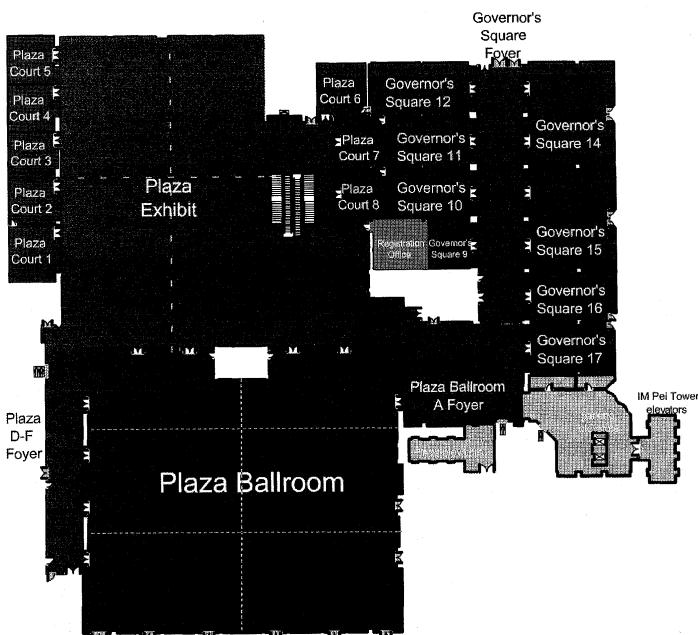
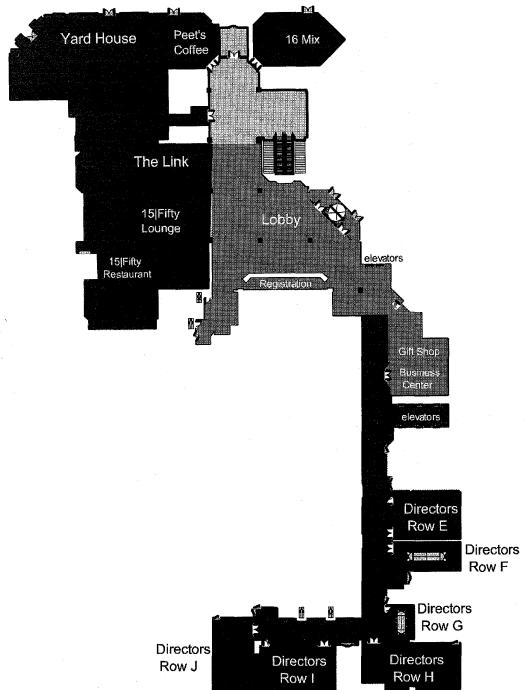


The 14th Conference of the
**North American Chapter of the Association for
Computational Linguistics:**
Human Language Technologies

CONFERENCE HANDBOOK



**CONCOURSE LEVEL
LOBBY LEVEL**



*Cover design by Aurelia Bunescu
Handbook assembled by Matt Post
Printing by Omnipress of Madison, Wisconsin*

Contents

Table of Contents	i
1 Conference Information	3
Message from the General Chair	3
Message from the Program Committee Co-Chairs	5
Organizing Committee	7
Program Committee	8
Meal Info	10
2 Tutorials: Sunday, May 31	11
T1: Hands-on Learning to Search for Structured Prediction	12
T2: Crowdsourcing for NLP	14
T3: The Logic of AMR	16
T4: Social Media Predictive Analytics	17
T5: Deep Learning and Continuous Representations for NLP	19
T6: Using FrameNet in NLP	21
Welcome Reception	22
3 Main Conference: Monday, June 1	23
Keynote Address: Lillian Lee	24
Session 1	25
Student Lunch	29
Session 2	30
Session 3	34
One Minute Madness!	38
Poster session 1A	38
Poster session 1B	45
SRW Poster Session	51

4 Main Conference: Tuesday, June 2	57
Session 4	59
Session 5	63
Session 6	67
Session 7	71
Demo Session A	75
Demo Session B	77
Poster session 2A (Short papers)	79
Poster session 2B (Short papers)	83
Social Event	88
5 Main Conference: Wednesday, June 3	89
Keynote Address: Fei-Fei Li	90
Session 8	91
NAACL Business Meeting	95
Session 9	96
Best Paper Plenary Session	100
6 Colocated conference: *SEM: Thursday–Friday, June 4–5	101
*SEM: 4th Joint Conference on Lexical and Computational Semantics	103
7 Workshops: Thursday–Friday, June 4–5	107
W1: 9th International Workshop on Semantic Evaluations	108
W2: 2nd Workshop on Argumentation Mining	118
W3: The 10th Workshop on Innovative Use of NLP for Building Educational Applications	120
W4: Fourth Workshop on Computational Linguistics for Literature	122
W5: Cognitive Modeling and Computational Linguistics	124
W6: The 3rd Workshop on EVENTS: Definition, Detection, Coreference, and Representation	125
W7: The 11th Workshop on Multiword Expressions	127
W8: Ninth Workshop on Syntax, Semantics and Structure in Statistical Translation	129
W9: Computational Linguistics and Clinical Psychology — From Linguistic Signal to Clinical Reality	131
W10: Extra-Propositional Aspects of Meaning in Computational Linguistics	133
W11: The 9th Linguistic Annotation Workshop	134
W12: The Third Workshop on Metaphor in NLP	136
W13: Third International Workshop on Natural Language Processing for Social Media	137
W14: Vector Space Modeling for NLP	138
8 Anti-harassment policy	141
Author Index	143
9 Local Guide	159



Big Data

Natural Language Processing • Graph Analytics • Statistical Modeling

Learn more at goldmansachs.com



© 2015 Goldman Sachs. All rights reserved.

YAHOO! LABS

Science-Driven Innovation

Powering products at massive scale with cutting-edge research

Yahoo Labs powers Yahoo's most critical products with innovative science.

- We serve as Yahoo's research arm—its incubator for bold new ideas and laboratory for rigorous experimentation.
- Yahoo Labs applies its scientific findings in powering products for Yahoo's users and enhancing value for its partners and advertisers.
- We create technology that is used by hundreds of millions of customers every day.
- Our forward-looking innovation also helps position Yahoo as an industry and scientific thought leader.

 labs.yahoo.com



Join Us!

SDL.com/Research

Los Angeles, CA

Cambridge, UK

WHAT WE

MAKE

BREAKS

BOUNDARIES

CAREERS

   /bloombergcareers

Make your mark.

The world's top investors, traders and leaders depend on our information and news. We need your ideas and passion to help us meet some of the biggest challenges around.

jobs.bloomberg.com

Bloomberg

©2015 Bloomberg L.P. All rights reserved. S571691131 031

Conference Information

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 – Lilian 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 Thamar 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

Message from the General Chair

research workshop; Matt Gerber, Catherine Havasi, and Finley Lacatusu (demo co-chairs); Annie Louis (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 Committee Co-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 presentatons.

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.

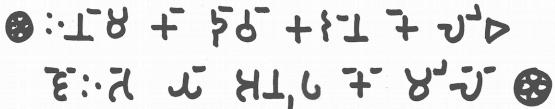
We would like to thank the ACL Business Manager Priscilla Rasmussen. She was our *go to* person who knew all details of the conference in and out. We are very grateful for her help.

Finally, this conference could not have happened without the efforts of the general chair, Rada

Message from the Program Committee Co-Chairs

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 Houston, 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, Johns 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

Meal Info

The following meals are provided as part of your registration fee:

- A full buffet breakfast will be provided each day in the Plaza Exhibit (foyer)
- Mid-Morning breaks include coffee and tea in the Plaza Exhibit (foyer)
- Mid-Afternoon breaks include coffee, tea, soda, water, and snacks in the Plaza Exhibit (foyer)
- A full dinner buffer is provided during the poster sessions on Monday and Tuesday evenings in the Plaza Exhibit (foyer)

Lunch is provided for students on Monday, but you are otherwise on your own.

2

Tutorials: Sunday, May 31

Overview

7:30–6:00	Registration	<i>Plaza Exhibit All</i>
7:30–9:00	Breakfast	<i>Plaza Exhibit All</i>
9:00–12:30	Morning Tutorials	
	Hands-on Learning to Search for Structured Prediction <i>Hal Daumé III, John Langford, Kai-Wei Chang, He He, and Sudha Rao</i>	<i>Governor's Square 15</i>
	Crowdsourcing for NLP <i>Chris Callison-Burch, Lyle Ungar, and Ellie Pavlick</i>	<i>Governor's Square 17</i>
	The Logic of AMR: Practical, Unified, Graph-Based Sentence Semantics for NLP <i>Nathan Schneider, Jeffrey Flanigan, and Tim O'Gorman</i>	<i>Governor's Square 16</i>
12:30–2:00	Lunch break	
2:00–5:30	Afternoon Tutorials	
	Social Media Predictive Analytics <i>Svitlana Volkova, Benjamin Van Durme, David Yarowsky, and Yoram Bachrach</i>	<i>Governor's Square 17</i>
	Deep Learning and Continuous Representations for Natural Language Processing <i>Wen-tau Yih, Xiaodong He, and Jianfeng Gao</i>	<i>Governor's Square 15</i>
	Getting the Roles Right: Using FrameNet in NLP <i>Collin Baker, Nathan Schneider, Miriam R L Petrucc, and Michael Ellsworth</i>	<i>Governor's Square 16</i>
6:00–9:00	Welcome Reception	<i>Foyer</i>

Tutorial 1

Hands-on Learning to Search for Structured Prediction

Hal Daumé III, John Langford, Kai-Wei Chang, He He, and Sudha Rao

Sunday, May 31, 2015, 9:00–12:30pm

Governor's Square 15

Many problems in natural language processing involve building outputs that are structured. The predominant approach to structured prediction is “global models” (such as conditional random fields), which have the advantage of clean underlying semantics at the cost of computational burdens and extreme difficulty in implementation. An alternative strategy is the “learning to search” (L2S) paradigm, in which the structured prediction task is cast as a sequential decision making process. One can then devise training-time algorithms that learn to make near optimal collective decisions. This paradigm has been gaining increasing traction over the past five years: most notably in dependency parsing (e.g., MaltParser, ClearNLP, etc.), but also much more broadly in less “sequential” tasks like entity/relation classification and even graph prediction problems found in social network analysis and computer vision.

Hal Daumé III is an associate professor in Computer Science at the University of Maryland, College Park. He holds joint appointments in UMIACS and Linguistics. His primary research interest is in developing new learning algorithms for prototypical problems that arise in the context of language processing and artificial intelligence. This includes topics like structured prediction, domain adaptation and unsupervised learning; as well as multilingual modeling and affect analysis.

John Langford is a machine learning research scientist. He is the author of the blog hunch.net and the principal developer of Vowpal Wabbit. John works at Microsoft Research New York, of which he was one of the founding members, and was previously affiliated with Yahoo! Research, Toyota Technological Institute, and IBM’s Watson Research Center. He studied Physics and Computer Science at the California Institute of Technology, earning a double bachelor’s degree in 1997, and received his Ph.D. in Computer Science from Carnegie Mellon University in 2002. He was the program co-chair for the 2012 International Conference on Machine Learning.

Kai-Wei Chang is a doctoral candidate in Computer Science at the University of Illinois at Urbana-Champaign. His research interests lie in designing practical machine learning techniques for large and complex data and applying them to real world applications. He has been working on various topics in Machine learning and Natural Language Processing, including large-scale learning, structured learning, coreference resolution, and relation extraction.

He He is a doctoral student in Computer Science at the University of Maryland, College Park. Her research interests focus on developing fast inference methods for structured prediction problems in machine learning and natural language processing, including features selection, dependency parsing and machine translation.

Sudha Rao is a doctoral student in Computer Science at the University of Maryland, College Park. Her research interests currently lie in exploring how semantic representations like Abstract Meaning Representation can be used in Machine Translation.

This tutorial has precisely one goal: an attendee should leave the tutorial with hands on experience writing small programs to perform structured prediction for a variety of tasks, like sequence labeling, dependency parsing and, time-permitting, more.

This proposed tutorial is unique (to our knowledge) among ACL tutorials in this regard: half of the time spent will be in the style of a “flipped classroom” in which attendees get hands on experience writing structured predictors on their own or in small groups. All course materials (software, exercises, hints, solutions, etc., will be made available at least two weeks prior to the event so that students can download the required data ahead of time; we will also bring copies on USB in case there is a problem with the internet).

The first half of the tutorial will be mostly “lecture” style, in which we will cover the basics of how learning to search works for structured prediction. The goal is to provide enough background information that students can understand how to write and debug their own predictors, but the emphasis will not be on how to build new machine learning algorithms. This will also include a brief tutorial on the basics of Vowpal Wabbit, to the extent necessary to understand its structured prediction interface. The second half of the tutorial will focus on hands-on exploration of structured prediction using the Vowpal Wabbit python “learning to search” interface; a preliminary python notebook explaining the interface can be viewed at <http://tinyurl.com/pyvwsearch>; an elaborated version of this notebook will serve as the backbone for the “hands on” part of the tutorial, paired with exercises.

Tutorial 2

Crowdsourcing for NLP

Chris Callison-Burch, Lyle Ungar, and Ellie Pavlick

Sunday, May 31, 2015, 9:00–12:30pm

Governor's Square 17

Crowdsourced applications to scientific problems is a hot research area, with over 10,000 publications in the past five years. Platforms such as Amazon's Mechanical Turk and CrowdFlower provide researchers with easy access to large numbers of workers. The crowd's vast supply of inexpensive, intelligent labor allows people to attack problems that were previously impractical and gives potential for detailed scientific inquiry of social, psychological, economic, and linguistic phenomena via massive sample sizes of human annotated data. We introduce crowdsourcing and describe how it is being used in both industry and academia. Crowdsourcing is valuable to computational linguists both (a) as a source of labeled training data for use in machine learning

Chris Callison-Burch is the Aravind K Joshi term assistant professor in the Computer and Information Science Department at the University of Pennsylvania. Before joining Penn, he was a research faculty member at the Center for Language and Speech Processing at Johns Hopkins University for 6 years. He was the Chair of the Executive Board of the North American chapter of the Association for Computational Linguistics (NAACL) from 2011-2013, and he has served on the editorial boards of the journals Transactions of the ACL (TACL) and Computational Linguistics. He is a Sloan Research Fellow, and he has received faculty research awards from Google, Microsoft and Facebook in addition to funding from DARPA and the NSF. Chris teaches a semester-long course on Crowdsourcing at Penn (<http://crowdsourcing-class.org/>).

Lyle Ungar is a Professor of Computer and Information Science at the University of Pennsylvania. He also holds appointments in several other departments in the Engineering, Medicine, and Business Schools. Dr. Ungar received a B.S. from Stanford University and a Ph.D. from M.I.T. He has published over 200 articles and is co-inventor on eight patents. His current research includes machine learning, data mining, and text mining, and uses social media to better understand the drivers of physical and mental well-being. Lyle's research group collects MTurk crowdsourced labels on natural language data such Facebook posts and tweets, which they use for a variety of NLP and psychology studies. Lyle (with collaborators) has given highly successful tutorials on information extraction, sentiment analysis, and spectral methods for NLP at conferences including NAACL, KDD, SIGIR, ICWSM, CIKM, and AAAI. He and his student gave a tutorial on crowdsourcing last year at the Joint Statistical Meetings (JSM).

Ellie Pavlick is a Ph.D. student at the University of Pennsylvania. Ellie received her B.A. in economics from the Johns Hopkins University, where she began working with Dr. Chris Callison-Burch on using crowdsourcing to create low-cost training data for statistical machine translation by hiring non-professional translators and post-editors. Her current research interests include entailment and paraphrase recognition, for which she has looked at using MTurk to provide more difficult linguistic annotations such as discriminating between fine-grained lexical entailment relations and identifying missing lexical triggers in FrameNet. Ellie TAed and helped design the curriculum for the Crowdsourcing and Human Computation course at Penn.

and (b) as a means of collecting computational social science data that link language use to underlying beliefs and behavior. We present case studies for both categories: (a) collecting labeled data for use in natural language processing tasks such as word sense disambiguation and machine translation and (b) collecting experimental data in the context of psychology; e.g. finding how word use varies with age, sex, personality, health, and happiness.

We will also cover tools and techniques for crowdsourcing. Effectively collecting crowdsourced data requires careful attention to the collection process, through selection of appropriately qualified workers, giving clear instructions that are understandable to non-experts, and performing quality control on the results to eliminate spammers who complete tasks randomly or carelessly in order to collect the small financial reward. We will introduce different crowdsourcing platforms, review privacy and institutional review board issues, and provide rules of thumb for cost and time estimates. Crowdsourced data also has a particular structure that raises issues in statistical analysis; we describe some of the key methods to address these issues.

Tutorial 3

The Logic of AMR: Practical, Unified, Graph-Based Sentence Semantics for NLP

Nathan Schneider, Jeffrey Flanigan, and Tim O’Gorman

Sunday, May 31, 2015, 9:00–12:30pm

Governor’s Square 16

The Abstract Meaning Representation formalism is rapidly emerging as an important practical form of structured sentence semantics which, thanks to the availability of large-scale annotated corpora, has potential as a convergence point for NLP research. This tutorial unmasks the design philosophy, data creation process, and existing algorithms for AMR semantics. It is intended for anyone interested in working with AMR data, including parsing text into AMRs, generating text from AMRs, and applying AMRs to tasks such as machine translation and summarization.

The goals of this tutorial are twofold. First, it will describe the nature and design principles behind the representation, and demonstrate that it can be practical for annotation. In Part I: The AMR Formalism, participants will be coached in the basics of annotation so that, when working with AMR data in the future, they will appreciate the benefits and limitations of the process by which it was created. Second, the tutorial will survey the state of the art for computation with AMRs. Part II: Algorithms and Applications will focus on the task of parsing English text into AMR graphs, which requires algorithms for alignment, for structured prediction, and for statistical learning. The tutorial will also address graph grammar formalisms that have been recently developed, and future applications such as AMR-based machine translation and summarization.

Nathan Schneider is an annotation schemer and computational modeler for natural language. He has been involved in the design of the AMR formalism since 2012, when he interned with Kevin Knight at ISI. His 2014 dissertation introduced a coarse-grained representation for lexical semantics that facilitates rapid annotation and is practical for broad-coverage statistical NLP. He has also worked on semantic parsing for the FrameNet representation and other forms of syntactic/semantic annotation and processing for social media text. For most of these projects, he led the design of the annotation scheme, guidelines, and workflows, and the training and supervision of annotators.

Tim O’Gorman is a third year Linguistics Ph.D. student at the University of Colorado – Boulder, working with Martha Palmer. He manages CU-Boulder’s AMR annotation team, and participated in the Fred Jelinek Memorial Workshop in Prague in 2014, working on mapping the Prague tectogrammatical layer to AMRs. His research areas include implicit arguments, semantic role projection, and linking sentence-level semantics to discourse.

Jeffrey Flanigan is a fifth year Ph.D. candidate at Carnegie Mellon University. He and his collaborators at CMU built the first broad-coverage AMR parser. He also participated in the Fred Jelinek Memorial Workshop in Prague in 2014, working on cross-lingual parsing and generation from AMR. His research areas include machine translation, generation, summarization, and semantic parsing.

Tutorial 4

Social Media Predictive Analytics

Svitlana Volkova, Benjamin Van Durme, David Yarowsky, and Yoram Bachrach

Sunday, May 31, 2015, 2:00–5:30pm

Governor's Square 17

The recent explosion of social media services like Twitter, Google+ and Facebook has led to an interest in social media predictive analytics – automatically inferring hidden information from the large amounts of freely available content. It has a number of applications, including: online targeted advertising, personalized marketing, large-scale passive polling and real-time live polling, personalized recommendation systems and search, and real-time healthcare analytics etc.

In this tutorial, we will describe how to build a variety of social media predictive analytics for inferring latent user properties from a Twitter network including demographic traits, personality, interests, emotions and opinions etc. Our methods will address several important aspects of social media such as: dynamic, streaming nature of the data, multi-relationality in social networks, data collection and annotation biases, data and model sharing, generalization of the existing models, data drift, and scalability to other languages.

We will start with an overview of the existing approaches for social media predictive analytics. We will describe the state-of-the-art static (batch) models and features. We will then present models for streaming (online) inference from single and multiple data streams; and formulate a latent attribute prediction task as a sequence-labeling problem. Finally, we present several techniques for dynamic (iterative) learning and prediction using active learning setup with rationale annotation and filtering. The tutorial will conclude with a practice session focusing on

Svitlana Volkova is a Ph.D. Candidate in Computer Science at the Center for Language and Speech Processing, Johns Hopkins University. She works on machine learning and natural language processing techniques for social media predictive analytics. She develops batch and streaming (dynamic) models for automatically inferring psycho-demographic profiles from social media data streams, fine-grained emotion detection and sentiment analysis for under-explored languages and dialects in microblogs, effective interactive and iterative rationale annotation via crowdsourcing.

Benjamin Van Durme is the Chief Lead of Text Research at the Human Language Technology Center of Excellence, and an Assistant Research Professor at the Center for Language and Speech Processing. He works on natural language processing (specifically computational semantics), predictive analytics in social media and streaming/randomized algorithms.

David Yarowsky is a Professor at the Center for Language and Speech Processing, Johns Hopkins University. His research interests include natural language processing and spoken language systems, machine translation, information retrieval, very large text databases and machine learning. His research focuses on word sense disambiguation, minimally supervised induction algorithms in NLP, and multilingual natural language processing.

Yoram Bachrach is a researcher in the Online Services and Advertising group at Microsoft Research Cambridge UK. His research area is artificial intelligence (AI), focusing on multi-agent systems and computational game theory. Computational game theory combines the theoretical foundations of economics and game theory with creative solutions from AI and computer science.

walk-through examples for predicting latent user properties e.g., political preferences, income, education level, life satisfaction and emotions emanating from user communications on Twitter.

Tutorial 5

Deep Learning and Continuous Representations for Natural Language Processing

Wen-tau Yih, Xiaodong He, and Jianfeng Gao

Sunday, May 31, 2015, 2:00–5:30pm

Scott Wen-tau Yih is a Researcher in the Machine Learning Group at Microsoft Research Redmond. His research interests include natural language processing, machine learning and information retrieval. Yih received his Ph.D. in computer science at the University of Illinois at Urbana-Champaign. His work on joint inference using integer linear programming (ILP) [Roth & Yih, 2004] helped the UIUC team win the CoNLL-05 shared task on semantic role labeling, and the approach has been widely adopted in the NLP community. After joining MSR in 2005, he has worked on email spam filtering, keyword extraction and search & ad relevance. His recent work focuses on continuous semantic representations using neural networks and matrix/tensor decomposition methods, with applications in lexical semantics, knowledge base embedding and question answering. Yih received the best paper award from CoNLL-2011 and has served as area chairs (HLT-NAACL-12, ACL-14) and program co-chairs (CEAS-09, CoNLL-14) in recent years.

Xiaodong He is a Researcher of Microsoft Research, Redmond, WA, USA. He is also an Affiliate Professor in Electrical Engineering at the University of Washington, Seattle, WA, USA. His research interests include deep learning, information retrieval, natural language understanding, machine translation, and speech recognition. Dr. He has published a book and more than 70 technical papers in these areas, and has given tutorials at international conferences in these fields. In benchmark evaluations, he and his colleagues have developed entries that obtained No. 1 place in the 2008 NIST Machine Translation Evaluation (NIST MT) and the 2011 International Workshop on Spoken Language Translation Evaluation (IWSLT), both in Chinese-English translation, respectively. He serves as Associate Editor of IEEE Signal Processing Magazine and IEEE Signal Processing Letters, as Guest Editors of IEEE TASLP for the Special Issue on Continuous-space and related methods in natural language processing, and Area Chair of NAACL2015. He also served as GE for several IEEE Journals, and served in organizing committees and program committees of major speech and language processing conferences in the past. He is a senior member of IEEE and a member of ACL.

