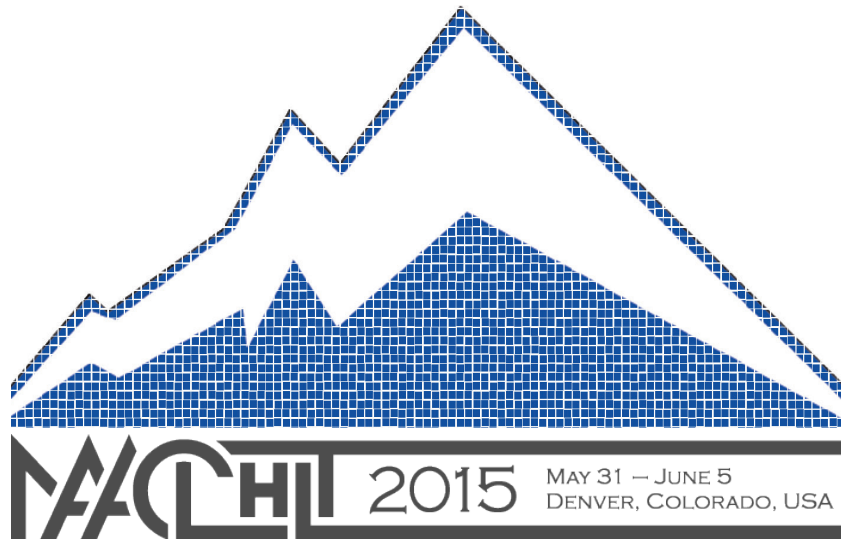


NAACL HLT 2015

**The 2015 Conference of the
North American Chapter of the
Association for Computational Linguistics:
Human Language Technologies**



Proceedings of the Conference

May 31 – June 5, 2015
Denver, Colorado, USA

Platinum

Bloomberg

Gold



Silver



Bronze

YAHOO! LABS

Supporters



Overseas student fellowship sponsors



©2015 The Association for Computational Linguistics

Order print-on-demand copies from:

Curran Associates
57 Morehouse Lane
Red Hook, New York 12571
USA
Tel: +1-845-758-0400
Fax: +1-845-758-2633
curran@proceedings.com

ISBN 978-1-941643-49-5

Message from the General Chair

Welcome to NAACL-HLT 2015 – at its 14th edition! Computational Linguistics has grown into one of the most exciting and diverse research communities, with an ever increasing number of researchers, many big and small companies, and a vibrant community of learners eager to get prepared to take on some of the fun and exciting challenges in the field. This year’s NAACL-HLT conference is a testimony to the vibrancy and vitality of this community.

Some of the highlights of this year’s program include two excellent invited speakers – Lillian Lee from Cornell and Fei-fei Li from Stanford – who will talk about the exciting research going on at the intersection of our field with social sciences and computer vision; many interesting paper presentations on cutting-edge research in computational linguistics, culminating with three best paper awards that will be presented in a plenary session during the last day of the conference; several excellent student-authored papers and dissertation proposals as part of the student research workshop; many exciting demos showing the latest in terms of developed systems available in our field; six tutorials on some of the most up-and-coming research topics in computational linguistics; several workshops on diverse topics ranging from multiword expressions and metaphors to clinical psychology and educational applications, including thirteen (!) one-day workshops and SEMEVAL as a two-day workshop; the fourth joint conference on lexical and computational semantics *SEM as a collocated conference; and, last but not least: a country line dance lesson!

As with any event of this scale, it would have not been possible without the hard work of a wonderful group of people. I would like to thank Priscilla Rasmussen for the zillions of bits and pieces that she has been doing on an everyday basis, to make sure that every single logistical detail of the forthcoming NAACL-HLT was ironed out. It is no overstatement to say that the success (and fun!) of this year’s conference is in large part due to Priscilla.

I am also grateful to Hal Daumé III for getting us started on this “NAACL-HLT 2015” journey, and being always willing to help with advice and information from his experience from previous years. Lucy Vanderwende and Daniel Marcu have also graciously agreed to “pass the baton” conversations that were very helpful and informative.

I was extremely fortunate to have the chance to work with the best committee ever: Joyce Chai and Anoop Sarkar (program chairs); Cornelia Caragea and Bing Liu (workshop co-chairs); Yang Liu and Tamar Solorio (tutorial co-chairs); Shibamouli Lahiri, Karen Mazidi and Alisa Zhila (student co-chairs) and Diana Inkpen and Smaranda Muresan (faculty advisors) for the student research workshop; Matt Gerber, Catherine Havasi, and Finley Lacatusu (demo co-chairs); Annie Louise (student volunteer coordinator); Kevin Cohen (local sponsorship chair); Saif Mohammad (publicity chair); Matt Post and Adam Lopez (publication co-chairs); Peter Ljunglöf (website chair); Aurelia Bunescu (handbook cover designer); Graeme Hirst and Joel Tetreault (treasurers); Asli Celikyilmaz and Julia Hockenmaier (sponsorship co-chairs).

I am also grateful to our sponsors for their generous contributions, which made the conference possible: A9, Baobab, Bloomberg, Digital Roots, Goldman Sachs, Google, IBM, Information Sciences Institute, National Science Foundation, Nuance, SDL, University of Washington Computational Linguistics, Yahoo Labs.

Finally, my gratitude goes to everyone else who contributed to the success of the conference: area chairs, workshop organizers, tutorial presenters, student mentors, and reviewers. And of course my deepest thanks to you, the attendees: you are the life and spirit of this entire conference.

Here is to an enjoyable NAACL-HLT 2015, and many more exciting conferences to come!

Rada Mihalcea, University of Michigan
NAACL 2015 General Chair

Message from the Program Chairs

Welcome to the 2015 Conference of the North American Chapter of the Association for Computational Linguistics – Human Language Technologies or NAACL HLT 2015 for short.

This year, we received the largest number of submissions in the history of NAACL: a total of 714 submissions with 402 long paper submissions and 312 short papers submissions. From these, 117 long papers (62 oral presentations and 55 poster presentations) and 69 short papers (24 oral presentations and 45 poster presentations) were accepted to appear at the conference.

The submissions to NAACL HLT 2015 were assigned to 18 technical areas including a new topic area called *Language and Vision*. This track was introduced with an intent to broaden research on natural language processing that is situated in a rich visual and perceptual context. We received 16 submissions for this area and seven of them will be presented at the conference.

For NAACL HLT 2015 we initiated a meta review process, where each paper received an analysis of the merits of the paper from the area chair's perspective that was based on the reviewer comments, the reviewer discussion and the author rebuttal. We found the meta reviews very helpful in consolidating the reviews and providing justifications for final decisions. As this was an experiment this year, the meta reviews were not sent to the authors.

Based on comments from reviewers, nominations from area chairs, and rankings from the best paper committee, three papers were selected to receive the best paper awards at the conference.

Continuing the tradition, NAACL HLT 2015 will feature 19 papers which were accepted for publication in the Transactions of the Association for Computational Linguistics (TACL). The TACL papers were split into 10 oral presentations and 9 poster presentations.

We are very pleased to have two exciting keynote talks: one by Professor Lillian Lee (Cornell University) and the other by Professor Fei-fei Li (Stanford University).

There are many people to thank for who have worked diligently to make NAACL HLT 2015 possible. Thanks to the 32 area chairs for their hard work on recruiting reviewers, managing reviews, leading discussions, and making recommendations. All the area chairs are listed in the Program Committee section of the Front Matter. Thanks to Chris Callison-Burch, David Mimno, Sameer Pradhan, and Philip Resnik for stepping in to serve as area co-chairs at the last minute when we were faced with an unexpectedly large number of submissions in some tracks.

Following what was done in the last NAACL conference, we used the paper assignment tool developed by Mark Dredze to assign papers to reviewers. Thanks to Mark Dredze and Jiang Guo for their hard work on assigning papers to reviewers based on their preferences. We had to especially rely on this tool this year because the distribution of submissions across areas was very different from past trends.

This program certainly would not be possible without the help of the 460 reviewers. Their names are listed in the Program Committee section. In particular, 116 reviewers from this list were recognized by the area chairs as best reviewers who have turned in exceptionally well-written and constructive reviews and who have actively engaged in the post-rebuttal discussions. The names of the best reviewers are

marked with * in the list of reviewers.

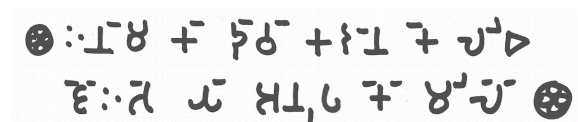
We are also indebted to the best paper award committee which consists of Claire Cardie, Daniel Gildea, Daniel Marcu, and Fernando Pereira. Their time and effort in recommending the best paper awards is much appreciated.

We also would like to thank Hal Daumé III, Kristina Toutanova, and Lucy Vanderwende for generously sharing their experience in organizing prior NAACL/ACL conferences and for their advice. We are grateful for the guidance and the support of the NAACL president Hal Daumé III, and the NAACL board. We also would like to thank the publication co-chairs Matt Post and Adam Lopez for putting together the proceedings and the conference handbook; and Paolo Gai and Rich Gerber from Softconf for always being responsive to our requests.

We would like to thank the ACL Business Manager Priscilla Rasmussen. She was our *go to* person who knew all details of the conference in and out. We are very grateful for her help.

Finally, this conference could not have happened without the efforts of the general chair, Rada Mihalcea. She made sure the various sections of NAACL organization worked well together. Her monthly newsletters informed all the organizers about what was being done by everyone else. We are very thankful for her leadership in the organization of NAACL HLT 2015.

We hope you will enjoy NAACL HLT 2015!



NAACL HLT 2015 Program Co-Chairs
Joyce Chai, Michigan State University
Anoop Sarkar, Simon Fraser University

Organizing Committee

General Chair

Rada Mihalcea, University of Michigan, USA

Local Arrangements Chair

Priscilla Rasmussen, ACL Business Manager

Program Committee Co-chairs

Joyce Chai, Michigan State University, USA

Anoop Sarkar, Simon Fraser University, Canada

Workshop Co-chairs

Cornelia Caragea, University of North Texas, USA

Bing Liu, University of Illinois at Chicago, USA

Tutorial Co-chairs

Yang Liu, University of Texas at Dallas, USA

Thamar Solorio, University of Alabama at Birmingham, USA

Student Research Workshop Co-chairs

Student Co-chairs

Shibamouli Lahiri, University of Michigan, USA

Karen Mazidi, University of North Texas, USA

Alisa Zhila, Instituto Politécnico Nacional (National Polytechnic Institute), Mexico

Faculty Advisors

Diana Inkpen, University of Ottawa, Canada

Smaranda Muresan, Columbia University, USA

Demo Co-chairs

Matt Gerber, University of Virginia, USA

Catherine Havasi, Massachusetts Institute of Technology, USA

Finley Lacatusu, Language Computer Corporation, USA

Student Volunteer Coordinator

Annie Louis, University of Edinburgh, UK

Reviewing Coordinators

Mark Dredze, Johns Hopkins University

Jiang Guo, Harbin Institute of Technology

Local Sponsorship Chair

Kevin B. Cohen, University of Colorado, USA

Publicity Chair

Saif M. Mohammad, National Research Council Canada

Publication Co-chairs

Matt Post, John Hopkins University, USA

Adam Lopez, University of Edinburgh, UK

Conference Handbook Editor

Matt Post, Johns Hopkins University

Website Chair

Peter Ljunglöf, University of Gothenburg and Chalmers University of Technology, Sweden

Business Manager

Priscilla Rasmussen

Program Committee

Program Committee Co-chairs

Joyce Chai, Michigan State University, USA
Anoop Sarkar, Simon Fraser University, Canada

Area Chairs

Dialogue and Interactive Systems

Jason D. Williams, Microsoft Research

Discourse and Pragmatics

Ani Nenkova, University of Pennsylvania
Vincent Ng, University of Texas at Dallas

