Diego
Amanda Stent, AT&T
David Suendermann, DHBW, Stuttgart, Germany
Andrew Kehler, UC San Diego
Becky Passonneau, Columbia
Hang Li, Huawei Technologies
Nancy Ide, Vassar
Piek Vossen, Freie Universitat Amsterdam
Philipp Cimiano, University of Bielefeld
Sabine Schulte im Walde, University of Stuttgart
Dekang Lin, Google
Chiori Hori, NICT, Japan
Keh-Yih Su, Behavior Design Corporation
Roland Kuhn, NRC
Dekai Wu, HKUST
Benjamin Snyder, University of Wisconsin-Madison
Thamar Solorio, University of Texas-Dallas
Ehud Reiter, University of Aberdeen
Massimiliano Ciaramita, Google
Ken Church, IBM
Carlo Strapparava, FBK
Tomaz Erjavec, Jožef Stefan Institute
Adam Przepiorkowski, Polish Academy of Sciences
Patrick Pantel, Microsoft
Owen Rambow, Columbia
Chris Dyer, CMU
Jason Eisner, Johns Hopkins
Jennifer Chu-Carroll, IBM
Bernardo Magnini, FBK
Lluis Marquez, Universitat Politecnica de Catalunya

Alessandro Moschitti, University of Trento
Claire Cardie, Cornell
Rada Mihalcea, University of North Texas
Dilek Hakkani-Tur, Microsoft
Walter Daelemans, University of Antwerp
Dan Roth, University of Illinois Urbana Champaign
Alex Koller, University of Potsdam
Ani Nenkova, University of Pennsylvania
Jamie Henderson, XRCE
Sadao Kurohashi, University of Kyoto
Yuji Matsumoto, Nara Institute of S&T
Heng Ji, CUNY
Marie-Francine Moens, KU Leuven
Hwee Tou Ng, NU Singapore

Program Committee:

Abend Omri, Abney Steven, Abu-Jbara Amjad, Agarwal Apoorv, Agirre Eneko, Aguado-de-Cea Guadalupe, Ahrenberg Lars, Akkaya Cem, Alfonseca Enrique, Alishahi Afra, Allauzen Alexander, Altun Yasemin, Androutsopoulos Ion, Araki Masahiro, Artiles Javier, Artzi Yoav, Asahara Masayuki, Asher Nicholas, Atserias Batalla Jordi, Attardi Giuseppe, Ayan Necip Fazil

Baker Collin, Baldridge Jason, Baldwin Timothy, Banchs Rafael E., Banea Carmen, Bangalore Srinivas, Baroni Marco, Barrault Loïc, Barreiro Anabela, Basili Roberto, Bateman John, Bechet Frederic, Beigman Klebanov Beata, Bel Núria, Benajiba Yassine, Bender Emily M., Bendersky Michael, Benotti Luciana, Bergler Sabine, Besacier Laurent, Bethard Steven, Bicknell Clinton, Biemann Chris, Bikel Dan, Birch Alexandra, Bisazza Arianna, Blache Philippe, Bloodgood Michael, Bod Rens, Boitet Christian, Bojar Ondrej, Bond Francis, Bontcheva Kalina, Bordino Ilaria, Bosch Sonja, Boschee Elizabeth, Botha Jan, Bouma Gosse, Boye Johan, Boyer Kristy, Bracewell David, Branco António, Brants Thorsten, Brew Chris, Briscoe Ted, Bu Fan, Buitelaar Paul, Bunesco Razvan, Busemann Stephan, Byrne Bill, Byron Donna

Cabrio Elena, Cahill Aoife, Cahill Lynne, Callison-Burch Chris, Calzolari Nicoletta, Campbell Nick, Cancedda Nicola, Cao Hailong, Caragea Cornelia, Carberry Sandra, Cardenosa Jesus, Cardie Claire, Carl Michael, Carpuat Marine, Carreras Xavier, Carroll John, Casacuberta Francisco, Caselli Tommaso, Cassidy Steve, Cassidy Taylor, Celikyilmaz Asli, Cerisara Christophe, Chambers Nate, Chang Jason, Chang Kai-Wei, Chang Ming-Wei, Chang Jing-Shin, Chelba Ciprian, Chen Wenliang, Chen Zheng, Chen Wenliang, Chen John, Chen Boxing, Chen David, Cheng Pu-Jen, Cherry Colin, Chiang David, Choi Yejin, Choi Key-Sun, Christodoulopoulos Christos, Chrupala Grzegorz, Chu-Carroll Jennifer, Clark Stephen, Clark Peter, Cohn Trevor, Collier Nigel, Conroy John, Cook Paul, Coppola Bonaventura, Corazza Anna, Core Mark, Costa-jussà Marta R., Cristea Dan, Croce Danilo, Culotta Aron, da Cunha Iria

Daelemans Walter, Dagan Ido, Daille Beatrice, Danescu-Niculescu-Mizil Cristian, Dang Hoa Trang, Danlos Laurence, Das Dipanjan, de Gispert Adrià, De La Clergerie Eric, de Marneffe Marie-Catherine, de Melo Gerard, Declerck Thierry, Delmonte Rodolfo, Demberg Vera, DeNero John, Deng Hongbo, Denis Pascal, Deoras Anoop, DeVault David, Di Eugenio Barbara, Di Fabrizio Giuseppe, Diab Mona, Diaz de Ilarraza Arantza, Diligenti Michelangelo, Dinarelli Marco, Dipper Stefanie, Do Quang, Downey Doug, Dragut Eduard, Dreyer Markus, Du Jinhua, Duh Kevin, Dymetman Marc

Eberle Kurt, Eguchi Koji, Eisele Andreas, Elhadad Michael, Erk Katrin, Esuli Andrea, Evert Stefan

Fader Anthony, Fan James, Fang Hui, Favre Benoit, Fazly Afsaneh, Federico Marcello, Feldman Anna, Feldman Naomi, Fellbaum Christiane, Feng Junlan, Fernandez Raquel, Filippova Katja, Finch Andrew, Fišer Darja, Fleck Margaret, Forcada Mikel, Foster Jennifer, Foster George, Frank Stella, Frank Stefan L., Frank Anette, Fraser Alexander

Gabrilovich Evgeniy, Gaizauskas Robert, Galley Michel, Gamon Michael, Ganitkevitch Juri, Gao Jianfeng, Gardent Claire, Garrido Guillermo, Gatt Albert, Gavrilidou Maria, Georgila Kallirroi, Gesmundo Andrea, Gildea Daniel, Gill Alastair, Gillenwater Jennifer, Gillick Daniel, Girju Roxana, Giuliano Claudio, Gliozzo Alfio, Goh Chooi-Ling, Goldberg Yoav, Goldwasser Dan, Goldwater Sharon, Gonzalo Julio, Grau Brigitte, Green Nancy, Greene Stephan, Grefenstette Gregory, Grishman Ralph, Guo Jiafeng, Gupta Rahul, Gurevych Iryna, Gustafson Joakim, Guthrie Louise, Gutiérrez Yoan

Habash Nizar, Hachey Ben, Haddow Barry, Hahn Udo, Hall David, Harabagiu Sanda, Hardmeier Christian, Hashimoto Chikara, Hayashi Katsuhiko, He Xiaodong, He Zhongjun, Heid Uli, Heinz Jeffrey, Henderson John, Hendrickx Iris, Hermjakob Ulf, Hirst Graeme, Hoang Hieu, Hockenmaier Julia, Hoffart Johannes, Hopkins Mark, Horak Ales, Hori Chiori, Hoste Veronique, Hovy Eduard, Hsieh Shu-Kai, Hsu Wen-Lian, Huang Xuanjing, Huang Minlie, Huang Liang, Huang Chu-Ren, Huang Xuanjing, Huang Liang, Huang Fei, Hwang Mei-Yuh

Iglesias Gonzalo, Ikbal Shajith, Ilisei Iustina, Inkpen Diana, Isabelle Pierre, Isahara Hitoshi, Ittycheriah Abe

Jaeger T. Florian, Jagarlamudi Jagadeesh, Jiampojarn Sittichai, Jiang Xing, Jiang Wenbin, Jiang Jing, Johansson Richard, Johnson Mark, Johnson Howard, Jurgens David

Kageura Kyo, Kan Min-Yen, Kanoulas Evangelos, Kanzaki Kyoko, Kawahara Daisuke, Keizer Simon, Kelleher John, Kempe Andre, Keshtkar Fazel, Khadivi Shahram, Kilgarriff Adam, King Tracy Holloway, Kit Chunyu, Knight Kevin, Koehn Philipp, Koeling Rob, Kolomiyets Oleksandr, Komatani Kazunori, Kondrak Grzegorz, Kong Fang, Kopp Stefan, Koppel Moshe, Kordoni Valia, Kozareva Zornitsa, Kozhevnikov Mikhail, Krahmer Emiel, Kremer Gerhard, Kudo Taku, Kuhlmann Marco, Kuhn Roland, Kumar Shankar, Kundu Gourab, Kurland Oren

Lam Wai, Lamar Michael, Lambert Patrik, Langlais Phillippe, Lapalme Guy, Lapata Mirella, Laws Florian, Leacock Claudia, Lee Yoong Keok, Lee Lin-shan, Lee Gary Geunbae, Lee Yoong Keok, Lee Sungjin, Lee John, Lefevre Fabrice, Lemon Oliver, Lenci Alessandro, Leong Ben, Leusch Gregor, Levenberg Abby, Levy Roger, Li Linlin, Li Fangtao, Li Yan, Li Haibo, Li Wenjie, Li Shoushan, Li Qi, Li Haizhou, Li Tao, Liao Shasha, Lin Dekang, Lin Ziheng, Lin Hui, Lin Ziheng, Lin Thomas, Litvak Marina, Liu Yang, Liu Bing, Liu Qun, Liu Ting, Liu Fei, Liu Zhiyuan, Liu Yiqun, Liu Chang, Liu Zhiyuan, Liu Jingjing, Liu Yiqun, Ljubešić Nikola, Lloret Elena, Lopez Adam, Lopez-Cozar Ramon, Louis Annie, Lu Wei, Lu Xiaofei, Lu Yue, Luca Dini, Luo Xiaoqiang, Lv Yajuan

Ma Yanjun, Macherey Wolfgang, Macherey Klaus, Madnani Nitin, Maegaard Bente, Magnini Bernardo, Maier Andreas, Manandhar Suresh, Marcu Daniel, Markantonatou Stella, Markert Katja, Marsi Erwin, Martin James H., Martinez David, Mason Rebecca, Matsubara Shigeki, Matsumoto