Jianfeng Gao is a Principal Researcher of Microsoft Research, Redmond, WA, USA. His research interests include Web search and information retrieval, natural language processing and statistical machine learning. Dr. Gao is the primary contributor of several key modeling technologies that help significantly boost the relevance of the Bing search engine. His research has also been applied to other MS products including Windows, Office and Ads. In benchmark evaluations, he and his colleagues have developed entries that obtained No. 1 place in the 2008 NIST Machine Translation Evaluation in Chinese-English translation. He was Associate Editor of ACM Trans on Asian Language Information Processing, (2007 to 2010), and was Member of the editorial board of Computational Linguistics (2006 – 2008). He also served as area chairs for ACL-IJCNLP2015, SIGIR2015, SIGIR2014, IJCAI2013, ACL2012, EMNLP2010, ACL-IJCNLP 2009, etc. Dr. Gao recently joined Deep Learning Technology Center at MSR-NExT, working on Enterprise Intelligence.

Deep learning techniques have demonstrated tremendous success in the speech and language processing community in recent years, establishing new state-of-the-art performance in speech recognition, language modeling, and have shown great potential for many other natural language processing tasks. The focus of this tutorial is to provide an extensive overview on recent deep learning approaches to problems in language or text processing, with particular emphasis on important real-world applications including language understanding, semantic representation modeling, question answering and semantic parsing, etc.

In this tutorial, we will first survey the latest deep learning technology, presenting both theoretical and practical perspectives that are most relevant to our topic. We plan to cover common methods of deep neural networks and more advanced methods of recurrent, recursive, stacking and convolutional networks. In addition, we will introduce recently proposed continuous-space representations for both semantic word embedding and knowledge base embedding, which are modeled by either matrix/tensor decomposition or neural networks.

Next, we will review general problems and tasks in text/language processing, and underline the distinct properties that differentiate language processing from other tasks such as speech and image object recognition. More importantly, we highlight the general issues of natural language processing, and elaborate on how new deep learning technologies are proposed and fundamentally address these issues. We then place particular emphasis on several important applications, including (1) machine translation, (2) semantic information retrieval and (3) semantic parsing and question answering. For each task, we will discuss what particular architectures of deep learning models are suitable given the nature of the task, and how learning can be performed efficiently and effectively using end-to-end optimization strategies.

Tutorial 6

Getting the Roles Right: Using FrameNet in NLP

Collin Baker, Nathan Schneider, Miriam R L Petrucc, and Michael Ellsworth

Sunday, May 31, 2015, 2:00–5:30pm

Governor's Square 16

The FrameNet lexical database (Fillmore & Baker 2010, Ruppenhofer et al. 2006, <http://framenet.icsi.berkeley.edu>), covers roughly 13,000 lexical units (word senses) for the core English lexicon, associating them with roughly 1,200 fully defined semantic frames; these frames and their roles cover the majority of event types in everyday, non-specialist text, and they are documented with 200,000 manually annotated examples. This tutorial will teach attendees what they need to know to start using the FrameNet lexical database as part of an NLP system. We will cover the basics of Frame Semantics, explain how the database was created, introduce the Python API and the state of the art in automatic frame semantic role labeling systems; and we will discuss FrameNet collaboration with commercial partners. Time permitting, we will present new research on frames and annotation of locative relations, as well as corresponding metaphorical uses, along with information about how frame semantic roles can aid the interpretation of metaphors.

Collin Baker has been Project Manager of the FrameNet Project since 2000. His research interests include FrameNets in other languages (Loenneker-Rodman & Baker 2009), aligning FrameNet to other lexical resources (Fellbaum & Baker 2013, Ferrandez et al 2010), linking to ontologies and reasoning (Scheffczyk et al. 2010), and the frame semantics of metaphor.

Nathan Schneider has worked on a coarse-grained representation for lexical semantics (2014 dissertation at Carnegie Mellon University) and the design of the Abstract Meaning Representation (AMR; Banarescu et al. 2014). Nathan helped develop the leading open-source frame-semantic parser for English, SEMAFOR (Das et al. 2010, 2014) (<http://demo.ark.cs.cmu.edu/parse>), as well as a Python interface to the FrameNet lexicon (with Chuck Wooters) that is part of the NLTK suite.

Miriam R. L. Petrucc received her PhD in Linguistics from the University of California, Berkeley. A key member of the team developing FrameNet almost since the project's founding, her research interests include semantics, knowledge base development, grammar and lexis, lexical semantics, Frame Semantics and Construction Grammar.

Michael Ellsworth has been involved with FrameNet for well over a decade. His chief focus is on semantic relations in FrameNet (Ruppenhofer et al. 2006), how they can be used for paraphrase (Ellsworth & Janin 2007), and mapping to other resources (Scheffczyk et al. 2006, Ferrandez et al. 2010). Increasingly, he has examined the connection of FrameNet to syntax and the Constructionicon (Torrent & Ellsworth 2013, Ziem & Ellsworth 2015), including in his pending dissertation on the constructions and frame semantics of emotion.

Welcome Reception

Sunday, May 31, 2015, 6:00pm – 9:00pm

Sheraton Denver Downtown Hotel (conference venue)
Foyer

Catch up with your colleagues at the **Welcome Reception!** It will be held immediately following the Tutorials on Sunday, May 31 at 6:00pm in the Plaza Exhibit (foyer) of the Sheraton Denver Downtown Hotel (the conference venue). Refreshments and a light dinner will be provided, and a cash bar will be available.

Main Conference: Monday, June 1

Overview

7:30–8:45	Registration and Breakfast	<i>Plaza Exhibit All</i>	
8:45–9:00	Welcome to NAACL HLT 2015	<i>Plaza Ballroom A, B, & C</i>	
9:00–10:10	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)	<i>Plaza Ballroom A, B, & C</i>	
10:10–10:40	Break	<i>Plaza Exhibit All</i>	
	Session 1		
10:40–12:20	Semantics (Long Papers) <i>Plaza Ballroom A & B</i>	Tagging, Chunking, Syntax and Parsing (Long + TACL Papers) <i>Plaza Ballroom D & E</i>	Information Retrieval, Text Categorization, Topic Modeling (Long Papers) <i>Plaza Ballroom F</i>
12:20–2:00	Lunch		
	Session 2		
2:00–3:15	Generation and Summarization (Long Papers) <i>Plaza Ballroom A & B</i>	Language and Vision (Long Papers) <i>Plaza Ballroom D & E</i>	NLP for Web, Social Media and Social Sciences (Long Papers) <i>Plaza Ballroom F</i>
	Session 3		
3:15–4:00	Generation and Summarization (Short Papers) <i>Plaza Ballroom A & B</i>	Information Extraction and Question Answering (Short Papers) <i>Plaza Ballroom D & E</i>	Machine Learning for NLP (Short Papers) <i>Plaza Ballroom F</i>
4:00–4:30	Break	<i>Plaza Exhibit All</i>	
4:30–6:00	One minute madness (Long + TACL papers)	<i>Plaza Ballroom A, B, & C</i>	
6:00–7:30	Poster session 1A: Long + TACL papers	<i>Plaza Ballroom A, B, & C</i>	
7:30–9:00	Poster session 1B: Long + TACL papers	<i>Plaza Ballroom A, B, & C</i>	
6:00–9:00	SRW Poster Session	<i>Plaza Ballroom A, B, & C</i>	

Keynote Address: Lillian Lee

“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

Monday, June 1, 2015, 9:00–10:10am

Plaza Ballroom A, B, & C

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.

Session 1 Overview – Monday, June 1, 2015

Track A	Track B	Track C	
<i>Semantics (Long Papers)</i> Plaza Ballroom A & B	<i>Tagging, Chunking, Syntax and Parsing (Long + TACL Papers)</i> Plaza Ballroom D & E	<i>Information Retrieval, Text Categorization, Topic Modeling (Long Papers)</i> Plaza Ballroom F	
Unsupervised Induction of Semantic Roles within a Reconstruction-Error Minimization Framework <i>Titov and Khodam</i>	Randomized Greedy Inference for Joint Segmentation, POS Tagging and Dependency Parsing <i>Zhang, Li, Barzilay, and Darwisch</i>	A Hybrid Generative/Discriminative Approach To Citation Prediction <i>Tanner and Charniak</i>	11:05
Predicate Argument Alignment using a Global Coherence Model <i>Wolfe, Dredze, and Van Durme</i>	An Incremental Algorithm for Transition-based CCG Parsing <i>Ambati, Deoskar, Johnson, and Steedman</i>	Weakly Supervised Slot Tagging with Partially Labeled Sequences from Web Search Click Logs <i>Kim, Jeong, Stratos, and Sarikaya</i>	11:30
Improving unsupervised vector-space thematic fit evaluation via role-filler prototype clustering <i>Greenberg, Sayeed, and Demberg</i>	Because Syntax Does Matter: Improving Predicate-Argument Structures Parsing with Syntactic Features <i>Ribeyre, Villemonte de la Clergerie, and Seddah</i>	Not All Character N-grams Are Created Equal: A Study in Authorship Attribution <i>Sapkota, Bethard, Montes, and Solorio</i>	11:55
A Compositional and Interpretable Semantic Space <i>Fyshe, Wehbe, Talukdar, Murphy, and Mitchell</i>	[TACL] Exploring Compositional Architectures and Word Vector Representations for Prepositional Phrase Attachment <i>Belinkov, Lei, Barzilay, and Globerson</i>	Effective Use of Word Order for Text Categorization with Convolutional Neural Networks <i>Johnson and Zhang</i>	12:20

Parallel Session 1

Session 1A: Semantics (Long Papers)

Plaza Ballroom A & B

Chair: Hoifung Poon

Unsupervised Induction of Semantic Roles within a Reconstruction-Error Minimization Framework

Ivan Titov and Ehsan Khoddam

10:40–11:05

We introduce a new approach to unsupervised estimation of feature-rich semantic role labeling models. Our model consists of two components: (1) an encoding component: a semantic role labeling model which predicts roles given a rich set of syntactic and lexical features; (2) a reconstruction component: a tensor factorization model which relies on roles to predict argument fillers. When the components are estimated jointly to minimize errors in argument reconstruction, the induced roles largely correspond to roles defined in annotated resources. Our method performs on par with most accurate role induction methods on English and German, even though, unlike these previous approaches, we do not incorporate any prior linguistic knowledge about the languages.

Predicate Argument Alignment using a Global Coherence Model

Travis Wolfe, Mark Dredze, and Benjamin Van Durme

11:05–11:30

We present a joint model for predicate argument alignment. We leverage multiple sources of semantic information, including temporal ordering constraints between events. These are combined in a max-margin framework to find a globally consistent view of entities and events across multiple documents, which leads to improvements on the state-of-the-art.

Improving unsupervised vector-space thematic fit evaluation via role-filler prototype clustering

Clayton Greenberg, Asad Sayeed, and Vera Demberg

11:30–11:55

Most recent unsupervised methods in vector space semantics for assessing thematic fit (e.g. Erk, 2007; Baroni and Lenci, 2010; Sayeed and Demberg, 2014) create prototypical role-fillers without performing word sense disambiguation. This leads to a kind of sparsity problem: candidate role-fillers for different senses of the verb end up being measured by the same “yardstick”, the single prototypical role-filler. In this work, we use three different feature spaces to construct robust unsupervised models of distributional semantics. We show that correlation with human judgements on thematic fit estimates can be improved consistently by clustering typical role-fillers and then calculating similarities of candidate role-fillers with these cluster centroids. The suggested methods can be used in any vector space model that constructs a prototype vector from a non-trivial set of typical vectors.

A Compositional and Interpretable Semantic Space

Alona Fyshe, Leila Wehbe, Partha P. Talukdar, Brian Murphy, and Tom M. Mitchell 11:55–12:20

Vector Space Models (VSMs) of Semantics are useful tools for exploring the semantics of single words, and the composition of words to make phrasal meaning. While many methods can estimate the meaning (i.e. vector) of a phrase, few do so in an interpretable way. We introduce a new method (CNNSE) that allows word and phrase vectors to adapt to the notion of composition. Our method learns a VSM that is both tailored to support a chosen semantic composition operation, and whose resulting features have an intuitive interpretation. Interpretability allows for the exploration of phrasal semantics, which we leverage to analyze performance on a behavioral task.

Session 1B: Tagging, Chunking, Syntax and Parsing (Long + TACL Papers)

Plaza Ballroom D & E

Chair: Noah A. Smith

Randomized Greedy Inference for Joint Segmentation, POS Tagging and Dependency Parsing*Yuan Zhang, Chengtao Li, Regina Barzilay, and Kareem Darwish*

10:40–11:05

In this paper, we introduce a new approach for joint segmentation, POS tagging and dependency parsing. While joint modeling of these tasks addresses the issue of error propagation inherent in traditional pipeline architectures, it also complicates the inference task. Past research has addressed this challenge by placing constraints on the scoring function. In contrast, we propose an approach that can handle arbitrarily complex scoring functions. Specifically, we employ a randomized greedy algorithm that jointly predicts segmentations, POS tags and dependency trees. Moreover, this architecture readily handles different segmentation tasks, such as morphological segmentation for Arabic and word segmentation for Chinese. The joint model outperforms the state-of-the-art systems on three datasets, obtaining 2.1% TedEval absolute gain against the best published results in the 2013 SPMRL shared task.

An Incremental Algorithm for Transition-based CCG Parsing*Bharat Ram Ambari, Tejaswini Deoskar, Mark Johnson, and Mark Steedman*

11:05–11:30

Incremental parsers have potential advantages for applications like language modeling for machine translation and speech recognition. We describe a new algorithm for incremental transition-based Combinatory Categorial Grammar parsing. As English CCGbank derivations are mostly right branching and non-incremental, we design our algorithm based on the dependencies resolved rather than the derivation. We introduce two new actions in the shift-reduce paradigm based on the idea of ‘revealing’ (Pareschi and Steedman, 1987) the required information during parsing. On the standard CCGbank test data, our algorithm achieved improvements of 0.88% in labeled and 2.0% in unlabeled F-score over a greedy non-incremental shift-reduce parser.

Because Syntax Does Matter: Improving Predicate-Argument Structures Parsing with Syntactic Features*Corentin Ribeyre, Eric Villemonte de la Clergerie, and Djamel Seddah*

11:30–11:55

Parsing full-fledged predicate-argument structures in a deep syntax framework requires graphs to be predicted. Using the DeepBank (Flickinger et al., 2012) and the Predicate-Argument Structure treebank (Miyao and Tsuji, 2005) as a test field, we show how transition-based parsers, extended to handle connected graphs, benefit from the use of topologically different syntactic features such as dependencies, tree fragments, spines or syntactic paths, bringing a much needed context to the parsing models, improving notably over long distance dependencies and elided coordinate structures. By confirming this positive impact on an accurate 2nd-order graph-based parser (Martins and Almeida, 2014), we establish a new state-of-the-art on these data sets.

[TACL] Exploring Compositional Architectures and Word Vector Representations for Prepositional Phrase Attachment*Yonatan Belinkov, Tao Lei, Regina Barzilay, and Amir Globerson*

11:55–12:20

Prepositional phrase (PP) attachment disambiguation is a known challenge in syntactic parsing. The lexical sparsity associated with PP attachments motivates research in word representations that can capture pertinent syntactic and semantic features of the word. One promising solution is to use word vectors induced from large amounts of raw text. However, state-of-the-art systems that employ such representations yield modest gains in PP attachment accuracy. In this paper, we show that word vector representations can yield significant PP attachment performance gains. This is achieved via a non-linear architecture that is discriminatively trained to maximize PP attachment accuracy. The architecture is initialized with word vectors trained from unlabeled data, and relearns those to maximize attachment accuracy. We obtain additional performance gains with alternative representations such as dependency-based word vectors. When tested on both English and Arabic datasets, our method outperforms both a strong SVM classifier and state-of-the-art parsers. For instance, we achieve 82.6% PP attachment accuracy on Arabic, while the Turbo and Charniak self-trained parsers obtain 76.7% and 80.8% respectively.

Session 1C: Information Retrieval, Text Categorization, Topic Modeling (Long Papers)

Plaza Ballroom F

Chair: Jordan Boyd-Graber

A Hybrid Generative/Discriminative Approach To Citation Prediction

Chris Tanner and Eugene Charniak

10:40–11:05

Text documents of varying nature (e.g., summary documents written by analysts or published, scientific papers) often cite others as a means of providing evidence to support a claim, attributing credit, or referring the reader to related work. We address the problem of predicting a document’s cited sources by introducing a novel, discriminative approach which combines a content-based generative model (LDA) with author-based features. Further, our classifier is able to learn the importance and quality of each topic within our corpus – which can be useful beyond this task – and preliminary results suggest its metric is competitive with other standard metrics (Topic Coherence). Our flagship system, Logit-Expanded, provides state-of-the-art performance on the largest corpus ever used for this task.

Weakly Supervised Slot Tagging with Partially Labeled Sequences from Web Search Click Logs

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

11:05–11:30

In this paper, we apply a weakly-supervised learning approach for slot tagging using conditional random fields by exploiting web search click logs. We extend the constrained lattice training of Täckström et al. (2013) to non-linear conditional random fields in which latent variables mediate between observations and labels. When combined with a novel initialization scheme that leverages unlabeled data, we show that our method gives significant improvement over strong supervised and weakly-supervised baselines.

Not All Character N-grams Are Created Equal: A Study in Authorship Attribution

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

11:30–11:55

Character n-grams have been identified as the most successful feature in both single-domain and cross-domain Authorship Attribution (AA), but the reasons for their discriminative value were not fully understood. We identify subgroups of character n-grams that correspond to linguistic aspects commonly claimed to be covered by these features: morpho-syntax, thematic content and style. We evaluate the predictiveness of each of these groups in two AA settings: a single domain setting and a cross-domain setting where multiple topics are present. We demonstrate that character \$n\$-grams that capture information about affixes and punctuation account for almost all of the power of character n-grams as features. Our study contributes new insights into the use of n-grams for future AA work and other classification tasks.

Effective Use of Word Order for Text Categorization with Convolutional Neural Networks

Rie Johnson and Tong Zhang

11:55–12:20

Convolutional neural network (CNN) is a neural network that can make use of the internal structure of data such as the 2D structure of image data. This paper studies CNN on text categorization to exploit the 1D structure (namely, word order) of text data for accurate prediction. Instead of using low-dimensional word vectors as input as is often done, we directly apply CNN to high-dimensional text data, which leads to directly learning embedding of small text regions for use in classification. In addition to a straightforward adaptation of CNN from image to text, a simple but new variation which employs bag-of-word conversion in the convolution layer is proposed. An extension to combine multiple convolution layers is also explored for higher accuracy. The experiments demonstrate the effectiveness of our approach in comparison with state-of-the-art methods.

Student Lunch

Monday, June 1, 2015, 12:30–2:00pm

Plaza Ballroom C

Join your fellow students for a free students-only lunch on Monday, June 1 at 12:30 in the Plaza Ballroom C at the hotel. Entrance tickets will be in your conference badges. This is a chance to get to know others who share similar interests and goals and who may become your lifelong colleagues.

Session 2 Overview – Monday, June 1, 2015

	Track A	Track B	Track C
	<i>Generation and Summarization (Long Papers)</i> Plaza Ballroom A & B	<i>Language and Vision (Long Papers)</i> Plaza Ballroom D & E	<i>NLP for Web, Social Media and Social Sciences (Long Papers)</i> Plaza Ballroom F
2:25	Transition-Based Syntactic Linearization <i>Liu, Zhang, Che, and Qin</i>	What's Cookin'? Interpreting Cooking Videos using Text, Speech and Vision <i>Malmaud, Huang, Rathod, Johnston, Rabinovich, and Murphy</i>	TopicCheck: Interactive Alignment for Assessing Topic Model Stability <i>Chuang, Roberts, Stewart, Weiss, Tingley, Grimmer, and Heer</i>
2:50	Extractive Summarisation Based on Keyword Profile and Language Model <i>Xu, Martin, and Mahidadia</i>	Combining Language and Vision with a Multimodal Skip-gram Model <i>Lazaridou, Pham, and Baroni</i>	Inferring latent attributes of Twitter users with label regularization <i>Mohammady Ardehaly and Culotta</i>
3:15	HEADS: Headline Generation as Sequence Prediction Using an Abstract Feature-Rich Space <i>Colmenares, Litvak, Mantrach, and Silvestri</i>	Discriminative Unsupervised Alignment of Natural Language Instructions with Corresponding Video Segments <i>Naim, Song, Liu, Huang, Kautz, Luo, and Gildea</i>	A Neural Network Approach to Context-Sensitive Generation of Conversational Responses <i>Sordoni, Galley, Auli, Brockett, Ji, Mitchell, Nie, Gao, and Dolan</i>

Parallel Session 2

Session 2A: Generation and Summarization (Long Papers)

Plaza Ballroom A & B

Chair: Fei Liu

Transition-Based Syntactic Linearization

Yijia Liu, Yue Zhang, Wanxiang Che, and Bing Qin

14:00–14:25

Syntactic linearization algorithms take a bag of input words and a set of optional constraints, and construct an output sentence and its syntactic derivation simultaneously. The search problem is NP-hard, and the current best results are achieved by bottom-up best-first search. One drawback of the method is low efficiency; and there is no theoretical guarantee that a full sentence can be found within bounded time. We propose an alternative algorithm that constructs output structures from left to right using beam-search. The algorithm is based on incremental parsing algorithms. We extend the transition system so that word ordering is performed in addition to syntactic parsing, resulting in a linearization system that runs in guaranteed quadratic time. In standard evaluations, our system runs an order of magnitude faster than a state-of-the-art baseline using best-first search, with improved accuracies.

Extractive Summarisation Based on Keyword Profile and Language Model

Han Xu, Eric Martin, and Ashesh Mahidadia

14:25–14:50

We present a statistical framework to extract information-rich citation sentences that summarise the main contributions of a scientific paper. In a first stage, we automatically discover salient keywords from a paper's citation summary, keywords that characterise its main contributions. In a second stage, exploiting the results of the first stage, we identify citation sentences that best capture the paper's main contributions. Experimental results show that our approach using methods rooted in quantitative statistics and information theory outperforms the current state-of-the-art systems in scientific paper summarisation.

HEADS: Headline Generation as Sequence Prediction Using an Abstract Feature-Rich Space

Carlos A. Colmenares, Marina Litvak, Amin Mantrach, and Fabrizio Silvestri

14:50–15:15

Automatic headline generation is a sub-task of document summarization with many reported applications. In this study we present a sequence-prediction technique for learning how editors title their news stories. The introduced technique models the problem as a discrete optimization task in a feature-rich space. In this space the global optimum can be found in polynomial time by means of dynamic programming. We train and test our model on an extensive corpus of financial news, and compare it against a number of baselines by using standard metrics from the document summarization domain, as well as some new ones proposed in this work. We also assess the readability and informativeness of the generated titles through human evaluation. The obtained results are very appealing and substantiate the soundness of the approach.

Session 2B: Language and Vision (Long Papers)

Plaza Ballroom D & E

Chair: *Yejin Choi*

What's Cookin'? Interpreting Cooking Videos using Text, Speech and Vision

Jonathan Malmaud, Jonathan Huang, Vivek Rathod, Nicholas Johnston, Andrew Rabinovich, and Kevin Murphy

14:00–14:25

We present a novel method for aligning a sequence of instructions to a video of someone carrying out a task. In particular, we focus on the cooking domain, where the instructions correspond to the recipe. Our technique relies on an HMM to align the recipe steps to the (automatically generated) speech transcript. We then refine this alignment using a state-of-the-art visual food detector, based on a deep convolutional neural network. We show that our technique outperforms simpler techniques based on keyword spotting. It also enables interesting applications, such as automatically illustrating recipes with keyframes, and searching within a video for events of interest.

Combining Language and Vision with a Multimodal Skip-gram Model

Angeliki Lazaridou, Nghia The Pham, and Marco Baroni

14:25–14:50

We extend the SKIP-GRAM model of Mikolov et al. (2013a) by taking visual information into account. Like SKIP-GRAM, our multimodal models (MMSKIP-GRAM) build vector-based word representations by learning to predict linguistic contexts in text corpora. However, for a restricted set of words, the models are also exposed to visual representations of the objects they denote (extracted from natural images), and must predict linguistic and visual features jointly. The MMSKIP-GRAM models achieve good performance on a variety of semantic benchmarks. Moreover, since they propagate visual information to all words, we use them to improve image labeling and retrieval in the zero-shot setup, where the test concepts are never seen during model training. Finally, the MMSKIP-GRAM models discover intriguing visual properties of abstract words, paving the way to realistic implementations of embodied theories of meaning.