Generation and Summarization

Chris Callison-Burch, University of Pennsylvania
Fei Liu, Carnegie Mellon University

Information Extraction and Question Answering

Alan Ritter, The Ohio State University
Bowen Zhou, IBM Research

Information Retrieval

David Smith, Northeastern University

Language and Vision

Julia Hockenmaier, University of Illinois at Urbana-Champaign

Language Resources and Evaluation

Will Lewis, Microsoft Research
Sameer Pradhan, Harvard University

Linguistic and Psycholinguistic Aspects of CL

William Schuler, The Ohio State University

Machine Learning for NLP

Amarnag Subramanya, Google Research
Tong Zhang, Baidu Inc. and Rutgers University

Machine Translation

Bill Byrne, Cambridge University
Hany Hassan, Microsoft Research
Kevin Knight, Information Sciences Institute, University of Southern California
Haitao Mi, IBM Research

NLP for Web, Social Media and Social Sciences

Brendan O'Connor, University of Massachusetts
Philip Resnik, University of Maryland

NLP-enabled Technology

Richard Socher, Stanford University

Phonology, Morphology, and Word Segmentation

Sami Virpioja, Lingsoft and Aalto University

Semantics

Dipanjan Das, Google

William Dolan, Microsoft Research

Luke Zettlemoyer, University of Washington

Sentiment Analysis and Opinion Mining

Saif M. Mohammad, National Research Council Canada

Jerry Zhu, University of Wisconsin-Madison

Spoken Language Processing

Xiaodong He, Microsoft Research

Tagging, Chunking, Syntax, and Parsing

Noah Smith, Carnegie Mellon University

Yue Zhang, Singapore University of Technology and Design

Text Categorization and Topic Models

Jordan Boyd-Graber, University of Colorado at Boulder

David Mimno, Cornell University

Primary Reviewers

The list of all reviewers for NAACL HLT 2015. The area chairs also picked the best reviewers in their track. These best reviewers are marked with an asterisk in the list which is alphabetical on the last name. Multiple asterisks mean that multiple area chairs chose that reviewer to be the best in their track.

Amjad Abu-Jbara, Mikhail Ageev, Eneko Agirre, *Gregory Aist, Jan Alexandersson, *James Allan, Jesse Anderton, Jacob Andreas, Nicholas Andrews, Gabor Angeli, *Yoav Artzi, Nicholas Asher, Javed Aslam, Michael Auli, amittai axelrod.

Anton Bakalov, Kirk Baker, *Timothy Baldwin, Miguel Ballesteros, David Bamman, *Srinivas Bangalore, *Mohit Bansal, *Marco Baroni, Roberto Basili, Daniel Bauer, Beata Beigman Klebanov, *Emily M. Bender, Michael Bendersky, Jonathan Berant, Taylor Berg-Kirkpatrick, Sabine Bergler, Delphine Bernhard, Pushpak Bhattacharyya, *Klinton Bicknell, Dan Bikel, Steven Bird, Anders Björkelund, *Graeme Blackwood, Eduardo Blanco, John Blitzer, Nathan Bodenstab, Bernd Bohnet, Johan Bos, Alexandre Bouchard, Kristy Boyer, Chris Brew, *Chris Brockett, Paul Buitelaar, Wray Buntine, *David Burkett, *Benjamin Börschinger.

Aoife Cahill, Nicoletta Calzolari, Marine Carpuat, Xavier Carreras, Marc-Allen Cartright, Damir Cavar, Asli Celikyilmaz, *Daniel Cer, Nathanael Chambers, Ming-Wei Chang, Wanxiang Che, Boxing Chen, Hsin-Hsi Chen, Chen Chen, Wenliang Chen, *Colin Cherry, *Jackie Chi Kit Cheung, *David Chiang, Christian Chiarcos, Jinho D. Choi, *Yejin Choi, Jennifer Chu-Carroll, Christopher Cieri, Philipp Cimini, *Stephen Clark, *Jonathan Clark, Martin Cmejrek, Shay B. Cohen, Trevor Cohn, *John Conroy, Paul Cook, Dan Cristea, Peter Culicover, Aron Culotta.

Robert Daland, Jeff Dalton, Pradipto Das, *Hal Daumé III, Adrià de Gispert, Vera Demberg, *Steve DeNeefe, Pascal Denis, Leon Derczynski, David DeVault, *Jacob Devlin, *Barbara Di Eugenio, Mona Diab, *Laura Dietz, *Doug Downey, Gabriel Doyle, Eduard Dragut, Mark Dredze, Markus Dreyer, Lan Du, *Kevin Duh, Ewan Dunbar, *Greg Durrett, Chris Dyer.

Sauleh Eetemadi, Markus Egg, Koji Eguchi, Patrick Ehlen, Andreas Eisele, Jacob Eisenstein, Jason Eisner, Desmond Elliott, Tamer Elsayed, *Micha Elsner, Ahmad Emami, *Katrin Erk.

Anthony Fader, James Fan, *Manaal Faruqi, Afsaneh Fazly, Marcello Federico, *Christian Federmann, Yang Feng, Vanessa Wei Feng, Minwei Feng, *Katja Filippova, Nicholas FitzGerald, Jeffrey Flanigan, Radu Florian, Sandiway Fong, *George Foster, Jennifer Foster, Diego Frassinelli.

**Michel Galley, Michael Gamon, Sudeep Gandhe, Juri Ganitkevitch, *Qin Gao, *Claire Gardent, Milica Gasic, Niyu Ge, George Giannakopoulos, Daniel Gildea, *Kevin Gimpel, Filip Ginter, *Roxana Girju, Peter B. Golbus, Yoav Goldberg, Dan Goldwasser, Sharon Goldwater, Jeff Good, Kyle Gorman, Cyril Goutte, Joao Graca, Stephan Greene, Edward Grefenstette, Stig-Arne Grönroos, Weiwei Guo, Iryna Gurevych, *Carlos Gómez-Rodríguez.

Nizar Habash, Eva Hajicova, John Hale, Keith Hall, *David Hall, Michael Hammond, *Sanda Harabagiu, *Eva Hasler, Claudia Hauff, Kenneth Heafield, John Henderson, Aurélie Herbelot, Jack Hessel, Derrick Higgins, Julia Hirschberg, *Graeme Hirst, Anna Hjalmarsson, Micah Hodosh, *Dirk Hovy, Eduard Hovy, *Yuening Hu, Zhongqiang Huang, Minlie Huang, Liang Huang, Ruihong Huang, *Mans Hulden, Rebecca Hwa.

Gonzalo Iglesias, Ryu Iida, Diana Inkpen, Amy Isard, Abe Ittycheriah.

Jagadeesh Jagarlamudi, *Jing Jiang, Richard Johansson, Mark Johnson, *Michael Johnston, Shafiq Joty, Dan Jurafsky, *David Jurgens.

*Min-Yen Kan, Pallika Kanani, *Saurabh Kataria, Daisuke Kawahara, *Andrew Kehler, *Frank Keller, Svetlana Kiritchenko, Alexandre Klementiev, Alistair Knott, Philipp Koehn, Oskar Kohonen, Kazunori Komatani, Terry Koo, Anna Korhonen, Alexander Kotov, Zornitsa Kozareva, Jayant Krishnamurthy, Kriste Krstovski, Canasai Kruengkrai, Lun-Wei Ku, *Marco Kuhlmann, *Roland Kuhn, Shankar Kumar, Jonathan K. Kummerfeld, *Tom Kwiatkowski.

Igor Labutov, D Terence Langendoen, *Guy Lapalme, **Mirella Lapata, Jey Han Lau, *Alon Lavie, *Hung-yi Lee, Sungjin Lee, Yoong Keok Lee, Moontae Lee, Maider Lehr, Tao Lei, Alessandro Lenci, *Omer Levy, Shoushan Li, Fangtao Li, Jiwei Li, Junyi Jessy Li, Jinyu Li, Percy Liang, Mark Liberman, *Shou-de Lin, *Chin-Yew Lin, Xiao Ling, *Yang Liu, Lema Liu, Jingjing Liu, Qun Liu, *Yang Liu, Adam Lopez, *Annie Louis, Bin Lu, Xiaoqiang Luo, Thang Luong.

Wei-Yun Ma, *Andrew Maas, Wolfgang Macherey, Nitin Madnani, Suresh Manandhar, Gideon Mann, *Christopher D. Manning, *Daniel Marcu, Katja Markert, André F. T. Martins, Yuval Marton, Sameer Maskey, *Michael Maxwell, Diana Maynard, Diana McCarthy, David McClosky, Ryan McDonald, *Bridget McInnes, Louise McNally, *Arul Menezes, Florian Metze, Timothy Miller, Eleni Miltsakaki, *Margaret Mitchell, Yusuke Miyao, Samaneh Moghaddam, Karo Moilanen, Christof Monz, *Taesun Moon, *Robert Moore, Roser Morante, Alessandro Moschitti, *Robert Munro, Brian Murphy.

Mikio Nakano, Preslav Nakov, Roberto Navigli, *Graham Neubig, Guenter Neumann, Hwee Tou Ng, Viet-An Nguyen, Malvina Nissim, Joakim Nivre.

Alice Oh, *Manabu Okumura, Miles Osborne, Myle Ott, Karolina Owczarzak.

Ulrike Pado, *Alexis Palmer, *Martha Palmer, Sinno J. Pan, Siddharth Patwardhan, Michael J. Paul, Virgil Pavlu, *Ted Pedersen, Gerald Penn, Jose Manuel Perea-Ortega, Slav Petrov, Olivier Pietquin, Tommi Pirinen, Emily Pitler, Massimo Poesio, Hoifung Poon, *Andrei Popescu-Belis, Matt Post, Emily Prud'hommeaux, Stephen Pulman, Matthew Purver.

Xian Qian, *Chris Quirk.

Altaf Rahman, Owen Rambow, Antoine Raux, Sujith Ravi, Kyle Rawlins, Siva Reddy, Roi Reichart, David Reitter, Sebastian Riedel, Jason Riesa, Stefan Riezler, German Rigau, *Ellen Riloff, *Laura Rimell, *Eric Ringger, *Brian Roark, Kirk Roberts, Hannah Rohde, Carolyn Rose, Andrew Rosenberg, Sara Rosenthal, Paolo Rosso, *Michael Roth, Teemu Ruokolainen, Irene Russo.

*Kenji Sagae, Alicia Sagae, Saurav Sahay, Rajhans Samdani, *Murat Saraclar, *Giorgio Satta, Asad Sayeed, David Schlangen, Nathan Schneider, Alexandra Schofield, *Lane Schwartz, H. Andrew Schwartz, Holger Schwenk, Hendra Setiawan, Izhak Shafran, Ekaterina Shutova, Khalil Sima'an, Sameer Singh, *Kairit Sirts, Gabriel Skantze, Linfeng Song, Caroline Sporleder, Vivek Srikumar, Manfred Stede, Mark Steedman, Amanda Stent, Brandon Stewart, Veselin Stoyanov, Weiwei Sun, *Mihai Surdeanu, Anders Sjøgaard.

Partha P. Talukdar, *Joel Tetreault, *Kapil Thadani, Jörg Tiedemann, Christoph Tillmann, Ivan Titov,

*Kristina Toutanova, David Traum, Reut Tsarfaty, Oren Tsur, Yoshimasa Tsuruoka, *Peter Turney,
*Oscar Täckström.

Jakob Uszkoreit.

Tim Van de Cruys, *Benjamin Van Durme, **Lucy Vanderwende, Yannick Versley, Aline Villavicencio.

*Xiaojun Wan, *Stephen Wan, Haifeng Wang, Lidan Wang, Lu Wang, Zhiguo Wang, *William Yang
Wang, *Taro Watanabe, Andy Way, *Bonnie Webber, David Weir, David Weiss, *Michael White,
Janyce Wiebe, Michael Wiegand, Shuly Wintner, Andreas Witt, *Kam-Fai Wong, Kristian Woodsend.