Yuji, Matsuzaki Takuya, Mauro Cettolo, Mauser Arne, May Jon, Mayfield James, Maynard Diana, McCarthy Diana, McClosky David, McCoy Kathy, McCrae John Philip, McNamee Paul, Meij Edgar, Mejova Yelena, Mellish Chris, Merlo Paola, Metze Florian, Metzler Donald, Meyers Adam, Mi Haitao, Mihalcea Rada, Miltsakaki Eleni, Minkov Einat, Mitchell Margaret, Miyao Yusuke, Mochihashi Daichi, Moens Marie-Francine, Mohammad Saif, Moilanen Karo, Monson Christian, Montes Manuel, Monz Christof, Moon Taesun, Moore Robert, Morante Roser, Morarescu Paul, Mueller Thomas, Munteanu Dragos, Murawaki Yugo, Muresan Smaranda, Myaeng Sung-Hyon, Mylonakis Markos

Nakagawa Tetsuji, Nakano Mikio, Nakazawa Toshiaki, Nakov Preslav, Naradowsky Jason, Naseem Tahira, Nastase Vivi, Navarro Borja, Navigli Roberto, Nazarenko Adeline, Nederhof Mark-Jan, Negri Matteo, Nenkova Ani, Neubig Graham, Neumann Guenter, Ng Vincent, Ngai Grace, Nguyen ThuyLinh, Nivre Joakim, Nowson Scott

Och Franz, Odijk Jan, Oflazer Kemal, Oh Jong-Hoon, Okazaki Naoaki, Oltramari Alessandro, Orasan Constantin, Osborne Miles, Osenova Petya, Ott Myle, Ovesdotter Alm Cecilia

Padó Sebastian, Palmer Martha, Palmer Alexis, Pang Bo, Pantel Patrick, Paraboni Ivandre, Pardo Thiago, Paris Cecile, Paroubek Patrick, Patwardhan Siddharth, Paul Michael, Paulik Matthias, Pearl Lisa, Pedersen Ted, Pedersen Bolette, Pedersen Ted, Peñas Anselmo, Penn Gerald, Perez-Rosas Veronica, Peters Wim, Petrov Slav, Petrovic Sasa, Piasecki Maciej, Pighin Daniele, Pinkal Manfred, Piperidis Stelios, Piskorski Jakub, Pitler Emily, Plank Barbara, Ponzetto Simone Paolo, Popescu Octavian, Popescu-Belis Andrei, Popović Maja, Potts Christopher, Pradhan Sameer, Prager John, Prasad Rashmi, Prószyński Gábor, Pulman Stephen, Punyakanok Vasin, Purver Matthew, Pustejovsky James

Qazvinian Vahed, Qian Xian, Qu Shaolin, Quarteroni Silvia, Quattoni Ariadna, Quirk Chris

Raaijmakers Stephan, Rahman Altaf, Rambow Owen, Rao Delip, Rappoport Ari, Ravi Sujith, Rayner Manny, Recasens Marta, Regneri Michaela, Reichart Roi, Reitter David, Resnik Philip, Riccardi Giuseppe, Riedel Sebastian, Riesa Jason, Rieser Verena, Riezler Stefan, Rigau German, Ringgaard Michael, Ritter Alan, Roark Brian, Rodriguez Horacio, Rohde Hannah, Rosenberg Andrew, Rosso Paolo, Rozovskaya Alla, Rus Vasile, Rusu Delia

Sagae Kenji, Sahakian Sam, Saint-Dizier Patrick, Samdani Rajhans, Sammons Mark, Sangal Rajeev, Saraclar Murat, Sarkar Anoop, Sassano Manabu, Satta Giorgio, Saurí Roser, Scaiano Martin, Schlangen David, Schmid Helmut, Schneider Nathan, Schulte im Walde Sabine, Schwenk Holger, Segond Frederique, Seki Yohei, Sekine Satoshi, Senellart Jean, Setiawan Hendra, Severyn Aliaksei, Shanker Vijay, Sharma Dipti, Sharoff Serge, Shi Shuming, Shi Xiaodong, Shi Shuming, Shutova Ekaterina, Si Xiance, Sidner Candace, Silva Mario J., Sima'an Khalil, Simard Michel, Skantze Gabriel, Small Kevin, Smith Noah A., Smith Nathaniel, Smrz Pavel, Smrz Pavel, Šnajder Jan, Snyder Benjamin, Søggaard Anders, Solorio Tamar, Somasundaran Swapna, Song Yangqiu, Spotovský Valentin, Sporleder Caroline, Sprugnoli Rachele, Srikumar Vivek, Stede Manfred, Steedman Mark, Steinberger Ralf, Stevenson Mark, Stone Matthew, Stoyanov Veselin, Strube Michael, Strzalkowski Tomek, Stymne Sara, Su Keh-Yih, Su Jian, Sun Ang, Surdeanu Mihai, Suzuki Hisami, Schwartz Roy, Szpakowicz Stan, Szpektor Idan

Täckström Oscar, Takamura Hiroya, Talukdar Partha, Tatu Marta, Taylor Sarah, Tenbrink Thora, Thater Stefan, Tiedemann Jörg, Tillmann Christoph, Titov Ivan, Toivonen Hannu, Tokunaga Takenobu,

Tonelli Sara, Toutanova Kristina, Tsarfaty Reut, Tsochantaridis Ioannis, Tsujii Jun'ichi, Tsukada Hajime, Tsuruoka Yoshimasa, Tufis Dan, Tur Gokhan, Turney Peter, Tymoshenko Kateryna

Uchimoto Kiyotaka, Udupa Raghavendra, Uryupina Olga, Utiyama Masao

Valitutti Alessandro, van den Bosch Antal, van der Plas Lonneke, Van Durme Benjamin, van Genabith Josef, Van Huyssteen Gerhard, van Noord Gertjan, Vandeghinste Vincent, Veale Tony, Velardi Paola, Verhagen Marc, Vetulani Zygmunt, Viethen Jette, Vieu Laure, Vilar David, Villavicencio Aline, Virpioja Sami, Voorhees Ellen, Vossen Piek, Vulić Ivan

Walker Marilyn, Wan Stephen, Wan Xiaojun, Wang Lu, Wang Chi, Wang Jun, Wang Haifeng, Wang Mengqiu, Wang Quan, Wang Wen, Ward Nigel, Washtell Justin, Watanabe Taro, Webber Bonnie, Wei Furu, Welty Chris, Wen Zhen, Wen Ji-Rong, Wen Zhen, Wicentowski Rich, Widdows Dominic, Wiebe Jan, Williams Jason, Wilson Theresa, Wintner Shuly, Wong Kam-Fai, Woodsend Kristian, Wooters Chuck, Wu Xianchao

Xiao Tong, Xiong Deyi, Xu Wei, Xu Jun, Xue Nianwen, Xue Xiaobing

Yan Rui, Yang Muyun, Yang Bishan, Yangarber Roman, Yano Tae, Yao Limin, Yates Alexander, Yatskar Mark, Yih Wen-tau, Yli-Jyrä Anssi, Yu Bei, Yvon François

Zabokrtsky Zdenek, Zanzotto Fabio Massimo, Zens Richard, Zettlemoyer Luke, Zeyrek Deniz, Zhang Yue, Zhang Min, Zhang Ruiqiang, Zhang Hao, Zhang Yue, Zhang Hui, Zhang Yi, Zhang Joy Ying, Zhanyi Liu, Zhao Hai, Zhao Tiejun, Zhao Jun, Zhao Shiqi, Zheng Jing, Zhou Guodong, Zhou Ming, Zhou Ke, Zhou Guodong, Zhou Ming, Zhou Guodong, Zhu Jingbo, Zhu Xiaodan, Zock Michael, Zukerman Ingrid, Zweigenbaum Pierre.

Invited Talk

When parsing makes things worse: An eye-tracking study of English compounds

Harald Baayen

Seminar für Sprachwissenschaft, Eberhard Karls University, Tuebingen

Abstract

Compounds differ in the degree to which they are semantically compositional (compare, e.g., "carwash", "handbag", "beefcake" and "humbug"). Since even relatively transparent compounds such as "carwash" may leave the uninitiated reader with uncertainty about the intended meaning (soap for washing cars? a place where you can get your car washed?), an efficient way of retrieving the meaning of a compound is to use the compound's form as an access key for its meaning.

However, in psychology, the view has become popular that at the earliest stage of lexical processing in reading, a morpho-orthographic decomposition into morphemes would necessarily take place. Theorists ascribing to obligatory decomposition appear to have some hash coding scheme in mind, with the constituents providing entry points to a form of table look-up (e.g., Taft & Forster, 1976).

Leaving aside the question of whether such a hash coding scheme would be computationally efficient as well as the question how the putative morpho-orthographic representations would be learned, my presentation focuses on the details of lexical processing as revealed by an eye-tracking study of the reading of English compounds in sentences.

A careful examination of the eye-tracking record with generalized additive modeling (Wood, 2006), combined with computational modeling using naive discrimination learning (Baayen, Milin, Filipovic, Hendrix, & Marelli, 2011) revealed that how far the eye moved into the compound is co-determined by the compound's lexical distributional properties, including the cosine similarity of the compound and its head in document vector space (as measured with latent semantic analysis, Landauer & Dumais, 1997). This indicates that compound processing is initiated already while the eye is fixating on the preceding word, and that even before the eye has landed on the compound, processes discriminating the meaning of the compound from the meaning of its head have already come into play.

Once the eye lands on the compound, two very different reading signatures emerge, which critically depend on the letter trigrams spanning the morpheme boundary (e.g., "ndb" and "dba" in "handbag"). From a discrimination learning perspective, these boundary trigrams provide the crucial (and only) orthographic cues for the compound's (idiosyncratic) meaning. If the boundary trigrams are sufficiently strongly associated with the compound's meaning, and if the eye lands early enough in the word, a single fixation suffices. Within 240 ms (of which 80 ms involve planning the next saccade) the compound's meaning is discriminated well enough to proceed to the next word.

However, when the boundary trigrams are only weakly associated with the compound's meaning, multiple fixations become necessary. In this case, without the availability of the critical orthographic cues, the eye-tracking record bears witness to the cognitive system engaging not only bottom-up processes from form to meaning, but also top-down guessing processes that are informed by the a-priori probability of the head and the cosine similarities of the compound and its constituents in semantic vector space.