Discriminative Unsupervised Alignment of Natural Language Instructions with Corresponding Video Segments

Iftekhar Naim, Young C. Song, Qiguang Liu, Liang Huang, Henry Kautz, Jiebo Luo, and Daniel Gildea

14:50–15:15

We address the problem of automatically aligning natural language sentences with corresponding video segments without any direct supervision. Most existing algorithms for integrating language with videos rely on hand-aligned parallel data, where each natural language sentence is manually aligned with its corresponding image or video segment. Recently, fully unsupervised alignment of text with video has been shown to be feasible using hierarchical generative models. In contrast to the previous generative models, we propose three latent-variable discriminative models for the unsupervised alignment task. The proposed discriminative models are capable of incorporating domain knowledge, by adding diverse and overlapping features. The results show that discriminative models outperform the generative models in terms of alignment accuracy.

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

Plaza Ballroom F

*Chair: Tong Zhang***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* 14:00–14:25

Content analysis, a widely-applied social science research method, is increasingly being supplemented by topic modeling. However, while the discourse on content analysis centers heavily on reproducibility, computer scientists often focus more on scalability and less on coding reliability, leading to growing skepticism on the usefulness of topic models for automated content analysis. In response, we introduce TopicCheck, an interactive tool for assessing topic model stability. Our contributions are threefold. First, from established guidelines on reproducible content analysis, we distill a set of design requirements on how to computationally assess the stability of an automated coding process. Second, we devise an interactive alignment algorithm for matching latent topics from multiple models, and enable sensitivity evaluation across a large number of models. Finally, we demonstrate that our tool enables social scientists to gain novel insights into three active research questions.

Inferring latent attributes of Twitter users with label regularization*Ehsan Mohammady Ardehaly and Aron Culotta* 14:25–14:50

Inferring latent attributes of online users has many applications in public health, politics, and marketing. Most existing approaches rely on supervised learning algorithms, which require manual data annotation and therefore are costly to develop and adapt over time. In this paper, we propose a lightly supervised approach based on label regularization to infer the age, ethnicity, and political orientation of Twitter users. Our approach learns from a heterogeneous collection of soft constraints derived from Census demographics, trends in baby names, and Twitter accounts that are emblematic of class labels. To counteract the imprecision of such constraints, we compare several constraint selection algorithms that optimize classification accuracy on a tuning set. We find that using no user-annotated data, our approach is within 2% of a fully supervised baseline for three of four tasks. Using a small set of labeled data for tuning further improves accuracy on all tasks.

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* 14:50–15:15

We present a novel response generation system that can be trained end to end on large quantities of unstructured Twitter conversations. A neural network architecture is used to address sparsity issues that arise when integrating contextual information into classic statistical models, allowing the system to take into account previous dialog utterances. Our dynamic-context generative models show consistent gains over both context-sensitive and non-context-sensitive Machine Translation and Information Retrieval baselines.

Session 3 Overview – Monday, June 1, 2015

	Track A	Track B	Track C
	<i>Generation and Summarization (Short Papers)</i> Plaza Ballroom A & B	<i>Information Extraction and Question Answering (Short Papers)</i> Plaza Ballroom D & E	<i>Machine Learning for NLP (Short Papers)</i> Plaza Ballroom F
3:30	How to Make a Frenemy: Multitape FSTs for Portman-teau Generation <i>Deri and Knight</i>	Entity Linking for Spoken Language <i>Benton and Dredze</i>	When and why are log-linear models self-normalizing? <i>Andreas and Klein</i>
3:45	Aligning Sentences from Standard Wikipedia to Simple Wikipedia <i>Hwang, Hajishirzi, Ostdorf, and Wu</i>	Spinning Straw into Gold: Using Free Text to Train Monolingual Alignment Models for Non-factoid Question Answering <i>Sharp, Jansen, Surdeanu, and Clark</i>	Deep Multilingual Correlation for Improved Word Embeddings <i>Lu, Wang, Bansal, Gimpel, and Livescu</i>
4:00	Inducing Lexical Style Properties for Paraphrase and Genre Differentiation <i>Pavlick and Nenkova</i>	Personalized Page Rank for Named Entity Disambiguation <i>Pershina, He, and Grishman</i>	Disfluency Detection with a Semi-Markov Model and Prosodic Features <i>Ferguson, Durrett, and Klein</i>

Parallel Session 3

Session 3A: Generation and Summarization (Short Papers)

Plaza Ballroom A & B

Chair: Fei Liu

How to Make a Frenemy: Multitape FSTs for Portmanteau Generation

Aliya Deri and Kevin Knight

15:15–15:30

A portmanteau is a type of compound word that fuses the sounds and meanings of two component words; for example, “frenemy” (friend + enemy) or “smog” (smoke + fog). We develop a system, including a novel multitape FST, that takes an input of two words and outputs possible portmanteaux. Our system is trained on a list of known portmanteaux and their component words, and achieves 45% exact matches in cross-validated experiments.

Aligning Sentences from Standard Wikipedia to Simple Wikipedia

William Hwang, Hannaneh Hajishirzi, Mari Ostendorf, and Wei Wu

15:30–15:45

This work improves monolingual sentence alignment for text simplification, specifically for text in standard and simple Wikipedia. We introduce a method that improves over past efforts by using a greedy (vs. ordered) search over the document and a word-level semantic similarity score based on Wiktionary (vs. WordNet) that also accounts for structural similarity through syntactic dependencies. Experiments show improved performance on a hand-aligned set, with the largest gain coming from structural similarity. Resulting datasets of manually and automatically aligned sentence pairs are made available.

Inducing Lexical Style Properties for Paraphrase and Genre Differentiation

Ellie Pavlick and Ani Nenkova

15:45–16:00

We present an intuitive and effective method for inducing style scores on words and phrases. We exploit signal in a phrase’s rate of occurrence across stylistically contrasting corpora, making our method simple to implement and efficient to scale. We show strong results both intrinsically, by correlation with human judgements, and extrinsically, in applications to genre analysis and paraphrasing.

Session 3B: Information Extraction and Question Answering (Short Papers)

Plaza Ballroom D & E

Chair: Yejin Choi

Entity Linking for Spoken Language

Adrian Benton and Mark Dredze

15:15–15:30

Research on entity linking has considered a broad range of text, including newswire, blogs and web documents in multiple languages. However, the problem of entity linking for spoken language remains unexplored. Spoken language obtained from automatic speech recognition systems poses different types of challenges for entity linking; transcription errors can distort the context, and named entities tend to have high error rates. We propose features to mitigate these errors and evaluate the impact of ASR errors on entity linking using a new corpus of entity linked broadcast news transcripts.

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

15:30–15:45

Monolingual alignment models have been shown to boost the performance of question answering systems by “bridging the lexical chasm” between questions and answers. The main limitation of these approaches is that they require semistructured training data in the form of question-answer pairs, which is difficult to obtain in specialized domains or low-resource languages. We propose two inexpensive methods for training alignment models solely using free text, by generating artificial question-answer pairs from discourse structures. Our approach is driven by two representations of discourse: a shallow sequential representation, and a deep one based on Rhetorical Structure Theory. We evaluate the proposed model on two corpora from different genres and domains: one from Yahoo! Answers and one from the biology domain, and two types of non-factoid questions: manner and reason. We show that these alignment models trained directly from discourse structures imposed on free text improve performance considerably over an information retrieval baseline and a neural network language model trained on the same data.

Personalized Page Rank for Named Entity Disambiguation

Maria Pershina, Yifan He, and Ralph Grishman

15:45–16:00

The task of Named Entity Disambiguation is to map entity mentions in the document to their correct entries in some knowledge base. We present a novel graph-based disambiguation approach based on Personalized PageRank (PPR) that combines local and global evidence for disambiguation and effectively filters out noise introduced by incorrect candidates. Experiments show that our method outperforms state-of-the-art approaches by achieving 91.7% in micro- and 89.9% in macroaccuracy on a dataset of 27.8K named entity mentions.

Session 3C: Machine Learning for NLP (Short Papers)

Plaza Ballroom F

Chair: Tong Zhang

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

15:15–15:30

Several techniques have recently been proposed for training “self-normalized” discriminative models. These attempt to find parameter settings for which unnormalized model scores approximate the true label probability. However, the theoretical properties of such techniques (and of self-normalization generally) have not been investigated. This paper examines the conditions under which we can expect self-normalization to work. We characterize a general class of distributions that admit self-normalization, and prove generalization bounds for procedures that minimize empirical normalizer variance. Motivated by these results, we describe a novel variant of an established procedure for training self-normalized models. The new procedure avoids computing normalizers for most training examples, and decreases training time by as much as factor of ten while preserving model quality.

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

15:30–15:45

Word embeddings have been found useful for many NLP tasks, including part-of-speech tagging, named entity recognition, and parsing. Adding multilingual context when learning embeddings can improve their quality, for example via canonical correlation analysis (CCA) on embeddings from two languages. In this paper, we extend this idea to learn deep non-linear transformations of word embeddings of the two languages, using the recently proposed deep canonical correlation analysis. The resulting embeddings, when evaluated on multiple word and bigram similarity tasks, consistently improve over monolingual embeddings and over embeddings transformed with linear CCA.

Disfluency Detection with a Semi-Markov Model and Prosodic Features*James Ferguson, Greg Durrett, and Dan Klein*

15:45–16:00

We present a discriminative model for detecting disfluencies in spoken language transcripts. Structurally, our model is a semi-Markov conditional random field with features targeting characteristics unique to speech repairs. This gives a significant performance improvement over standard chain-structured CRFs that have been employed in past work. We then incorporate prosodic features over silences and relative word duration into our semi-CRF model, resulting in further performance gains; moreover, these features are not easily replaced by discrete prosodic indicators such as ToBI breaks. Our final system, the semi-CRF with prosodic information, achieves an F-score of 85.4, which is 1.3 F1 better than the best prior reported F-score on this dataset.

One Minute Madness!

Prior to the poster session, TACL and long-paper poster presenters will be given one minute each to pitch their paper. The poster session will immediately follow these presentations along with a buffet dinner.

Chair: Joel Tetreault

Poster session 1A

Time: 6:00–7:30

Location: Plaza Ballroom A, B, & C

Robust Morphological Tagging with Word Representations

Thomas Müller and Hinrich Schütze

We present a comparative investigation of word representations for part-of-speech and morphological tagging, focusing on scenarios with considerable differences between training and test data where a robust approach is necessary. Instead of adapting the model towards a specific domain we aim to build a robust model across domains. We developed a test suite for robust tagging consisting of six languages and different domains. We find that representations similar to Brown clusters perform best for POS tagging and that word representations based on linguistic morphological analyzers perform best for morphological tagging.

Multiview LSA: Representation Learning via Generalized CCA

Pushpendre Rastogi, Benjamin Van Durme, and Raman Arora

Multiview LSA (MVLSA) is a generalization of Latent Semantic Analysis (LSA) that supports the fusion of arbitrary views of data and relies on Generalized Canonical Correlation Analysis (GCCA). We present an algorithm for fast approximate computation of GCCA, which when coupled with methods for handling missing values, is general enough to approximate some recent algorithms for inducing vector representations of words. Experiments across a comprehensive collection of test-sets show our approach to be competitive with the state of the art.

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

Casey Kennington, Ryu Iida, Takenobu Tokunaga, and David Schlangen

A large part of human communication involves referring to entities in the world and often these entities are objects that are visually present for the interlocutors. A system that aims to resolve such references needs to tackle a complex task: objects and their visual features need to be determined, the referring expressions must be recognised, and extra-linguistic information such as eye gaze or pointing gestures need to be incorporated. Systems that can make use of such information sources exist, but have so far only been tested under very constrained settings, such as WOZ interactions. In this paper, we apply to a more complex domain a reference resolution model that works incrementally (i.e., word by word), grounds words with visually present properties of objects (such as shape and size), and can incorporate extra-linguistic information. We find that the model works well compared to previous work on the same data, despite using fewer features. We conclude that the model shows potential for use in a real-time interactive dialogue system.

Digital Leafleting: Extracting Structured Data from Multimedia Online Flyers

Emilia Apostolova, Payam Pourashraf, and Jeffrey Sack

Marketing materials such as flyers and other infographics are a vast online resource. In a number of industries, such as the commercial real estate industry, they are in fact the only authoritative source of information. Companies attempting to organize commercial real estate inventories spend a significant amount of resources on manual data entry of this information. In this work, we propose a method for extracting structured data from free-form commercial real estate flyers in PDF and HTML formats. We modeled the problem as text categorization and Named Entity Recognition (NER) tasks and applied a supervised machine learning approach (Support Vector Machines). Our dataset consists of more than 2,200 commercial real estate flyers and associated manually entered structured data, which was used to automatically create training datasets. Traditionally, text categorization and NER approaches are based on textual information only. However, information in visu-

ally rich formats such as PDF and HTML is often conveyed by a combination of textual and visual features. Large fonts, visually salient colors, and positioning often indicate the most relevant pieces of information. We applied novel features based on visual characteristics in addition to traditional text features and show that performance improved significantly for both the text categorization and NER tasks.

Towards a standard evaluation method for grammatical error detection and correction

Mariano Felice and Ted Briscoe

We present a novel evaluation method for grammatical error correction that addresses problems with previous approaches and scores systems in terms of improvement on the original text. Our method evaluates corrections at the token level using a globally optimal alignment between the source, a system hypothesis and a reference. Unlike the M² Scorer, our method provides scores for both detection and correction and is sensitive to different types of edit operations.

Constraint-Based Models of Lexical Borrowing

Yulia Tsvetkov, Waleed Ammar, and Chris Dyer

Linguistic borrowing is the phenomenon of transferring linguistic constructions (lexical, phonological, morphological, and syntactic) from a “donor” language to a “recipient” language as a result of contacts between communities speaking different languages. Borrowed words are found in all languages, and—in contrast to cognate relationships—borrowing relationships may exist across unrelated languages (for example, about 40% of Swahili’s vocabulary is borrowed from Arabic). In this paper, we develop a model of morpho-phonological transformations across languages with features based on universal constraints from Optimality Theory (OT). Compared to several standard—but linguistically naïve—baselines, our OT-inspired model obtains good performance with only a few dozen training examples, making this a cost-effective strategy for sharing lexical information across languages.

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

A key challenge of designing coherent semantic ontology for spoken language understanding is to consider inter-slot relations. In practice, however, it is difficult for domain experts and professional annotators to define a coherent slot set, while considering various lexical, syntactic, and semantic dependencies. In this paper, we exploit the typed syntactic dependency theory for unsupervised induction and filling of semantics slots in spoken dialogue systems. More specifically, we build two knowledge graphs: a slot-based semantic graph, and a word-based lexical graph. To jointly consider word-to-word, word-to-slot, and slot-to-slot relations, we use a random walk inference algorithm to combine the two knowledge graphs, guided by dependency grammars. The experiments show that considering inter-slot relations is crucial for generating a more coherent and compete slot set, resulting in a better spoken language understanding model, while enhancing the interpretability of semantic slots.

Interpreting Compound Noun Phrases Using Web Search Queries

Marius Pasca

A weakly-supervised method is applied to anonymized queries to extract lexical interpretations of compound noun phrases (e.g., “fortune 500 companies”). The interpretations explain the subsuming role (“listed in”) that modifiers (“fortune 500”) play relative to heads (“companies”) within the noun phrases. Experimental results over evaluation sets of noun phrases from multiple sources demonstrate that interpretations extracted from queries have encouraging coverage and precision. The top interpretation extracted is deemed relevant for more than 70% of the noun phrases.

Expanding Paraphrase Lexicons by Exploiting Lexical Variants

Atsushi Fujita and Pierre Isabelle

This study tackles the problem of paraphrase acquisition: achieving high coverage as well as accuracy. Our method first induces paraphrase patterns from given seed paraphrases, exploiting the generality of paraphrases exhibited by pairs of lexical variants, e.g., “amendment” and “amending,” in a fully empirical way. It then searches monolingual corpora for new paraphrases that match the patterns. This can extract paraphrases comprising words that are completely different from those of the given seeds. In experiments, our method expanded seed sets by factors of 42 to 206, gaining 84% to 208% more coverage than a previous method that generalizes only identical word forms. Human evaluation through a paraphrase substitution test demonstrated that the newly acquired paraphrases retained reasonable quality, given substantially high-quality seeds.

Lexicon-Free Conversational Speech Recognition with Neural Networks

Andrew Maas, Ziang Xie, Dan Jurafsky, and Andrew Ng

We present an approach to speech recognition that uses only a neural network to map acoustic input to characters, a character-level language model, and a beam search decoding procedure. This approach eliminates much of the complex infrastructure of modern speech recognition systems, making it possible to directly train a speech recognizer using errors generated by spoken language understanding tasks. The system naturally handles out of vocabulary words and spoken word fragments. We demonstrate our approach using the challenging Switchboard telephone conversation transcription task, achieving a word error rate competitive with existing baseline systems. To our knowledge, this is the first entirely neural-network-based system to achieve strong speech transcription results on a conversational speech task. We analyze qualitative differences between transcriptions produced by our lexicon-free approach and transcriptions produced by a standard speech recognition system. Finally, we evaluate the impact of large context neural network character language models as compared to standard n-gram models within our framework.

A Linear-Time Transition System for Crossing Interval Trees

Emily Pitler and Ryan McDonald

We define a restricted class of non-projective trees that 1) covers many natural language sentences; and 2) can be parsed exactly with a generalization of the popular arc-eager system for projective trees (Nivre, 2003). Crucially, this generalization only adds constant overhead in run-time and space keeping the parser's total run-time linear in the worst case. In empirical experiments, our proposed transition-based parser is more accurate on average than both the arc-eager system or the swap-based system, an unconstrained non-projective transition system with a worst-case quadratic runtime (Nivre, 2009).

Data-driven sentence generation with non-isomorphic trees

Miguel Ballesteros, Bernd Bohnet, Simon Mille, and Leo Wanner

Abstract structures from which the generation naturally starts often do not contain any functional nodes, while surface-syntactic structures or a chain of tokens in a linearized tree contain all of them. Therefore, data-driven linguistic generation needs to be able to cope with the projection between non-isomorphic structures that differ in their topology and number of nodes. So far, such a projection has been a challenge in data-driven generation and was largely avoided. We present a fully stochastic generator that is able to cope with projection between non-isomorphic structures. The generator, which starts from PropBank-like structures, consists of a cascade of SVM-classifier based submodules that map in a series of transitions the input structures onto sentences. The generator has been evaluated for English on the Penn-Treebank and for Spanish on the multi-layered Ancora-UPF corpus.

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

Sujay Kumar Jauhar, Chris Dyer, and Eduard Hovy

Words are polysemous. However, most approaches to representation learning for lexical semantics assign a single vector to every surface word type. Meanwhile, lexical ontologies such as WordNet provide a source of complementary knowledge to distributional information, including a word sense inventory. In this paper we propose two novel and general approaches for generating sense-specific word embeddings that are grounded in an ontology. The first applies graph smoothing as a post-processing step to tease the vectors of different senses apart, and is applicable to any vector space model. The second adapts predictive maximum likelihood models that learn word embeddings with latent variables representing senses grounded in an specified ontology. Empirical results on lexical semantic tasks show that our approaches effectively captures information from both the ontology and distributional statistics. Moreover, in most cases our sense-specific models outperform other models we compare against.

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

Michael Haas and Yannick Versley

Sentiment analysis has undergone a shift from document-level analysis, where labels express the sentiment of a whole document or whole sentence, to subsentential approaches, which assess the contribution of individual phrases, in particular including the composition of sentiment terms and phrases such as negators and intensifiers. Starting from a small sentiment treebank modeled after the Stanford Sentiment Treebank of Socher et al. (2013), we investigate suitable methods to perform compositional sentiment classification for German in a data-scarce setting, harnessing cross-lingual methods as well as existing general-domain lexical resources.

Using Summarization to Discover Argument Facets in Online Ideological Dialog*Amita Misra, Pranav Anand, Jean E. Fox Tree, and Marilyn Walker*

More and more of the information available on the web is dialogic, and a significant portion of it takes place in online forum conversations about current social and political topics. We aim to develop tools to summarize what these conversations are about. What are the CENTRAL PROPOSITIONS associated with different stances on an issue; what are the abstract objects under discussion that are central to a speaker's argument? How can we recognize that two CENTRAL PROPOSITIONS realize the same FACET of the argument? We hypothesize that the CENTRAL PROPOSITIONS are exactly those arguments that people find most salient, and use human summarization as a probe for discovering them. We describe our corpus of human summaries of opinionated dialogs, then show how we can identify similar repeated arguments, and group them into FACETS across many discussions of a topic. We define a new task, ARGUMENT FACET SIMILARITY (AFS), and show that we can predict AFS with a .54 correlation score, versus an ngram system baseline of .39 and a semantic textual similarity system baseline of .45.

Incorporating Word Correlation Knowledge into Topic Modeling*Pengtao Xie, Diyi Yang, and Eric Xing*

This paper studies how to incorporate the external word correlation knowledge to improve the coherence of topic modeling. Existing topic models assume words are generated independently and lack the mechanism to utilize the rich similarity relationships among words to learn coherent topics. To solve this problem, we build a Markov Random Field (MRF) regularized Latent Dirichlet Allocation (LDA) model, which define a MRF on the latent topic layer of LDA to encourage words labeled as similar to share the same topic label. Under our model, the topic assignment of each word is not independent, but rather affected by the topic labels of its correlated words. Similar words have better chance to be put into the same topic due to the regularization of MRF, hence the coherence of topics can be boosted. In addition, our model can accommodate the subtlety that whether two words are similar depends on which topic they appear in, which allows word with multiple senses to be put into different topics properly. We derive a variational inference method to infer the posterior probabilities and learn model parameters and present techniques to deal with the hard-to-compute partition function in MRF. Experiments on two datasets demonstrate the effectiveness of our model.

Active Learning with Rationales for Text Classification*Manali Sharma, Di Zhuang, and Mustafa Bilgic*

We present a simple and yet effective approach that can incorporate rationales elicited from annotators into the training of any off-the-shelf classifier. We show that our simple approach is effective for multinomial naive Bayes, logistic regression, and support vector machines. We additionally present an active learning method tailored specifically for the learning with rationales framework.

The Unreasonable Effectiveness of Word Representations for Twitter Named Entity Recognition*Colin Cherry and Hongyu Guo*

Named entity recognition (NER) systems trained on newswire perform very badly when tested on Twitter. Signals that were reliable in copy-edited text disappear almost entirely in Twitter's informal chatter, requiring the construction of specialized models. Using well-understood techniques, we set out to improve Twitter NER performance when given a small set of annotated training tweets. To leverage unlabeled tweets, we build Brown clusters and word vectors, enabling generalizations across distributionally similar words. To leverage annotated newswire data, we employ an importance weighting scheme. Taken all together, we establish a new state-of-the-art on two common test sets. Though it is well-known that word representations are useful for NER, supporting experiments have thus far focused on newswire data. We emphasize the effectiveness of representations on Twitter NER, and demonstrate that their inclusion can improve performance by up to 20 F1.

Inferred Temporally-Anchored Spatial Knowledge from Semantic Roles*Eduardo Blanco and Alakananda Vempala*

This paper presents a framework to infer spatial knowledge from verbal semantic role representations. First, we generate potential spatial knowledge deterministically. Second, we determine whether it can be inferred and a degree of certainty. Inferences capture that something is located or is not located somewhere, and temporally anchor this information. An annotation effort shows that inferences are ubiquitous and intuitive to humans.

Is Your Anchor Going Up or Down? Fast and Accurate Supervised Topic Models*Thang Nguyen, Jordan Boyd-Graber, Jeffrey Lund, Kevin Sepi, and Eric Ringger*

Topic models provide insights into document collections, and their supervised extensions also capture associated document level metadata such as sentiment. However, inferring such models from data is often slow and cannot scale to big data. We build upon the “anchor” method for learning topic models to capture the relationship between metadata and latent topics by extending the vector space representation of word cooccurrence to include metadata specific dimensions. These additional dimensions reveal new anchor words that reflect specific combinations of metadata and topic. We show that these new latent representations predict sentiment as accurately as supervised topic models, and we find these representations more quickly without sacrificing interpretability.