*Fei Xia, Yunqing Xia, Bing Xiang, Tong Xiao, Peng Xu, Wei Xu, Jia Xu, Nianwen Xue.

Hui Yang, Bishan Yang, Muyun Yang, Yi Yang, *Xuchen Yao, Mahsa Yarmohammadi, *Wen-tau Yih,
Dani Yogatama, Jianxing Yu, *François Yvon.

Fabio Massimo Zanzotto, Alessandra Zarccone, *Rabih Zbib, Richard Zens, Ke Zhai, Min Zhang, Lei
Zhang, Hao Zhang, Hai Zhao, Guodong Zhou, Xiaodan Zhu, Chengqing Zong, *Geoffrey Zweig.

Invited Talk: “Big data pragmatics!”, or, “Putting the ACL in computational social science”, or, if you think these title alternatives could turn people on, turn people off, or otherwise have an effect, this talk might be for you.

Lillian Lee

Cornell University

Abstract

What effect does language have on people?

You might say in response, "Who are you to discuss this problem?" and you would be right to do so; this is a Major Question that science has been tackling for many years. But as a field, I think natural language processing and computational linguistics have much to contribute to the conversation, and I hope to encourage the community to further address these issues.

This talk will focus on the effect of phrasing, emphasizing aspects that go beyond just the selection of one particular word over another. The issues we'll consider include: Does the way in which something is worded in and of itself have an effect on whether it is remembered or attracts attention, beyond its content or context? Can we characterize how different sides in a debate frame their arguments, in a way that goes beyond specific lexical choice (e.g., "pro-choice" vs. "pro-life")? The settings we'll explore range from movie quotes that achieve cultural prominence; to posts on Facebook, Wikipedia, Twitter, and the arXiv; to framing in public discourse on the inclusion of genetically-modified organisms in food.

Joint work with Lars Backstrom, Justin Cheng, Eunsol Choi, Cristian Danescu-Niculescu-Mizil, Jon Kleinberg, Bo Pang, Jennifer Spindel, and Chenhao Tan.

Biography

Lillian Lee is a professor of computer science and of information science at Cornell University, and the co-Editor-in-Chief, together with Michael Collins, of Transactions of the ACL. Her research interests include natural language processing and computational social science. She is the recipient of the inaugural Best Paper Award at HLT-NAACL 2004 (joint with Regina Barzilay), a citation in “Top Picks: Technology Research Advances of 2004” by Technology Research News (also joint with Regina Barzilay), and an Alfred P. Sloan Research Fellowship; and in 2013, she was named a Fellow of the Association for the Advancement of Artificial Intelligence (AAAI). Her group’s work has received several mentions in the popular press, including The New York Times, NPR’s All Things Considered, and NBC’s The Today Show, and one of her co-authored papers was publicly called “boring” by Youtubers Rhett and Link, in a video viewed over 1.8 million times.

Invited Talk: A Quest for Visual Intelligence in Computers

Fei-Fei Li

Stanford University

Abstract

More than half of the human brain is involved in visual processing. While it took mother nature billions of years to evolve and deliver us a remarkable human visual system, computer vision is one of the youngest disciplines of AI, born with the goal of achieving one of the loftiest dreams of AI. The central problem of computer vision is to turn millions of pixels of a single image into interpretable and actionable concepts so that computers can understand pictures just as well as humans do, from objects, to scenes, activities, events and beyond. Such technology will have a fundamental impact in almost every aspect of our daily life and the society as a whole, ranging from e-commerce, image search and indexing, assistive technology, autonomous driving, digital health and medicine, surveillance, national security, robotics and beyond. In this talk, I will give an overview of what computer vision technology is about and its brief history. I will then discuss some of the recent work from my lab towards large scale object recognition and visual scene story telling. I will particularly emphasize on what we call the "three pillars" of AI in our quest for visual intelligence: data, learning and knowledge. Each of them is critical towards the final solution, yet dependent on the other. This talk draws upon a number of projects ongoing at the Stanford Vision Lab.

Biography

Dr. Fei-Fei Li is an Associate Professor in the Computer Science Department at Stanford, and the Director of the Stanford Artificial Intelligence Lab and the Stanford Vision Lab. Her research areas are in machine learning, computer vision and cognitive and computational neuroscience, with an emphasis on Big Data analysis. Dr. Fei-Fei Li has published more than 100 scientific articles in top-tier journals and conferences, including Nature, PNAS, Journal of Neuroscience, CVPR, ICCV, NIPS, ECCV, IJCV, IEEE-PAMI, etc. Dr. Fei-Fei Li obtained her B.A. degree in physics from Princeton in 1999 with High Honors, and her PhD degree in electrical engineering from California Institute of Technology (Caltech) in 2005. She joined Stanford in 2009 as an assistant professor, and was promoted to associate professor with tenure in 2012. Prior to that, she was on faculty at Princeton University (2007-2009) and University of Illinois Urbana-Champaign (2005-2006). Dr. Fei-Fei Li is a speaker at TED2015 main conference, a recipient of the 2014 IBM Faculty Fellow Award, 2011 Alfred Sloan Faculty Award, 2012 Yahoo Labs FREP award, 2009 NSF CAREER award, the 2006 Microsoft Research New Faculty Fellowship and a number of Google Research awards. Work from Fei-Fei's lab have been featured in a number of popular press magazines and newspapers including New York Times, Wired Magazine, and New Scientists.

Table of Contents

<i>Unsupervised Induction of Semantic Roles within a Reconstruction-Error Minimization Framework</i> Ivan Titov and Ehsan Khoddam	1
<i>Predicate Argument Alignment using a Global Coherence Model</i> Travis Wolfe, Mark Dredze and Benjamin Van Durme	11
<i>Improving unsupervised vector-space thematic fit evaluation via role-filler prototype clustering</i> Clayton Greenberg, Asad Sayeed and Vera Demberg	21
<i>A Compositional and Interpretable Semantic Space</i> Alona Fyshe, Leila Wehbe, Partha P. Talukdar, Brian Murphy and Tom M. Mitchell	32
<i>Randomized Greedy Inference for Joint Segmentation, POS Tagging and Dependency Parsing</i> Yuan Zhang, Chengtao Li, Regina Barzilay and Kareem Darwish	42
<i>An Incremental Algorithm for Transition-based CCG Parsing</i> Bharat Ram Ambati, Tejaswini Deoskar, Mark Johnson and Mark Steedman	53
<i>Because Syntax Does Matter: Improving Predicate-Argument Structures Parsing with Syntactic Features</i> Corentin Ribeyre, Eric Villemonte de la Clergerie and Djamé Seddah	64
<i>A Hybrid Generative/Discriminative Approach To Citation Prediction</i> Chris Tanner and Eugene Charniak	75
<i>Weakly Supervised Slot Tagging with Partially Labeled Sequences from Web Search Click Logs</i> Young-Bum Kim, Minwoo Jeong, Karl Stratos and Ruhi Sarikaya	84
<i>Not All Character N-grams Are Created Equal: A Study in Authorship Attribution</i> Upendra Sapkota, Steven Bethard, Manuel Montes and Thamar Solorio	93
<i>Effective Use of Word Order for Text Categorization with Convolutional Neural Networks</i> Rie Johnson and Tong Zhang	103
<i>Transition-Based Syntactic Linearization</i> Yijia Liu, Yue Zhang, Wanxiang Che and Bing Qin	113
<i>Extractive Summarisation Based on Keyword Profile and Language Model</i> Han Xu, Eric Martin and Ashesh Mahidadia	123
<i>HEADS: Headline Generation as Sequence Prediction Using an Abstract Feature-Rich Space</i> Carlos A. Colmenares, Marina Litvak, Amin Mantrach and Fabrizio Silvestri	133
<i>What's Cookin'? Interpreting Cooking Videos using Text, Speech and Vision</i> Jonathan Malmaud, Jonathan Huang, Vivek Rathod, Nicholas Johnston, Andrew Rabinovich and Kevin Murphy	143

<i>Combining Language and Vision with a Multimodal Skip-gram Model</i> Angeliki Lazaridou, Nghia The Pham and Marco Baroni	153
<i>Discriminative Unsupervised Alignment of Natural Language Instructions with Corresponding Video Segments</i> Iftekhhar Naim, Young C. Song, Qiguang Liu, Liang Huang, Henry Kautz, Jiebo Luo and Daniel Gildea	164
<i>TopicCheck: Interactive Alignment for Assessing Topic Model Stability</i> Jason Chuang, Margaret E. Roberts, Brandon M. Stewart, Rebecca Weiss, Dustin Tingley, Justin Grimmer and Jeffrey Heer	175
<i>Inferring latent attributes of Twitter users with label regularization</i> Ehsan Mohammady Ardehaly and Aron Culotta	185
<i>A Neural Network Approach to Context-Sensitive Generation of Conversational Responses</i> Alessandro Sordoni, Michel Galley, Michael Auli, Chris Brockett, Yangfeng Ji, Margaret Mitchell, Jian-Yun Nie, Jianfeng Gao and Bill Dolan	196
<i>How to Make a Frenemy: Multitape FSTs for Portmanteau Generation</i> Aliya Deri and Kevin Knight	206
<i>Aligning Sentences from Standard Wikipedia to Simple Wikipedia</i> William Hwang, Hannaneh Hajishirzi, Mari Ostendorf and Wei Wu	211
<i>Inducing Lexical Style Properties for Paraphrase and Genre Differentiation</i> Ellie Pavlick and Ani Nenkova	218
<i>Entity Linking for Spoken Language</i> Adrian Benton and Mark Dredze	225
<i>Spinning Straw into Gold: Using Free Text to Train Monolingual Alignment Models for Non-factoid Question Answering</i> Rebecca Sharp, Peter Jansen, Mihai Surdeanu and Peter Clark	231
<i>Personalized Page Rank for Named Entity Disambiguation</i> Maria Pershina, Yifan He and Ralph Grishman	238
<i>When and why are log-linear models self-normalizing?</i> Jacob Andreas and Dan Klein	244
<i>Deep Multilingual Correlation for Improved Word Embeddings</i> Ang Lu, Weiran Wang, Mohit Bansal, Kevin Gimpel and Karen Livescu	250
<i>Disfluency Detection with a Semi-Markov Model and Prosodic Features</i> James Ferguson, Greg Durrett and Dan Klein	257
<i>Empty Category Detection With Joint Context-Label Embeddings</i> Xun Wang, Katsuhito Sudoh and Masaaki Nagata	263

<i>Incrementally Tracking Reference in Human/Human Dialogue Using Linguistic and Extra-Linguistic Information</i>	
Casey Kennington, Ryu Iida, Takenobu Tokunaga and David Schlangen	272
<i>Digital Leafleting: Extracting Structured Data from Multimedia Online Flyers</i>	
Emilia Apostolova, Payam Pourashraf and Jeffrey Sack	283
<i>Multi-Target Machine Translation with Multi-Synchronous Context-free Grammars</i>	
Graham Neubig, Philip Arthur and Kevin Duh	293
<i>Sign constraints on feature weights improve a joint model of word segmentation and phonology</i>	
Mark Johnson, Joe Pater, Robert Staubs and Emmanuel Dupoux	303
<i>Semi-Supervised Word Sense Disambiguation Using Word Embeddings in General and Specific Domains</i>	
Kaveh Taghipour and Hwee Tou Ng	314
<i>Continuous Space Representations of Linguistic Typology and their Application to Phylogenetic Inference</i>	
Yugo Murawaki	324
<i>Interpreting Compound Noun Phrases Using Web Search Queries</i>	
Marius Pasca	335
<i>Lexicon-Free Conversational Speech Recognition with Neural Networks</i>	
Andrew Maas, Ziang Xie, Dan Jurafsky and Andrew Ng	345
<i>I Can Has Cheezburger? A Nonparanormal Approach to Combining Textual and Visual Information for Predicting and Generating Popular Meme Descriptions</i>	
William Yang Wang and Miaomiao Wen	355
<i>A Transition-based Algorithm for AMR Parsing</i>	
Chuan Wang, Nianwen Xue and Sameer Pradhan	366
<i>The Geometry of Statistical Machine Translation</i>	
Aurelien Waite and Bill Byrne	376
<i>Data-driven sentence generation with non-isomorphic trees</i>	
Miguel Ballesteros, Bernd Bohnet, Simon Mille and Leo Wanner	387
<i>Latent Domain Word Alignment for Heterogeneous Corpora</i>	
Hoang Cuong and Khalil Sima'an	398
<i>Extracting Human Temporal Orientation from Facebook Language</i>	
H. Andrew Schwartz, Gregory Park, Maarten Sap, Evan Weingarten, Johannes Eichstaedt, Margaret Kern, David Stillwell, Michal Kosinski, Jonah Berger, Martin Seligman and Lyle Ungar	409
<i>An In-depth Analysis of the Effect of Text Normalization in Social Media</i>	
Tyler Baldwin and Yunyao Li	420