These results challenge theories positing obligatory decomposition with hash coding, as hash coding predicts insensitivity to semantic transparency, contrary to fact. Our results also challenge theories positing blind look-up based on compounds' orthographic forms. Although this might be computationally efficient, the eye can't help seeing parts of the whole. In summary, reality is much more complex, with deep pre-arrival parafoveal processing followed by either efficient discrimination driven by the boundary

trigrams (within 140 ms), or by an inefficient decompositional process (requiring an additional 200 ms) that seeks to make sense of the conjunction of head and modifier.

References

Baayen, R. H., Kuperman, V., Shaoul, C., Milin, P., Kliegl, R. & Ramscar, M. (submitted), Decomposition makes things worse. A discrimination learning approach to the time course of understanding compounds in reading.

Baayen, R. H., Milin, P., Filipovic Durdjevic, D., Hendrix, P. & Marelli, M. (2011), An amorphous model for morphological processing in visual comprehension based on naive discriminative learning, *Psychological Review*, 118, 3, 438-481.

Landauer, T.K. & Dumais, S.T. (1997), A Solution to Plato's Problem: The Latent Semantic Analysis theory of acquisition, induction and representation of knowledge, *Psychological Review*, 104, 2, 211-240.

Taft, M. & Forster, K. I. (1976), Lexical Storage and Retrieval of Polymorphemic and Polysyllabic Words, *Journal of Verbal Learning and Verbal Behavior*, 15, 607-620.

Wood, S. N. (2006), *Generalized Additive Models*, Chapman & Hall/CRC, New York.

Invited Talk

The Natural Language Interface of Graph Search

Lars Rasmussen

Facebook Inc

Abstract

The backbone of the Facebook social network service is an enormous graph representing hundreds of types of nodes and thousands of types of edges. Among these nodes are over 1 billion users and 250 billion photos. The edges connecting these nodes have exceeded 1 trillion and continue to grow at an incredible rate. Retrieving information from such a graph has been a formidable and exciting task. Now it is possible for you to find, in an aggregated manner, restaurants in a city that your friends have visited, or photos of people who have attended college with you, and explore many other nuanced connections between the nodes and edges in our graph given that such information is visible to you.

Graph Search Beta, launched early this year, is a personalized semantic search engine that allows users to express their intent in natural language. It seeks answers through the traversal of relevant graph edges and ranks results by various signals extracted from our data. You can find “tv shows liked by people who study linguistics“ by issuing this query verbatim and, for the entertainment value, compare the results with “tv shows liked by people who study computer science“. Our system is built to be robust to many varied inputs, such as grammatically incorrect user queries or traditional keyword searches. Our query suggestions are always constructed in natural language, expressing the precise intention interpreted by our system. This means users would know in advance whether the system has correctly understood their intent before selecting any suggestion. The system also assists users with auto-completions, demonstrating what kinds of queries it can understand.

The development of the natural language interface encountered an array of challenging problems. The grammar structure needed to incorporate semantic information in order to translate an unstructured query into a structured semantic function, and also use syntactic information to return grammatically meaningful suggestions. The system required not only the recognition of entities in a query, but also the resolution of entities to database entries based on proximity of the entity and user nodes. Semantic parsing aimed to rank potential semantics including those that may match the immediate purpose of the query along with other refinements of the original intent. The ambiguous nature of natural language led us to consider how to interpret certain queries in the most sensible way. The need for speed demanded state-of-the-art parsing algorithms tailored for our system. In this talk, I will introduce the audience to Graph Search Beta, share our experience in developing the technical components of the natural language interface, and bring up topics that may be of interesting research value to the NLP community.

Invited Talk

Individual Differences in Language and Executive Processes: How the Brain Keeps Track of Variables

Chantel S. Prat

University of Washington

Abstract

Language comprehension is a complex cognitive process which requires tracking and integrating multiple variables. Thus, it is not surprising that language abilities (e.g., reading comprehension) vary widely even in the college population, and that language and general cognitive abilities (e.g., working memory capacity) co-vary. Although it has been widely accepted that improvements in general cognitive abilities enable (or give rise to) increased linguistic skills, the fact that individuals who develop bilingually outperform monolinguals in tests of executive functioning provides evidence of a situation in which a particular language experience gives rise to improvements in general cognitive processes. In this talk, I will describe two converging lines of research investigating individual differences in working memory capacity and reading ability in monolinguals and improved executive functioning in bilinguals. Results from these investigations suggest that the functioning of the fronto-striatal loops can explain the relation between language and non-linguistic executive functioning in both populations. I then discuss evidence suggesting that this system may function to track and route “variables” into prefrontal control structures.

Table of Contents

<i>A Shift-Reduce Parsing Algorithm for Phrase-based String-to-Dependency Translation</i> Yang Liu	1
<i>Integrating Translation Memory into Phrase-Based Machine Translation during Decoding</i> Kun Wang, Chengqing Zong and Keh-Yih Su	11
<i>Training Nondeficient Variants of IBM-3 and IBM-4 for Word Alignment</i> Thomas Schoenemann	22
<i>Modelling Annotator Bias with Multi-task Gaussian Processes: An Application to Machine Translation Quality Estimation</i> Trevor Cohn and Lucia Specia	32
<i>Smoothed marginal distribution constraints for language modeling</i> Brian Roark, Cyril Allauzen and Michael Riley	43
<i>Grounded Language Learning from Video Described with Sentences</i> Haonan Yu and Jeffrey Mark Siskind	53
<i>Plurality, Negation, and Quantification: Towards Comprehensive Quantifier Scope Disambiguation</i> Mehdi Manshadi, Daniel Gildea and James Allen	64
<i>Joint Event Extraction via Structured Prediction with Global Features</i> Qi Li, Heng Ji and Liang Huang	73
<i>Language-Independent Discriminative Parsing of Temporal Expressions</i> Gabor Angeli and Jakob Uszkoreit	83
<i>Graph-based Local Coherence Modeling</i> Camille Guinaudeau and Michael Strube	93
<i>Recognizing Rare Social Phenomena in Conversation: Empowerment Detection in Support Group Chatrooms</i> Elijah Mayfield, David Adamson and Carolyn Penstein Rosé	104
<i>Decentralized Entity-Level Modeling for Coreference Resolution</i> Greg Durrett, David Hall and Dan Klein	114
<i>Chinese Parsing Exploiting Characters</i> Meishan Zhang, Yue Zhang, Wanxiang Che and Ting Liu	125
<i>A Transition-Based Dependency Parser Using a Dynamic Parsing Strategy</i> Francesco Sartorio, Giorgio Satta and Joakim Nivre	135
<i>General binarization for parsing and translation</i> Matthias Büchse, Alexander Koller and Heiko Vogler	145
<i>Distortion Model Considering Rich Context for Statistical Machine Translation</i> Isao Goto, Masao Utiyama, Eiichiro Sumita, Akihiro Tamura and Sadao Kurohashi	155
<i>Word Alignment Modeling with Context Dependent Deep Neural Network</i> Nan Yang, Shujie Liu, Mu Li, Ming Zhou and Nenghai Yu	166

<i>Microblogs as Parallel Corpora</i>	
Wang Ling, Guang Xiang, Chris Dyer, Alan Black and Isabel Trancoso	176
<i>Improved Bayesian Logistic Supervised Topic Models with Data Augmentation</i>	
Jun Zhu, Xun Zheng and Bo Zhang	187
<i>Fast and Robust Compressive Summarization with Dual Decomposition and Multi-Task Learning</i>	
Miguel Almeida and Andre Martins	196
<i>Unsupervised Transcription of Historical Documents</i>	
Taylor Berg-Kirkpatrick, Greg Durrett and Dan Klein	207
<i>Adapting Discriminative Reranking to Grounded Language Learning</i>	
Joohyun Kim and Raymond Mooney	218
<i>Universal Conceptual Cognitive Annotation (UCCA)</i>	
Omri Abend and Ari Rappoport	228
<i>Linking Tweets to News: A Framework to Enrich Short Text Data in Social Media</i>	
Weiwei Guo, Hao Li, Heng Ji and Mona Diab	239
<i>A computational approach to politeness with application to social factors</i>	
Cristian Danescu-Niculescu-Mizil, Moritz Sudhof, Dan Jurafsky, Jure Leskovec and Christopher Potts	250
<i>Modeling Thesis Clarity in Student Essays</i>	
Isaac Persing and Vincent Ng	260
<i>Translating Italian connectives into Italian Sign Language</i>	
Camillo Lugaresi and Barbara Di Eugenio	270
<i>Stop-probability estimates computed on a large corpus improve Unsupervised Dependency Parsing</i>	
David Mareček and Milan Straka	281
<i>Transfer Learning for Constituency-Based Grammars</i>	
Yuan Zhang, Regina Barzilay and Amir Globerson	291
<i>A Context Free TAG Variant</i>	
Ben Swanson, Elif Yamangil, Eugene Charniak and Stuart Shieber	302
<i>Fast and Adaptive Online Training of Feature-Rich Translation Models</i>	
Spence Green, Sida Wang, Daniel Cer and Christopher D. Manning	311
<i>Advancements in Reordering Models for Statistical Machine Translation</i>	
Minwei Feng, Jan-Thorsten Peter and Hermann Ney	322
<i>A Markov Model of Machine Translation using Non-parametric Bayesian Inference</i>	
Yang Feng and Trevor Cohn	333
<i>Scaling Semi-supervised Naive Bayes with Feature Marginals</i>	
Michael Lucas and Doug Downey	343
<i>Learning Latent Personas of Film Characters</i>	
David Bamman, Brendan O'Connor and Noah A. Smith	352