Grounded Semantic Parsing for Complex Knowledge Extraction

Ankur P. Parikh, Hoifung Poon, and Kristina Toutanova

Recently, there has been increasing interest in learning semantic parsers with indirect supervision, but existing work focuses almost exclusively on question answering. Separately, there have been active pursuits in leveraging databases for distant supervision in information extraction, yet such methods are often limited to binary relations and none can handle nested events. In this paper, we generalize distant supervision to complex knowledge extraction, by proposing the first approach to learn a semantic parser for extracting nested event structures without annotated examples, using only a database of such complex events and unannotated text. The key idea is to model the annotations as latent variables, and incorporate a prior that favors semantic parses containing known events. Experiments on the GENIA event extraction dataset show that our approach can learn from and extract complex biological pathway events. Moreover, when supplied with just five example words per event type, it becomes competitive even among supervised systems, outperforming 19 out of 24 teams that participated in the original shared task.

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

Chen Li, Yang Liu, and Lin Zhao

Some state-of-the-art summarization systems use integer linear programming (ILP) based methods that aim to maximize the important concepts covered in the summary. These concepts are often obtained by selecting bigrams from the documents. In this paper, we improve such bigram based ILP summarization methods from different aspects. First we use syntactic information to select more important bigrams. Second, to estimate the importance of the bigrams, in addition to the internal features based on the test documents (e.g., document frequency, bigram positions), we propose to extract features by leveraging multiple external resources (such as word embedding from additional corpus, Wikipedia, Dbpedia, Word-Net, SentiWordNet). The bigram weights are then trained discriminatively in a joint learning model that predicts the bigram weights and selects the summary sentences in the ILP framework at the same time. We demonstrate that our system consistently outperforms the prior ILP method on different TAC data sets, and performs competitively compared to other previously reported best results. We also conducted various analysis to show the contribution of different components.

Transforming Dependencies into Phrase Structures

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

We present a new algorithm for transforming dependency parse trees into phrase-structure parse trees. We cast the problem as structured prediction and learn a statistical model. Our algorithm is faster than traditional phrase-structure parsing and achieves 90.4% English parsing accuracy and 82.4% Chinese parsing accuracy, near to the state of the art on both benchmarks.

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

Jianfu Chen, Polina Kuznetsova, David Warren, and Yejin Choi

We present a new approach to harvesting a large-scale, high quality image-caption corpus that makes a better use of already existing web data with no additional human efforts. The key idea is to focus on Déjà Image-Captions: naturally existing image descriptions that are repeated almost verbatim — by more than one individual for different images. The resulting corpus provides association structure between 4 million images with 180K unique captions, capturing a rich spectrum of everyday narratives including figurative and pragmatic language. Exploring the use of the new corpus, we also present new conceptual tasks of visually situated paraphrasing, creative image captioning, and creative visual paraphrasing.

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

Attapol Rutherford and Nianwen Xue

Discourse relation classification is an important component for automatic discourse parsing and natural language understanding. The performance bottleneck of a discourse parser comes from implicit discourse relations, whose discourse connectives are not overtly present. Explicit discourse connectives can potentially be exploited to collect more training data to collect more data and boost the performance. However, using them indiscriminately has been shown to hurt the performance because not all discourse connectives can be dropped arbitrarily. Based on this insight, we investigate the interaction between discourse connectives and the discourse relations and propose the criteria for selecting the discourse connectives that can be dropped independently of the context without changing the interpretation of the discourse. Extra training data collected only by the freely ommissible connectives improve the performance of the system without additional features.

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

Arvind Neelakantan and Ming-Wei Chang

Most of previous work in knowledge base (KB) completion has focused on the problem of relation extraction. In this work, we focus on the task of inferring missing entity type instances in a KB, a fundamental task for KB competition yet receives little attention. Due to the novelty of this task, we construct a large-scale dataset and design an automatic evaluation methodology. Our knowledge base completion method uses information within the existing KB and external information from Wikipedia. We show that individual methods trained with a global objective that considers unobserved cells from both the entity and the type side gives consistently higher quality predictions compared to baseline methods. We also perform manual evaluation on a small subset of the data to verify the effectiveness of our knowledge base completion methods and the correctness of our proposed automatic evaluation method.

Pragmatic Neural Language Modelling in Machine Translation

Paul Baltescu and Phil Blunsom

This paper presents an in-depth investigation on integrating neural language models in translation systems. Scaling neural language models is a difficult task, but crucial for real-world applications. This paper evaluates the impact on end-to-end MT quality of both new and existing scaling techniques. We show when explicitly normalising neural models is necessary and what optimisation tricks one should use in such scenarios. We also focus on scalable training algorithms and investigate noise contrastive estimation and diagonal contexts as sources for further speed improvements. We explore the trade-offs between neural models and back-off n-gram models and find that neural models make strong candidates for natural language applications in memory constrained environments, yet still lag behind traditional models in raw translation quality. We conclude with a set of recommendations one should follow to build a scalable neural language model for MT.

English orthography is not “close to optimal”

Garrett Nicolai and Grzegorz Kondrak

In spite of the apparent irregularity of the English spelling system, Chomsky and Halle (1968) characterize it as “near optimal”. We investigate this assertion using computational techniques and resources. We design an algorithm to generate word spellings that maximize both phonemic transparency and morphological consistency. Experimental results demonstrate that the constructed system is much closer to optimality than the traditional English orthography.

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

The Bechdel test is a sequence of three questions designed to assess the presence of women in movies. Many believe that because women are seldom represented in film as strong leaders and thinkers, viewers associate weaker stereotypes with women. In this paper, we present a computational approach to automate the task of finding whether a movie passes or fails the Bechdel test. This allows us to study the key differences in language use and in the importance of roles of women in movies that pass the test versus the movies that fail the test. Our experiments confirm that in movies that fail the test, women are in fact portrayed as less-central and less-important characters.

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

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

The past 10 years of event ordering research has focused on learning partial orderings over document events and time expressions. The most popular corpus, the TimeBank, contains a small subset of the possible ordering graph. Many evaluations follow suit by only testing certain pairs of events (e.g., only main verbs of neighboring sentences). This has led most research to focus on specific learners for partial labelings. This paper attempts to nudge the discussion from identifying some relations to all relations. We present new experiments on strongly connected event graphs that contain 10 times more relations per document than the TimeBank. We also describe a shift away from the single learner to a sieve-based architecture that naturally blends multiple learners into a precision-ranked cascade of sieves. Each sieve adds labels to the event graph one at a time, and earlier sieves inform later ones through transitive closure. This paper thus describes innovations in both approach and task. We experiment on the densest event graphs to date and show a 14% gain over state-of-the-art.

[TACL] Unsupervised Discovery of Biographical Structure from Text

David Bamman and Noah A. Smith

We present a method for discovering abstract event classes in biographies, based on a probabilistic latent-variable model. Taking as input timestamped text, we exploit latent correlations among events to learn a set of event classes (such as "Born", "Graduates high school", and "Becomes citizen"), along with the typical times in a person's life when those events occur. In a quantitative evaluation at the task of predicting a person's age for a given event, we find that our generative model outperforms a strong linear regression baseline, along with simpler variants of the model that ablate some features. The abstract event classes that we learn allow us to perform a large-scale analysis of 242,970 Wikipedia biographies. Though it is known that women are greatly underrepresented on Wikipedia – not only as editors (Wikipedia, 2011) but also as subjects of articles (Reagle and Rhue, 2011) – we find that there is a bias in their characterization as well, with biographies of women containing significantly more emphasis on events of marriage and divorce than biographies of men.

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

Jonathan H. Clark, Chris Dyer, and Alon Lavie

Linear models, which support efficient learning and inference, are the workhorses of statistical machine translation; however, linear decision rules are less attractive from a modeling perspective. In this work, we introduce a technique for learning arbitrary, rule-local, nonlinear feature transforms that improve model expressivity, but do not sacrifice the efficient inference and learning associated with linear models. To demonstrate the value of our technique, we discard the customary log transform of lexical probabilities and drop the phrasal translation probability in favor of raw counts. We observe that our algorithm learns a variation of a log transform that leads to better translation quality compared to the explicit log transform. We conclude that non-linear responses play an important role in SMT, an observation that we hope will inform the efforts of feature engineers.

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

Xian Qian and Yang Liu

We show that the decoding problem in generalized Higher Order Conditional Random Fields (CRFs) can be decomposed into two parts: one is a tree labeling problem that can be solved in linear time using dynamic programming; the other is a supermodular quadratic pseudo-Boolean maximization problem, which can be solved in cubic time using a minimum cut algorithm. We use dual decomposition to force their agreement. Experimental results on Twitter named entity recognition and sentence dependency tagging tasks show that our method outperforms spanning tree based dual decomposition.

[TACL] SPRITE: Generalizing Topic Models with Structured Priors

Michael J. Paul and Mark Dredze

We introduce SPRITE, a family of topic models that incorporates structure into model priors as a function of underlying components. The structured priors can be constrained to model topic hierarchies, factorizations, correlations, and supervision, allowing SPRITE to be tailored to particular settings. We demonstrate this flexibility by constructing a SPRITE-based model to jointly infer topic hierarchies and author perspective, which we apply to corpora of political debates and online reviews. We show that the model learns intuitive topics, outperforming several other topic models at predictive tasks.

[TACL] Learning Strictly Local Subsequential Functions

Jane Chandee, Remi Eyraud, and Jeffrey Heinz

We define two proper subclasses of subsequential functions based on the concept of Strict Locality (McNaughton and Papert, 1971; Rogers and Pullum, 2011; Rogers et al., 2013) for formal languages. They are called Input and Output Strictly Local (ISL and OSL). We provide an automata-theoretic characterization of the ISL class and theorems establishing how the classes are related to each other and to Strictly Local languages. We give evidence that local phonological and morphological processes belong to these classes. Finally we provide a learning algorithm which provably identifies the class of ISL functions in the limit from positive data in polynomial time and data. We demonstrate this learning result on appropriately synthesized artificial corpora. We leave a similar learning result for OSL functions for future work and suggest future directions for addressing non-local phonological processes.

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

Word sense induction (WSI) seeks to automatically discover the senses of a word in a corpus via unsupervised methods. We propose a sense-topic model for WSI, which treats sense and topic as two separate latent variables to be inferred jointly. Topics are informed by the entire document, while senses are informed by the local context surrounding the ambiguous word. We also discuss unsupervised ways of enriching the original corpus in order to improve model performance, including using neural word embeddings and external corpora to expand the context of each data instance. We demonstrate significant improvements over the previous state-of-the-art, achieving the best results reported to date on the SemEval-2013 WSI task.

Poster session 1B

Time: 7:30–9:00**Location:** Plaza Ballroom A, B, & C**Modeling Word Meaning in Context with Substitute Vectors***Oren Melamud, Ido Dagan, and Jacob Goldberger*

Context representations are a key element in distributional models of word meaning. In contrast to typical representations based on neighboring words, a recently proposed approach suggests to represent a context of a target word by a substitute vector, comprising the potential fillers for the target word slot in that context. In this work we first propose a variant of substitute vectors, which we find particularly suitable for measuring context similarity. Then, we propose a novel model for representing word meaning in context based on this context representation. Our model outperforms state-of-the-art results on lexical substitution tasks in an unsupervised setting.

LCCT: A Semi-supervised Model for Sentiment Classification*Min Yang, Wenting Tu, Ziyu Lu, Wenpeng Yin, and Kam-Pui Chow*

Analyzing public opinions towards products, services and social events is an important but challenging task. An accurate sentiment analyzer should take both lexicon-level information and corpus-level information into account. It also needs to exploit the domain-specific knowledge and utilize the common knowledge shared across domains. In addition, we want the algorithm being able to deal with missing labels and learning from incomplete sentiment lexicons. This paper presents a LCCT (Lexicon-based and Corpus-based, Co-Training) model for semi-supervised sentiment classification. The proposed method combines the idea of lexicon-based learning and corpus-based learning in a unified co-training framework. It is capable of incorporating both domain-specific and domain-independent knowledge. Extensive experiments show that it achieves very competitive classification accuracy, even with a small portion of labeled data. Comparing to state-of-the-art sentiment classification methods, the LCCT approach exhibits significantly better performances on a variety of datasets in both English and Chinese.

Empty Category Detection With Joint Context-Label Embeddings*Xun Wang, Katsuhiro Sudoh, and Masaaki Nagata*

This paper presents a novel technique for empty category (EC) detection using distributed word representations. We formulate it as an annotation task. We explore the hidden layer of a neural network by mapping both the distributed representations of the contexts of ECs and EC types to a low dimensional space. In the testing phase, using the model learned from annotated data, we project the context of a possible EC position to the same space and further compare it with the representations of EC types. The closest EC type is assigned

to the candidate position. Experiments on Chinese Treebank prove the effectiveness of the proposed method. We improve the precision by about 6 point on a subset of Chinese Treebank, which is a new state-of-the-art performance on CTB.

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

José Camacho-Collados, Mohammad Taher Pilehvar, and Roberto Navigli

The semantic representation of individual word senses and concepts is of fundamental importance to several applications in Natural Language Processing. To date, concept modeling techniques have in the main based their representation either on lexicographic resources, such as WordNet, or on encyclopedic resources, such as Wikipedia. We propose a vector representation technique that combines the complementary knowledge of both these types of resource. Thanks to its use of explicit semantics combined with a novel cluster-based dimensionality reduction and an effective weighting scheme, our representation attains state-of-the-art performance on multiple datasets in two standard benchmarks: word similarity and sense clustering. We are releasing our vector representations at <http://lcl.uniroma1.it/nasari/>.

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

Graham Neubig, Philip Arthur, and Kevin Duh

We propose a method for simultaneously translating from a single source language to multiple target languages T1, T2, etc. The motivation behind this method is that if we only have a weak language model for T1 and translations in T1 and T2 are associated, we can use the information from a strong language model over T2 to disambiguate the translations in T1, providing better translation results. As a specific framework to realize multi-target translation, we expand the formalism of synchronous context-free grammars to handle multiple targets, and describe methods for rule extraction, scoring, pruning, and search with these models. Experiments find that multi-target translation with a strong language model in a similar second target language can provide gains of up to 0.8-1.5 BLEU points.

Using Zero-Resource Spoken Term Discovery for Ranked Retrieval

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

Research on ranked retrieval of spoken content has assumed the existence of some automated (word or phonetic) transcription. Recently, however, methods have been demonstrated for matching spoken terms to spoken content without the need for language-tuned transcription. This paper describes the first application of such techniques to ranked retrieval, evaluated using a newly created test collection. Both the queries and the collection to be searched are based on Gujarati produced naturally by native speakers; relevance assessment was performed by other native speakers of Gujarati. Ranked retrieval is based on fast acoustic matching that identifies a deeply nested set of matching speech regions, coupled with ways of combining evidence from those matching regions. Results indicate that the resulting ranked lists may be useful for some practical similarity-based ranking tasks.

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

Mark Johnson, Joe Pater, Robert Staubs, and Emmanuel Dupoux

This paper describes a joint model of word segmentation and phonological alternations, which takes unsegmented utterances as input and infers word segmentations and underlying phonological representations. The model is a Maximum Entropy or log-linear model, which can express a probabilistic version of Optimality Theory (OT; Prince 2004), a standard phonological framework. The features in our model are inspired by OT's Markedness and Faithfulness constraints. Following the OT principle that such features indicate "violations", we require their weights to be non-positive. We apply our model to a modified version of the Buckeye corpus (Pitt 2007) in which the only phonological alternations are deletions of word-final /d/ and /t/ segments. The model sets a new state-of-the-art for this corpus for word segmentation, identification of underlying forms, and identification of /d/ and /t/ deletions. We also show that the OT-inspired sign constraints on feature weights are crucial for accurate identification of deleted /d/s; without them our model posits approximately 10 times more deleted underlying /d/s than appear in the manually annotated data.

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

Kaveh Taghipour and Hwee Tou Ng

One of the weaknesses of current supervised word sense disambiguation (WSD) systems is that they only treat a word as a discrete entity. However, a continuous-space representation of words (word embeddings) can provide valuable information and thus improve generalization accuracy. Since word embeddings are typically obtained from unlabeled data using unsupervised methods, this method can be seen as a semi-supervised word sense disambiguation approach. This paper investigates two ways of incorporating word embeddings in a word sense disambiguation setting and evaluates these two methods on some SensEval/SemEval lexical sample and all-words tasks and also a domain-specific lexical sample task. The obtained results show that such representations consistently improve the accuracy of the selected supervised WSD system. Moreover, our experiments on a domain-specific dataset show that our supervised baseline system beats the best knowledge-based systems by a large margin.

Model Invertibility Regularization: Sequence Alignment With or Without Parallel Data

Tomer Levinboim, Ashish Vaswani, and David Chiang

We present Model Invertibility Regularization MIR, a method that jointly trains two directional sequence alignment models, one in each direction, and takes into account the invertibility of the alignment task. By coupling the two models through their parameters (as opposed to through their inferences, as in Liang et al.'s Alignment by Agreement (ABA), and Ganchev et al.'s Posterior Regularization (PostCAT)), our method seamlessly extends to all IBM-style word alignment models as well as to alignment without parallel data. Our proposed algorithm is mathematically sound and inherits convergence guarantees from EM. We evaluate MIR on two tasks: (1) On word alignment, applying MIR on fertility based models we attain higher F-scores than ABA and PostCAT. (2) On Japanese-to-English back-transliteration without parallel data, applied to the decipherment model of Ravi and Knight, MIR learns sparser models that close the gap in whole-name error rate by 33% relative to a model trained on parallel data, and further, beats a previous approach by Mylonakis et al.

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

Yugo Murawaki

For phylogenetic inference, linguistic typology is a promising alternative to lexical evidence because it allows us to compare an arbitrary pair of languages. A challenging problem with typology-based phylogenetic inference is that the changes of typological features over time are less intuitive than those of lexical features. In this paper, we work on reconstructing typologically natural ancestors To do this, we leverage dependencies among typological features. We first represent each language by continuous latent components that capture feature dependencies. We then combine them with a typology evaluator that distinguishes typologically natural languages from other possible combinations of features. We perform phylogenetic inference in the continuous space and use the evaluator to ensure the typological naturalness of inferred ancestors. We show that the proposed method reconstructs known language families more accurately than baseline methods. Lastly, assuming the monogenesis hypothesis, we attempt to reconstruct a common ancestor of the world's languages.

Diamonds in the Rough: Event Extraction from Imperfect Microblog Data

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

We introduce a distantly supervised event extraction approach that extracts complex event templates from microblogs. We show that this near real-time data source is more challenging than news because it contains information that is both approximate (e.g., with values that are close but different from the gold truth) and ambiguous (due to the brevity of the texts), impacting both the evaluation and extraction methods. For the former, we propose a novel, "soft", F1 metric that incorporates similarity between extracted fillers and the gold truth, giving partial credit to different but similar values. With respect to extraction methodology, we propose two extensions to the distant supervision paradigm: to address approximate information, we allow positive training examples to be generated from information that is similar but not identical to gold values; to address ambiguity, we aggregate contexts across tweets discussing the same event. We evaluate our contributions on the complex domain of earthquakes, with events with up to 20 arguments. Our results indicate that, despite their simplicity, our contributions yield a statistically-significant improvement of 25% (relative) over a strong distantly-supervised system. The dataset containing the knowledge base, relevant tweets and manual annotations is publicly available.

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

The advent of social media has brought Internet memes, a unique social phenomenon, to the front stage of the Web. Embodied in the form of images with text descriptions, little do we know about the “language of memes”. In this paper, we statistically study the correlations among popular memes and their wordings, and generate meme descriptions from raw images. To do this, we take a multimodal approach—we propose a robust nonparanormal model to learn the stochastic dependencies among the image, the candidate descriptions, and the popular votes. In experiments, we show that combining text and vision helps identifying popular meme descriptions; that our nonparanormal model is able to learn dense and continuous vision features jointly with sparse and discrete text features in a principled manner, outperforming various competitive baselines; that our system can generate meme descriptions using a simple pipeline.

Unsupervised Dependency Parsing: Let's Use Supervised Parsers

Phong Le and Willem Zuidema

We present a self-training approach to unsupervised dependency parsing that reuses existing supervised and unsupervised parsing algorithms. Our approach, called ‘iterated reranking’ (IR), starts with dependency trees generated by an unsupervised parser, and iteratively improves these trees using the richer probability models used in supervised parsing that are in turn trained on these trees. Our system achieves 1.8% accuracy higher than the state-of-the-part parser of Spitskovsky et al. (2013) on the WSJ corpus.

A Transition-based Algorithm for AMR Parsing

Chuan Wang, Nianwen Xue, and Sameer Pradhan

We present a two-stage framework to parse a sentence into its Abstract Meaning Representation (AMR). We first use a dependency parser to generate a dependency tree for the sentence. In the second stage, we design a novel transition-based algorithm that transforms the dependency tree to an AMR graph. There are several advantages with this approach. First, the dependency parser can be trained on a training set much larger than the training set for the tree-to-graph algorithm, resulting in a more accurate AMR parser overall. Our parser yields an improvement of 5% absolute in F-measure over the best previous result. Second, the actions that we design are linguistically intuitive and capture the regularities in the mapping between the dependency structure and the AMR of a sentence. Third, our parser runs in nearly linear time in practice in spite of a worst-case complexity of $O(n^{<\sup>2})$.

The Geometry of Statistical Machine Translation

Aurelien Waite and Bill Byrne

Most modern statistical machine translation systems are based on linear statistical models. One extremely effective method for estimating the model parameters is minimum error rate training (MERT), which is an efficient form of line optimisation adapted to the highly non-linear objective functions used in machine translation. We describe a polynomial-time generalisation of line optimisation that computes the error surface over a plane embedded in parameter space. The description of this algorithm relies on convex geometry, which is the mathematics of polytopes and their faces. Using this geometric representation of MERT we investigate whether the optimisation of linear models is tractable in general. Previous work on finding optimal solutions in MERT (Galley and Quirk, 2011) established a worst-case complexity that was exponential in the number of sentences, in contrast we show that exponential dependence in the worst-case complexity is mainly in the number of features. Although our work is framed with respect to MERT, the convex geometric description is also applicable to other error-based training methods for linear models. We believe our analysis has important ramifications because it suggests that the current trend in building statistical machine translation systems by introducing a very large number of sparse features is inherently not robust.

Unsupervised Multi-Domain Adaptation with Feature Embeddings

Yi Yang and Jacob Eisenstein

Representation learning is the dominant technique for unsupervised domain adaptation, but existing approaches have two major weaknesses. First, they often require the specification of “pivot features” that generalize across domains, which are selected by task-specific heuristics. We show that a novel but simple feature embedding approach provides better performance, by exploiting the feature template structure common in NLP problems. Second, unsupervised domain adaptation is typically treated as a task of moving from a single source to a single target domain. In reality, test data may be diverse, relating to the training data in some ways but not others. We propose an alternative formulation, in which each instance has a vector of domain attributes, can be used to learn distill the domain-invariant properties of each feature.

Latent Domain Word Alignment for Heterogeneous Corpora*Hoang Cuong and Khalil Sima'an*

This work focuses on the insensitivity of existing word alignment models to domain differences, which often yields suboptimal results on large heterogeneous data. A novel latent domain word alignment model is proposed, which induces domain-conditioned lexical and alignment statistics. We propose to train the model on a heterogeneous corpus under partial supervision, using a small number of seed samples from different domains. The seed samples allow estimating sharper, domain-conditioned word alignment statistics for sentence pairs. Our experiments show that the derived domain-conditioned statistics, once combined together, produce notable improvements both in word alignment accuracy and in translation accuracy of their resulting SMT systems.

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*

People vary widely in their temporal orientation — how often they emphasize the past, present, and future — and this affects their finances, health, and happiness. Traditionally, temporal orientation has been assessed by self-report questionnaires. In this paper, we develop a novel behavior-based assessment using human language on Facebook. We first create a past, present, and future message classifier, engineering features and evaluating a variety of classification techniques. Our message classifier achieves an accuracy of 71.8%, compared with 52.8% from the most frequent class and 58.6% from a model based entirely on time expression features. We quantify a users' overall temporal orientation based on their distribution of messages and validate it against known human correlates: conscientiousness, age, and gender. We then explore social scientific questions, finding novel associations with the factors openness to experience, satisfaction with life, depression, IQ, and one's number of friends. Further, demonstrating how one can track orientation over time, we find differences in future orientation around birthdays.