<i>Using Summarization to Discover Argument Facets in Online Ideological Dialog</i>	
Amita Misra, Pranav Anand, Jean E. Fox Tree and Marilyn Walker	430
<i>Active Learning with Rationales for Text Classification</i>	
Manali Sharma, Di Zhuang and Mustafa Bilgic	441
<i>Inferring Temporally-Anchored Spatial Knowledge from Semantic Roles</i>	
Eduardo Blanco and Alakananda Vempala	452
<i>A Dynamic Programming Algorithm for Tree Trimming-based Text Summarization</i>	
Masaaki Nishino, Norihito Yasuda, Tsutomu Hirao, Shin-ichi Minato and Masaaki Nagata ...	462
<i>Modeling Word Meaning in Context with Substitute Vectors</i>	
Oren Melamud, Ido Dagan and Jacob Goldberger	472
<i>Corpus-based discovery of semantic intensity scales</i>	
Chaitanya Shivade, Marie-Catherine de Marneffe, Eric Fosler-Lussier and Albert M. Lai	483
<i>Dialogue focus tracking for zero pronoun resolution</i>	
Sudha Rao, Allyson Ettinger, Hal Daumé III and Philip Resnik	494
<i>Déjà Image-Captions: A Corpus of Expressive Descriptions in Repetition</i>	
Jianfu Chen, Polina Kuznetsova, David Warren and Yejin Choi	504
<i>Inferring Missing Entity Type Instances for Knowledge Base Completion: New Dataset and Methods</i>	
Arvind Neelakantan and Ming-Wei Chang	515
<i>Robust Morphological Tagging with Word Representations</i>	
Thomas Müller and Hinrich Schuetze	526
<i>English orthography is not "close to optimal"</i>	
Garrett Nicolai and Grzegorz Kondrak	537
<i>LCCT: A Semi-supervised Model for Sentiment Classification</i>	
Min Yang, Wenting Tu, Ziyu Lu, Wenpeng Yin and Kam-Pui Chow	546
<i>Multiview LSA: Representation Learning via Generalized CCA</i>	
Pushpendre Rastogi, Benjamin Van Durme and Raman Arora	556
<i>NASARI: a Novel Approach to a Semantically-Aware Representation of Items</i>	
José Camacho-Collados, Mohammad Taher Pilehvar and Roberto Navigli	567
<i>Towards a standard evaluation method for grammatical error detection and correction</i>	
Mariano Felice and Ted Briscoe	578
<i>Using Zero-Resource Spoken Term Discovery for Ranked Retrieval</i>	
Jerome White, Douglas Oard, Aren Jansen, Jiaul Paik and Rashmi Sankepally	588
<i>Constraint-Based Models of Lexical Borrowing</i>	
Yulia Tsvetkov, Waleed Ammar and Chris Dyer	598

<i>Model Invertibility Regularization: Sequence Alignment With or Without Parallel Data</i>	
Tomer Levinboim, Ashish Vaswani and David Chiang	609
<i>Jointly Modeling Inter-Slot Relations by Random Walk on Knowledge Graphs for Unsupervised Spoken Language Understanding</i>	
Yun-Nung Chen, William Yang Wang and Alexander Rudnicky	619
<i>Expanding Paraphrase Lexicons by Exploiting Lexical Variants</i>	
Atsushi Fujita and Pierre Isabelle	630
<i>Diamonds in the Rough: Event Extraction from Imperfect Microblog Data</i>	
Ander Intxaurreondo, Eneko Agirre, Oier Lopez de Lacalle and Mihai Surdeanu	641
<i>Unsupervised Dependency Parsing: Let's Use Supervised Parsers</i>	
Phong Le and Willem Zuidema	651
<i>A Linear-Time Transition System for Crossing Interval Trees</i>	
Emily Pitler and Ryan McDonald	662
<i>Unsupervised Multi-Domain Adaptation with Feature Embeddings</i>	
Yi Yang and Jacob Eisenstein	672
<i>Ontologically Grounded Multi-sense Representation Learning for Semantic Vector Space Models</i>	
Sujay Kumar Jauhar, Chris Dyer and Eduard Hovy	683
<i>Subsentential Sentiment on a Shoestring: A Crosslingual Analysis of Compositional Classification</i>	
Michael Haas and Yannick Versley	694
<i>Cost Optimization in Crowdsourcing Translation: Low cost translations made even cheaper</i>	
Mingkun Gao, Wei Xu and Chris Callison-Burch	705
<i>Multitask Learning for Adaptive Quality Estimation of Automatically Transcribed Utterances</i>	
José G. C. de Souza, Hamed Zamani, Matteo Negri, Marco Turchi and Falavigna Daniele ...	714
<i>Incorporating Word Correlation Knowledge into Topic Modeling</i>	
Pengtao Xie, Diyi Yang and Eric Xing	725
<i>The Unreasonable Effectiveness of Word Representations for Twitter Named Entity Recognition</i>	
Colin Cherry and Hongyu Guo	735
<i>Is Your Anchor Going Up or Down? Fast and Accurate Supervised Topic Models</i>	
Thang Nguyen, Jordan Boyd-Graber, Jeffrey Lund, Kevin Seppi and Eric Ringger	746
<i>Grounded Semantic Parsing for Complex Knowledge Extraction</i>	
Ankur P. Parikh, Hoifung Poon and Kristina Toutanova	756
<i>Sentiment after Translation: A Case-Study on Arabic Social Media Posts</i>	
Mohammad Salameh, Saif Mohammad and Svetlana Kiritchenko	767

<i>Using External Resources and Joint Learning for Bigram Weighting in ILP-Based Multi-Document Summarization</i>	
Chen Li, Yang Liu and Lin Zhao	778
<i>Transforming Dependencies into Phrase Structures</i>	
Lingpeng Kong, Alexander M. Rush and Noah A. Smith	788
<i>Improving the Inference of Implicit Discourse Relations via Classifying Explicit Discourse Connectives</i>	
Attapol Rutherford and Nianwen Xue	799
<i>Solving Hard Coreference Problems</i>	
Haoruo Peng, Daniel Khashabi and Dan Roth	809
<i>Pragmatic Neural Language Modelling in Machine Translation</i>	
Paul Baltescu and Phil Blunsom	820
<i>Key Female Characters in Film Have More to Talk About Besides Men: Automating the Bechdel Test</i>	
Apoorv Agarwal, Jiehan Zheng, Shruti Kamath, Sriramkumar Balasubramanian and Shirin Ann Dey	830
<i>Semantic Grounding in Dialogue for Complex Problem Solving</i>	
Xiaolong Li and Kristy Boyer	841
<i>Learning Knowledge Graphs for Question Answering through Conversational Dialog</i>	
Ben Hixon, Peter Clark and Hannaneh Hajishirzi	851
<i>Sentence segmentation of aphasic speech</i>	
Kathleen C. Fraser, Naama Ben-David, Graeme Hirst, Naida Graham and Elizabeth Rochon ..	862
<i>Semantic parsing of speech using grammars learned with weak supervision</i>	
Judith Gaspers, Philipp Cimiano and Britta Wrede	872
<i>Early Gains Matter: A Case for Preferring Generative over Discriminative Crowdsourcing Models</i>	
Paul Felt, Kevin Black, Eric Ringger, Kevin Seppi and Robbie Haertel	882
<i>Optimizing Multivariate Performance Measures for Learning Relation Extraction Models</i>	
Gholamreza Haffari, Ajay Nagesh and Ganesh Ramakrishnan	892
<i>Convolutional Neural Network for Paraphrase Identification</i>	
Wenpeng Yin and Hinrich Schütze	901
<i>Representation Learning Using Multi-Task Deep Neural Networks for Semantic Classification and Information Retrieval</i>	
Xiaodong Liu, Jianfeng Gao, Xiaodong He, Li Deng, Kevin Duh and Ye-Yi Wang	912
<i>Inflection Generation as Discriminative String Transduction</i>	
Garrett Nicolai, Colin Cherry and Grzegorz Kondrak	922
<i>Penalized Expectation Propagation for Graphical Models over Strings</i>	
Ryan Cotterell and Jason Eisner	932

<i>Joint Generation of Transliterations from Multiple Representations</i>	
Lei Yao and Grzegorz Kondrak	943
<i>Prosodic boundary information helps unsupervised word segmentation</i>	
Bogdan Ludusan, Gabriel Synnaeve and Emmanuel Dupoux	953
<i>So similar and yet incompatible: Toward the automated identification of semantically compatible words</i>	
Germán Kruszewski and Marco Baroni	964
<i>Do Supervised Distributional Methods Really Learn Lexical Inference Relations?</i>	
Omer Levy, Steffen Remus, Chris Biemann and Ido Dagan	970
<i>A Word Embedding Approach to Predicting the Compositionality of Multiword Expressions</i>	
Bahar Salehi, Paul Cook and Timothy Baldwin	977
<i>Word Embedding-based Antonym Detection using Thesauri and Distributional Information</i>	
Masataka Ono, Makoto Miwa and Yutaka Sasaki	984
<i>A Comparison of Word Similarity Performance Using Explanatory and Non-explanatory Texts</i>	
Lifeng Jin and William Schuler	990
<i>Morphological Modeling for Machine Translation of English-Iraqi Arabic Spoken Dialogs</i>	
Katrin Kirchhoff, Yik-Cheung Tam, Colleen Richey and Wen Wang	995
<i>Continuous Adaptation to User Feedback for Statistical Machine Translation</i>	
Frédéric Blain, Fethi Bougares, Amir Hazem, Loïc Barrault and Holger Schwenk	1001
<i>Normalized Word Embedding and Orthogonal Transform for Bilingual Word Translation</i>	
Chao Xing, Dong Wang, Chao Liu and Yiye Lin	1006
<i>Fast and Accurate Preordering for SMT using Neural Networks</i>	
Adrià de Gispert, Gonzalo Iglesias and Bill Byrne	1012
<i>APRO: All-Pairs Ranking Optimization for MT Tuning</i>	
Markus Dreyer and Yuanzhe Dong	1018
<i>Paradigm classification in supervised learning of morphology</i>	
Malin Ahlberg, Markus Forsberg and Mans Hulden	1024
<i>Shift-Reduce Constituency Parsing with Dynamic Programming and POS Tag Lattice</i>	
Haitao Mi and Liang Huang	1030
<i>Unsupervised Code-Switching for Multilingual Historical Document Transcription</i>	
Dan Garrette, Hannah Alpert-Abrams, Taylor Berg-Kirkpatrick and Dan Klein	1036
<i>Matching Citation Text and Cited Spans in Biomedical Literature: a Search-Oriented Approach</i>	
Arman Cohan, Luca Soldaini and Nazli Goharian	1042
<i>Effective Feature Integration for Automated Short Answer Scoring</i>	
Keisuke Sakaguchi, Michael Heilman and Nitin Madnani	1049