<i>Scalable Decipherment for Machine Translation via Hash Sampling</i> Sujith Ravi	362
<i>Automatic Interpretation of the English Possessive</i> Stephen Tratz and Eduard Hovy	372
<i>Is a 204 cm Man Tall or Small ? Acquisition of Numerical Common Sense from the Web</i> Katsuma Narisawa, Yotaro Watanabe, Junta Mizuno, Naoaki Okazaki and Kentaro Inui	382
<i>Probabilistic Domain Modelling With Contextualized Distributional Semantic Vectors</i> Jackie Chi Kit Cheung and Gerald Penn	392
<i>Extracting bilingual terminologies from comparable corpora</i> Ahmet Aker, Monica Paramita and Rob Gaizauskas	402
<i>The Haves and the Have-Nots: Leveraging Unlabelled Corpora for Sentiment Analysis</i> Kashyap Popat, Balamurali A.R, Pushpak Bhattacharyya and Gholamreza Haffari	412
<i>Large-scale Semantic Parsing via Schema Matching and Lexicon Extension</i> Qingqing Cai and Alexander Yates	423
<i>Fast and Accurate Shift-Reduce Constituent Parsing</i> Muhua Zhu, Yue Zhang, Wenliang Chen, Min Zhang and Jingbo Zhu	434
<i>Nonconvex Global Optimization for Latent-Variable Models</i> Matthew R. Gormley and Jason Eisner	444
<i>Parsing with Compositional Vector Grammars</i> Richard Socher, John Bauer, Christopher D. Manning and Ng Andrew Y.	455
<i>Discriminative state tracking for spoken dialog systems</i> Angeliki Metallinou, Dan Bohus and Jason Williams	466
<i>Leveraging Synthetic Discourse Data via Multi-task Learning for Implicit Discourse Relation Recognition</i> Man Lan, Yu Xu and Zhengyu Niu	476
<i>Combining Intra- and Multi-sentential Rhetorical Parsing for Document-level Discourse Analysis</i> Shafiq Joty, Giuseppe Carenini, Raymond Ng and Yashar Mehdad	486
<i>Improving pairwise coreference models through feature space hierarchy learning</i> Emmanuel Lassalle and Pascal Denis	497
<i>Feature-Based Selection of Dependency Paths in Ad Hoc Information Retrieval</i> K. Tamsin Maxwell, Jon Oberlander and W. Bruce Croft	507
<i>Coordination Structures in Dependency Treebanks</i> Martin Popel, David Mareček, Jan Štěpánek, Daniel Zeman and Zdeněk Žabokrtský	517
<i>GlossBoot: Bootstrapping Multilingual Domain Glossaries from the Web</i> Flavio De Benedictis, Stefano Faralli and Roberto Navigli	528
<i>Collective Annotation of Linguistic Resources: Basic Principles and a Formal Model</i> Ulle Endriss and Raquel Fernández	539

<i>ParGramBank: The ParGram Parallel Treebank</i>	
Sebastian Sulger, Miriam Butt, Tracy Holloway King, Paul Meurer, Tibor Laczkó, György Rákosi, Cheikh Bamba Dione, Helge Dyvik, Victoria Rosén, Koenraad De Smedt, Agnieszka Patejuk, Ozlem Cetinoglu, I Wayan Arka and Meladel Mistica	550
<i>Identifying Bad Semantic Neighbors for Improving Distributional Thesauri</i>	
Olivier Ferret	561
<i>Models of Semantic Representation with Visual Attributes</i>	
Carina Silberer, Vittorio Ferrari and Mirella Lapata	572
<i>Real-World Semi-Supervised Learning of POS-Taggers for Low-Resource Languages</i>	
Dan Garrette, Jason Mielens and Jason Baldridge	583
<i>Using subcategorization knowledge to improve case prediction for translation to German</i>	
Marion Weller, Alexander Fraser and Sabine Schulte im Walde	593
<i>Name-aware Machine Translation</i>	
Haibo Li, Jing Zheng, Heng Ji, Qi Li and Wen Wang	604
<i>Decipherment Complexity in 1:1 Substitution Ciphers</i>	
Malte Nuhn and Hermann Ney	615
<i>Non-Monotonic Sentence Alignment via Semisupervised Learning</i>	
Xiaojun Quan, Chunyu Kit and Yan Song	622
<i>Bootstrapping Entity Translation on Weakly Comparable Corpora</i>	
Taesung Lee and Seung-won Hwang	631
<i>Transfer Learning Based Cross-lingual Knowledge Extraction for Wikipedia</i>	
Zhigang Wang, Zhixing Li, Juanzi Li, Jie Tang and Jeff Z. Pan	641
<i>Bridging Languages through Etymology: The case of cross language text categorization</i>	
Vivi Nastase and Carlo Strapparava	651
<i>Creating Similarity: Lateral Thinking for Vertical Similarity Judgments</i>	
Tony Veale and Guofu Li	660
<i>Discovering User Interactions in Ideological Discussions</i>	
Arjun Mukherjee and Bing Liu	671
<i>Multilingual Affect Polarity and Valence Prediction in Metaphor-Rich Texts</i>	
Zornitsa Kozareva	682
<i>Large tagset labeling using Feed Forward Neural Networks. Case study on Romanian Language</i>	
Tiberiu Boros, Radu Ion and Dan Tufis	692
<i>Learning to lemmatise Polish noun phrases</i>	
Adam Radziszewski	701
<i>Using Conceptual Class Attributes to Characterize Social Media Users</i>	
Shane Bergsma and Benjamin Van Durme	710
<i>The Impact of Topic Bias on Quality Flaw Prediction in Wikipedia</i>	
Oliver Fersckhe, Iryna Gurevych and Marc Rittberger	721

<i>Mining Informal Language from Chinese Microtext: Joint Word Recognition and Segmentation</i> Aobo Wang and Min-Yen Kan	731
<i>Generating Synthetic Comparable Questions for News Articles</i> Oleg Rokhlenko and Idan Szpektor	742
<i>Punctuation Prediction with Transition-based Parsing</i> Dongdong Zhang, Shuangzhi Wu, Nan Yang and Mu Li	752
<i>Discriminative Learning with Natural Annotations: Word Segmentation as a Case Study</i> Wenbin Jiang, Meng Sun, Yajuan Lü, Yating Yang and Qun Liu	761
<i>Graph-based Semi-Supervised Model for Joint Chinese Word Segmentation and Part-of-Speech Tagging</i> Xiaodong Zeng, Derek F. Wong, Lidia S. Chao and Isabel Trancoso	770
<i>An Infinite Hierarchical Bayesian Model of Phrasal Translation</i> Trevor Cohn and Gholamreza Haffari	780
<i>Additive Neural Networks for Statistical Machine Translation</i> Jiemao Liu, Taro Watanabe, Eiichiro Sumita and Tiejun Zhao	791
<i>Hierarchical Phrase Table Combination for Machine Translation</i> Conghui Zhu, Taro Watanabe, Eiichiro Sumita and Tiejun Zhao	802
<i>Shallow Local Multi-Bottom-up Tree Transducers in Statistical Machine Translation</i> Fabienne Braune, Nina Seemann, Daniel Quernheim and Andreas Maletti	811
<i>Enlisting the Ghost: Modeling Empty Categories for Machine Translation</i> Bing Xiang, Xiaoqiang Luo and Bowen Zhou	822
<i>A Multi-Domain Translation Model Framework for Statistical Machine Translation</i> Rico Sennrich, Holger Schwenk and Walid Aransa	832
<i>Part-of-Speech Induction in Dependency Trees for Statistical Machine Translation</i> Akihiro Tamura, Taro Watanabe, Eiichiro Sumita, Hiroya Takamura and Manabu Okumura ...	841
<i>Statistical Machine Translation Improves Question Retrieval in Community Question Answering via Matrix Factorization</i> Guangyou Zhou, Fang Liu, Yang Liu, Shizhu He and Jun Zhao	852
<i>Improved Lexical Acquisition through DPP-based Verb Clustering</i> Roi Reichart and Anna Korhonen	862
<i>Semantic Frames to Predict Stock Price Movement</i> Boyi Xie, Rebecca J. Passonneau, Leon Wu and Germán G. Creamer	873
<i>Density Maximization in Context-Sense Metric Space for All-words WSD</i> Koichi Tanigaki, Mitsuteru Shiba, Tatsuji Munaka and Yoshinori Sagisaka	884
<i>The Role of Syntax in Vector Space Models of Compositional Semantics</i> Karl Moritz Hermann and Phil Blunsom	894
<i>Margin-based Decomposed Amortized Inference</i> Gourab Kundu, Vivek Srikumar and Dan Roth	905

<i>Semi-Supervised Semantic Tagging of Conversational Understanding using Markov Topic Regression</i> Asli Celikyilmaz, Dilek Hakkani-Tur, Gokhan Tur and Ruhi Sarikaya	914
<i>Parsing Graphs with Hyperedge Replacement Grammars</i> David Chiang, Jacob Andreas, Daniel Bauer, Karl Moritz Hermann, Bevan Jones and Kevin Knight	924
<i>Grounded Unsupervised Semantic Parsing</i> Hoifung Poon	933
<i>Automatic detection of deception in child-produced speech using syntactic complexity features</i> Maria Yancheva and Frank Rudzicz	944
<i>Sentiment Relevance</i> Christian Scheible and Hinrich Schütze	954
<i>Predicting and Eliciting Addressee’s Emotion in Online Dialogue</i> Takayuki Hasegawa, Nobuhiro Kaji, Naoki Yoshinaga and Masashi Toyoda	964
<i>Utterance-Level Multimodal Sentiment Analysis</i> Veronica Perez-Rosas, Rada Mihalcea and Louis-Philippe Morency	973
<i>Probabilistic Sense Sentiment Similarity through Hidden Emotions</i> Mitra Mohtarami, Man Lan and Chew Lim Tan	983
<i>A user-centric model of voting intention from Social Media</i> Vasileios Lampos, Daniel Preoŧiuc-Pietro and Trevor Cohn	993
<i>Using Supervised Bigram-based ILP for Extractive Summarization</i> Chen Li, Xian Qian and Yang Liu	1004
<i>Summarization Through Submodularity and Dispersion</i> Anirban Dasgupta, Ravi Kumar and Sujith Ravi	1014
<i>Subtree Extractive Summarization via Submodular Maximization</i> Hajime Morita, Ryohei Sasano, Hiroya Takamura and Manabu Okumura	1023
<i>The effect of non-tightness on Bayesian estimation of PCFGs</i> Shay B. Cohen and Mark Johnson	1033
<i>Integrating Multiple Dependency Corpora for Inducing Wide-coverage Japanese CCG Resources</i> Sumire Uematsu, Takuya Matsuzaki, Hiroki Hanaoka, Yusuke Miyao and Hideki Mima	1042
<i>Transition-based Dependency Parsing with Selectional Branching</i> Jinho D. Choi and Andrew McCallum	1052
<i>Bilingually-Guided Monolingual Dependency Grammar Induction</i> Kai Liu, Yajuan Lü, Wenbin Jiang and Qun Liu	1063
<i>Joint Word Alignment and Bilingual Named Entity Recognition Using Dual Decomposition</i> Mengqiu Wang, Wanxiang Che and Christopher D. Manning	1073
<i>Resolving Entity Morphs in Censored Data</i> Hongzhao Huang, Zhen Wen, Dian Yu, Heng Ji, Yizhou Sun, Jiawei Han and He Li	1083