Cost Optimization in Crowdsourcing Translation: Low cost translations made even cheaper*Mingkun Gao, Wei Xu, and Chris Callison-Burch*

Crowdsourcing makes it possible to create translations at much lower cost than hiring professional translators. However, it is still expensive to obtain the millions of translations that are needed to train statistical machine translation systems. We propose two mechanisms to reduce the cost of crowdsourcing while maintaining high translation quality. First, we develop a method to reduce redundant translations. We train a linear model to evaluate the translation quality on a sentence-by-sentence basis, and fit a threshold between acceptable and unacceptable translations. Unlike past work, which always paid for a fixed number of translations for each source sentence and then chose the best from them, we can stop earlier and pay less when we receive a translation that is good enough. Second, we introduce a method to reduce the pool of translators by quickly identifying bad translators after they have translated only a few sentences. This also allows us to rank translators, so that we re-hire only good translators to reduce cost.

An In-depth Analysis of the Effect of Text Normalization in Social Media*Tyler Baldwin and Yunyao Li*

Recent years have seen increased interest in text normalization in social media, as the informal writing styles found in Twitter and other social media data often cause problems for NLP applications. Unfortunately, most current approaches narrowly regard the normalization task as a “one size fits all” task of replacing non-standard words with their standard counterparts. In this work we build a taxonomy of normalization edits and present a study of normalization to examine its effect on three different downstream applications (dependency parsing, named entity recognition, and text-to-speech synthesis). The results suggest that how the normalization task should be viewed is highly dependent on the targeted application. The results also show that normalization must be thought of as more than word replacement in order to produce results comparable to those seen on clean text.

Multitask Learning for Adaptive Quality Estimation of Automatically Transcribed Utterances*José G. C. de Souza, Hamed Zamani, Matteo Negri, Marco Turchi, and Falavigna Daniele*

We investigate the problem of predicting the quality of automatic speech recognition (ASR) output under the following rigid constraints: i) reference transcriptions are not available, ii) confidence information about the system that produced the transcriptions is not accessible, and iii) training and test data come from multiple domains. To cope with these constraints (typical of the constantly increasing amount of automatic transcrip-

tions that can be found on the Web), we propose a domain-adaptive approach based on multitask learning. Different algorithms and strategies are evaluated with English data coming from four domains, showing that the proposed approach can cope with the limitations of previously proposed single task learning methods.

A Dynamic Programming Algorithm for Tree Trimming-based Text Summarization

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

Tree trimming is the problem of extracting an optimal subtree from an input tree, and sentence extraction and sentence compression methods can be formulated and solved as tree trimming problems. Previous approaches require integer linear programming (ILP) solvers to obtain exact solutions. The problem of this approach is that ILP solvers are black-boxes and have no theoretical guarantee as to their computation complexity. We propose a dynamic programming (DP) algorithm for tree trimming problems whose running time is $O(NL \log N)$, where N is the number of tree nodes and L is the length limit. Our algorithm exploits the zero-suppressed binary decision diagram (ZDD), a data structure that represents a family of sets as a directed acyclic graph, to represent the set of subtrees in a compact form; the structure of ZDD permits the application of DP to obtain exact solutions, and our algorithm is applicable to different tree trimming problems. Moreover, experiments show that our algorithm is faster than state-of-the-art ILP solvers, and that it scales well to handle large summarization problems.

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

Mohammad Salameh, Saif Mohammad, and Svetlana Kiritchenko

When text is translated from one language into another, sentiment is preserved to varying degrees. In this paper, we use Arabic social media posts as stand-in for source language text, and determine loss in sentiment predictability when they are translated into English, manually and automatically. As benchmarks, we use manually and automatically determined sentiment labels of the Arabic texts. We show that sentiment analysis of English translations of Arabic texts produces competitive results, w.r.t. Arabic sentiment analysis. We discover that even though translation significantly reduces the human ability to recover sentiment, automatic sentiment systems are still able to capture sentiment information from the translations.

Corpus-based discovery of semantic intensity scales

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

Gradable terms such as brief, lengthy and extended illustrate varying degrees of a scale and can therefore participate in comparative constructs. Knowing the set of words that can be compared on the same scale and the associated ordering between them (brief < lengthy < extended) is very useful for a variety of lexical semantic tasks. Current techniques to derive such an ordering rely on WordNet to determine which words belong on the same scale and are limited to adjectives. Here we describe an extension to recent work: we investigate a fully automated pipeline to extract gradable terms from a corpus, group them into clusters reflecting the same scale and establish an ordering among them. This methodology reduces the amount of required handcrafted knowledge, and can infer gradability of words independent of their part of speech. Our approach infers an ordering for adjectives with comparable performance to previous work, but also for adverbs with an accuracy of 71%. We find that the technique is useful for inferring such rankings among words across different domains, and present an example using biomedical text.

Dialogue focus tracking for zero pronoun resolution

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

We take a novel approach to zero pronoun resolution in Chinese: our model explicitly tracks the flow of focus in a discourse. Our approach, which generalizes to deictic references, is not reliant on the presence of overt noun phrase antecedents to resolve to, and allows us to address the large percentage of “non-anaphoric” pronouns filtered out in other approaches. We furthermore train our model using readily available parallel Chinese/English corpora, allowing for training without hand-annotated data. Our results demonstrate improvements on two test sets, as well as the usefulness of linguistically motivated features.

Solving Hard Coreference Problems

Haoruo Peng, Daniel Khashabi, and Dan Roth

Coreference resolution is a key problem in natural language understanding that still escapes reliable solutions. One fundamental difficulty has been that of resolving instances involving pronouns since they often require deep language understanding and use of background knowledge. In this paper we propose an algorithmic solution that involves a new representation for the knowledge required to address hard coreference problems,

along with a constrained optimization framework that uses this knowledge in coreference decision making. Our representation, Predicate Schemas, is instantiated with knowledge acquired in an unsupervised way, and is compiled automatically into constraints that impact the coreference decision. We present a general coreference resolution system that significantly improves state-of-the-art performance on hard, Winograd-style, pronoun resolution cases, while still performing at the state-of-the-art level on standard coreference resolution datasets.

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

Little work from the Natural Language Processing community has targeted the role of quantities in Natural Language Understanding. This paper takes some key steps towards facilitating reasoning about quantities expressed in natural language. We investigate two different tasks of numerical reasoning. First, we consider Quantity Entailment, a new task formulated to understand the role of quantities in general textual inference tasks. Second, we consider the problem of automatically understanding and solving elementary school math word problems. In order to address these quantitative reasoning problems we first develop a computational approach which we show to successfully recognize and normalize textual expressions of quantities. We then use these capabilities to further develop algorithms to assist reasoning in the context of the aforementioned tasks.

[TACL] Learning Constraints for Information Structure Analysis of Scientific Documents*Yufan Guo, Roi Reichart, and Anna Korhonen*

Inferring the information structure of scientific documents is useful for many NLP applications. Existing approaches to this task require substantial human effort. We propose a framework for constraint learning that reduces human involvement considerably. Our model uses topic models to identify latent topics and their key linguistic features in input documents, induces constraints from this information and maps sentences to their dominant information structure categories through a constrained unsupervised model. When the induced constraints are combined with a fully unsupervised model, the resulting model challenges existing lightly supervised feature-based models as well as unsupervised models that use manually constructed declarative knowledge. Our results demonstrate that useful declarative knowledge can be learned from data with very limited human involvement.

SRW Poster Session

Time: 6:00–9:00

Location: Plaza Ballroom A, B, & C

Cache-Augmented Latent Topic Language Models for Speech Retrieval*Jonathan Wintrode*

We aim to improve speech retrieval performance by augmenting traditional N-gram language models with different types of topic context. We present a latent topic model framework that treats documents as arising from an underlying topic sequence combined with a cache-based repetition model. We analyze our proposed model both for its ability to capture word repetition via the cache and for its suitability as a language model for speech recognition and retrieval. We show this model, augmented with the cache, captures intuitive repetition behavior across languages and exhibits lower perplexity than regular LDA on held out data in multiple languages. Lastly, we show that our joint model improves speech retrieval performance beyond N-grams or latent topics alone, when applied to a term detection task in all languages considered.

Reliable Lexical Simplification for Non-Native Speakers*Gustavo Paetzold*

Lexical Simplification is the task of modifying the lexical content of complex sentences in order to make them simpler. Due to the lack of reliable resources available for the task, most existing approaches have difficulties producing simplifications which are grammatical and that preserve the meaning of the original text. In order to improve on the state-of-the-art of this task, we propose user studies with non-native speakers, which will result in new, sizeable datasets, as well as novel ways of performing Lexical Simplification. The results of our first experiments show that new types of classifiers, along with the use of additional resources such as spoken text language models, produce the state-of-the-art results for the Lexical Simplification task of SemEval-2012.

Analyzing Newspaper Crime Reports for Identification of Safe Transit Paths

Vasu Sharma, Rajat Kulshreshtha, Puneet Singh, Nishant Agrawal, and Akshay Kumar

In the present situation, where every day we come across one or the other crime, it is of prime importance to ensure ones safety. Any measure which could ensure ones safety is certain to pay for itself. An automated system which can suggest the safest path between any two locations in the city avoiding crime prone areas will be of utmost importance for the society. In this paper, we propose a method to find the safest path between two locations, based on the geographical model of crime intensities. We consider the police records and news articles for finding crime density of different areas of the city. It is essential to consider news articles as there is a significant delay in updating police crime records. We address this problem by updating the crime intensities based on current news feeds. Based on the updated crime intensities, we identify the safest path. It is this real time updation of crime intensities which makes our model way better than the models that are presently in use. Our model would also inform the user of crime sprees and spurt of crimes in a particular area thereby ensuring that user avoids these crime hot spots.

Relation extraction pattern ranking using word similarity

Konstantinos Lambrou-Latreille

Our thesis proposal aims at integrating word similarity measures in pattern ranking for relation extraction bootstrapping algorithms. We note that although many contributions have been done on pattern ranking schemas, few explored the use of word-level semantic similarity. Our hypothesis is that word similarity would allow better pattern comparison and better pattern ranking, resulting in less semantic drift commonly problematic in bootstrapping algorithms. In this paper, as a first step into this research, we explore different pattern representations, various existing pattern ranking approaches and some word similarity measures. We also present a methodology and evaluation approach to test our hypothesis.

Towards a Better Semantic Role Labeling of Complex Predicates

Glorianna Jagfeld and Lonneke van der Plas

We propose a way to automatically improve the annotation of verbal complex predicates in PropBank which until now has been treating language mostly in a compositional manner. In order to minimize the manual re-annotation effort, we build on the recently introduced concept of aliasing complex predicates to existing PropBank rolesets which encompass the same meaning and argument structure. We suggest to find aliases automatically by applying a multilingual distributional model that uses the translations of simple and complex predicates as features. Furthermore, we set up an annotation effort to obtain a frequency balanced, realistic test set for this task. Our method reaches an accuracy of 44% on this test set and 72% for the more frequent test items in a lenient evaluation, which is not far from the upper bounds from human annotation.

Exploring Relational Features and Learning under Distant Supervision for Information Extraction Tasks

Ajay Nagesh

Information Extraction (IE) has become an indispensable tool in our quest to handle the data deluge of the information age. IE can broadly be classified into Named-entity Recognition (NER) and Relation Extraction (RE). In this thesis, we view the task of IE as finding patterns in unstructured data, which can either take the form of features and/or be specified by constraints. In NER, we study the categorization of complex relational features and outline methods to learn feature combinations through induction. We demonstrate the efficacy of induction techniques in learning : i) rules for the identification of named entities in text — the novelty is the application of induction techniques to learn in a very expressive declarative rule language ii) a richer sequence labeling model — enabling optimal learning of discriminative features. In RE, our investigations are in the paradigm of distant supervision, which facilitates the creation of large albeit noisy training data. We devise an inference framework in which constraints can be easily specified in learning relation extractors. In addition, we reformulate the learning objective in a max-margin framework. To the best of our knowledge, our formulation is the first to optimize multi-variate non-linear performance measures such as $F\text{-}\beta$ for a latent variable structure prediction task.

Entity/Event-Level Sentiment Detection and Inference

Lingjia Deng

Most of the work in sentiment analysis and opinion mining focuses on extracting explicit sentiments. Opinions may be expressed implicitly via inference rules over explicit sentiments. In this thesis, we incorporate the inference rules as constraints in joint prediction models, to develop an entity/event-level sentiment analysis system which aims at detecting both explicit and implicit sentiments expressed among entities and events in

the text, especially focusing on but not limited to sentiments toward events that positively or negatively affect entities (+/-effect events).

Initial Steps for Building a Lexicon of Adjectives with Scalemates

Bryan Wilkinson

This paper describes work in progress to use clustering to create a lexicon of words that engage in the lexicosemantic relationship known as grading. While other resources like thesauri and taxonomies exist detailing relationships such as synonymy, antonymy, and hyponymy, we do not know of any thorough resource for grading. This work focuses on identifying the words that may participate in this relationship, paving the way for the creation of a true grading lexicon later.

A Preliminary Evaluation of the Impact of Syntactic Structure in Semantic Textual Similarity and Semantic Relatedness Tasks

Ngoc Phuoc An Vo and Octavian Popescu

The well related tasks of evaluating the Semantic Textual Similarity and Semantic Relatedness have been under a special attention in NLP community. Many different approaches have been proposed, implemented and evaluated at different levels, such as lexical similarity, word/string/POS tags overlapping, semantic modeling (LSA, LDA), etc. However, at the level of syntactic structure, it is not clear how significant it contributes to the overall accuracy. In this paper, we make a preliminary evaluation of the impact of the syntactic structure in the tasks by running and analyzing the results from several experiments regarding to how syntactic structure contributes to solving the tasks.

Benchmarking Machine Translated Sentiment Analysis for Arabic Tweets

Eshrag Refaei and Verena Rieser

Traditional approaches to Sentiment Analysis (SA) rely on large annotated data sets or wide-coverage sentiment lexica, and as such often perform poorly on under-resourced languages. This paper presents empirical evidence of an efficient SA approach using freely available machine translation (MT) systems to translate Arabic tweets to English, which we then label for sentiment using a state-of-the-art English SA system. We show that this approach significantly outperforms a number of standard approaches on a gold-standard heldout data set, and performs equally well compared to more cost-intense methods with 76% accuracy. This confirms MT-based SA as a cheap and effective alternative to building a fully fledged SA system when dealing with under-resourced languages.

Learning Kernels for Semantic Clustering: A Deep Approach

Ignacio Arroyo-Fernández

In this thesis proposal we present a novel semantic embedding method, which aims at consistently performing semantic clustering at sentence level. Taking into account special aspects of Vector Space Models (VSMs), we propose to learn reproducing kernels in classification tasks. By this way, capturing spectral features from data is possible. These features make it theoretically plausible to model semantic similarity criteria in Hilbert spaces, i.e. the embedding spaces. We could improve the semantic assessment over embeddings, which are criterion-derived representations from traditional semantic vectors. The learned kernel could be easily transferred to clustering methods, where the Multi-Class Imbalance Problem is considered (e.g. semantic clustering of definitions of terms).

Narrowing the Loop: Integration of Resources and Linguistic Dataset Development with Interactive Machine Learning

Seid Muhie Yimam

This thesis proposal sheds light on the role of interactive machine learning and implicit user feedback for manual annotation tasks and semantic writing aid applications. First we focus on the cost-effective annotation of training data using an interactive machine learning approach by conducting an experiment for sequence tagging of German named entity recognition. To show the effectiveness of the approach, we further carry out a sequence tagging task on Amharic part-of-speech and are able to significantly reduce time used for annotation. The second research direction is to systematically integrate different NLP resources for our new semantic writing aid tool using again an interactive machine learning approach to provide contextual paraphrase suggestions. We develop a baseline system where three lexical resources are combined to provide paraphrasing in context and show that combining resources is a promising direction.

Relation Extraction from Community Generated Question-Answer Pairs

Denis Savenkov, Wei-Lwun Lu, Jeff Dalton, and Eugene Agichtein

Community question answering (CQA) websites contain millions of question and answer (QnA) pairs that represent real users' interests. Traditional methods for relation extraction from natural language text operate over individual sentences. However answer text is sometimes hard to understand without knowing the question, e.g., it may not name the subject or relation of the question. This work presents a novel model for relation extraction from CQA data, which uses discourse of QnA pairs to predict relations between entities mentioned in question and answer sentences. Experiments on 2 publicly available datasets demonstrate that the model can extract from ~20% to ~40% additional relation triples, not extracted by existing sentence-based models.

Detecting Translation Direction: A Cross-Domain Study

Sauleh Etemadi and Kristina Toutanova

Parallel corpora are constructed by taking a document authored in one language and translating it into another language. However, the information about the authored and translated sides of the corpus is usually not preserved. When available, this information can be used to improve statistical machine translation. Existing statistical methods for translation direction detection have low accuracy when applied to the realistic out-of-domain setting, especially when the input texts are short. Our contributions in this work are three fold: 1) We develop a multi-corpus parallel dataset with translation direction labels at the sentence level, 2) we perform a comparative evaluation of previously introduced features for translation direction detection in a cross-domain setting and 3) we generalize a previously introduced type of features to outperform the best previously proposed features in detecting translation direction and achieve 0.80 precision with 0.85 recall.

Improving the Translation of Discourse Markers for Chinese into English

David Steele

Discourse markers (DMs) are ubiquitous cohesive devices used to connect what is said or written. However, across languages there is divergence in their usage, placement, and frequency, which is considered to be a major problem for machine translation (MT). This paper presents an overview of a proposed thesis, exploring the difficulties around DMs in MT, with a focus on Chinese and English. The thesis will examine two main areas: modelling cohesive devices within sentences and modelling discourse relations (DRs) across sentences. Initial experiments have shown promising results for building a prediction model that uses linguistically inspired features to help improve word alignments with respect to the implicit use of cohesive devices, which in turn leads to improved hierarchical phrase based MT.

Discourse and Document-level Information for Evaluating Language Output Tasks

Carolina Scartton

Evaluating the quality of language output tasks such as Machine Translation (MT) and Automatic Summarisation (AS) is a challenging topic in Natural Language Processing (NLP). Recently, techniques focusing only on the use of outputs of the systems and source information have been investigated. In MT, this is referred to as Quality Estimation (QE), an approach based on using machine learning techniques to predict the quality of unseen data, generalising from a few labelled data points. Traditional QE research addresses sentence-level QE evaluation and prediction, disregarding document-level information. Document-level QE requires a different set up from sentence-level, which makes the study of appropriate quality scores, features and models necessary. Our aim is to explore document-level QE of MT, focusing on discourse information. However, the findings of this research can improve other NLP tasks, such as AS.

Speeding Document Annotation with Topic Models

Forough Poursabzi-Sangdeh and Jordan Boyd-Graber

Document classification and topic models are useful tools for managing and understanding large corpora. Topic models are used to uncover underlying semantic and structure of document collections. Categorizing large collection of documents requires hand-labeled training data, which is time consuming and needs human expertise. We believe engaging user in the process of document labeling helps reduce annotation time and address user needs. We present an interactive tool for document labeling. We use topic models to help users in this procedure. Our preliminary results show that users can more effectively and efficiently apply labels to documents using topic model information.

Lifelong Machine Learning for Topic Modeling and Beyond

Zhiyuan Chen

Machine learning has been popularly used in numerous natural language processing tasks. However, most ma-

chine learning models are built using a single dataset. This is often referred to as one-shot learning. Although this one-shot learning paradigm is very useful, it will never make an NLP system understand the natural language because it does not accumulate knowledge learned in the past and make use of the knowledge in future learning and problem solving. In this thesis proposal, I first present a survey of lifelong machine learning (LML). I then narrow down to one specific NLP task, i.e., topic modeling. I propose several approaches to apply lifelong learning idea in topic modeling. Such capability is essential to make an NLP system versatile and holistic.

Semantics-based Graph Approach to Complex Question-Answering

Tomasz Jurczyk and Jinho D. Choi

This paper suggests an architectural approach of representing knowledge graph for complex question-answering. There are four kinds of entity relations added to our knowledge graph: syntactic dependencies, semantic role labels, named entities, and coreference links, which can be effectively applied to answer complex questions. As a proof of concept, we demonstrate how our knowledge graph can be used to solve complex questions such as arithmetics. Our experiment shows a promising result on solving arithmetic questions, achieving the 3-folds cross-validation score of 71.75%.

Recognizing Textual Entailment using Dependency Analysis and Machine Learning

Nidhi Sharma, Richa Sharma, and Kanad K. Biswas

This paper presents a machine learning system that uses dependency-based features and lexical features for recognizing textual entailment. The system first generates a frame-based structured representation of the sentences using dependency relations obtained from Stanford Dependency parser. This structured representation is then used to evaluate a set of lexical and syntactic features. These features are simple and intuitive; easy to comprehend and evaluate. The system evaluates the feature values automatically without any manual effort or intervention. The performance of the proposed system is evaluated by performing experiments on RTE1, RTE2 and RTE3 datasets. Decision trees are found to outperform other classifiers while SVM classifier using an RBF kernel shows comparable performance. Additionally, a comparative study of the current system with other ML-based systems for RTE to check the performance of the proposed system is also carried out. The dependency-based heuristics and lexical features from the current system have resulted in significant improvement in accuracy over existing state-of-art ML-based solutions for RTE. A task-based analysis of the current system for RTE1 dataset is also performed which, shows improved performance as compared to similar study performed by different researchers in the past.

Bilingual lexicon extraction for a distant language pair using a small parallel corpus

Ximena Gutierrez-Vasques

The aim of this thesis proposal is to perform bilingual lexicon extraction for cases in which small parallel corpora are available and it is not easy to obtain monolingual corpus for at least one of the languages. Moreover, the languages are typologically distant and there is no bilingual seed lexicon available. We focus on the language pair Spanish-Nahuatl, we propose to work with morpheme based representations in order to reduce the sparseness and to facilitate the task of finding lexical correspondences between a highly agglutinative language and a fusional one. We take into account contextual information but instead of using a precompiled seed dictionary, we use the distribution and dispersion of the positions of the morphological units as cues to compare the contextual vectors and obtaining the translation candidates.

Morphological Paradigms: Computational Structure and Unsupervised Learning

Jackson Lee

This thesis explores the computational structure of morphological paradigms from the perspective of unsupervised learning. Three topics are studied: (i) stem identification, (ii) paradigmatic similarity, and (iii) paradigm induction. All the three topics progress in terms of the scope of data in question. The first and second topics explore structure when morphological paradigms are given, first within a paradigm and then across paradigms. The third topic asks where morphological paradigms come from in the first place, and explores strategies of paradigm induction from child-directed speech. This research is of interest to linguists and natural language processing researchers, for both theoretical questions and applied areas.

Computational Exploration to Linguistic Structures of Future: Classification and Categorization

Aiming Ni, Jinho D. Choi, Jason Shepard, and Phillip Wolff

Many languages, like English, lack a future tense, which makes the automatic identification of future reference a challenge. In this research we extend Latent Dirichlet allocation (LDA) for use in the identification of future referring sentences. Building off a set of hand designed rules, we trained a ADAGRAD classifier to be able automatically detect sentences referring to the future. Uni-bi-trigram and syntactic rule mixed feature was found to provide the highest accuracy. Latent Dirichlet Allocation (LDA) indicated the existence of four major categories future orientation. Lastly, the results of these analyses were found to correlate with a range of behavioral measures, offering evidence in support of the psychological reality of the categories.