<i>Socially-Informed Timeline Generation for Complex Events</i>	
Lu Wang, Claire Cardie and Galen Marchetti	1055
<i>Movie Script Summarization as Graph-based Scene Extraction</i>	
Philip John Gorinski and Mirella Lapata	1066
<i>Toward Abstractive Summarization Using Semantic Representations</i>	
Fei Liu, Jeffrey Flanigan, Sam Thomson, Norman Sadeh and Noah A. Smith	1077
<i>Encoding World Knowledge in the Evaluation of Local Coherence</i>	
Muyu Zhang, Vanessa Wei Feng, Bing Qin, Graeme Hirst, Ting Liu and Jingwen Huang ...	1087
<i>Chinese Event Coreference Resolution: An Unsupervised Probabilistic Model Rivaling Supervised Resolvers</i>	
Chen Chen and Vincent Ng	1097
<i>Removing the Training Wheels: A Coreference Dataset that Entertains Humans and Challenges Computers</i>	
Anupam Guha, Mohit Iyyer, Danny Bouman and Jordan Boyd-Graber	1108
<i>Injecting Logical Background Knowledge into Embeddings for Relation Extraction</i>	
Tim Rocktäschel, Sameer Singh and Sebastian Riedel	1119
<i>Unsupervised Entity Linking with Abstract Meaning Representation</i>	
Xiaoman Pan, Taylor Cassidy, Ulf Hermjakob, Heng Ji and Kevin Knight	1130
<i>Idest: Learning a Distributed Representation for Event Patterns</i>	
Sebastian Krause, Enrique Alfonseca, Katja Filippova and Daniele Pighin	1140
<i>High-Order Low-Rank Tensors for Semantic Role Labeling</i>	
Tao Lei, Yuan Zhang, Lluís Màrquez, Alessandro Moschitti and Regina Barzilay	1150
<i>Lexical Event Ordering with an Edge-Factored Model</i>	
Omri Abend, Shay B. Cohen and Mark Steedman	1161
<i>Bag-of-Words Forced Decoding for Cross-Lingual Information Retrieval</i>	
Felix Hieber and Stefan Riezler	1172
<i>Accurate Evaluation of Segment-level Machine Translation Metrics</i>	
Yvette Graham, Timothy Baldwin and Nitika Mathur	1183
<i>Leveraging Small Multilingual Corpora for SMT Using Many Pivot Languages</i>	
Raj Dabre, Fabien Cromieres, Sadao Kurohashi and Pushpak Bhattacharyya	1192
<i>Why Read if You Can Scan? Trigger Scoping Strategy for Biographical Fact Extraction</i>	
Dian Yu, Heng Ji, Sujian Li and Chin-Yew Lin	1203
<i>Lachmannian Archetype Reconstruction for Ancient Manuscript Corpora</i>	
Armin Hoenen	1209

<i>Distributed Representations of Words to Guide Bootstrapped Entity Classifiers</i> Sonal Gupta and Christopher D. Manning	1215
<i>Multi-Task Word Alignment Triangulation for Low-Resource Languages</i> Tomer Levinboim and David Chiang	1221
<i>Automatic cognate identification with gap-weighted string subsequences.</i> Taraka Rama	1227
<i>Short Text Understanding by Leveraging Knowledge into Topic Model</i> Shansong Yang, Weiming Lu, Dezhi Yang, Liang Yao and Baogang Wei	1232
<i>Unsupervised Most Frequent Sense Detection using Word Embeddings</i> Sudha Bhingardive, Dharendra Singh, Rudramurthy V, Hanumant Redkar and Pushpak Bhat- tacharyya	1238
<i>Chain Based RNN for Relation Classification</i> Javid Ebrahimi and Dejing Dou	1244
<i>LR Parsing for LCFRS</i> Laura Kallmeyer and Wolfgang Maier	1250
<i>Mining for unambiguous instances to adapt part-of-speech taggers to new domains</i> Dirk Hovy, Barbara Plank, Héctor Martínez Alonso and Anders Søgaard	1256
<i>Clustering Sentences with Density Peaks for Multi-document Summarization</i> Yang Zhang, Yunqing Xia, Yi Liu and Wenmin Wang	1262
<i>Development of the Multilingual Semantic Annotation System</i> Scott Piao, Francesca Bianchi, Carmen Dayrell, Angela D’Egidio and Paul Rayson	1268
<i>Unsupervised Sparse Vector Densification for Short Text Similarity</i> Yangqiu Song and Dan Roth	1275
<i>#WhyIStayed, #WhyILeft: Microblogging to Make Sense of Domestic Abuse</i> Nicolas Schrading, Cecilia Ovesdotter Alm, Raymond Ptucha and Christopher Homan	1281
<i>Morphological Word-Embeddings</i> Ryan Cotterell and Hinrich Schütze	1287
<i>Recognizing Social Constructs from Textual Conversation</i> Somak Aditya, Chitta Baral, Nguyen Ha Vo, Joohyung Lee, Jieping Ye, Zaw Naung, Barry Lump- kin, Jenny Hastings, Richard Scherl, Dawn M. Sweet and Daniela Inglezan	1293
<i>Two/Too Simple Adaptations of Word2Vec for Syntax Problems</i> Wang Ling, Chris Dyer, Alan W Black and Isabel Trancoso	1299
<i>Estimating Numerical Attributes by Bringing Together Fragmentary Clues</i> Hiroya Takamura and Jun’ichi Tsujii	1305

<i>Unsupervised POS Induction with Word Embeddings</i>	
Chu-Cheng Lin, Waleed Ammar, Chris Dyer and Lori Levin	1311
<i>Improving Update Summarization via Supervised ILP and Sentence Reranking</i>	
Chen Li, Yang Liu and Lin Zhao	1317
<i>MPQA 3.0: An Entity/Event-Level Sentiment Corpus</i>	
Lingjia Deng and Janyce Wiebe	1323
<i>Everyone Likes Shopping! Multi-class Product Categorization for e-Commerce</i>	
Zornitsa Kozareva	1329
<i>GPU-Friendly Local Regression for Voice Conversion</i>	
Taylor Berg-Kirkpatrick and Dan Klein	1334
<i>Response-based Learning for Machine Translation of Open-domain Database Queries</i>	
Carolin Haas and Stefan Riezler	1339
<i>Context-Dependent Automatic Response Generation Using Statistical Machine Translation Techniques</i>	
Andrew Shin, Ryohei Sasano, Hiroya Takamura and Manabu Okumura	1345
<i>Multilingual Open Relation Extraction Using Cross-lingual Projection</i>	
Manaal Faruqi and Shankar Kumar	1351
<i>Learning to parse with IAA-weighted loss</i>	
Héctor Martínez Alonso, Barbara Plank, Arne Skjærholt and Anders Søgaard	1357
<i>Exploiting Text and Network Context for Geolocation of Social Media Users</i>	
Afshin Rahimi, Duy Vu, Trevor Cohn and Timothy Baldwin	1362
<i>Discriminative Phrase Embedding for Paraphrase Identification</i>	
Wenpeng Yin and Hinrich Schütze	1368
<i>Combining Word Embeddings and Feature Embeddings for Fine-grained Relation Extraction</i>	
Mo Yu, Matthew R. Gormley and Mark Dredze	1374
<i>CASSA: A Context-Aware Synonym Simplification Algorithm</i>	
Ricardo Baeza-Yates, Luz Rello and Julia Dembowski	1380
<i>Simple task-specific bilingual word embeddings</i>	
Stephan Gouws and Anders Søgaard	1386
<i>Sampling Techniques for Streaming Cross Document Coreference Resolution</i>	
Luke Shrimpton, Victor Lavrenko and Miles Osborne	1391
<i>On the Automatic Learning of Sentiment Lexicons</i>	
Aliaksei Severyn and Alessandro Moschitti	1397
<i>Large-Scale Native Language Identification with Cross-Corpus Evaluation</i>	
Shervin Malmasi and Mark Dras	1403

<i>Unediting: Detecting Disfluencies Without Careful Transcripts</i> Victoria Zayats, Mari Ostendorf and Hannaneh Hajishirzi	1410
<i>Type-Driven Incremental Semantic Parsing with Polymorphism</i> Kai Zhao and Liang Huang	1416
<i>Template Kernels for Dependency Parsing</i> Hillel Taub-Tabib, Yoav Goldberg and Amir Globerson	1422
<i>Embedding a Semantic Network in a Word Space</i> Richard Johansson and Luis Nieto Piña	1428
<i>Random Walks and Neural Network Language Models on Knowledge Bases</i> Josu Goikoetxea, Aitor Soroa and Eneko Agirre	1434
<i>Identification and Characterization of Newsworthy Verbs in World News</i> Benjamin Nye and Ani Nenkova	1440
<i>Enhancing Sumerian Lemmatization by Unsupervised Named-Entity Recognition</i> Yudong Liu, Clinton Burkhart, James Hearne and Liang Luo	1446
<i>Extracting Information about Medication Use from Veterinary Discussions</i> Haibo Ding and Ellen Riloff	1452
<i>Reserating the awesometastic: An automatic extension of the WordNet taxonomy for novel terms</i> David Jurgens and Mohammad Taher Pilehvar	1459
<i>Cross-lingual Text Classification Using Topic-Dependent Word Probabilities</i> Daniel Andrade, Kunihiko Sadamasa, Akihiro Tamura and Masaaki Tsuchida	1466
<i>Testing and Comparing Computational Approaches for Identifying the Language of Framing in Political News</i> Eric Baumer, Elisha Elovic, Ying Qin, Francesca Polletta and Geri Gay	1472
<i>Echoes of Persuasion: The Effect of Euphony in Persuasive Communication</i> Marco Guerini, Gözde Özbal and Carlo Strapparava	1483
<i>Translating Videos to Natural Language Using Deep Recurrent Neural Networks</i> Subhashini Venugopalan, Huijuan Xu, Jeff Donahue, Marcus Rohrbach, Raymond Mooney and Kate Saenko	1494
<i>Learning to Interpret and Describe Abstract Scenes</i> Luis Gilberto Mateos Ortiz, Clemens Wolff and Mirella Lapata	1505
<i>A Comparison of Update Strategies for Large-Scale Maximum Expected BLEU Training</i> Joern Wuebker, Sebastian Muehr, Patrick Lehnen, Stephan Peitz and Hermann Ney	1516
<i>Learning Translation Models from Monolingual Continuous Representations</i> Kai Zhao, Hany Hassan and Michael Auli	1527

<i>A Corpus and Model Integrating Multiword Expressions and Supersenses</i>	
Nathan Schneider and Noah A. Smith	1537
<i>Good News or Bad News: Using Affect Control Theory to Analyze Readers' Reaction Towards News Articles</i>	
Areej Alhothali and Jesse Hoey	1548
<i>Do We Really Need Lexical Information? Towards a Top-down Approach to Sentiment Analysis of Product Reviews</i>	
Yulia Otmakhova and Hyopil Shin	1559
<i>How to Memorize a Random 60-Bit String</i>	
Marjan Ghazvininejad and Kevin Knight	1569
<i>A Bayesian Model for Joint Learning of Categories and their Features</i>	
Lea Frermann and Mirella Lapata	1576
<i>Shared common ground influences information density in microblog texts</i>	
Gabriel Doyle and Michael Frank	1587
<i>Hierarchical syntax improves reading time prediction</i>	
Marten van Schijndel and William Schuler	1597
<i>Retrofitting Word Vectors to Semantic Lexicons</i>	
Manaal Faruqi, Jesse Dodge, Sujay Kumar Jauhar, Chris Dyer, Eduard Hovy and Noah A. Smith	
1606	
<i>"You're Mr. Leowski, I'm the Dude": Inducing Address Term Formality in Signed Social Networks</i>	
Vinodh Krishnan and Jacob Eisenstein	1616
<i>Unsupervised Morphology Induction Using Word Embeddings</i>	
Radu Soricut and Franz Och	1627

Conference Program

Monday, June 1, 2015

07:30–08:45 *Registration and Breakfast*

08:45–09:00 *Welcome to NAACL HLT 2015*
Rada Mihalcea

09:00–10:10 *Invited Talk: "Big data pragmatics!", or, "Putting the ACL in computational social science", or, if you think these title alternatives could turn people on, turn people off, or otherwise have an effect, this talk might be for you*
Lillian Lee