<i>Learning to Extract International Relations from Political Context</i> Brendan O'Connor, Brandon M. Stewart and Noah A. Smith	1094
<i>Graph Propagation for Paraphrasing Out-of-Vocabulary Words in Statistical Machine Translation</i> Majid Razmara, Maryam Siahbani, Reza Haffari and Anoop Sarkar	1105
<i>Online Relative Margin Maximization for Statistical Machine Translation</i> Vladimir Eidelman, Yuval Marton and Philip Resnik	1116
<i>Handling Ambiguities of Bilingual Predicate-Argument Structures for Statistical Machine Translation</i> Feifei Zhai, Jiajun Zhang, Yu Zhou and Chengqing Zong	1127
<i>Reconstructing an Indo-European Family Tree from Non-native English Texts</i> Ryo Nagata and Edward Whittaker	1137
<i>Word Association Profiles and their Use for Automated Scoring of Essays</i> Beata Beigman Klebanov and Michael Flor	1148
<i>Adaptive Parser-Centric Text Normalization</i> Congle Zhang, Tyler Baldwin, Howard Ho, Benny Kimelfeld and Yunyao Li	1159
<i>A Random Walk Approach to Selectional Preferences Based on Preference Ranking and Propagation</i> Zhenhua Tian, Hengheng Xiang, Ziqi Liu and Qinghua Zheng	1169
<i>ImpAr: A Deterministic Algorithm for Implicit Semantic Role Labelling</i> Egoitz Laparra and German Rigau	1180
<i>Cross-lingual Transfer of Semantic Role Labeling Models</i> Mikhail Kozhevnikov and Ivan Titov	1190
<i>DERivBase: Inducing and Evaluating a Derivational Morphology Resource for German</i> Britta Zeller, Jan Šnajder and Sebastian Padó	1201
<i>Crowdsourcing Interaction Logs to Understand Text Reuse from the Web</i> Martin Potthast, Matthias Hagen, Michael Völske and Benno Stein	1212
<i>SPred: Large-scale Harvesting of Semantic Predicates</i> Tiziano Flati and Roberto Navigli	1222
<i>Towards Robust Abstractive Multi-Document Summarization: A Caseframe Analysis of Centrality and Domain</i> Jackie Chi Kit Cheung and Gerald Penn	1233
<i>HEADY: News headline abstraction through event pattern clustering</i> Enrique Alfonseca, Daniele Pighin and Guillermo Garrido	1243
<i>Conditional Random Fields for Responsive Surface Realisation using Global Features</i> Nina Dethlefs, Helen Hastie, Heriberto Cuayáhuitl and Oliver Lemon	1254
<i>Two-Neighbor Orientation Model with Cross-Boundary Global Contexts</i> Hendra Setiawan, Bowen Zhou, Bing Xiang and Libin Shen	1264
<i>Cut the noise: Mutually reinforcing reordering and alignments for improved machine translation</i> Karthik Visweswariah, Mitesh M. Khapra and Ananthkrishnan Ramanathan	1275

<i>Vector Space Model for Adaptation in Statistical Machine Translation</i> Boxing Chen, Roland Kuhn and George Foster	1285
<i>From Natural Language Specifications to Program Input Parsers</i> Tao Lei, Fan Long, Regina Barzilay and Martin Rinard	1294
<i>Entity Linking for Tweets</i> Xiaohua Liu, Yitong Li, Haocheng Wu, Ming Zhou, Furu Wei and Yi Lu	1304
<i>Identification of Speakers in Novels</i> Hua He, Denilson Barbosa and Grzegorz Kondrak	1312
<i>Language Acquisition and Probabilistic Models: keeping it simple</i> Aline Villavicencio, Marco Idiart, Robert Berwick and Igor Malioutov	1321
<i>A Two Level Model for Context Sensitive Inference Rules</i> Oren Melamud, Jonathan Berant, Ido Dagan, Jacob Goldberger and Idan Szpektor	1331
<i>Align, Disambiguate and Walk: A Unified Approach for Measuring Semantic Similarity</i> Mohammad Taher Pilehvar, David Jurgens and Roberto Navigli	1341
<i>Linking and Extending an Open Multilingual Wordnet</i> Francis Bond and Ryan Foster	1352
<i>FrameNet on the Way to Babel: Creating a Bilingual FrameNet Using Wiktionary as Interlingual Connection</i> Silvana Hartmann and Iryna Gurevych	1363
<i>Dirt Cheap Web-Scale Parallel Text from the Common Crawl</i> Jason R. Smith, Herve Saint-Amand, Magdalena Plamada, Philipp Koehn, Chris Callison-Burch and Adam Lopez	1374
<i>A Sentence Compression Based Framework to Query-Focused Multi-Document Summarization</i> Lu Wang, Hema Raghavan, Vittorio Castelli, Radu Florian and Claire Cardie	1384
<i>Domain-Independent Abstract Generation for Focused Meeting Summarization</i> Lu Wang and Claire Cardie	1395
<i>A Statistical NLG Framework for Aggregated Planning and Realization</i> Ravi Kondadadi, Blake Howald and Frank Schilder	1406
<i>Models of Translation Competitions</i> Mark Hopkins and Jonathan May	1416
<i>Learning a Phrase-based Translation Model from Monolingual Data with Application to Domain Adaptation</i> Jiajun Zhang and Chengqing Zong	1425
<i>SenseSpotting: Never let your parallel data tie you to an old domain</i> Marine Carpuat, Hal Daume III, Katharine Henry, Ann Irvine, Jagadeesh Jagarlamudi and Rachel Rudinger	1435
<i>BRAINSUP: Brainstorming Support for Creative Sentence Generation</i> Gözde Özbal, Daniele Pighin and Carlo Strapparava	1446

<i>Grammatical Error Correction Using Integer Linear Programming</i> Yuanbin Wu and Hwee Tou Ng	1456
<i>Text-Driven Toponym Resolution using Indirect Supervision</i> Michael Speriosu and Jason Baldridge	1466
<i>Argument Inference from Relevant Event Mentions in Chinese Argument Extraction</i> Peifeng Li, Qiaoming Zhu and Guodong Zhou	1477
<i>Fine-grained Semantic Typing of Emerging Entities</i> Ndapandula Nakashole, Tomasz Tylenda and Gerhard Weikum	1488
<i>Embedding Semantic Similarity in Tree Kernels for Domain Adaptation of Relation Extraction</i> Barbara Plank and Alessandro Moschitti	1498
<i>A joint model of word segmentation and phonological variation for English word-final /t/-deletion</i> Benjamin Börschinger, Mark Johnson and Katherine Demuth	1508
<i>Compositional-ly Derived Representations of Morphologically Complex Words in Distributional Semantics</i> Angeliki Lazaridou, Marco Marelli, Roberto Zamparelli and Marco Baroni	1517
<i>Unsupervised Consonant-Vowel Prediction over Hundreds of Languages</i> Young-Bum Kim and Benjamin Snyder	1527
<i>Improving Text Simplification Language Modeling Using Un-simplified Text Data</i> David Kauchak	1537
<i>Combining Referring Expression Generation and Surface Realization: A Corpus-Based Investigation of Architectures</i> Sina Zarrieß and Jonas Kuhn	1547
<i>Named Entity Recognition using Cross-lingual Resources: Arabic as an Example</i> Kareem Darwish	1558
<i>Beam Search for Solving Substitution Ciphers</i> Malte Nuhn, Julian Schamper and Hermann Ney	1568
<i>Social Text Normalization using Contextual Graph Random Walks</i> Hany Hassan and Arul Menezes	1577
<i>Integrating Phrase-based Reordering Features into a Chart-based Decoder for Machine Translation</i> ThuyLinh Nguyen and Stephan Vogel	1587
<i>Machine Translation Detection from Monolingual Web-Text</i> Yuki Arase and Ming Zhou	1597
<i>Paraphrase-Driven Learning for Open Question Answering</i> Anthony Fader, Luke Zettlemoyer and Oren Etzioni	1608
<i>Aid is Out There: Looking for Help from Tweets during a Large Scale Disaster</i> István Varga, Motoki Sano, Kentaro Torisawa, Chikara Hashimoto, Kiyonori Ohtake, Takao Kawai, Jong-Hoon Oh and Stijn De Saeger	1619
<i>A Bayesian Model for Joint Unsupervised Induction of Sentiment, Aspect and Discourse Representations</i> Angeliki Lazaridou, Ivan Titov and Caroline Sporleder	1630

<i>Joint Inference for Fine-grained Opinion Extraction</i> Bishan Yang and Claire Cardie	1640
<i>Linguistic Models for Analyzing and Detecting Biased Language</i> Marta Recasens, Cristian Danescu-Niculescu-Mizil and Dan Jurafsky	1650
<i>Evaluating a City Exploration Dialogue System with Integrated Question-Answering and Pedestrian Navigation</i> Srinivasan Janarthanam, Oliver Lemon, Phil Bartie, Tiphaine Dalmas, Anna Dickinson, Xingkun Liu, William Mackaness and Bonnie Webber	1660
<i>Lightly Supervised Learning of Procedural Dialog Systems</i> Svitlana Volkova, Pallavi Choudhury, Chris Quirk, Bill Dolan and Luke Zettlemoyer	1669
<i>Public Dialogue: Analysis of Tolerance in Online Discussions</i> Arjun Mukherjee, Vivek Venkataraman, Bing Liu and Sharon Meraz	1680
<i>Offspring from Reproduction Problems: What Replication Failure Teaches Us</i> Antske Fokkens, Marieke van Erp, Marten Postma, Ted Pedersen, Piek Vossen and Nuno Freire	1691
<i>Evaluating Text Segmentation using Boundary Edit Distance</i> Chris Fournier	1702
<i>Crowd Prefers the Middle Path: A New IAA Metric for Crowdsourcing Reveals Turker Biases in Query Segmentation</i> Rohan Ramanath, Monojit Choudhury, Kalika Bali and Rishiraj Saha Roy	1713
<i>Deceptive Answer Prediction with User Preference Graph</i> Fangtao Li, Yang Gao, Shuchang Zhou, Xiance Si and Decheng Dai	1723
<i>Why-Question Answering using Intra- and Inter-Sentential Causal Relations</i> Jong-Hoon Oh, Kentaro Torisawa, Chikara Hashimoto, Motoki Sano, Stijn De Saeger and Kiyonori Ohtake	1733
<i>Question Answering Using Enhanced Lexical Semantic Models</i> Wen-tau Yih, Ming-Wei Chang, Christopher Meek and Andrzej Pastusiak	1744
<i>Syntactic Patterns versus Word Alignment: Extracting Opinion Targets from Online Reviews</i> Kang Liu, Liheng Xu and Jun Zhao	1754
<i>Mining Opinion Words and Opinion Targets in a Two-Stage Framework</i> Liheng Xu, Kang Liu, Siwei Lai, Yubo Chen and Jun Zhao	1764
<i>Connotation Lexicon: A Dash of Sentiment Beneath the Surface Meaning</i> Song Feng, Jun Seok Kang, Polina Kuznetsova and Yejin Choi	1774