4

Main Conference: Tuesday, June 2

Overview

7:30–9:00	Registration and Breakfast Session 4	<i>Plaza Exhibit All</i>
9:00–10:40	Dialogue and Spoken Language Processing (Long Papers) <i>Plaza Ballroom A & B</i>	Machine Learning for NLP (Long Papers) <i>Plaza Ballroom D & E</i>
		Phonology, Morphology and Word Segmentation (Long Papers) <i>Plaza Ballroom F</i>
10:40–11:15	Break	<i>Plaza Exhibit All</i>
	Session 5	
11:15–12:30	Semantics (Short Papers) <i>Plaza Ballroom A & B</i>	Machine Translation (Short Papers) <i>Plaza Ballroom D & E</i>
		Morphology, Syntax, Multilinguality, and Applications (Short Papers) <i>Plaza Ballroom F</i>
12:30–2:00	Lunch	
	Session 6	
2:00–3:15	Generation and Summarization (Long Papers) <i>Plaza Ballroom A & B</i>	Discourse and Coreference (Long Papers) <i>Plaza Ballroom D & E</i>
		Information Extraction and Question Answering (Long Papers) <i>Plaza Ballroom F</i>
3:15–3:45	Break	<i>Plaza Exhibit All</i>
	Session 7	
3:45–5:00	Semantics (Long + TACL Papers) <i>Plaza Ballroom A & B</i>	Information Extraction and Question Answering (Long + TACL Papers) <i>Plaza Ballroom D & E</i>
		Machine Translation (Long Papers) <i>Plaza Ballroom F</i>
5:00–6:30	Poster session 2A (Short papers)	<i>Plaza Ballroom A, B, & C</i>
5:00–6:30	Demo Session A	<i>Plaza Ballroom A, B, & C</i>
6:30–8:00	Poster session 2B (Short papers)	<i>Plaza Ballroom A, B, & C</i>

Main Conference

6:30 – 8:00 **Demo Session B**
8:00 – 11:00 **Social Event**

Plaza Ballroom A, B, & C
Grand Ballroom

Session 4 Overview – Tuesday, June 2, 2015

Track A	Track B	Track C	
<i>Dialogue and Spoken Language Processing (Long Papers)</i> Plaza Ballroom A & B	<i>Machine Learning for NLP (Long Papers)</i> Plaza Ballroom D & E	<i>Phonology, Morphology and Word Segmentation (Long Papers)</i> Plaza Ballroom F	9:25
Semantic Grounding in Dialogue for Complex Problem Solving <i>Li and Boyer</i>	Early Gains Matter: A Case for Preferring Generative over Discriminative Crowdsourcing Models <i>Felt, Black, Ringger, Seppi, and Haertel</i>	Inflection Generation as Discriminative String Transduction <i>Nicolai, Cherry, and Kondrak</i>	9:50
Learning Knowledge Graphs for Question Answering through Conversational Dialog <i>Hixon, Clark, and Hajishirzi</i>	Optimizing Multivariate Performance Measures for Learning Relation Extraction Models <i>Haffari, Nagesh, and Ramakrishnan</i>	Penalized Expectation Propagation for Graphical Models over Strings <i>Cotterell and Eisner</i>	10:15
Sentence segmentation of aphasic speech <i>Fraser, Ben-David, Hirst, Graham, and Rochon</i>	Convolutional Neural Network for Paraphrase Identification <i>Yin and Schütze</i>	Joint Generation of Translations from Multiple Representations <i>Yao and Kondrak</i>	10:40
Semantic parsing of speech using grammars learned with weak supervision <i>Gaspers, Cimiano, and Wrede</i>	Representation Learning Using Multi-Task Deep Neural Networks for Semantic Classification and Information Retrieval <i>Liu, Gao, He, Deng, Duh, and Wang</i>	Prosodic boundary information helps unsupervised word segmentation <i>Ludusan, Synnaeve, and Dupoux</i>	

Parallel Session 4

Session 4A: Dialogue and Spoken Language Processing (Long Papers)

Plaza Ballroom A & B

Chair: Marilyn Walker

Semantic Grounding in Dialogue for Complex Problem Solving

Xiaolong Li and Kristy Boyer

09:00–09:25

Dialogue systems that support users in complex problem solving must interpret user utterances within the context of a dynamically changing, user-created problem solving artifact. This paper presents a novel approach to semantic grounding of noun phrases within tutorial dialogue for computer programming. Our approach performs joint segmentation and labeling of the noun phrases to link them to attributes of entities within the problem-solving environment. Evaluation results on a corpus of tutorial dialogue for Java programming demonstrate that a Conditional Random Field model performs well, achieving an accuracy of 89.3% for linking semantic segments to the correct entity attributes. This work is a step toward enabling dialogue systems to support users in increasingly complex problem-solving tasks.

Learning Knowledge Graphs for Question Answering through Conversational Dialog

Ben Hixon, Peter Clark, and Hannaneh Hajishirzi

09:25–09:50

We describe how a question-answering system can learn about its domain from conversational dialogs. Our system learns to relate concepts in science questions to propositions in a fact corpus, stores new concepts and relations in a knowledge graph (KG), and uses the graph to solve questions. We are the first to acquire knowledge for question-answering from open, natural language dialogs without a fixed ontology or domain model that predetermines what users can say. Our relation-based strategies complete more successful dialogs than a query expansion baseline, our task-driven relations are more effective for solving science questions than relations from general knowledge sources, and our method is practical enough to generalize to other domains.

Sentence segmentation of aphasic speech

Kathleen C. Fraser, Naama Ben-David, Graeme Hirst, Naida Graham, and Elizabeth Rochon
09:50–10:15

Automatic analysis of impaired speech for screening or diagnosis is a growing research field; however there are still many barriers to a fully automated approach. When automatic speech recognition is used to obtain the speech transcripts, sentence boundaries must be inserted before most measures of syntactic complexity can be computed. In this paper, we consider how language impairments can affect segmentation methods, and compare the results of computing syntactic complexity metrics on automatically and manually segmented transcripts. We find that the important boundary indicators and the resulting segmentation accuracy can vary depending on the type of impairment observed, but that results on patient data are generally similar to control data. We also find that a number of syntactic complexity metrics are robust to the types of segmentation errors that are typically made.

Semantic parsing of speech using grammars learned with weak supervision

Judith Gaspers, Philipp Cimiano, and Britta Wrede

10:15–10:40

Semantic grammars can be applied both as a language model for a speech recognizer and for semantic parsing, e.g. in order to map the output of a speech recognizer into formal meaning representations. Semantic speech recognition grammars are, however, typically created manually or learned in a supervised fashion, requiring extensive manual effort in both cases. Aiming to reduce this effort, in this paper we investigate the induction of semantic speech recognition grammars under weak supervision. We present empirical results, indicating that the induced grammars support semantic parsing of speech with a rather low loss in performance when compared to parsing of input without recognition errors. Further, we show improved parsing performance compared to applying n-gram models as language models and demonstrate how our semantic speech recognition grammars can be enhanced by weights based on occurrence frequencies, yielding an improvement in parsing performance over applying unweighted grammars.

Session 4B: Machine Learning for NLP (Long Papers)

Plaza Ballroom D & E

Chair: Amarnag Subramanya

Early Gains Matter: A Case for Preferring Generative over Discriminative Crowdsourcing Models*Paul Felt, Kevin Black, Eric Ringger, Kevin Seppi, and Robbie Haertel*

09:00–09:25

In modern practice, labeling a dataset often involves aggregating annotator judgments obtained from crowdsourcing. State-of-the-art aggregation is performed via inference on probabilistic models, some of which are data-aware, meaning that they leverage features of the data (e.g., words in a document) in addition to annotator judgments. Previous work largely prefers discriminatively trained conditional models. This paper demonstrates that a data-aware crowdsourcing model incorporating a generative multinomial data model enjoys a strong competitive advantage over its discriminative log-linear counterpart in the typical crowdsourcing setting. That is, the generative approach is better except when the annotators are highly accurate in which case simple majority vote is often sufficient. Additionally, we present a novel mean-field variational inference algorithm for the generative model that significantly improves on the previously reported state-of-the-art for that model. We validate our conclusions on six text classification datasets with both human-generated and synthetic annotations.

Optimizing Multivariate Performance Measures for Learning Relation Extraction Models*Gholamreza Haffari, Ajay Nagesh, and Ganesh Ramakrishnan*

09:25–09:50

We describe a novel max-margin learning approach to optimize non-linear performance measures for distantly-supervised relation extraction models. Our approach can be generally used to learn latent variable models under multivariate non-linear performance measures, such as F_β -score. Our approach interleaves Concave-Convex Procedure (CCCP) for populating latent variables with dual decomposition to factorize the original hard problem into smaller independent sub-problems. The experimental results demonstrate that our learning algorithm is more effective than the ones commonly used in the literature for distant supervision of information extraction models. On several data conditions, we show that our method outperforms the baseline and results in up to 8.5% improvement in the F_1 -score.

Convolutional Neural Network for Paraphrase Identification*Wenpeng Yin and Hinrich Schütze*

09:50–10:15

We present a new deep learning architecture Bi-CNN-MI for paraphrase identification (PI). Based on the insight that PI requires comparing two sentences on multiple levels of granularity, we learn multigranular sentence representations using convolutional neural network (CNN) and model interaction features at each level. These features are then the input to a logistic classifier for PI. All parameters of the model (for embeddings, convolution and classification) are directly optimized for PI. To address the lack of training data, we pretrain the network in a novel way using a language modeling task. Results on the MSRP corpus surpass that of previous NN competitors.

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*

10:15–10:40

Methods of deep neural networks (DNNs) have recently demonstrated superior performance on a number of natural language processing tasks. However, in most previous work, the models are learned based on either unsupervised objectives, which does not directly optimize the desired task, or single-task supervised objectives, which often suffer from insufficient training data. We develop a multi-task DNN for learning representations across multiple tasks, not only leveraging large amounts of cross-task data, but also benefiting from a regularization effect that leads to more general representations to help tasks in new domains. Our multi-task DNN approach combines tasks of multiple-domain classification (for query classification) and information retrieval (ranking for web search), and demonstrates significant gains over strong baselines in a comprehensive set of domain adaptation and other multi-task learning experiments.

Session 4C: Phonology, Morphology and Word Segmentation (Long Papers)

Plaza Ballroom F

Chair: Hal Daumé III

Inflection Generation as Discriminative String Transduction

Garrett Nicolai, Colin Cherry, and Grzegorz Kondrak

09:00–09:25

We approach the task of morphological inflection generation as discriminative string transduction. Our supervised system learns to generate word-forms from lemmas accompanied by morphological tags, and refines them by referring to the other forms within a paradigm. Results of experiments on six diverse languages with varying amounts of training data demonstrate that our approach improves the state of the art in terms of predicting inflected word-forms.

Penalized Expectation Propagation for Graphical Models over Strings

Ryan Cotterell and Jason Eisner

09:25–09:50

We present penalized expectation propagation, a novel algorithm for approximate inference in graphical models. Expectation propagation is a variant of loopy belief propagation that keeps messages tractable by projecting them back into a given family of functions. Our extension speeds up the method by using a structured-sparsity penalty to prefer simpler messages within the family. In the case of string-valued random variables, penalized EP lets us work with an expressive non-parametric function family based on variable-length n-gram models. On phonological inference problems, we obtain substantial speedup over previous related algorithms with no significant loss in accuracy.

Joint Generation of Transliterations from Multiple Representations

Lei Yao and Grzegorz Kondrak

09:50–10:15

Machine transliteration is often referred to as phonetic translation. We show that transliterations incorporate information from both spelling and pronunciation, and propose an effective model for joint transliteration generation from both representations. We further generalize this model to include transliterations from other languages, and enhance it with re-ranking and lexicon features. We demonstrate substantial improvements in transliteration accuracy on several datasets.

Prosodic boundary information helps unsupervised word segmentation

Bogdan Ludusan, Gabriel Synnaeve, and Emmanuel Dupoux

10:15–10:40

It is well known that prosodic information is used by infants in early language acquisition. In particular, prosodic boundaries have been shown to help infants with sentence and word-level segmentation. In this study, we extend an unsupervised method for word segmentation to include information about prosodic boundaries. The boundary information used was either derived from oracle data (hand-annotated), or extracted automatically with a system that employs only acoustic cues for boundary detection. The approach was tested on two different languages, English and Japanese, and the results show that boundary information helps word segmentation in both cases. The performance gain obtained for two typologically distinct languages shows the robustness of prosodic information for word segmentation. Furthermore, the improvements are not limited to the use of oracle information, similar performances being obtained also with automatically extracted boundaries.

Session 5 Overview – Tuesday, June 2, 2015

Track A	Track B	Track C	
<i>Semantics (Short Papers)</i> Plaza Ballroom A & B	<i>Machine Translation (Short Papers)</i> Plaza Ballroom D & E	<i>Morphology, Syntax, Multilinguality, and Applications (Short Papers)</i> Plaza Ballroom F	
So similar and yet incompatible: Toward the automated identification of semantically compatible words <i>Kruszewski and Baroni</i>	Morphological Modeling for Machine Translation of English-Iraqi Arabic Spoken Dialogs <i>Kirchhoff, Tam, Richey, and Wang</i>	Paradigm classification in supervised learning of morphology <i>Ahlberg, Forsberg, and Hulden</i>	11:30
Do Supervised Distributional Methods Really Learn Lexical Inference Relations? <i>Levy, Remus, Biemann, and Dagan</i>	Continuous Adaptation to User Feedback for Statistical Machine Translation <i>Blain, Bougares, Hazem, Barrault, and Schwenk</i>	Shift-Reduce Constituency Parsing with Dynamic Programming and POS Tag Lattice <i>Mi and Huang</i>	11:45
A Word Embedding Approach to Predicting the Compositionality of Multi-word Expressions <i>Salehi, Cook, and Baldwin</i>	Normalized Word Embedding and Orthogonal Transform for Bilingual Word Translation <i>Xing, Wang, Liu, and Lin</i>	Unsupervised Code-Switching for Multilingual Historical Document Transcription <i>Garrette, Alpert-Abrams, Berg-Kirkpatrick, and Klein</i>	12:00
Word Embedding-based Antonym Detection using Thesauri and Distributional Information <i>Ono, Miwa, and Sasaki</i>	Fast and Accurate Preordering for SMT using Neural Networks <i>de Gispert, Iglesias, and Byrne</i>	Matching Citation Text and Cited Spans in Biomedical Literature: a Search-Oriented Approach <i>Cohan, Soldaini, and Goharian</i>	12:15
A Comparison of Word Similarity Performance Using Explanatory and Non-explanatory Texts <i>Jin and Schuler</i>	APRO: All-Pairs Ranking Optimization for MT Tuning <i>Dreyer and Dong</i>	Effective Feature Integration for Automated Short Answer Scoring <i>Sakaguchi, Heilman, and Madnani</i>	12:30

Parallel Session 5

Session 5A: Semantics (Short Papers)

Plaza Ballroom A & B

Chair: Daniel Gildea

So similar and yet incompatible: Toward the automated identification of semantically compatible words

Germán Kruszewski and Marco Baroni

11:15–11:30

We introduce the challenge of detecting semantically compatible words, that is, words that can potentially refer to the same thing (cat and hindrance are compatible, cat and dog are not), arguing for its central role in many semantic tasks. We present a publicly available data-set of human compatibility ratings, and a neural-network model that takes distributional embeddings of words as input and learns alternative embeddings that perform the compatibility detection task quite well.

Do Supervised Distributional Methods Really Learn Lexical Inference Relations?

Omer Levy, Steffen Remus, Chris Biemann, and Ido Dagan

11:30–11:45

Distributional representations of words have been recently used in supervised settings for recognizing lexical inference relations between word pairs, such as hypernymy and entailment. We investigate a collection of these state-of-the-art methods, and show that they do not actually learn a relation between two words. Instead, they learn an independent property of a single word in the pair: whether that word is a “prototypical hypernym”.

A Word Embedding Approach to Predicting the Compositionality of Multiword Expressions

Bahar Salehi, Paul Cook, and Timothy Baldwin

11:45–12:00

This paper presents the first attempt to use word embeddings to predict the compositionality of multiword expressions. We consider both single- and multi-prototype word embeddings. Experimental results show that, in combination with a back-off method based on string similarity, word embeddings outperform a method using count-based distributional similarity. Our best results are competitive with, or superior to, state-of-the-art methods over three standard compositionality datasets, which include two types of multiword expressions and two languages.

Word Embedding-based Antonym Detection using Thesauri and Distributional Information

Masataka Ono, Makoto Miwa, and Yutaka Sasaki

12:00–12:15

This paper proposes a novel approach to train word embeddings to capture antonyms. Word embeddings have shown to capture synonyms and analogies. Such word embeddings, however, cannot capture antonyms since they depend on the distributional hypothesis. Our approach utilizes supervised synonym and antonym information from thesauri, as well as distributional information from large-scale unlabelled text data. The evaluation results on the GRE antonym question task show that our model outperforms the state-of-the-art systems and it can answer the antonym questions in the F-score of 89%.

A Comparison of Word Similarity Performance Using Explanatory and Non-explanatory Texts

Lifeng Jin and William Schuler

12:15–12:30

Vectorial representations derived from large current events datasets such as Google News have been shown to perform well on word similarity tasks. This paper shows vectorial representations derived from substantially smaller explanatory text datasets such as English Wikipedia and Simple English Wikipedia preserve enough lexical semantic information to make these kinds of category judgments with equal or better accuracy. Analysis shows these results are driven by a prevalence of commonsense facts in explanatory text. These positive results for small datasets suggest vectors derived from slower but more accurate deep parsers may be practical for lexical semantic applications.

Session 5B: Machine Translation (Short Papers)

Plaza Ballroom D & E

Chair: David Chiang

Morphological Modeling for Machine Translation of English-Iraqi Arabic Spoken Dialogs
Katrin Kirchhoff, Yik-Cheung Tam, Colleen Richey, and Wen Wang 11:15–11:30

This paper addresses the problem of morphological modeling in statistical speech-to-speech translation for English to Iraqi Arabic. An analysis of user data from a real-time MT-based dialog system showed that generating correct verbal inflections is a key problem for this language pair. We approach this problem by enriching the training data with morphological information derived from source-side dependency parses. We analyze the performance of several parsers as well as the effect on different types of translation models. Our method achieves an improvement of more than a full BLEU point and a significant increase in verbal inflection accuracy; at the same time, it is computationally inexpensive and does not rely on target-language linguistic tools.

Continuous Adaptation to User Feedback for Statistical Machine Translation*Frédéric Blain, Fethi Bougares, Amir Hazem, Loïc Barrault, and Holger Schwenk* 11:30–11:45

This paper gives a detailed experiment feedback of different approaches to adapt a statistical machine translation system towards a targeted translation project, using only small amounts of parallel in-domain data. The experiments were performed by professional translators under realistic conditions of work using a computer assisted translation tool. We analyze the influence of these adaptations on the translator productivity and on the overall post-editing effort. We show that significant improvements can be obtained by using the presented adaptation techniques.

Normalized Word Embedding and Orthogonal Transform for Bilingual Word Translation*Chao Xing, Dong Wang, Chao Liu, and Yiye Lin* 11:45–12:00

Word embedding has been found to be highly powerful to translate words from one language to another by a simple linear transform. However, we found some inconsistency among the objective functions of the embedding and the transform learning, as well as the distance measuring. This paper proposes a solution which normalizes the word vectors on a hypersphere and constrains the linear transform as an orthogonal transform. The experimental results confirmed that the proposed solution can offer better performance on a word similarity task and an English-to-Spanish word translation task.

Fast and Accurate Preordering for SMT using Neural Networks*Adrià de Gispert, Gonzalo Iglesias, and Bill Byrne* 12:00–12:15

We propose the use of neural networks to model source-side preordering for faster and better statistical machine translation. The neural network trains a logistic regression model to predict whether two sibling nodes of the source-side parse tree should be swapped in order to obtain a more monotonic parallel corpus, based on samples extracted from the word-aligned parallel corpus. For multiple language pairs and domains, we show that this yields the best reordering performance against other state-of-the-art techniques, resulting in improved translation quality and very fast decoding.

APRO: All-Pairs Ranking Optimization for MT Tuning*Markus Dreyer and Yuanzhe Dong* 12:15–12:30

We present APRO, a new method for machine translation tuning that can handle large feature sets. As opposed to other popular methods (e.g., MERT, MIRA, PRO), which involve randomness and require multiple runs to obtain a reliable result, APRO gives the same result on any run, given initial feature weights. APRO follows the pairwise ranking approach of PRO (Hopkins and May, 2011), but instead of ranking a small sampled subset of pairs from the k-best list, APRO efficiently ranks all pairs. By obviating the need for manually determined sampling settings, we obtain more reliable results. APRO converges more quickly than PRO and gives similar or better translation results.

Session 5C: Morphology, Syntax, Multilinguality, and Applications (Short Papers)

Plaza Ballroom F

Chair: William Lewis

Paradigm classification in supervised learning of morphology

Malin Ahlberg, Markus Forsberg, and Mans Hulden

11:15–11:30

Supervised morphological paradigm learning by identifying and aligning the longest common subsequence found in inflection tables has recently been proposed as a simple yet competitive way to induce morphological patterns. We combine this non-probabilistic strategy of inflection table generalization with a discriminative classifier to permit the reconstruction of complete inflection tables of unseen words. Our system learns morphological paradigms from labeled examples of inflection patterns (inflection tables) and then produces inflection tables from unseen lemmas or base forms. We evaluate the approach on datasets covering 11 different languages and show that this approach results in consistently higher accuracies vis-à-vis other methods on the same task, thus indicating that the general method is a viable approach to quickly creating high-accuracy morphological resources.

Shift-Reduce Constituency Parsing with Dynamic Programming and POS Tag Lattice

Haitao Mi and Liang Huang

11:30–11:45

We present the first dynamic programming (DP) algorithm for shift-reduce constituency parsing, which extends the DP idea of Huang and Sagae (2010) to context-free grammars. To alleviate the propagation of errors from part-of-speech tagging, we also extend the parser to take a tag lattice instead of a fixed tag sequence. Experiments on both English and Chinese treebanks show that our DP parser significantly improves parsing quality over non-DP baselines, and achieves the best accuracies among empirical linear-time parsers.

Unsupervised Code-Switching for Multilingual Historical Document Transcription

Dan Garrette, Hannah Alpert-Abrams, Taylor Berg-Kirkpatrick, and Dan Klein 11:45–12:00

Transcribing documents from the printing press era, a challenge in its own right, is more complicated when documents interleave multiple languages—a common feature of 16th century texts. Additionally, many of these documents predate consistent orthographic conventions, making the task even harder. We extend the state-of-the-art historical OCR model of Berg-Kirkpatrick et al. (2013) to handle word-level code-switching between multiple languages. Further, we enable our system to handle spelling variability, including now-obsolete shorthand systems used by printers. Our results show average relative character error reductions of 14% across a variety of historical texts.

Matching Citation Text and Cited Spans in Biomedical Literature: a Search-Oriented Approach

Arman Cohan, Luca Soldaini, and Nazli Goharian

12:00–12:15

Citation sentences (citances) to a reference article have been extensively studied for summarization tasks. However, citances might not accurately represent the content of the cited article, as they often fail to capture the context of the reported findings and can be affected by epistemic value drift. Following the intuition behind the TAC (Text Analysis Conference) 2014 Biomedical Summarization track, we propose a system that identifies text spans in the reference article that are related to a given citance. We refer to this problem as citance-reference spans matching. We approach the problem as a retrieval task; in this paper, we detail a comparison of different citance reformulation methods and their combinations. While our results show improvement over the baseline (up to 25.9%), their absolute magnitude implies that there is ample room for future improvement.

Effective Feature Integration for Automated Short Answer Scoring

Kéisuke Sakaguchi, Michael Heilman, and Nitin Madnani

12:15–12:30

A major opportunity for NLP to have a real-world impact is in helping educators score student writing, particularly content-based writing (i.e., the task of automated short answer scoring). A major challenge in this enterprise is that scored responses to a particular question (i.e., labeled data) are valuable for modeling but limited in quantity. Additional information from the scoring guidelines for humans, such as exemplars for each score level and descriptions of key concepts, can also be used. Here, we explore methods for integrating scoring guidelines and labeled responses, and we find that stacked generalization (Wolpert, 1992) improves performance, especially for small training sets.