10:10–10:40 *Break*

10:40–12:20 **Session 1A: Semantics (Long Papers)**

10:40–11:05 *Unsupervised Induction of Semantic Roles within a Reconstruction-Error Minimization Framework*
Ivan Titov and Ehsan Khoddam

11:05–11:30 *Predicate Argument Alignment using a Global Coherence Model*
Travis Wolfe, Mark Dredze and Benjamin Van Durme

11:30–11:55 *Improving unsupervised vector-space thematic fit evaluation via role-filler prototype clustering*
Clayton Greenberg, Asad Sayeed and Vera Demberg

11:55–12:20 *A Compositional and Interpretable Semantic Space*
Alona Fyshe, Leila Wehbe, Partha P. Talukdar, Brian Murphy and Tom M. Mitchell

Monday, June 1, 2015 (continued)

10:40–12:20 Session 1B: Tagging, Chunking, Syntax and Parsing (Long + TACL Papers)

10:40–11:05 *Randomized Greedy Inference for Joint Segmentation, POS Tagging and Dependency Parsing*

Yuan Zhang, Chengtao Li, Regina Barzilay and Kareem Darwish

11:05–11:30 *An Incremental Algorithm for Transition-based CCG Parsing*

Bharat Ram Ambati, Tejaswini Deoskar, Mark Johnson and Mark Steedman

11:30–11:55 *Because Syntax Does Matter: Improving Predicate-Argument Structures Parsing with Syntactic Features*

Corentin Ribeyre, Eric Villemonte de la Clergerie and Djamé Seddah

11:55–12:20 *[TACL] Exploring Compositional Architectures and Word Vector Representations for Prepositional Phrase Attachment*

Yonatan Belinkov, Tao Lei, Regina Barzilay, Amir Globerson

10:40–12:20 Session 1C: Information Retrieval, Text Categorization, Topic Modeling (Long Papers)

10:40–11:05 *A Hybrid Generative/Discriminative Approach To Citation Prediction*

Chris Tanner and Eugene Charniak

11:05–11:30 *Weakly Supervised Slot Tagging with Partially Labeled Sequences from Web Search Click Logs*

Young-Bum Kim, Minwoo Jeong, Karl Stratos and Ruhi Sarikaya

11:30–11:55 *Not All Character N-grams Are Created Equal: A Study in Authorship Attribution*

Uendra Sapkota, Steven Bethard, Manuel Montes and Thamar Solorio

11:55–12:20 *Effective Use of Word Order for Text Categorization with Convolutional Neural Networks*

Rie Johnson and Tong Zhang

12:20–14:00 Lunch

Monday, June 1, 2015 (continued)

14:00–15:15 Session 2A: Generation and Summarization (Long Papers)

- 14:00–14:25 *Transition-Based Syntactic Linearization*
Yijia Liu, Yue Zhang, Wanxiang Che and Bing Qin
- 14:25–14:50 *Extractive Summarisation Based on Keyword Profile and Language Model*
Han Xu, Eric Martin and Ashesh Mahidadia
- 14:50–15:15 *HEADS: Headline Generation as Sequence Prediction Using an Abstract Feature-Rich Space*
Carlos A. Colmenares, Marina Litvak, Amin Mantrach and Fabrizio Silvestri

14:00–15:15 Session 2B: Language and Vision (Long Papers)

- 14:00–14:25 *What's Cookin'? Interpreting Cooking Videos using Text, Speech and Vision*
Jonathan Malmaud, Jonathan Huang, Vivek Rathod, Nicholas Johnston, Andrew Rabinovich and Kevin Murphy
- 14:25–14:50 *Combining Language and Vision with a Multimodal Skip-gram Model*
Angeliki Lazaridou, Nghia The Pham and Marco Baroni
- 14:50–15:15 *Discriminative Unsupervised Alignment of Natural Language Instructions with Corresponding Video Segments*
Iftekhhar Naim, Young C. Song, Qiguang Liu, Liang Huang, Henry Kautz, Jiebo Luo and Daniel Gildea

14:00–15:15 Session 2C: NLP for Web, Social Media and Social Sciences (Long Papers)

- 14:00–14:25 *TopicCheck: Interactive Alignment for Assessing Topic Model Stability*
Jason Chuang, Margaret E. Roberts, Brandon M. Stewart, Rebecca Weiss, Dustin Tingley, Justin Grimmer and Jeffrey Heer
- 14:25–14:50 *Inferring latent attributes of Twitter users with label regularization*
Ehsan Mohammady Ardehaly and Aron Culotta
- 14:50–15:15 *A Neural Network Approach to Context-Sensitive Generation of Conversational Responses*
Alessandro Sordoni, Michel Galley, Michael Auli, Chris Brockett, Yangfeng Ji, Margaret Mitchell, Jian-Yun Nie, Jianfeng Gao and Bill Dolan

Monday, June 1, 2015 (continued)

15:15–16:00 Session 3A: Generation and Summarization (Short Papers)

15:15–15:30 *How to Make a Frenemy: Multitape FSTs for Portmanteau Generation*
Aliya Deri and Kevin Knight

15:30–15:45 *Aligning Sentences from Standard Wikipedia to Simple Wikipedia*
William Hwang, Hannaneh Hajishirzi, Mari Ostendorf and Wei Wu

15:45–16:00 *Inducing Lexical Style Properties for Paraphrase and Genre Differentiation*
Ellie Pavlick and Ani Nenkova

15:15–16:00 Session 3B: Information Extraction and Question Answering (Short Papers)

15:15–15:30 *Entity Linking for Spoken Language*
Adrian Benton and Mark Dredze

15:30–15:45 *Spinning Straw into Gold: Using Free Text to Train Monolingual Alignment Models for Non-factoid Question Answering*
Rebecca Sharp, Peter Jansen, Mihai Surdeanu and Peter Clark

15:45–16:00 *Personalized Page Rank for Named Entity Disambiguation*
Maria Pershina, Yifan He and Ralph Grishman

15:15–16:00 Session 3C: Machine Learning for NLP (Short Papers)

15:15–15:30 *When and why are log-linear models self-normalizing?*
Jacob Andreas and Dan Klein

15:30–15:45 *Deep Multilingual Correlation for Improved Word Embeddings*
Ang Lu, Weiran Wang, Mohit Bansal, Kevin Gimpel and Karen Livescu

15:45–16:00 *Disfluency Detection with a Semi-Markov Model and Prosodic Features*
James Ferguson, Greg Durrett and Dan Klein

16:00–16:30 Break

Monday, June 1, 2015 (continued)

16:30–18:00 One minute madness (Long + TACL papers)

Session P1A: 18:00–19:30 Poster session 1A: Long + TACL papers

Empty Category Detection With Joint Context-Label Embeddings

Xun Wang, Katsuhito Sudoh and Masaaki Nagata

Incrementally Tracking Reference in Human/Human Dialogue Using Linguistic and Extra-Linguistic Information

Casey Kennington, Ryu Iida, Takenobu Tokunaga and David Schlangen

Digital Leafleting: Extracting Structured Data from Multimedia Online Flyers

Emilia Apostolova, Payam Pourashraf and Jeffrey Sack

Multi-Target Machine Translation with Multi-Synchronous Context-free Grammars

Graham Neubig, Philip Arthur and Kevin Duh

Sign constraints on feature weights improve a joint model of word segmentation and phonology

Mark Johnson, Joe Pater, Robert Staubs and Emmanuel Dupoux

Semi-Supervised Word Sense Disambiguation Using Word Embeddings in General and Specific Domains

Kaveh Taghipour and Hwee Tou Ng

Continuous Space Representations of Linguistic Typology and their Application to Phylogenetic Inference

Yugo Murawaki

Interpreting Compound Noun Phrases Using Web Search Queries

Marius Pasca

Lexicon-Free Conversational Speech Recognition with Neural Networks

Andrew Maas, Ziang Xie, Dan Jurafsky and Andrew Ng

I Can Has Cheezburger? A Nonparanormal Approach to Combining Textual and Visual Information for Predicting and Generating Popular Meme Descriptions

William Yang Wang and Miaomiao Wen

A Transition-based Algorithm for AMR Parsing

Chuan Wang, Nianwen Xue and Sameer Pradhan

Monday, June 1, 2015 (continued)

The Geometry of Statistical Machine Translation

Aurelien Waite and Bill Byrne

Data-driven sentence generation with non-isomorphic trees

Miguel Ballesteros, Bernd Bohnet, Simon Mille and Leo Wanner

Latent Domain Word Alignment for Heterogeneous Corpora

Hoang Cuong and Khalil Sima'an

Extracting Human Temporal Orientation from Facebook Language

H. Andrew Schwartz, Gregory Park, Maarten Sap, Evan Weingarten, Johannes Eichstaedt, Margaret Kern, David Stillwell, Michal Kosinski, Jonah Berger, Martin Seligman and Lyle Ungar

An In-depth Analysis of the Effect of Text Normalization in Social Media

Tyler Baldwin and Yunyao Li

Using Summarization to Discover Argument Facets in Online Ideological Dialog

Amita Misra, Pranav Anand, Jean E. Fox Tree and Marilyn Walker

Active Learning with Rationales for Text Classification

Manali Sharma, Di Zhuang and Mustafa Bilgic

Inferring Temporally-Anchored Spatial Knowledge from Semantic Roles

Eduardo Blanco and Alakananda Vempala

A Dynamic Programming Algorithm for Tree Trimming-based Text Summarization

Masaaki Nishino, Norihito Yasuda, Tsutomu Hiraio, Shin-ichi Minato and Masaaki Nagata

Modeling Word Meaning in Context with Substitute Vectors

Oren Melamud, Ido Dagan and Jacob Goldberger

Corpus-based discovery of semantic intensity scales

Chaitanya Shivade, Marie-Catherine de Marneffe, Eric Fosler-Lussier and Albert M. Lai

Dialogue focus tracking for zero pronoun resolution

Sudha Rao, Allyson Ettinger, Hal Daumé III and Philip Resnik

Déjà Image-Captions: A Corpus of Expressive Descriptions in Repetition

Jianfu Chen, Polina Kuznetsova, David Warren and Yejin Choi

Monday, June 1, 2015 (continued)

Inferring Missing Entity Type Instances for Knowledge Base Completion: New Dataset and Methods

Arvind Neelakantan and Ming-Wei Chang

Robust Morphological Tagging with Word Representations

Thomas Müller and Hinrich Schuetze

English orthography is not "close to optimal"

Garrett Nicolai and Grzegorz Kondrak

LCCT: A Semi-supervised Model for Sentiment Classification

Min Yang, Wenting Tu, Ziyu Lu, Wenpeng Yin and Kam-Pui Chow

[TACL] Unsupervised Discovery of Biographical Structure from Text

David Bamman and Noah A. Smith

[TACL] 2-Slave Dual Decomposition for Generalized High Order CRFs

Xian Qian and Yang Liu

[TACL] Learning Strictly Local Subsequential Functions

Jane Chandlee, Remi Eyraud and Jeffrey Heinz

[TACL] Learning Constraints for Information Structure Analysis of Scientific Documents

Yufan Guo, Roi Reichart, and Anna Korhonen

Session P1B: 19:30–21:00 Poster session 1B: Long + TACL papers

Multiview LSA: Representation Learning via Generalized CCA

Pushpendre Rastogi, Benjamin Van Durme and Raman Arora

NASARI: a Novel Approach to a Semantically-Aware Representation of Items

José Camacho-Collados, Mohammad Taher Pilehvar and Roberto Navigli

Towards a standard evaluation method for grammatical error detection and correction

Mariano Felice and Ted Briscoe

Using Zero-Resource Spoken Term Discovery for Ranked Retrieval

Jerome White, Douglas Oard, Aren Jansen, Jiaul Paik and Rashmi Sankepally

Monday, June 1, 2015 (continued)