Conference Program

Monday August 5, 2013

(7:30 - 17:00) Registration

(9:00 - 9:30) Opening session

(9:30) Invited Talk 1: Harald Baayen

(10:30) Coffee Break

Oral Presentations

(11:00 -12:15) LP 1a

11:00 *A Shift-Reduce Parsing Algorithm for Phrase-based String-to-Dependency Translation*

Yang Liu

11:25 *Integrating Translation Memory into Phrase-Based Machine Translation during Decoding*

Kun Wang, Chengqing Zong and Keh-Yih Su

11:50 *Training Nondeficient Variants of IBM-3 and IBM-4 for Word Alignment*

Thomas Schoenemann

(11:00 -12:15) LP 1b

11:00 *Modelling Annotator Bias with Multi-task Gaussian Processes: An Application to Machine Translation Quality Estimation*

Trevor Cohn and Lucia Specia

11:25 *Smoothed marginal distribution constraints for language modeling*

Brian Roark, Cyril Allauzen and Michael Riley

11:50 *Grounded Language Learning from Video Described with Sentences*

Haonan Yu and Jeffrey Mark Siskind

Monday August 5, 2013 (continued)

(11:00 -12:15) LP 1c

- 11:00 *Plurality, Negation, and Quantification: Towards Comprehensive Quantifier Scope Disambiguation*
Mehdi Manshadi, Daniel Gildea and James Allen
- 11:25 *Joint Event Extraction via Structured Prediction with Global Features*
Qi Li, Heng Ji and Liang Huang
- 11:50 *Language-Independent Discriminative Parsing of Temporal Expressions*
Gabor Angeli and Jakob Uszkoreit

(11:00 -12:15) LP 1d

- 11:00 *Graph-based Local Coherence Modeling*
Camille Guinaudeau and Michael Strube
- 11:25 *Recognizing Rare Social Phenomena in Conversation: Empowerment Detection in Support Group Chatrooms*
Elijah Mayfield, David Adamson and Carolyn Penstein Rosé
- 11:50 *Decentralized Entity-Level Modeling for Coreference Resolution*
Greg Durrett, David Hall and Dan Klein

(11:00 -12:15) LP 1e

- 11:00 *Chinese Parsing Exploiting Characters*
Meishan Zhang, Yue Zhang, Wanxiang Che and Ting Liu
- 11:25 *A Transition-Based Dependency Parser Using a Dynamic Parsing Strategy*
Francesco Sartorio, Giorgio Satta and Joakim Nivre
- 11:50 *General binarization for parsing and translation*
Matthias Büchse, Alexander Koller and Heiko Vogler

Monday August 5, 2013 (continued)

(12:15) Lunch break

(13:45 -15:00) LP 2a

13:45 *Distortion Model Considering Rich Context for Statistical Machine Translation*
Isao Goto, Masao Utiyama, Eiichiro Sumita, Akihiro Tamura and Sadao Kurohashi

14:10 *Word Alignment Modeling with Context Dependent Deep Neural Network*
Nan Yang, Shujie Liu, Mu Li, Ming Zhou and Nenghai Yu

14:35 *Microblogs as Parallel Corpora*
Wang Ling, Guang Xiang, Chris Dyer, Alan Black and Isabel Trancoso

(13:45 -15:00) LP 2b

13:45 *Improved Bayesian Logistic Supervised Topic Models with Data Augmentation*
Jun Zhu, Xun Zheng and Bo Zhang

14:10 *Fast and Robust Compressive Summarization with Dual Decomposition and Multi-Task Learning*
Miguel Almeida and Andre Martins

14:35 *Unsupervised Transcription of Historical Documents*
Taylor Berg-Kirkpatrick, Greg Durrett and Dan Klein

(13:45 -15:00) LP 2c

13:45 *Adapting Discriminative Reranking to Grounded Language Learning*
Joohyun Kim and Raymond Mooney

14:10 *Universal Conceptual Cognitive Annotation (UCCA)*
Omri Abend and Ari Rappoport

14:35 *Linking Tweets to News: A Framework to Enrich Short Text Data in Social Media*
Weiwei Guo, Hao Li, Heng Ji and Mona Diab

Monday August 5, 2013 (continued)

(13:45 -15:00) LP 2d

- 13:45 *A computational approach to politeness with application to social factors*
Cristian Danescu-Niculescu-Mizil, Moritz Sudhof, Dan Jurafsky, Jure Leskovec and Christopher Potts
- 14:10 *Modeling Thesis Clarity in Student Essays*
Isaac Persing and Vincent Ng
- 14:35 *Translating Italian connectives into Italian Sign Language*
Camillo Lugaresi and Barbara Di Eugenio

(13:45 -15:00) LP 2e

- 13:45 *Stop-probability estimates computed on a large corpus improve Unsupervised Dependency Parsing*
David Mareček and Milan Straka
- 14:10 *Transfer Learning for Constituency-Based Grammars*
Yuan Zhang, Regina Barzilay and Amir Globerson
- 14:35 *A Context Free TAG Variant*
Ben Swanson, Elif Yamangil, Eugene Charniak and Stuart Shieber

(15:00 -16:15) LP 3a

- 15:00 *Fast and Adaptive Online Training of Feature-Rich Translation Models*
Spence Green, Sida Wang, Daniel Cer and Christopher D. Manning
- 15:25 *Advancements in Reordering Models for Statistical Machine Translation*
Minwei Feng, Jan-Thorsten Peter and Hermann Ney
- 15:50 *A Markov Model of Machine Translation using Non-parametric Bayesian Inference*
Yang Feng and Trevor Cohn

Monday August 5, 2013 (continued)

(15:00 -16:15) LP 3b

- 15:00 *Scaling Semi-supervised Naive Bayes with Feature Marginals*
Michael Lucas and Doug Downey
- 15:25 *Learning Latent Personas of Film Characters*
David Bamman, Brendan O'Connor and Noah A. Smith
- 15:50 *Scalable Decipherment for Machine Translation via Hash Sampling*
Sujith Ravi

(15:00 -16:15) LP 3c

- 15:00 *Automatic Interpretation of the English Possessive*
Stephen Tratz and Eduard Hovy
- 15:25 *Is a 204 cm Man Tall or Small ? Acquisition of Numerical Common Sense from the Web*
Katsuma Narisawa, Yotaro Watanabe, Junta Mizuno, Naoaki Okazaki and Kentaro Inui
- 15:50 *Probabilistic Domain Modelling With Contextualized Distributional Semantic Vectors*
Jackie Chi Kit Cheung and Gerald Penn

(15:00 -16:15) LP 3d

- 15:00 *Extracting bilingual terminologies from comparable corpora*
Ahmet Aker, Monica Paramita and Rob Gaizauskas
- 15:25 *The Haves and the Have-Nots: Leveraging Unlabelled Corpora for Sentiment Analysis*
Kashyap Popat, Balamurali A.R, Pushpak Bhattacharyya and Gholamreza Haffari
- 15:50 *Large-scale Semantic Parsing via Schema Matching and Lexicon Extension*
Qingqing Cai and Alexander Yates

Monday August 5, 2013 (continued)

(15:00 -16:15) LP 3e

- 15:00 *Fast and Accurate Shift-Reduce Constituent Parsing*
Muhua Zhu, Yue Zhang, Wenliang Chen, Min Zhang and Jingbo Zhu
- 15:25 *Nonconvex Global Optimization for Latent-Variable Models*
Matthew R. Gormley and Jason Eisner
- 15:50 *Parsing with Compositional Vector Grammars*
Richard Socher, John Bauer, Christopher D. Manning and Ng Andrew Y.

(16:15) Coffee Break

(18:30 - 19:45) Poster Session A

LP - Dialogue and Interactive Systems

Discriminative state tracking for spoken dialog systems
Angeliki Metallinou, Dan Bohus and Jason Williams

LP- Discourse, Coreference and Pragmatics

Leveraging Synthetic Discourse Data via Multi-task Learning for Implicit Discourse Relation Recognition
Man Lan, Yu Xu and Zhengyu Niu

Combining Intra- and Multi-sentential Rhetorical Parsing for Document-level Discourse Analysis
Shafiq Joty, Giuseppe Carenini, Raymond Ng and Yashar Mehdad

Improving pairwise coreference models through feature space hierarchy learning
Emmanuel Lassalle and Pascal Denis

Monday August 5, 2013 (continued)

LP - Information Retrieval

Feature-Based Selection of Dependency Paths in Ad Hoc Information Retrieval

K. Tamsin Maxwell, Jon Oberlander and W. Bruce Croft

LP - Language Resources

Coordination Structures in Dependency Treebanks

Martin Popel, David Mareček, Jan Štěpánek, Daniel Zeman and Zdeněk Žabokrtský

GlossBoot: Bootstrapping Multilingual Domain Glossaries from the Web

Flavio De Benedictis, Stefano Faralli and Roberto Navigli

Collective Annotation of Linguistic Resources: Basic Principles and a Formal Model

Ulle Endriss and Raquel Fernández

ParGramBank: The ParGram Parallel Treebank

Sebastian Sulger, Miriam Butt, Tracy Holloway King, Paul Meurer, Tibor Laczkó, György Rákosi, Cheikh Bamba Dione, Helge Dyvik, Victoria Rosén, Koenraad De Smedt, Agnieszka Patejuk, Ozlem Cetinoglu, I Wayan Arka and Meladel Mistica