Session 6 Overview – Tuesday, June 2, 2015

Track A	Track B	Track C
<i>Generation and Summarization (Long Papers)</i> Plaza Ballroom A & B	<i>Discourse and Coreference (Long Papers)</i> Plaza Ballroom D & E	<i>Information Extraction and Question Answering (Long Papers)</i> Plaza Ballroom F
Socially-Informed Timeline Generation for Complex Events <i>Wang, Cardie, and Marchetti</i>	Encoding World Knowledge in the Evaluation of Local Coherence <i>Zhang, Feng, Qin, Hirst, Liu, and Huang</i>	Injecting Logical Background Knowledge into Embeddings for Relation Extraction <i>Rocktäschel, Singh, and Riedel</i>
Movie Script Summarization as Graph-based Scene Extraction <i>Gorinski and Lapata</i>	Chinese Event Coreference Resolution: An Unsupervised Probabilistic Model Rivaling Supervised Resolvers <i>Chen and Ng</i>	Unsupervised Entity Linking with Abstract Meaning Representation <i>Pan, Cassidy, Hermjakob, Ji, and Knight</i>
Toward Abstractive Summarization Using Semantic Representations <i>Liu, Flanigan, Thomson, Sadeh, and Smith</i>	Removing the Training Wheels: A Coreference Dataset that Entertains Humans and Challenges Computers <i>Guha, Iyyer, Bouman, and Boyd-Graber</i>	Idest: Learning a Distributed Representation for Event Patterns <i>Krause, Alfonseca, Filippova, and Pighin</i>

2:25

2:50

3:15

Parallel Session 6

Session 6A: Generation and Summarization (Long Papers)

Plaza Ballroom A & B

Chair: Wei Xu

Socially-Informed Timeline Generation for Complex Events

Lu Wang, Claire Cardie, and Galen Marchetti

14:00–14:25

Existing timeline generation systems for complex events consider only information from traditional media, ignoring the rich social context provided by user-generated content that reveals representative public interests or insightful opinions. We instead aim to generate socially-informed timelines that contain both news article summaries and selected user comments. We present an optimization framework designed to balance topical cohesion between the article and comment summaries along with their informativeness and coverage of the event. Automatic evaluations on real-world datasets that cover four complex events show that our system produces more informative timelines than state-of-the-art systems. In human evaluation, the associated comment summaries are furthermore rated more insightful than editor's picks and comments ranked highly by users.

Movie Script Summarization as Graph-based Scene Extraction

Philip John Gorinski and Mirella Lapata

14:25–14:50

In this paper we study the task of movie script summarization, which we argue could enhance script browsing, give readers a rough idea of the script's plotline, and speed up reading time. We formalize the process of generating a shorter version of a screenplay as the task of finding an optimal chain of scenes. We develop a graph-based model that selects a chain by jointly optimizing its logical progression, diversity, and importance. Human evaluation based on a question-answering task shows that our model produces summaries which are more informative compared to competitive baselines.

Toward Abstractive Summarization Using Semantic Representations

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

14:50–15:15

We present a novel abstractive summarization framework that draws on the recent development of a treebank for the Abstract Meaning Representation (AMR). In this framework, the source text is parsed to a set of AMR graphs, the graphs are transformed into a summary graph, and then text is generated from the summary graph. We focus on the graph-to-graph transformation that reduces the source semantic graph into a summary graph, making use of an existing AMR parser and assuming the eventual availability of an AMR-to-text generator. The framework is data-driven, trainable, and not specifically designed for a particular domain. Experiments on gold-standard AMR annotations and system parses show promising results. Code is available at: <https://github.com/summarization>

Session 6B: Discourse and Coreference (Long Papers)

Plaza Ballroom D & E

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

Previous work on text coherence was primarily based on matching multiple mentions of the same entity in different parts of the text; therefore, it misses the contribution from semantically related but not necessarily coreferential entities (e.g., Gates and Microsoft). In this paper, we capture such semantic relatedness by leveraging world knowledge (e.g., Gates is the person who created Microsoft), and use two existing evaluation frameworks. First, in the unsupervised framework, we introduce semantic relatedness as an enrichment to the original graph-based model of Guinaudeau and Strube (2013). In addition, we incorporate semantic relatedness as additional features into the popular entity-based model of Barzilay and Lapata (2008). Across both frameworks, our enriched model with semantic relatedness outperforms the original methods, especially on short documents.

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

14:25–14:50

Recent work has successfully leveraged the semantic information extracted from lexical knowledge bases such as WordNet and FrameNet to improve English event coreference resolvers. The lack of comparable resources in other languages, however, has made the design of high-performance non-English event coreference resolvers, particularly those employing unsupervised models, very difficult. We propose a generative model for the under-studied task of Chinese event coreference resolution that rivals its supervised counterparts in performance when evaluated on the ACE 2005 corpus.

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

14:50–15:15

Coreference is a core NLP problem. However, newswire data, the primary source of existing coreference data, lack the richness necessary to truly solve coreference. We present a new domain with denser references—quiz bowl questions—that is challenging and enjoyable to humans, and we use the quiz bowl community to develop a new coreference dataset, together with an annotation framework that can tag any text data with coreferences and named entities. We also successfully integrate active learning into this annotation pipeline to collect documents maximally useful to coreference models. State-of-the-art coreference systems underperform a simple classifier on our new dataset, motivating non-newswire data for future coreference research.

Session 6C: Information Extraction and Question Answering (Long Papers)

Plaza Ballroom F

Chair: Lucy Vanderwende

Injecting Logical Background Knowledge into Embeddings for Relation Extraction

Tim Rocktäschel, Sameer Singh, and Sebastian Riedel

14:00–14:25

Matrix factorization approaches to relation extraction provide several attractive features: they support distant supervision, handle open schemas, and leverage unlabeled data. Unfortunately, these methods share a shortcoming with all other distantly supervised approaches: they cannot learn to extract target relations without existing data in the knowledge base, and likewise, these models are inaccurate for relations with sparse data. Rule-based extractors, on the other hand, can be easily extended to novel relations and improved for existing but inaccurate relations, through first-order formulae that capture auxiliary domain knowledge. However, usually a large set of such formulae is necessary to achieve generalization. In this paper, we introduce a paradigm for learning low-dimensional embeddings of entity-pairs and relations that combine the advantages of matrix factorization with first-order logic domain knowledge. We introduce simple approaches for estimating such embeddings, as well as a novel training algorithm to jointly optimize over factual and first-order logic information. Our results show that this method is able to learn accurate extractors with little or no distant supervision alignments, while at the same time generalizing to textual patterns that do not appear in the formulae.

Unsupervised Entity Linking with Abstract Meaning Representation

Xiaoman Pan, Taylor Cassidy, Ulf Hermjakob, Heng Ji, and Kevin Knight

14:25–14:50

Most successful Entity Linking (EL) methods aim to link mentions to their referent entities in a structured Knowledge Base (KB) by comparing their respective contexts, often using similarity measures. While the KB structure is given, current methods have suffered from impoverished information representations on the mention side. In this paper, we demonstrate the effectiveness of Abstract Meaning Representation (AMR) (Banarescu et al., 2013) to select high quality sets of entity “collaborators” to feed a simple similarity measure (Jaccard) to link entity mentions. Experimental results show that AMR captures contextual properties discriminative enough to make linking decisions, without the need for EL training data, and that system with AMR parsing output outperforms hand labeled traditional semantic roles as context representation for EL. Finally, we show promising preliminary results for using AMR to select sets of “coherent” entity mentions for collective entity linking.

Idest: Learning a Distributed Representation for Event Patterns

Sebastian Krause, Enrique Alfonseca, Katja Filippova, and Daniele Pighin

14:50–15:15

This paper describes Idest, a new method for learning paraphrases of event patterns. It is based on a new neural network architecture that only relies on the weak supervision signal that comes from the news published on the same day and mention the same real-world entities. It can generalize across extractions from different dates to produce a robust paraphrase model for event patterns that can also capture meaningful representations for rare patterns. We compare it with two state-of-the-art systems and show that it can attain comparable quality when trained on a small dataset. Its generalization capabilities also allow it to leverage much more data, leading to substantial quality improvements.

Session 7 Overview – Tuesday, June 2, 2015

Track A	Track B	Track C	
<i>Semantics (Long + TACL Papers)</i> Plaza Ballroom A & B	<i>Information Extraction and Question Answering (Long + TACL Papers)</i> Plaza Ballroom D & E	<i>Machine Translation (Long Papers)</i> Plaza Ballroom F	
High-Order Low-Rank Tensors for Semantic Role Labeling <i>Lei, Zhang, Márquez, Moschitti, and Barzilay</i>	Lexical Event Ordering with an Edge-Factored Model <i>Abend, Cohen, and Steedman</i>	Bag-of-Words Forced Decoding for Cross-Lingual Information Retrieval <i>Hieber and Riezler</i>	4:10
[TACL] Large-scale Semantic Parsing without Question-Answer Pairs <i>Reddy, Lapata, and Steedman</i>	[TACL] Entity disambiguation with web links <i>Chisholm and Hachey</i>	Accurate Evaluation of Segment-level Machine Translation Metrics <i>Graham, Baldwin, and Mathur</i>	4:35
[TACL] A Large Scale Evaluation of Distributional Semantic Models: Parameters, Interactions and Model Selection <i>Lapesa and Evert</i>	[TACL] A Joint Model for Entity Analysis: Coreference, Typing, and Linking <i>Durrett and Klein</i>	Leveraging Small Multilingual Corpora for SMT Using Many Pivot Languages <i>Dabre, Cromieres, Kurohashi, and Bhattacharyya</i>	5:00

Parallel Session 7

Session 7A: Semantics (Long + TACL Papers)

Plaza Ballroom A & B

Chair: Xiaodong He

High-Order Low-Rank Tensors for Semantic Role Labeling

Tao Lei, Yuan Zhang, Lluís Márquez, Alessandro Moschitti, and Regina Barzilay 15:45–16:10

This paper introduces a tensor-based approach to semantic role labeling (SRL). The motivation behind the approach is to automatically induce a compact feature representation for words and their relations, tailoring them to the task. In this sense, our dimensionality reduction method provides a clear alternative to the traditional feature engineering approach used in SRL. To capture meaningful interactions between the argument, predicate, their syntactic path and the corresponding role label, we compress each feature representation first to a lower dimensional space prior to assessing their interactions. This corresponds to using an overall cross-product feature representation and maintaining associated parameters as a four-way low-rank tensor. The tensor parameters are optimized for the SRL performance using standard online algorithms. Our tensor-based approach rivals the best performing system on the CoNLL-2009 shared task. In addition, we demonstrate that adding the representation tensor to a competitive tensor-free model yields 2% absolute increase in F-score.

[TACL] Large-scale Semantic Parsing without Question-Answer Pairs

Siva Reddy, Mirella Lapata, and Mark Steedman 16:10–16:35

In this paper we introduce a novel semantic parsing approach to query Freebase in natural language without requiring manual annotations or question-answer pairs. Our key insight is to represent natural language via semantic graphs whose topology shares many commonalities with Freebase. Given this representation, we conceptualize semantic parsing as a graph matching problem. Our model converts sentences to semantic graphs using CCG and subsequently grounds them to Freebase guided by denotations as a form of weak supervision. Evaluation experiments on a subset of the Free917 and WebQuestions benchmark datasets show our semantic parser improves over the state of the art.

[TACL] A Large Scale Evaluation of Distributional Semantic Models: Parameters, Interactions and Model Selection

Gabriella Lapesa and Stefan Evert 16:35–17:00

This paper presents the results of a large-scale evaluation study of window-based Distributional Semantic Models on a wide variety of tasks. Our study combines a broad coverage of model parameters with a model selection methodology that is robust to overfitting and able to capture parameter interactions. We show that our strategy allows us to identify parameter configurations that achieve good performance across different datasets and tasks.

Session 7B: Information Extraction and Question Answering (Long + TACL Papers)

Plaza Ballroom D & E

*Chair: Vincent Ng***Lexical Event Ordering with an Edge-Factored Model***Omri Abend, Shay B. Cohen, and Mark Steedman*

15:45–16:10

Extensive lexical knowledge is necessary for temporal analysis and planning tasks. We address in this paper a lexical setting that allows for the straightforward incorporation of rich features and structural constraints. We explore a lexical event ordering task, namely determining the likely temporal order of events based solely on the identity of their predicates and arguments. We propose an “edge-factored” model for the task that decomposes over the edges of the event graph. We learn it using the structured perceptron. As lexical tasks require large amounts of text, we do not attempt manual annotation and instead use the textual order of events in a domain where this order is aligned with their temporal order, namely cooking recipes.

[TACL] Entity disambiguation with web links*Andrew Chisholm and Ben Hachey*

16:10–16:35

Entity disambiguation with Wikipedia relies on structured information from redirect pages, article text, inter-article links, and categories. We explore whether web links can replace a curated encyclopaedia, obtaining entity prior, name, context, and coherence models from a corpus of web pages with links to Wikipedia. Experiments compare web link models to Wikipedia models on well-known CoNLL and TAC data sets. Results show that using 34 million web links approaches Wikipedia performance. Combining web link and Wikipedia models produces the best-known disambiguation accuracy of 88.7 on standard newswire test data.

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

16:35–17:00

We present a joint model of three core tasks in the entity analysis stack: coreference resolution (within-document clustering), named entity recognition (coarse semantic typing), and entity linking (matching to Wikipedia entities). Our model is formally a structured conditional random field. Unary factors encode local features from strong baselines for each task. We then add binary and ternary factors to capture cross-task interactions, such as the constraint that coreferent mentions have the same semantic type. On the ACE 2005 and OntoNotes datasets, we achieve state-of-the-art results for all three tasks. Moreover, joint modeling improves performance on each task over strong independent baselines.

Session 7C: Machine Translation (Long Papers)

Plaza Ballroom F

Chair: Hany Hassan

Bag-of-Words Forced Decoding for Cross-Lingual Information Retrieval

Felix Hieber and Stefan Riezler

15:45–16:10

Current approaches to cross-lingual information retrieval (CLIR) rely on standard retrieval models into which query translations by statistical machine translation (SMT) are integrated at varying degree. In this paper, we present an attempt to turn this situation on its head: Instead of the retrieval aspect, we emphasize the translation component in CLIR. We perform search by using an SMT decoder in forced decoding mode to produce a bag-of-words representation of the target documents to be ranked. The SMT model is extended by retrieval-specific features that are optimized jointly with standard translation features for a ranking objective. We find significant gains over the state-of-the-art in a large-scale evaluation on cross-lingual search in the domains patents and Wikipedia.

Accurate Evaluation of Segment-level Machine Translation Metrics

Yvette Graham, Timothy Baldwin, and Nitika Mathur

16:10–16:35

Evaluation of segment-level machine translation metrics is currently hampered by: (1) low inter-annotator agreement levels in human assessments; (2) lack of an effective mechanism for evaluation of translations of equal quality; and (3) lack of methods of significance testing improvements over a baseline. In this paper, we provide solutions to each of these challenges and outline a new human evaluation methodology aimed specifically at assessment of segment-level metrics. We replicate the human evaluation component of WMT-13 and reveal that the current state-of-the-art performance of segment-level metrics is better than previously believed. Three segment-level metrics — Meteor, nLepor and sentBLEU-moses — are found to correlate with human assessment at a level not significantly outperformed by any other metric in both the individual language pair assessment for Spanish to English and the aggregated set of 9 language pairs.

Leveraging Small Multilingual Corpora for SMT Using Many Pivot Languages

Raj Dabre, Fabien Cromieres, Sadao Kurohashi, and Pushpak Bhattacharyya

16:35–17:00

We present our work on leveraging multilingual parallel corpora of small sizes for Statistical Machine Translation between Japanese and Hindi using multiple pivot languages. In our setting, the source and target part of the corpus remains the same, but we show that using several different pivot to extract phrase pairs from these source and target parts lead to large BLEU improvements. We focus on a variety of ways to exploit phrase tables generated using multiple pivots to support a direct source-target phrase table. Our main method uses the Multiple Decoding Paths (MDP) feature of Moses, which we empirically verify as the best compared to the other methods we used. We compare and contrast our various results to show that one can overcome the limitations of small corpora by using as many pivot languages as possible in a multilingual setting. Most importantly, we show that such pivoting aids in learning of additional phrase pairs which are not learned when the direct source-target corpus is small. We obtained improvements of up to 3 BLEU points using multiple pivots for Japanese to Hindi translation compared to when only one pivot is used. To the best of our knowledge, this work is also the first of its kind to attempt the simultaneous utilization of 7 pivot languages at decoding time.

Demo Session A

Time: 5:00–6:30

Location: Plaza Ballroom A, B, & C

Two Practical Rhetorical Structure Theory Parsers

Mihai Surdeanu, Tom Hicks, and Marco Antonio Valenzuela-Escarcega

We describe the design, development, and API for two discourse parsers for Rhetorical Structure Theory. The two parsers use the same underlying framework, but one uses features that rely on dependency syntax, produced by a fast shift-reduce parser, whereas the other uses a richer feature space, including both constituent- and dependency-syntax and coreference information, produced by the Stanford CoreNLP toolkit. Both parsers obtain state-of-the-art performance, and use a very simple API consisting of, minimally, two lines of Scala code. We accompany this code with a visualization library that runs the two parsers in parallel, and displays the two generated discourse trees side by side, which provides an intuitive way of comparing the two parsers.

Analyzing and Visualizing Coreference Resolution Errors

Sebastian Martschat, Thierry Göckel, and Michael Strube

We present a toolkit for coreference resolution error analysis. It implements a recently proposed analysis framework and contains rich components for analyzing and visualizing recall and precision errors.

hyp: A Toolkit for Representing, Manipulating, and Optimizing Hypergraphs

Markus Dreyer and Jonathan Graehl

We present hyp, an open-source toolkit for the representation, manipulation, and optimization of weighted directed hypergraphs. hyp provides compose, project, invert functionality, k-best path algorithms, the inside and outside algorithms, and more. Finite-state machines are modeled as a special case of directed hypergraphs. hyp consists of a C++ API, as well as a command line tool, and is available for download at github.com/sdl-research/hyp.

Enhancing Instructor-Student and Student-Student Interactions with Mobile Interfaces and Summarization

Wencan Luo, Xiangmin Fan, Muhsin Menekse, Jingtao Wang, and Diane Litman

Educational research has demonstrated that asking students to respond to reflection prompts can increase interaction between instructors and students, which in turn can improve both teaching and learning especially in large classrooms. However, administering an instructor's prompts, collecting the students' responses, and summarizing these responses for both instructors and students is challenging and expensive. To address these challenges, we have developed an application called CourseMIRROR (Mobile In-situ Reflections and Review with Optimized Rubrics). CourseMIRROR uses a mobile interface to administer prompts and collect reflective responses for a set of instructor-assigned course lectures. After collection, CourseMIRROR automatically summarizes the reflections with an extractive phrase summarization method, using a clustering algorithm to rank extracted phrases by student coverage. Finally, CourseMIRROR presents the phrase summary to both instructors and students to help them understand the difficulties and misunderstandings encountered.

RExtractor: a Robust Information Extractor

Vincent Kríž and Barbora Hladka

The RExtractor system is an information extractor that processes input documents by natural language processing tools and consequently queries the parsed sentences to extract a knowledge base of entities and their relations. The extraction queries are designed manually using a tool that enables natural graphical representation of queries over dependency trees. A workflow of the system is designed to be language and domain independent. We demonstrate RExtractor on Czech and English legal documents.

An AMR parser for English, French, German, Spanish and Japanese and a new AMR-annotated corpus

Lucy Vanderwende, Arul Menezes, and Chris Quirk

In this demonstration, we will present our online parser that allows users to submit any sentence and obtain an analysis following the specification of AMR (Banarescu et al., 2014) to a large extent. This AMR analysis is generated by a small set of rules that convert a native Logical Form analysis provided by a pre-existing parser into the AMR format. While we demonstrate the performance of our AMR parser on data sets annotated by the LDC, we will focus attention in the demo on the following two areas: 1) we will make available AMR

annotations for the data sets that were used to develop our parser, to serve as a supplement to the LDC data sets, and 2) we will demonstrate AMR parsers for German, French, Spanish and Japanese that make use of the same small set of LF-to-AMR conversion rules.

ICE: Rapid Information Extraction Customization for NLP Novices

Yifan He and Ralph Grishman

We showcase ICE, an Integrated Customization Environment for Information Extraction. ICE is an easy tool for non-NLP experts to rapidly build customized IE systems for a new domain.

AMRICA: an AMR Inspector for Cross-language Alignments

Naomi Saphra and Adam Lopez

Abstract Meaning Representation (AMR), an annotation scheme for natural language semantics, has drawn attention for its simplicity and representational power. Because AMR annotations are not designed for human readability, we present AMRICA, a visual aid for exploration of AMR annotations. AMRICA can visualize an AMR or the difference between two AMRs to help users diagnose interannotator disagreement or errors from an AMR parser. AMRICA can also automatically align and visualize the AMRs of a sentence and its translation in a parallel text. We believe AMRICA will simplify and streamline exploratory research on cross-lingual AMR corpora.

Ckylark: A More Robust PCFG-LA Parser

Yusuke Oda, Graham Neubig, Sakriani Sakti, Tomoki Toda, and Satoshi Nakamura

This paper describes Ckylark, a PCFG-LA style phrase structure parser that is more robust than other parsers in the genre. PCFG-LA parsers are known to achieve highly competitive performance, but sometimes the parsing process fails completely, and no parses can be generated. Ckylark introduces three new techniques that prevent possible causes for parsing failure: outputting intermediate results when coarse-to-fine analysis fails, smoothing lexicon probabilities, and scaling probabilities to avoid underflow. An experiment shows that this allows millions of sentences can be parsed without any failures, in contrast to other publicly available PCFG-LA parsers. Ckylark is implemented in C++, and is available open-source under the LGPL license.

ELCO3: Entity Linking with Corpus Coherence Combining Open Source Annotators

Pablo Ruiz, Thierry Poibeau, and Frédérique Mélanie

Entity Linking (EL) systems' performance is uneven across corpora or depending on entity types. To help overcome this issue, we propose an EL workflow that combines the outputs of several open source EL systems, and selects annotations via weighted voting. The results are displayed on a UI that allows the users to navigate the corpus and to evaluate annotation quality based on several metrics.

SETS: Scalable and Efficient Tree Search in Dependency Graphs

Juhani Luotolahti, Jenna Kanerva, Sampo Pyysalo, and Filip Ginter

We present a syntactic analysis query toolkit geared specifically towards massive dependency parsebanks and morphologically rich languages. The query language allows arbitrary tree queries, including negated branches, and is suitable for querying analyses with rich morphological annotation. Treebanks of over a million words can be comfortably queried on a low-end netbook, and a parsebank with over 100M words on a single consumer-grade server. We also introduce a web-based interface for interactive querying. All contributions are available under open licenses.

Visualizing Deep-Syntactic Parser Output

Juan Soler-Company, Miguel Ballesteros, Bernd Bohnet, Simon Mille, and Leo Wanner

"Deep-syntactic" dependency structures bridge the gap between the surface-syntactic structures as produced by state-of-the-art dependency parsers and semantic logical forms in that they abstract away from surface-syntactic idiosyncrasies, but still keep the linguistic structure of a sentence. They have thus a great potential for such downstream applications as machine translation and summarization. In this demo paper, we propose an online version of a deep-syntactic parser that outputs deep-syntactic structures from plain sentences and visualizes them using the Brat tool. Along with the deep-syntactic structures, the user can also inspect the visual presentation of the surface-syntactic structures that serve as input to the deep-syntactic parser and that are produced by the joint tagger and syntactic transition-based parser ran in the pipeline before the deep-syntactic parser.

WOLFE: An NLP-friendly Declarative Machine Learning Stack*Sameer Singh, Tim Rocktäschel, Luke Hewitt, Jason Naradowsky, and Sebastian Riedel*

Developing machine learning algorithms for natural language processing (NLP) applications is inherently an iterative process, involving a continuous refinement of the choice of model, engineering of features, selection of inference algorithms, search for the right hyper-parameters, and error analysis. Existing probabilistic program languages (PPLs) only provide partial solutions; most of them do not support commonly used models such as matrix factorization or neural networks, and do not facilitate interactive and iterative programming that is crucial for rapid development of these models. In this demo we introduce WOLFE, a stack designed to facilitate the development of NLP applications: (1) the WOLFE language allows the user to concisely define complex models, enabling easy modification and extension, (2) the WOLFE interpreter transforms declarative machine learning code into automatically differentiable terms or, where applicable, into factor graphs that allow for complex models to be applied to real-world applications, and (3) the WOLFE IDE provides a number of different visual and interactive elements, allowing intuitive exploration and editing of the data representations, the underlying graphical models, and the execution of the inference algorithms.