Constraint-Based Models of Lexical Borrowing

Yulia Tsvetkov, Waleed Ammar and Chris Dyer

Model Invertibility Regularization: Sequence Alignment With or Without Parallel Data

Tomer Levinboim, Ashish Vaswani and David Chiang

Jointly Modeling Inter-Slot Relations by Random Walk on Knowledge Graphs for Unsupervised Spoken Language Understanding

Yun-Nung Chen, William Yang Wang and Alexander Rudnicky

Expanding Paraphrase Lexicons by Exploiting Lexical Variants

Atsushi Fujita and Pierre Isabelle

Diamonds in the Rough: Event Extraction from Imperfect Microblog Data

Ander Intxaurreondo, Eneko Agirre, Oier Lopez de Lacalle and Mihai Surdeanu

Unsupervised Dependency Parsing: Let's Use Supervised Parsers

Phong Le and Willem Zuidema

A Linear-Time Transition System for Crossing Interval Trees

Emily Pitler and Ryan McDonald

Unsupervised Multi-Domain Adaptation with Feature Embeddings

Yi Yang and Jacob Eisenstein

Ontologically Grounded Multi-sense Representation Learning for Semantic Vector Space Models

Sujay Kumar Jauhar, Chris Dyer and Eduard Hovy

Subsentential Sentiment on a Shoestring: A Crosslingual Analysis of Compositional Classification

Michael Haas and Yannick Versley

Cost Optimization in Crowdsourcing Translation: Low cost translations made even cheaper

Mingkun Gao, Wei Xu and Chris Callison-Burch

Multitask Learning for Adaptive Quality Estimation of Automatically Transcribed Utterances

José G. C. de Souza, Hamed Zamani, Matteo Negri, Marco Turchi and Falavigna Daniele

Incorporating Word Correlation Knowledge into Topic Modeling

Pengtao Xie, Diyi Yang and Eric Xing

Monday, June 1, 2015 (continued)

The Unreasonable Effectiveness of Word Representations for Twitter Named Entity Recognition

Colin Cherry and Hongyu Guo

Is Your Anchor Going Up or Down? Fast and Accurate Supervised Topic Models

Thang Nguyen, Jordan Boyd-Graber, Jeffrey Lund, Kevin Seppi and Eric Ringger

Grounded Semantic Parsing for Complex Knowledge Extraction

Ankur P. Parikh, Hoifung Poon and Kristina Toutanova

Sentiment after Translation: A Case-Study on Arabic Social Media Posts

Mohammad Salameh, Saif Mohammad and Svetlana Kiritchenko

Using External Resources and Joint Learning for Bigram Weighting in ILP-Based Multi-Document Summarization

Chen Li, Yang Liu and Lin Zhao

Transforming Dependencies into Phrase Structures

Lingpeng Kong, Alexander M. Rush and Noah A. Smith

Improving the Inference of Implicit Discourse Relations via Classifying Explicit Discourse Connectives

Attapol Rutherford and Nianwen Xue

Solving Hard Coreference Problems

Haoruo Peng, Daniel Khashabi and Dan Roth

Pragmatic Neural Language Modelling in Machine Translation

Paul Baltescu and Phil Blunsom

Key Female Characters in Film Have More to Talk About Besides Men: Automating the Bechdel Test

Apoorv Agarwal, Jiehan Zheng, Shruti Kamath, Sriramkumar Balasubramanian and Shirin Ann Dey

[TACL] Dense Event Ordering with a Multi-Pass Architecture

Nathanael Chambers, Taylor Cassidy, Bill McDowell, and Steven Bethard

[TACL] Locally Non-Linear Learning for Statistical Machine Translation via Discretization and Structured Regularization

Jonathan H. Clark, Chris Dyer, and Alon Lavie

[TACL] SPRITE: Generalizing Topic Models with Structured Priors

Michael J. Paul and Mark Dredze

Monday, June 1, 2015 (continued)

[TACL] Reasoning about Quantities in Natural Language
Subhro Roy, Tim Vieira, and Dan Roth

[TACL] A sense-topic model for WSI with unsupervised data enrichment
Jing Wang, Mohit Bansal, Kevin Gimpel, Brian D. Ziebart, and Clement T. Yu

Tuesday, June 2, 2015

07:30–09:00 *Registration and Breakfast*

09:00–10:40 **Session 4A: Dialogue and Spoken Language Processing (Long Papers)**

09:00–09:25 *Semantic Grounding in Dialogue for Complex Problem Solving*
Xiaolong Li and Kristy Boyer

09:25–09:50 *Learning Knowledge Graphs for Question Answering through Conversational Dialogue*
Ben Hixon, Peter Clark and Hannaneh Hajishirzi

09:50–10:15 *Sentence segmentation of aphasic speech*
Kathleen C. Fraser, Naama Ben-David, Graeme Hirst, Naida Graham and Elizabeth Rochon

10:15–10:40 *Semantic parsing of speech using grammars learned with weak supervision*
Judith Gaspers, Philipp Cimiano and Britta Wrede

Tuesday, June 2, 2015 (continued)

09:00–10:40 Session 4B: Machine Learning for NLP (Long Papers)

09:00–09:25 *Early Gains Matter: A Case for Preferring Generative over Discriminative Crowdsourcing Models*

Paul Felt, Kevin Black, Eric Ringger, Kevin Seppi and Robbie Haertel

09:25–09:50 *Optimizing Multivariate Performance Measures for Learning Relation Extraction Models*

Gholamreza Haffari, Ajay Nagesh and Ganesh Ramakrishnan

09:50–10:15 *Convolutional Neural Network for Paraphrase Identification*

Wenpeng Yin and Hinrich Schütze

10:15–10:40 *Representation Learning Using Multi-Task Deep Neural Networks for Semantic Classification and Information Retrieval*

Xiaodong Liu, Jianfeng Gao, Xiaodong He, Li Deng, Kevin Duh and Ye-Yi Wang

09:00–10:40 Session 4C: Phonology, Morphology and Word Segmentation (Long Papers)

09:00–09:25 *Inflection Generation as Discriminative String Transduction*

Garrett Nicolai, Colin Cherry and Grzegorz Kondrak

09:25–09:50 *Penalized Expectation Propagation for Graphical Models over Strings*

Ryan Cotterell and Jason Eisner

09:50–10:15 *Joint Generation of Transliterations from Multiple Representations*

Lei Yao and Grzegorz Kondrak

10:15–10:40 *Prosodic boundary information helps unsupervised word segmentation*

Bogdan Ludusan, Gabriel Synnaeve and Emmanuel Dupoux

10:40–11:15 Break

Tuesday, June 2, 2015 (continued)

11:15–12:30 Session 5A: Semantics (Short Papers)

- 11:15–11:30 *So similar and yet incompatible: Toward the automated identification of semantically compatible words*
Germán Kruszewski and Marco Baroni
- 11:30–11:45 *Do Supervised Distributional Methods Really Learn Lexical Inference Relations?*
Omer Levy, Steffen Remus, Chris Biemann and Ido Dagan
- 11:45–12:00 *A Word Embedding Approach to Predicting the Compositionality of Multiword Expressions*
Bahar Salehi, Paul Cook and Timothy Baldwin
- 12:00–12:15 *Word Embedding-based Antonym Detection using Thesauri and Distributional Information*
Masataka Ono, Makoto Miwa and Yutaka Sasaki
- 12:15–12:30 *A Comparison of Word Similarity Performance Using Explanatory and Non-explanatory Texts*
Lifeng Jin and William Schuler

11:15–12:30 Session 5B: Machine Translation (Short Papers)

- 11:15–11:30 *Morphological Modeling for Machine Translation of English-Iraqi Arabic Spoken Dialogs*
Katrin Kirchhoff, Yik-Cheung Tam, Colleen Richey and Wen Wang
- 11:30–11:45 *Continuous Adaptation to User Feedback for Statistical Machine Translation*
Frédéric Blain, Fethi Bougares, Amir Hazem, Loïc Barrault and Holger Schwenk
- 11:45–12:00 *Normalized Word Embedding and Orthogonal Transform for Bilingual Word Translation*
Chao Xing, Dong Wang, Chao Liu and Yiye Lin
- 12:00–12:15 *Fast and Accurate Preordering for SMT using Neural Networks*
Adrià de Gispert, Gonzalo Iglesias and Bill Byrne
- 12:15–12:30 *APRO: All-Pairs Ranking Optimization for MT Tuning*
Markus Dreyer and Yuanzhe Dong

Tuesday, June 2, 2015 (continued)

11:15–12:30 Session 5C: Morphology, Syntax, Multilinguality, and Applications (Short Papers)

11:15–11:30 *Paradigm classification in supervised learning of morphology*

Malin Ahlberg, Markus Forsberg and Mans Hulden

11:30–11:45 *Shift-Reduce Constituency Parsing with Dynamic Programming and POS Tag Lattice*

Haitao Mi and Liang Huang

11:45–12:00 *Unsupervised Code-Switching for Multilingual Historical Document Transcription*

Dan Garrette, Hannah Alpert-Abrams, Taylor Berg-Kirkpatrick and Dan Klein

12:00–12:15 *Matching Citation Text and Cited Spans in Biomedical Literature: a Search-Oriented Approach*

Arman Cohan, Luca Soldaini and Nazli Goharian

12:15–12:30 *Effective Feature Integration for Automated Short Answer Scoring*

Keisuke Sakaguchi, Michael Heilman and Nitin Madnani

12:30–14:00 Lunch

14:00–15:15 Session 6A: Generation and Summarization (Long Papers)

14:00–14:25 *Socially-Informed Timeline Generation for Complex Events*

Lu Wang, Claire Cardie and Galen Marchetti

14:25–14:50 *Movie Script Summarization as Graph-based Scene Extraction*

Philip John Gorinski and Mirella Lapata

14:50–15:15 *Toward Abstractive Summarization Using Semantic Representations*

Fei Liu, Jeffrey Flanigan, Sam Thomson, Norman Sadeh and Noah A. Smith

Tuesday, June 2, 2015 (continued)

14:00–15:15 Session 6B: Discourse and Coreference (Long Papers)

14:00–14:25 *Encoding World Knowledge in the Evaluation of Local Coherence*
Muyu Zhang, Vanessa Wei Feng, Bing Qin, Graeme Hirst, Ting Liu and Jingwen Huang

14:25–14:50 *Chinese Event Coreference Resolution: An Unsupervised Probabilistic Model Rivaling Supervised Resolvers*
Chen Chen and Vincent Ng

14:50–15:15 *Removing the Training Wheels: A Coreference Dataset that Entertains Humans and Challenges Computers*
Anupam Guha, Mohit Iyyer, Danny Bouman and Jordan Boyd-Graber

14:00–15:15 Session 6C: Information Extraction and Question Answering (Long Papers)

14:00–14:25 *Injecting Logical Background Knowledge into Embeddings for Relation Extraction*
Tim Rocktäschel, Sameer Singh and Sebastian Riedel

14:25–14:50 *Unsupervised Entity Linking with Abstract Meaning Representation*
Xiaoman Pan, Taylor Cassidy, Ulf Hermjakob, Heng Ji and Kevin Knight

14:50–15:15 *Idest: Learning a Distributed Representation for Event Patterns*
Sebastian Krause, Enrique Alfonseca, Katja Filippova and Daniele Pighin

15:15–15:45 Break

Tuesday, June 2, 2015 (continued)

15:45–17:00 Session 7A: Semantics (Long + TACL Papers)

15:45–16:10 *High-Order Low-Rank Tensors for Semantic Role Labeling*
Tao Lei, Yuan Zhang, Lluís Màrquez, Alessandro Moschitti and Regina Barzilay

16:10–16:35 *[TACL] Large-scale Semantic Parsing without Question-Answer Pairs*
Siva Reddy, Mirella Lapata, and Mark Steedman

16:35–17:00 *[TACL] A Large Scale Evaluation of Distributional Semantic Models: Parameters, Interactions and Model Selection*
Gabriella Lapesa and Stefan Evert