LP - Lexical Semantics and Ontologies

Identifying Bad Semantic Neighbors for Improving Distributional Thesauri

Olivier Ferret

Models of Semantic Representation with Visual Attributes

Carina Silberer, Vittorio Ferrari and Mirella Lapata

Monday August 5, 2013 (continued)

LP - Low Resource Language Processing

Real-World Semi-Supervised Learning of POS-Taggers for Low-Resource Languages

Dan Garrette, Jason Mielens and Jason Baldridge

LP - Machine Translation: Methods, Applications and Evaluations

Using subcategorization knowledge to improve case prediction for translation to German

Marion Weller, Alexander Fraser and Sabine Schulte im Walde

Name-aware Machine Translation

Haibo Li, Jing Zheng, Heng Ji, Qi Li and Wen Wang

Decipherment Complexity in 1:1 Substitution Ciphers

Malte Nuhn and Hermann Ney

Non-Monotonic Sentence Alignment via Semisupervised Learning

Xiaojun Quan, Chunyu Kit and Yan Song

LP - Multilinguality

Bootstrapping Entity Translation on Weakly Comparable Corpora

Taesung Lee and Seung-won Hwang

Transfer Learning Based Cross-lingual Knowledge Extraction for Wikipedia

Zhigang Wang, Zhixing Li, Juanzi Li, Jie Tang and Jeff Z. Pan

Bridging Languages through Etymology: The case of cross language text categorization

Vivi Nastase and Carlo Strapparava

Monday August 5, 2013 (continued)

LP - NLP Applications

Creating Similarity: Lateral Thinking for Vertical Similarity Judgments

Tony Veale and Guofu Li

Discovering User Interactions in Ideological Discussions

Arjun Mukherjee and Bing Liu

LP - NLP and Creativity

Multilingual Affect Polarity and Valence Prediction in Metaphor-Rich Texts

Zornitsa Kozareva

LP - NLP for the Languages of Central and Eastern Europe and the Balkans

Large tagset labeling using Feed Forward Neural Networks. Case study on Romanian Language

Tiberiu Boros, Radu Ion and Dan Tufis

Learning to lemmatise Polish noun phrases

Adam Radziszewski

LP - NLP for the Web and Social Media

Using Conceptual Class Attributes to Characterize Social Media Users

Shane Bergsma and Benjamin Van Durme

The Impact of Topic Bias on Quality Flaw Prediction in Wikipedia

Oliver Ferschke, Iryna Gurevych and Marc Rittberger

Mining Informal Language from Chinese Microtext: Joint Word Recognition and Segmentation

Aobo Wang and Min-Yen Kan

Generating Synthetic Comparable Questions for News Articles

Oleg Rokhlenko and Idan Szpektor

Monday August 5, 2013 (continued)

LP - Spoken Language Processing

Punctuation Prediction with Transition-based Parsing

Dongdong Zhang, Shuangzhi Wu, Nan Yang and Mu Li

LP - Word Segmentation

Discriminative Learning with Natural Annotations: Word Segmentation as a Case Study

Wenbin Jiang, Meng Sun, Yajuan Lü, Yating Yang and Qun Liu

Graph-based Semi-Supervised Model for Joint Chinese Word Segmentation and Part-of-Speech Tagging

Xiaodong Zeng, Derek F. Wong, Lidia S. Chao and Isabel Trancoso

(19:45 - 21:00) Poster Session B

LP - Machine Translation: Statistical Models

An Infinite Hierarchical Bayesian Model of Phrasal Translation

Trevor Cohn and Gholamreza Haffari

Additive Neural Networks for Statistical Machine Translation

lemao liu, Taro Watanabe, Eiichiro Sumita and Tiejun Zhao

Hierarchical Phrase Table Combination for Machine Translation

Conghui Zhu, Taro Watanabe, Eiichiro Sumita and Tiejun Zhao

Shallow Local Multi-Bottom-up Tree Transducers in Statistical Machine Translation

Fabienne Braune, Nina Seemann, Daniel Quernheim and Andreas Maletti

Enlisting the Ghost: Modeling Empty Categories for Machine Translation

Bing Xiang, Xiaoqiang Luo and Bowen Zhou

A Multi-Domain Translation Model Framework for Statistical Machine Translation

Rico Sennrich, Holger Schwenk and Walid Aransa

Part-of-Speech Induction in Dependency Trees for Statistical Machine Translation

Akihiro Tamura, Taro Watanabe, Eiichiro Sumita, Hiroya Takamura and Manabu Okumura

Monday August 5, 2013 (continued)

LP - Question Answering

Statistical Machine Translation Improves Question Retrieval in Community Question Answering via Matrix Factorization

Guangyou Zhou, Fang Liu, Yang Liu, Shizhu He and Jun Zhao

LP - Semantics

Improved Lexical Acquisition through DPP-based Verb Clustering

Roi Reichart and Anna Korhonen

Semantic Frames to Predict Stock Price Movement

Boyi Xie, Rebecca J. Passonneau, Leon Wu and Germán G. Creamer

Density Maximization in Context-Sense Metric Space for All-words WSD

Koichi Tanigaki, Mitsuteru Shiba, Tatsuji Munaka and Yoshinori Sagisaka

The Role of Syntax in Vector Space Models of Compositional Semantics

Karl Moritz Hermann and Phil Blunsom

Margin-based Decomposed Amortized Inference

Gourab Kundu, Vivek Srikumar and Dan Roth

Semi-Supervised Semantic Tagging of Conversational Understanding using Markov Topic Regression

Asli Celikyilmaz, Dilek Hakkani-Tur, Gokhan Tur and Ruhi Sarikaya

Parsing Graphs with Hyperedge Replacement Grammars

David Chiang, Jacob Andreas, Daniel Bauer, Karl Moritz Hermann, Bevan Jones and Kevin Knight

Grounded Unsupervised Semantic Parsing

Hoifung Poon

Monday August 5, 2013 (continued)

LP - Sentiment Analysis, Opinion Mining and Text Classification

Automatic detection of deception in child-produced speech using syntactic complexity features

Maria Yancheva and Frank Rudzicz

Sentiment Relevance

Christian Scheible and Hinrich Schütze

Predicting and Eliciting Addressee's Emotion in Online Dialogue

Takayuki Hasegawa, Nobuhiro Kaji, Naoki Yoshinaga and Masashi Toyoda

Utterance-Level Multimodal Sentiment Analysis

Veronica Perez-Rosas, Rada Mihalcea and Louis-Philippe Morency

Probabilistic Sense Sentiment Similarity through Hidden Emotions

Mitra Mohtarami, Man Lan and Chew Lim Tan

LP - Statistical and Machine Learning Methods in NLP

A user-centric model of voting intention from Social Media

Vasileios Lampos, Daniel Preoțiuc-Pietro and Trevor Cohn

LP - Summarization and Generation

Using Supervised Bigram-based ILP for Extractive Summarization

Chen Li, Xian Qian and Yang Liu

Summarization Through Submodularity and Dispersion

Anirban Dasgupta, Ravi Kumar and Sujith Ravi

Subtree Extractive Summarization via Submodular Maximization

Hajime Morita, Ryohei Sasano, Hiroya Takamura and Manabu Okumura

Monday August 5, 2013 (continued)

LP - Syntax and Parsing

The effect of non-tightness on Bayesian estimation of PCFGs

Shay B. Cohen and Mark Johnson

Integrating Multiple Dependency Corpora for Inducing Wide-coverage Japanese CCG Resources

Sumire Uematsu, Takuya Matsuzaki, Hiroki Hanaoka, Yusuke Miyao and Hideki Mima

Transition-based Dependency Parsing with Selectional Branching

Jinho D. Choi and Andrew McCallum

Bilingually-Guided Monolingual Dependency Grammar Induction

Kai Liu, Yajuan Lü, Wenbin Jiang and Qun Liu

LP - Tagging and Chunking

Joint Word Alignment and Bilingual Named Entity Recognition Using Dual Decomposition

Mengqiu Wang, Wanxiang Che and Christopher D. Manning

Resolving Entity Morphs in Censored Data

Hongzhao Huang, Zhen Wen, Dian Yu, Heng Ji, Yizhou Sun, Jiawei Han and He Li

LP - Text Mining and Information Extraction

Learning to Extract International Relations from Political Context

Brendan O'Connor, Brandon M. Stewart and Noah A. Smith

Tuesday August 6, 2013

(7:30 - 17:00) Registration

(9:00) Industrial Lecture: Lars Rasmussen (Facebook)

(10:00) Best Paper Award

(10:30) Coffee Break

Oral Presentations

(11:00 -12:15) LP 5a

11:00 *Graph Propagation for Paraphrasing Out-of-Vocabulary Words in Statistical Machine Translation*

Majid Razmara, Maryam Siahbani, Reza Haffari and Anoop Sarkar

11:25 *Online Relative Margin Maximization for Statistical Machine Translation*

Vladimir Eidelman, Yuval Marton and Philip Resnik

11:50 *Handling Ambiguities of Bilingual Predicate-Argument Structures for Statistical Machine Translation*

Feifei Zhai, Jiajun Zhang, Yu Zhou and Chengqing Zong

(11:00 -12:15) LP 5b

11:00 *Reconstructing an Indo-European Family Tree from Non-native English Texts*

Ryo Nagata and Edward Whittaker

11:25 *Word Association Profiles and their Use for Automated Scoring of Essays*

Beata Beigman Klebanov and Michael Flor

11:50 *Adaptive Parser-Centric Text Normalization*

Congle Zhang, Tyler Baldwin, Howard Ho, Benny Kimelfeld and Yunyao Li

Tuesday August 6, 2013 (continued)

(11:00 -12:15) LP 5c

- 11:00 *A Random Walk Approach to Selectional Preferences Based on Preference Ranking and Propagation*
Zhenhua Tian, Hengheng Xiang, Ziqi Liu and Qinghua Zheng
- 11:25 *ImpAr: A Deterministic Algorithm for Implicit Semantic Role Labelling*
Egoitz Laparra and German Rigau
- 11:50 *Cross-lingual Transfer of Semantic Role Labeling Models*
Mikhail Kozhevnikov and Ivan Titov

(11:00 -12:15) LP 5d

- 11:00 *DERivBase: Inducing and Evaluating a Derivational Morphology Resource for German*
Britta Zeller, Jan Šnajder and Sebastian Padó
- 11:25 *Crowdsourcing Interaction Logs to Understand Text Reuse from the Web*
Martin Potthast, Matthias Hagen, Michael Völske and Benno Stein
- 11:50 *SPred: Large-scale Harvesting of Semantic Predicates*
Tiziano Flati and Roberto Navigli