15:45–17:00 Session 7B: Information Extraction and Question Answering (Long + TACL Papers)

15:45–16:10 *Lexical Event Ordering with an Edge-Factored Model*
Omri Abend, Shay B. Cohen and Mark Steedman

16:10–16:35 *[TACL] Entity disambiguation with web links*
Andrew Chisholm and Ben Hachey

16:35–17:00 *[TACL] A Joint Model for Entity Analysis: Coreference, Typing, and Linking*
Greg Durrett and Dan Klein

15:45–17:00 Session 7C: Machine Translation (Long Papers)

15:45–16:10 *Bag-of-Words Forced Decoding for Cross-Lingual Information Retrieval*
Felix Hieber and Stefan Riezler

16:10–16:35 *Accurate Evaluation of Segment-level Machine Translation Metrics*
Yvette Graham, Timothy Baldwin and Nitika Mathur

16:35–17:00 *Leveraging Small Multilingual Corpora for SMT Using Many Pivot Languages*
Raj Dabre, Fabien Cromieres, Sadao Kurohashi and Pushpak Bhattacharyya

Tuesday, June 2, 2015 (continued)

Session P2A: 17:00–18:30 Poster session 2A: Short papers

Why Read if You Can Scan? Trigger Scoping Strategy for Biographical Fact Extraction

Dian Yu, Heng Ji, Sujian Li and Chin-Yew Lin

Lachmannian Archetype Reconstruction for Ancient Manuscript Corpora

Armin Hoenen

Distributed Representations of Words to Guide Bootstrapped Entity Classifiers

Sonal Gupta and Christopher D. Manning

Multi-Task Word Alignment Triangulation for Low-Resource Languages

Tomer Levinboim and David Chiang

Automatic cognate identification with gap-weighted string subsequences.

Taraka Rama

Short Text Understanding by Leveraging Knowledge into Topic Model

Shansong Yang, Weiming Lu, Dezhi Yang, Liang Yao and Baogang Wei

Unsupervised Most Frequent Sense Detection using Word Embeddings

Sudha Bhingardive, Dharendra Singh, Rudramurthy V, Hanumant Redkar and Pushpak Bhattacharyya

Chain Based RNN for Relation Classification

Javid Ebrahimi and Dejing Dou

LR Parsing for LCFRS

Laura Kallmeyer and Wolfgang Maier

Mining for unambiguous instances to adapt part-of-speech taggers to new domains

Dirk Hovy, Barbara Plank, Héctor Martínez Alonso and Anders Søgaard

Clustering Sentences with Density Peaks for Multi-document Summarization

Yang Zhang, Yunqing Xia, Yi Liu and Wenmin Wang

Development of the Multilingual Semantic Annotation System

Scott Piao, Francesca Bianchi, Carmen Dayrell, Angela D'Egidio and Paul Rayson

Tuesday, June 2, 2015 (continued)

Unsupervised Sparse Vector Densification for Short Text Similarity

Yangqiu Song and Dan Roth

#WhyIStayed, #WhyILeft: Microblogging to Make Sense of Domestic Abuse

Nicolas Schrading, Cecilia Ovesdotter Alm, Raymond Ptucha and Christopher Homan

Morphological Word-Embeddings

Ryan Cotterell and Hinrich Schütze

Recognizing Social Constructs from Textual Conversation

Somak Aditya, Chitta Baral, Nguyen Ha Vo, Joohyung Lee, Jieping Ye, Zaw Naung, Barry Lumpkin, Jenny Hastings, Richard Scherl, Dawn M. Sweet and Daniela In-clezan

Two/Too Simple Adaptations of Word2Vec for Syntax Problems

Wang Ling, Chris Dyer, Alan W Black and Isabel Trancoso

Estimating Numerical Attributes by Bringing Together Fragmentary Clues

Hiroya Takamura and Jun'ichi Tsujii

Unsupervised POS Induction with Word Embeddings

Chu-Cheng Lin, Waleed Ammar, Chris Dyer and Lori Levin

Improving Update Summarization via Supervised ILP and Sentence Reranking

Chen Li, Yang Liu and Lin Zhao

MPQA 3.0: An Entity/Event-Level Sentiment Corpus

Lingjia Deng and Janyce Wiebe

Everyone Likes Shopping! Multi-class Product Categorization for e-Commerce

Zornitsa Kozareva

GPU-Friendly Local Regression for Voice Conversion

Taylor Berg-Kirkpatrick and Dan Klein

Tuesday, June 2, 2015 (continued)

Session P2B: 18:30–20:00 Poster session 2B: Short papers

Response-based Learning for Machine Translation of Open-domain Database Queries

Carolin Haas and Stefan Riezler

Context-Dependent Automatic Response Generation Using Statistical Machine Translation Techniques

Andrew Shin, Ryohei Sasano, Hiroya Takamura and Manabu Okumura

Multilingual Open Relation Extraction Using Cross-lingual Projection

Manaal Faruqui and Shankar Kumar

Learning to parse with IAA-weighted loss

Héctor Martínez Alonso, Barbara Plank, Arne Skjærholt and Anders Søgaard

Exploiting Text and Network Context for Geolocation of Social Media Users

Afshin Rahimi, Duy Vu, Trevor Cohn and Timothy Baldwin

Discriminative Phrase Embedding for Paraphrase Identification

Wenpeng Yin and Hinrich Schütze

Combining Word Embeddings and Feature Embeddings for Fine-grained Relation Extraction

Mo Yu, Matthew R. Gormley and Mark Dredze

CASSA: A Context-Aware Synonym Simplification Algorithm

Ricardo Baeza-Yates, Luz Rello and Julia Dembowski

Simple task-specific bilingual word embeddings

Stephan Gouws and Anders Søgaard

Sampling Techniques for Streaming Cross Document Coreference Resolution

Luke Shrimpton, Victor Lavrenko and Miles Osborne

On the Automatic Learning of Sentiment Lexicons

Aliaksei Severyn and Alessandro Moschitti

Large-Scale Native Language Identification with Cross-Corpus Evaluation

Shervin Malmasi and Mark Dras

Tuesday, June 2, 2015 (continued)

Unediting: Detecting Disfluencies Without Careful Transcripts

Victoria Zayats, Mari Ostendorf and Hannaneh Hajishirzi

Type-Driven Incremental Semantic Parsing with Polymorphism

Kai Zhao and Liang Huang

Template Kernels for Dependency Parsing

Hillel Taub-Tabib, Yoav Goldberg and Amir Globerson

Embedding a Semantic Network in a Word Space

Richard Johansson and Luis Nieto Piña

Random Walks and Neural Network Language Models on Knowledge Bases

Josu Goikoetxea, Aitor Soroa and Eneko Agirre

Identification and Characterization of Newsworthy Verbs in World News

Benjamin Nye and Ani Nenkova

Enhancing Sumerian Lemmatization by Unsupervised Named-Entity Recognition

Yudong Liu, Clinton Burkhart, James Hearne and Liang Luo

Extracting Information about Medication Use from Veterinary Discussions

Haibo Ding and Ellen Riloff

Reserating the awesometastic: An automatic extension of the WordNet taxonomy for novel terms

David Jurgens and Mohammad Taher Pilehvar

Cross-lingual Text Classification Using Topic-Dependent Word Probabilities

Daniel Andrade, Kunihiro Sadamasa, Akihiro Tamura and Masaaki Tsuchida

Wednesday, June 3, 2015

07:30–09:00 *Registration and Breakfast*

09:00–10:10 *Invited Talk: A Quest for Visual Intelligence in Computers*
Fei-fei Li

10:10–10:40 *Break*

10:40–11:55 **Session 8A: NLP for Web, Social Media and Social Sciences (Long + TACL Papers)**

10:40–11:05 *Testing and Comparing Computational Approaches for Identifying the Language of Framing in Political News*
Eric Baumer, Elisha Elovic, Ying Qin, Francesca Polletta and Geri Gay

11:05–11:30 *[TACL] Extracting Lexically Divergent Paraphrases from Twitter*
Wei Xu, Alan Ritter, Chris Callison-Burch, William B. Dolan, and Yangfeng Ji

11:30–11:55 *Echoes of Persuasion: The Effect of Euphony in Persuasive Communication*
Marco Guerini, Gözde Özbal and Carlo Strapparava

10:40–11:55 **Session 8B: Language and Vision (Long + TACL Papers)**

10:40–11:05 *Translating Videos to Natural Language Using Deep Recurrent Neural Networks*
Subhashini Venugopalan, Huijuan Xu, Jeff Donahue, Marcus Rohrbach, Raymond Mooney and Kate Saenko

11:05–11:30 *[TACL] A Bayesian Model of Grounded Color Semantics*
Brian McMahan and Matthew Stone

11:30–11:55 *Learning to Interpret and Describe Abstract Scenes*
Luis Gilberto Mateos Ortiz, Clemens Wolff and Mirella Lapata

Wednesday, June 3, 2015 (continued)

10:40–11:55 Session 8C: Machine Translation (Long + TACL Papers)

10:40–11:05 *A Comparison of Update Strategies for Large-Scale Maximum Expected BLEU Training*

Joern Wuebker, Sebastian Muehr, Patrick Lehen, Stephan Peitz and Hermann Ney

11:05–11:30 *[TACL] Gappy Pattern Matching on GPUs for On-Demand Extraction of Hierarchical Translation Grammars*

Hua He, Jimmy Lin, and Adam Lopez

11:30–11:55 *Learning Translation Models from Monolingual Continuous Representations*

Kai Zhao, Hany Hassan and Michael Auli

11:55–13:00 Lunch

13:00–14:00 NAACL Business Meeting

14:00–15:15 Session 9A: Lexical Semantics and Sentiment Analysis (Long Papers)

14:00–14:25 *A Corpus and Model Integrating Multiword Expressions and Supersenses*

Nathan Schneider and Noah A. Smith

14:25–14:50 *Good News or Bad News: Using Affect Control Theory to Analyze Readers' Reaction Towards News Articles*

Areej Alhothali and Jesse Hoey

14:50–15:15 *Do We Really Need Lexical Information? Towards a Top-down Approach to Sentiment Analysis of Product Reviews*

Yulia Otmakhova and Hyopil Shin

Wednesday, June 3, 2015 (continued)

14:00–15:15 Session 9B: NLP-enabled Technology (Long + TACL Papers)

14:00–14:25 *How to Memorize a Random 60-Bit String*

Marjan Ghazvininejad and Kevin Knight

14:25–14:50 *[TACL] Building a State-of-the-Art Grammatical Error Correction System*

Alla Rozovskaya and Dan Roth

14:50–15:15 *[TACL] Predicting the Difficulty of Language Proficiency Tests*

Lisa Beinborn, Torsten Zesch, and Iryna Gurevych

14:00–15:15 Session 9C: Linguistic and Psycholinguistic Aspects of CL (Long Papers)

14:00–14:25 *A Bayesian Model for Joint Learning of Categories and their Features*

Lea Frermann and Mirella Lapata

14:25–14:50 *Shared common ground influences information density in microblog texts*

Gabriel Doyle and Michael Frank

14:50–15:15 *Hierarchic syntax improves reading time prediction*

Marten van Schijndel and William Schuler

15:15–15:45 Break

Wednesday, June 3, 2015 (continued)

15:45–17:15 Best Paper Plenary Session

15:45–16:15 *Retrofitting Word Vectors to Semantic Lexicons*

Manaal Faruqui, Jesse Dodge, Sujay Kumar Jauhar, Chris Dyer, Eduard Hovy and Noah A. Smith

16:15–16:45 *“You’re Mr. Lebowski, I’m the Dude” : Inducing Address Term Formality in Signed Social Networks*

Vinodh Krishnan and Jacob Eisenstein

16:45–17:15 *Unsupervised Morphology Induction Using Word Embeddings*

Radu Soricut and Franz Och

17:15–17:30 Best paper awards and Closing Remarks