(11:00 -12:15) LP 5e

- 11:00 *Towards Robust Abstractive Multi-Document Summarization: A Caseframe Analysis of Centrality and Domain*
Jackie Chi Kit Cheung and Gerald Penn
- 11:25 *HEADY: News headline abstraction through event pattern clustering*
Enrique Alfonseca, Daniele Pighin and Guillermo Garrido
- 11:50 *Conditional Random Fields for Responsive Surface Realisation using Global Features*
Nina Dethlefs, Helen Hastie, Heriberto Cuayáhuitl and Oliver Lemon

Tuesday August 6, 2013 (continued)

(12:15) Lunch break

(13:45 -15:00) LP 6a

13:45 *Two-Neighbor Orientation Model with Cross-Boundary Global Contexts*
Hendra Setiawan, Bowen Zhou, Bing Xiang and Libin Shen

14:10 *Cut the noise: Mutually reinforcing reordering and alignments for improved machine translation*
Karthik Visweswariah, Mitesh M. Khapra and Ananthakrishnan Ramanathan

14:35 *Vector Space Model for Adaptation in Statistical Machine Translation*
Boxing Chen, Roland Kuhn and George Foster

(13:45 -15:00) LP 6b

13:45 *From Natural Language Specifications to Program Input Parsers*
Tao Lei, Fan Long, Regina Barzilay and Martin Rinard

14:10 *Entity Linking for Tweets*
Xiaohua Liu, Yitong Li, Haocheng Wu, Ming Zhou, Furu Wei and Yi Lu

14:35 *Identification of Speakers in Novels*
Hua He, Denilson Barbosa and Grzegorz Kondrak

(13:45 -15:00) LP 6c

13:45 *Language Acquisition and Probabilistic Models: keeping it simple*
Aline Villavicencio, Marco Idiart, Robert Berwick and Igor Malioutov

14:10 *A Two Level Model for Context Sensitive Inference Rules*
Oren Melamud, Jonathan Berant, Ido Dagan, Jacob Goldberger and Idan Szpektor

14:35 *Align, Disambiguate and Walk: A Unified Approach for Measuring Semantic Similarity*
Mohammad Taher Pilehvar, David Jurgens and Roberto Navigli

Tuesday August 6, 2013 (continued)

(13:45 -15:00) LP 6d

- 13:45 *Linking and Extending an Open Multilingual Wordnet*
Francis Bond and Ryan Foster
- 14:10 *FrameNet on the Way to Babel: Creating a Bilingual FrameNet Using Wiktionary as Interlingual Connection*
Silvana Hartmann and Iryna Gurevych
- 14:35 *Dirt Cheap Web-Scale Parallel Text from the Common Crawl*
Jason R. Smith, Herve Saint-Amand, Magdalena Plamada, Philipp Koehn, Chris Callison-Burch and Adam Lopez

(13:45 -15:00) LP 6e

- 13:45 *A Sentence Compression Based Framework to Query-Focused Multi-Document Summarization*
Lu Wang, Hema Raghavan, Vittorio Castelli, Radu Florian and Claire Cardie
- 14:10 *Domain-Independent Abstract Generation for Focused Meeting Summarization*
Lu Wang and Claire Cardie
- 14:35 *A Statistical NLG Framework for Aggregated Planning and Realization*
Ravi Kondadadi, Blake Howald and Frank Schilder

(15:00 -16:15) LP 7a

- 15:00 *Models of Translation Competitions*
Mark Hopkins and Jonathan May
- 15:25 *Learning a Phrase-based Translation Model from Monolingual Data with Application to Domain Adaptation*
Jiajun Zhang and Chengqing Zong
- 15:50 *SenseSpotting: Never let your parallel data tie you to an old domain*
Marine Carpuat, Hal Daume III, Katharine Henry, Ann Irvine, Jagadeesh Jagarlamudi and Rachel Rudinger

Tuesday August 6, 2013 (continued)

(15:00 -16:15) LP 7b

15:00 *BRAINSUP: Brainstorming Support for Creative Sentence Generation*
Gözde Özbal, Daniele Pighin and Carlo Strapparava

15:25 *Grammatical Error Correction Using Integer Linear Programming*
Yuanbin Wu and Hwee Tou Ng

15:50 *Text-Driven Toponym Resolution using Indirect Supervision*
Michael Speriosu and Jason Baldridge

(15:00 -16:15) LP 7c

15:00 *Argument Inference from Relevant Event Mentions in Chinese Argument Extraction*
Peifeng Li, Qiaoming Zhu and Guodong Zhou

15:25 *Fine-grained Semantic Typing of Emerging Entities*
Ndapandula Nakashole, Tomasz Tylanda and Gerhard Weikum

15:50 *Embedding Semantic Similarity in Tree Kernels for Domain Adaptation of Relation Extraction*
Barbara Plank and Alessandro Moschitti

(15:00 -16:15) LP 7d

15:00 *A joint model of word segmentation and phonological variation for English word-final /t/-deletion*
Benjamin Börschinger, Mark Johnson and Katherine Demuth

15:25 *Compositional-ly Derived Representations of Morphologically Complex Words in Distributional Semantics*
Angeliki Lazaridou, Marco Marelli, Roberto Zamparelli and Marco Baroni

15:50 *Unsupervised Consonant-Vowel Prediction over Hundreds of Languages*
Young-Bum Kim and Benjamin Snyder

Tuesday August 6, 2013 (continued)

(15:00 -16:15) LP 7e

15:00 *Improving Text Simplification Language Modeling Using Unsimplified Text Data*
David Kauchak

15:25 *Combining Referring Expression Generation and Surface Realization: A Corpus-Based Investigation of Architectures*
Sina Zarrieß and Jonas Kuhn

15:50 *Named Entity Recognition using Cross-lingual Resources: Arabic as an Example*
Kareem Darwish

(16:15) Coffee Break

(18:30) Banquet

Wednesday August 7, 2013

(9:30) Invited Talk 3: Chantal Prat

(10:30) Coffee Break

Oral Presentations

(11:00 -12:15) LP 9a

11:00 *Beam Search for Solving Substitution Ciphers*
Malte Nuhn, Julian Schamper and Hermann Ney

11:25 *Social Text Normalization using Contextual Graph Random Walks*
Hany Hassan and Arul Menezes

11:50 *Integrating Phrase-based Reordering Features into a Chart-based Decoder for Machine Translation*
ThuyLinh Nguyen and Stephan Vogel

Wednesday August 7, 2013 (continued)

(11:00 -12:15) LP 9b

- 11:00 *Machine Translation Detection from Monolingual Web-Text*
Yuki Arase and Ming Zhou
- 11:25 *Paraphrase-Driven Learning for Open Question Answering*
Anthony Fader, Luke Zettlemoyer and Oren Etzioni
- 11:50 *Aid is Out There: Looking for Help from Tweets during a Large Scale Disaster*
István Varga, Motoki Sano, Kentaro Torisawa, Chikara Hashimoto, Kiyonori Ohtake,
Takao Kawai, Jong-Hoon Oh and Stijn De Saeger

(11:00 -12:15) LP 9c

- 11:00 *A Bayesian Model for Joint Unsupervised Induction of Sentiment, Aspect and Discourse Representations*
Angeliki Lazaridou, Ivan Titov and Caroline Sporleder
- 11:25 *Joint Inference for Fine-grained Opinion Extraction*
Bishan Yang and Claire Cardie
- 11:50 *Linguistic Models for Analyzing and Detecting Biased Language*
Marta Recasens, Cristian Danescu-Niculescu-Mizil and Dan Jurafsky

(11:00 -12:15) LP 9d

- 11:00 *Evaluating a City Exploration Dialogue System with Integrated Question-Answering and Pedestrian Navigation*
Srinivasan Janarthanam, Oliver Lemon, Phil Bartie, Tiphaine Dalmas, Anna Dickinson,
Xingkun Liu, William Mackaness and Bonnie Webber
- 11:25 *Lightly Supervised Learning of Procedural Dialog Systems*
Svitlana Volkova, Pallavi Choudhury, Chris Quirk, Bill Dolan and Luke Zettlemoyer
- 11:50 *Public Dialogue: Analysis of Tolerance in Online Discussions*
Arjun Mukherjee, Vivek Venkataraman, Bing Liu and Sharon Meraz

Wednesday August 7, 2013 (continued)

(12:15) Lunch break

(13:30) ACL Business Meeting

(15:00 -16:15) LP 10a

15:00 *Offspring from Reproduction Problems: What Replication Failure Teaches Us*
Antske Fokkens, Marieke van Erp, Marten Postma, Ted Pedersen, Piek Vossen and Nuno Freire

15:25 *Evaluating Text Segmentation using Boundary Edit Distance*
Chris Fournier

15:50 *Crowd Prefers the Middle Path: A New IAA Metric for Crowdsourcing Reveals Turker Biases in Query Segmentation*
Rohan Ramanath, Monojit Choudhury, Kalika Bali and Rishiraj Saha Roy

(15:00 -16:15) LP 10b

15:00 *Deceptive Answer Prediction with User Preference Graph*
Fangtao Li, Yang Gao, Shuchang Zhou, Xiance Si and Decheng Dai

15:25 *Why-Question Answering using Intra- and Inter-Sentential Causal Relations*
Jong-Hoon Oh, Kentaro Torisawa, Chikara Hashimoto, Motoki Sano, Stijn De Saeger and Kiyonori Ohtake

15:50 *Question Answering Using Enhanced Lexical Semantic Models*
Wen-tau Yih, Ming-Wei Chang, Christopher Meek and Andrzej Pastusiak

Wednesday August 7, 2013 (continued)

(15:00 -16:15) LP 10c

15:00 *Syntactic Patterns versus Word Alignment: Extracting Opinion Targets from Online Reviews*

Kang Liu, Liheng Xu and Jun Zhao

15:25 *Mining Opinion Words and Opinion Targets in a Two-Stage Framework*

Liheng Xu, Kang Liu, Siwei Lai, Yubo Chen and Jun Zhao

15:50 *Connotation Lexicon: A Dash of Sentiment Beneath the Surface Meaning*

Song Feng, Jun Seok Kang, Polina Kuznetsova and Yejin Choi

(16:15) Coffee Break

(18:30) Lifetime Achievement Award Session

(19:15) Closing Session

(19:30) End