Demo Session B

Time: 6:30–8:00

Location: Plaza Ballroom A, B, & C

Lean Question Answering over Freebase from Scratch*Xuchen Yao*

For the task of question answering (QA) over Freebase on the WebQuestions dataset (Berant et al., 2013), we found that 85% of all questions (in the training set) can be directly answered via a single binary relation. Thus we turned this task into slot-filling for \langle question topic, relation, answer \rangle tuples: predicting relations to get answers given a question’s topic. We design efficient data structures to identify question topics organically from 46 million Freebase topic names, without employing any NLP processing tools. Then we present a lean QA system that runs in real time (in offline batch testing it answered two thousand questions in 51 seconds on a laptop). The system also achieved 7.8% better F1 score (harmonic mean of average precision and recall) than the previous state of the art.

A Web Application for Automated Dialect Analysis*Sravana Reddy and James Stanford*

Sociolinguists are regularly faced with the task of measuring phonetic features from speech, which involves manually transcribing audio recordings – a major bottleneck to analyzing large collections of data. We harness automatic speech recognition to build an online end-to-end web application where users upload untranscribed speech collections and receive formant measurements of the vowels in their data. We demonstrate this tool by using it to automatically analyze President Barack Obama’s vowel pronunciations.

An Open-source Framework for Multi-level Semantic Similarity Measurement*Mohammad Taher Pilehvar and Roberto Navigli*

We present an open source, freely available Java implementation of Align, Disambiguate, and Walk (ADW), a state-of-the-art approach for measuring semantic similarity based on the Personalized PageRank algorithm. A pair of linguistic items, such as phrases or sentences, are first disambiguated using an alignment-based disambiguation technique and then modeled using random walks on the WordNet graph. ADW provides three main advantages: (1) it is applicable to all types of linguistic items, from word senses to texts; (2) it is all-in-one, i.e., it does not need any additional resource, training or tuning; and (3) it has proven to be highly reliable at different lexical levels and multiple evaluation benchmarks. We are releasing the source code at <https://github.com/pilehvar/adw/>. We also provide at <http://lcl.uniroma1.it/adw/> a Web interface and a Java API that can be seamlessly integrated into other NLP systems requiring semantic similarity measurement.

Brahmi-Net: A transliteration and script conversion system for languages of the Indian subcontinent*Anoop Kunchukuttan, Ratish Puduppully, and Pushpak Bhattacharyya*

We present Brahmi-Net - an online system for transliteration and script conversion for all major Indian language pairs (306 pairs). The system covers 13 Indo-Aryan languages, 4 Dravidian languages and English. For training the transliteration systems, we mined parallel transliteration corpora from parallel translation cor-

pora using an unsupervised method and trained statistical transliteration systems using the mined corpora. Languages which do not have a parallel corpora are supported by transliteration through a bridge language. Our script conversion system supports conversion between all Brahmi-derived scripts as well as ITRANS romanization scheme. For this, we leverage co-ordinated Unicode ranges between Indic scripts and use an extended ITRANS encoding for transliterating between English and Indic scripts. The system also provides top-k transliterations and simultaneous transliteration into multiple output languages. We provide a Python as well as REST API to access these services. The API and the mined transliteration corpus are made available for research use under an open source license.

A Concrete Chinese NLP Pipeline

Nanyun Peng, Francis Ferraro, Mo Yu, Nicholas Andrews, Jay DeYoung, Max Thomas, Matthew R. Gormley, Travis Wolfe, Craig Harman, Benjamin Van Durme, and Mark Dredze

Natural language processing research increasingly relies on the output of a variety of syntactic and semantic analytics. Yet integrating output from multiple analytics into a single framework can be time consuming and slow research progress. We present a Chinese Concrete NLP Pipeline: an NLP stack built using a series of open source systems integrated based on the Concrete data schema. Our pipeline includes data ingest, word segmentation, part of speech tagging, parsing, named entity recognition, relation extraction and cross document coreference resolution. Additionally, we integrate a tool for visualizing these annotations as well as allowing for the manual annotation of new data. We release our pipeline to the research community to facilitate work on Chinese language tasks that require rich linguistic annotations.

CroVeWA: Crosslingual Vector-Based Writing Assistance

Hubert Soyer, Goran Topić, Pontus Stenetorp, and Akiko Aizawa

We present an interactive web-based writing assistance system that is based on recent advances in crosslingual compositional distributed semantics. Given queries in Japanese or English, our system can retrieve semantically related sentences from high quality English corpora. By employing crosslingually constrained vector space models to represent phrases, our system naturally sidesteps several difficulties that would arise from direct word-to-text matching, and is able to provide novel functionality like the visualization of semantic relationships between phrases interlingually and intralingually.

Online Readability and Text Complexity Analysis with TextEvaluator

Diane Napolitano, Kathleen Sheehan, and Robert Mundkowsky

We have developed the TextEvaluator system for providing text complexity and Common Core-aligned readability information. Detailed text complexity information is provided by eight component scores, presented in such a way as to aid in the user's understanding of the overall readability metric, which is provided as a holistic score on a scale of 100 to 2000. The user may select a targeted US grade level and receive additional analysis relative to it. This and other capabilities are accessible via a feature-rich front-end, located at <http://texteval-pilot.ets.org/TextEvaluator/>.

Natural Language Question Answering and Analytics for Diverse and Interlinked Datasets

Dezhao Song, Frank Schilder, Charese Smiley, and Chris Brew

Previous systems for natural language questions over complex linked datasets require the user to enter a complete and well-formed question, and present the answers as raw lists of entities. Using a feature-based grammar with a full formal semantics, we have developed a system that is able to support rich autosuggest, and to deliver dynamically generated analytics for each result that it returns.

WriteAhead2: Mining Lexical Grammar Patterns for Assisted Writing

Jim Chang and Jason Chang

This paper describes WriteAhead2, an interactive writing environment that provides lexical and grammatical suggestions for second language learners, and helps them write fluently and avoid common writing errors. The method involves learning phrase templates from dictionary examples, and extracting grammar patterns with example phrases from an academic corpus. At run-time, as the user types word after word, the actions trigger a list after list of suggestions. Each successive list contains grammar patterns and examples, most relevant to the half-baked sentence. WriteAhead2 facilitates steady, timely, and spot-on interactions between learner writers and relevant information for effective assisted writing. Preliminary experiments show that WriteAhead2 has the potential to induce better writing and improve writing skills.

Question Answering System using Multiple Information Source and Open Type Answer Merge

Seonyeong Park, Soonchoul Kwon, Byungsoo Kim, Sangdo Han, Hyosup Shim, and Gary Geunbae Lee

This paper presents a multi-strategy and multi-source question answering (QA) system that can use multiple strategies to both answer natural language (NL) questions and respond to keywords. We use multiple information sources including curated knowledge base, raw text, auto-generated triples, and NL processing results. We develop open semantic answer type detector for answer merging and improve previous developed single QA modules such as knowledge base based QA, information retrieval based QA.

Using Word Semantics To Assist English as a Second Language Learners

Mahmoud Azab, Chris Hokamp, and Rada Mihalcea

We introduce an interactive interface that aims to help English as a Second Language (ESL) students overcome language related hindrances while reading a text. The interface allows the user to find supplementary information on selected difficult words. The interface is empowered by our lexical substitution engine that provides context-based synonyms for difficult words. We also provide a practical solution for a real-world usage scenario. We demonstrate using the lexical substitution engine – as a browser extension that can annotate and disambiguate difficult words on any webpage.

Poster session 2A (Short papers)

Time: 5:00–6:30

Location: Plaza Ballroom A, B, & C

Context-Dependent Automatic Response Generation Using Statistical Machine Translation Techniques

Andrew Shin, Ryohei Sasano, Hiroya Takamura, and Manabu Okumura

Developing a system that can automatically respond to a user's utterance has recently become a topic of research in natural language processing. However, most works on the topic take into account only a single preceding utterance to generate a response. Recent works demonstrate that the application of statistical machine translation (SMT) techniques towards monolingual dialogue setting, in which a response is treated as a translation of a stimulus, has a great potential, and we exploit the approach to tackle the context-dependent response generation task. We attempt to extract relevant and significant information from the wider contextual scope of the conversation, and incorporate it into the SMT techniques. We also discuss the advantages and limitations of this approach through our experimental results.

Distributed Representations of Words to Guide Bootstrapped Entity Classifiers

Sonal Gupta and Christopher D. Manning

Bootstrapped classifiers iteratively generalize from a few seed examples or prototypes to other examples of target labels. However, sparseness of language and limited supervision make the task difficult. We address this problem by using distributed vector representations of words to aid the generalization. We use the word vectors to expand entity sets used for training classifiers in a bootstrapped pattern-based entity extraction system. Our experiments show that the classifiers trained with the expanded sets perform better on entity extraction from four online forums, with 30% F1 improvement on one forum. The results suggest that distributed representations can provide good directions for generalization in a bootstrapping system.

Multilingual Open Relation Extraction Using Cross-lingual Projection

Manaal Faruqui and Shankar Kumar

Open domain relation extraction systems identify relation and argument phrases in a sentence without relying on any underlying schema. However, current state-of-the-art relation extraction systems are available only for English because of their heavy reliance on linguistic tools such as part-of-speech taggers and dependency parsers. We present a cross-lingual annotation projection method for language independent relation extraction. We evaluate our method on a manually annotated test set and present results on three typologically different languages. We release these manual annotations and extracted relations in ten languages from Wikipedia.

Exploiting Text and Network Context for Geolocation of Social Media Users

Afshin Rahimi, Duy Vu, Trevor Cohn, and Timothy Baldwin

Research on automatically geolocating social media users has conventionally been based on the text content of posts from a given user or the social network of the user, with very little crossover between the two, and no benchmarking of the two approaches over comparable datasets. We bring the two threads of research together in first proposing a text-based method based on adaptive grids, followed by a hybrid network- and text-based method. Evaluating over three Twitter datasets, we show that the empirical difference between text- and network-based methods is not great, and that hybridisation of the two is superior to the component methods, especially in contexts where the user graph is not well connected. We achieve state-of-the-art results on all three datasets.

Unsupervised Most Frequent Sense Detection using Word Embeddings

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

An acid test for any new Word Sense Disambiguation (WSD) algorithm is its performance against the Most Frequent Sense (MFS). The field of WSD has found the MFS baseline very hard to beat. Clearly, if WSD researchers had access to MFS values, their striving to better this heuristic will push the WSD frontier. However, getting MFS values requires sense marked corpus in enormous amount, which is out of bounds for most languages, even if their WordNets are available. In this paper, we propose an unsupervised method for MFS detection from the untagged corpora, which exploits word embeddings. *We compare the word embeddings of a word with all its sense embeddings and obtain the predominant sense with the highest similarity.* We observe significant performance gain for Hindi WSD over the WordNet First Sense (WFS) baseline. As for English, the SemCor baseline is bettered for those words whose frequency is greater than 2. Our approach is language and domain independent.

Chain Based RNN for Relation Classification

Javid Ebrahimi and Dejing Dou

We present a novel approach for relation classification, using a recursive neural network (RNN), based on the shortest path between two entities in a dependency graph. Previous works on RNN are based on constituency-based parsing because phrasal nodes in a parse tree can capture compositionality in a sentence. Compared with constituency-based parse trees, dependency graphs can represent relations more compactly. This is particularly important in sentences with distant entities, where the parse tree spans words that are not relevant to the relation. In such cases RNN cannot be trained effectively in a timely manner. However, due to the lack of phrasal nodes in dependency graphs, application of RNN is not straightforward. In order to tackle this problem, we utilize dependency constituent units called chains. Our experiments on two relation classification datasets show that Chain based RNN provides a shallower network, which performs considerably faster and achieves better classification results.

CASSA: A Context-Aware Synonym Simplification Algorithm

Ricardo Baeza-Yates, Luz Rello, and Julia Dembowski

We present a new context-aware method for lexical simplification that uses two free language resources and real web frequencies. We compare it with the state-of-the-art method for lexical simplification in Spanish and the established simplification baseline, that is, the most frequent synonym. Our method improves upon the other methods in the detection of complex words, in meaning preservation, and in simplicity. Although we use Spanish, the method can be extended to other languages since it does not require alignment of parallel corpora.

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

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

We present a simple, yet effective approach to adapt part-of-speech (POS) taggers to new domains. Our approach only requires a dictionary and large amounts of unlabeled target data. The idea is to use the dictionary to mine the unlabeled target data for unambiguous word sequences, thus effectively collecting labeled target data. We add the mined instances to available labeled newswire data to train a POS tagger for the target domain. The induced models significantly improve tagging accuracy on held-out test sets across three domains (Twitter, spoken language, and search queries). We also present results for Dutch, Spanish and Portuguese Twitter data, and provide two novel manually-annotated test sets.

On the Automatic Learning of Sentiment Lexicons

Aliaksei Severyn and Alessandro Moschitti

This paper describes a simple and principled approach to automatically construct sentiment lexicons using distant supervision. We induce the sentiment association scores for the lexicon items from a model trained on

a weakly supervised corpora. Our empirical findings show that features extracted from such a machine-learned lexicon outperform models using manual or other automatically constructed sentiment lexicons. Finally, our system achieves the state-of-the-art in Twitter Sentiment Analysis tasks from Semeval-2013 and ranks 2nd best in Semeval-2014 according to the average rank.

Development of the Multilingual Semantic Annotation System

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

This paper reports on our research to generate multilingual semantic lexical resources and develop multilingual semantic annotation software, which assigns each word in running text to a semantic category based on a lexical semantic classification scheme. Such tools have an important role in developing intelligent multilingual NLP, text mining and ICT systems. In this work, we aim to extend an existing English semantic annotation tool to cover a range of languages, namely Italian, Chinese and Brazilian Portuguese, by bootstrapping new semantic lexical resources via automatically translating existing English semantic lexicons into these languages. We used a set of bilingual dictionaries and word lists for this purpose. In our experiment, with minor manual improvement of the automatically generated semantic lexicons, the prototype tools based on the new lexicons achieved an average lexical coverage of 79.86% and an average annotation precision of 71.42% (if only precise annotations are considered) or 84.64% (if partially correct annotations are included) on the three languages. Our experiment demonstrates that it is feasible to rapidly develop prototype semantic annotation tools for new languages by automatically bootstrapping new semantic lexicons based on existing ones.

Unediting: Detecting Disfluencies Without Careful Transcripts

Victoria Zayats, Mari Ostendorf, and Hannaneh Hajishirzi

Speech transcripts often only capture semantic content, omitting disfluencies that can be useful for analyzing social dynamics of a discussion. This work describes steps in building a model that can recover a large fraction of locations where disfluencies were present, by transforming carefully annotated text to match the standard transcription style, introducing a two-stage model for handling different types of disfluencies, and applying semi-supervised learning. Experiments show improvement in disfluency detection on Supreme Court oral arguments, nearly 23% improvement in F1.

Type-Driven Incremental Semantic Parsing with Polymorphism

Kai Zhao and Liang Huang

Semantic parsing has made significant progress, but most current semantic parsers are extremely slow (CKY-based) and rather primitive in representation. We introduce three new techniques to tackle these problems. First, we design the first linear-time incremental shift-reduce-style semantic parsing algorithm which is more efficient than conventional cubic-time bottom-up semantic parsers. Second, our parser, being type-driven instead of syntax-driven, uses type-checking to decide the direction of reduction, which eliminates the need for a syntactic grammar such as CCG. Third, to fully exploit the power of type-driven semantic parsing beyond simple types (such as entities and truth values), we borrow from programming language theory the concepts of subtype polymorphism and parametric polymorphism to enrich the type system in order to better guide the parsing. Our system learns very accurate parses in GeoQuery, Jobs and Atis domains.

Template Kernels for Dependency Parsing

Hillel Taub-Tabib, Yoav Goldberg, and Amir Globerson

A common approach to dependency parsing is scoring a parse via a linear function of a set of indicator features. These features are typically manually constructed from templates that are applied to parts of the parse tree. The templates define which properties of a part should combine to create features. Existing approaches consider only a small subset of the possible combinations, due to statistical and computational efficiency considerations. In this work we present a novel kernel which facilitates efficient parsing with feature representations corresponding to a much larger set of combinations. We integrate the kernel into a parse reranking system and demonstrate its effectiveness on four languages from the CoNLL-X shared task.

Embedding a Semantic Network in a Word Space

Richard Johansson and Luis Nieto Piña

We present a framework for using continuous-space vector representations of word meaning to derive new vectors representing the meaning of senses listed in a semantic network. It is a post-processing approach that can be applied to several types of word vector representations. It uses two ideas: first, that vectors for polysemous words can be decomposed into a mix of sense vectors; secondly, that the vector for a sense is kept

similar to those of its neighbors in the network. This leads to a constrained optimization problem, and we present an approximation for the case when the distance function is the squared Euclidean. We applied this algorithm on a Swedish semantic network, and we evaluate the quality of the resulting sense representations extrinsically by showing that they give large improvements when used in a classifier that creates lexical units for FrameNet frames.

Identification and Characterization of Newsworthy Verbs in World News

Benjamin Nye and Ani Nenkova

We present a data-driven technique for acquiring domain-level importance of verbs from the analysis of abstract/article pairs of world news articles. We show that existing lexical resources capture some the semantic characteristics for important words in the domain. We develop a novel characterization of the association between verbs and personal story narratives, which is descriptive of verbs avoided in summaries for this domain.

Enhancing Sumerian Lemmatization by Unsupervised Named-Entity Recognition

Yudong Liu, Clinton Burkhart, James Hearne, and Liang Luo

Lemmatization for the Sumerian language, compared to the modern languages, is much more challenging due to that it is a long dead language, highly skilled language experts are extremely scarce and more and more Sumerian texts are coming out. This paper describes how our unsupervised Sumerian named-entity recognition (NER) system helps to improve the lemmatization of the Cuneiform Digital Library Initiative (CDLI), a specialist database of cuneiform texts, from the Ur III period. Experiments show that a promising improvement in personal name annotation in such texts and a substantial reduction in expert annotation effort can be achieved by leveraging our system with minimal seed annotation.

Improving Update Summarization via Supervised ILP and Sentence Reranking

Chen Li, Yang Liu, and Lin Zhao

Integer Linear Programming (ILP) based summarization methods have been widely adopted recently because of their state-of-the-art performance. This paper proposes two new modifications in this framework for update summarization. Our key idea is to use discriminative models with a set of features to measure both the salience and the novelty of words and sentences. First, these features are used in a supervised model to predict the weights of the concepts used in the ILP model. Second, we generate preliminary sentence candidates in the ILP model and then rerank them using sentence level features. We evaluate our method on different TAC update summarization data sets, and the results show that our system performs competitively compared to the best TAC systems based on the ROUGE evaluation metric.

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

David Jurgens and Mohammad Taher Pilehvar

This paper presents CROWN, an automatically constructed extension of WordNet that augments its taxonomy with novel lemmas from Wiktionary. CROWN fills the important gap in WordNet's lexicon for slang, technical, and rare lemmas, and more than doubles its current size. In two evaluations, we demonstrate that the construction procedure is accurate and has a significant impact on a WordNet-based algorithm encountering novel lemmas.

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

Zornitsa Kozareva

Online shopping caters the needs of millions of users on a daily basis. To build an accurate system that can retrieve relevant products for a query like "MB252 with travel bags" one requires product and query categorization mechanisms, which classify the text as "Home&Garden >Kitchen&Dining >Kitchen Appliances >Blenders". One of the biggest challenges in e-Commerce is that providers like Amazon, e-Bay, Google, Yahoo! and Walmart organize products into different product taxonomies making it hard and time-consuming for sellers to categorize goods for each shopping platform. To address this challenge, we propose an automatic product categorization mechanism, which for a given product title assigns the correct product category from a taxonomy. We conducted an empirical evaluation on 445,408 product titles and used a rich product taxonomy of 319 categories organized into 6 levels. We compared performance against multiple algorithms and found that the best performing system reaches .88 f-score.

Cross-lingual Text Classification Using Topic-Dependent Word Probabilities

Daniel Andrade, Kunihiko Sadamasa, Akihiro Tamura, and Masaaki Tsuchida

Cross-lingual text classification is a major challenge in natural language processing, since often training data is available in only one language (target language), but not available for the language of the document we want to classify (source language). Here, we propose a method that only requires a bilingual dictionary to bridge the language gap. Our proposed probabilistic model allows us to estimate translation probabilities that are conditioned on the whole source document. The assumption of our probabilistic model is that each document can be characterized by a distribution over topics that help to solve the translation ambiguity of single words. Using the derived translation probabilities, we then calculate the expected word frequency of each word type in the target language. Finally, these expected word frequencies can be used to classify the source text with any classifier that was trained using only target language documents. Our experiments confirm the usefulness of our proposed method.

Poster session 2B (Short papers)

Time: 6:30–8:00

Location: Plaza Ballroom A, B, & C

Why Read if You Can Scan? Trigger Scoping Strategy for Biographical Fact Extraction *Dian Yu, Heng Ji, Sujian Li, and Chin-Yew Lin*

The rapid growth of information sources brings a unique challenge to biographical information extraction: how to find specific facts without having to read all the words. An effective solution is to follow the human scanning strategy which keeps a specific keyword in mind and searches within a specific scope. In this paper, we mimic a scanning process to extract biographical facts: we use event and relation triggers as keywords, identify their scopes and apply type constraints to extract answers within the scope of a trigger. Experiments demonstrate that our approach outperforms state-of-the-art methods up to 26% absolute gain in F-score without using any syntactic analysis or external knowledge bases.

Response-based Learning for Machine Translation of Open-domain Database Queries *Carolin Haas and Stefan Riezler*

Response-based learning allows to adapt a statistical machine translation (SMT) system to an extrinsic task by extracting supervision signals from task-specific feedback. In this paper, we elicit response signals for SMT adaptation by executing semantic parses of translated queries against the Freebase database. The challenge of our work lies in scaling semantic parsers to the lexical diversity of open-domain databases. We find that parser performance on incorrect English sentences, which is standardly ignored in parser evaluation, is key in model selection. In our experiments, the biggest improvements in F1-score for returning the correct answer from a semantic parse for a translated query are achieved by selecting a parser that is carefully enhanced by paraphrases and synonyms.

Lachmannian Archetype Reconstruction for Ancient Manuscript Corpora *Armin Hoeneni*

Two goals are targeted by computer philology for ancient manuscript corpora: firstly, making an edition, that is roughly speaking one text version representing the whole corpus, which contains variety induced through copy errors and other processes and secondly, producing a stemma. A stemma is a graph-based visualization of the copy history with manuscripts as nodes and copy events as edges. Its root, the so-called archetype, is the supposed original text or urtext from which all subsequent copies are made. Our main contribution is to present one of the first computational approaches to automatic archetype reconstruction and to introduce the first text-based evaluation for automatically produced archetypes. We compare a philologically generated archetype with one generated by bio-informatic software.

Multi-Task Word Alignment Triangulation for Low-Resource Languages *Tomer Levinboim and David Chiang*

We present a multi-task learning approach that jointly trains three word alignment models over disjoint bitexts of three languages - source, target and pivot. Our approach builds upon model triangulation, following Wang et al., which approximates a source-target model by combining source-pivot and pivot-target models. We develop a MAP-EM algorithm that uses triangulation as a prior, and show how to extend it to a multi-task setting. On a low-resource Czech-English corpus, using French as the pivot, our multi-task learning approach more than doubles the gains in both F- and BLEU scores compared to the interpolation approach of Wang et al. Further

experiments reveal that the choice of pivot language does not significantly effect performance.

Learning to parse with IAA-weighted loss

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

Natural language processing (NLP) annotation projects employ guidelines to maximize inter-annotator agreement (IAA), and models are estimated assuming that there is one single ground truth. However, not all disagreement is noise, and in fact some of it may contain valuable linguistic information. We integrate such information in the training of a cost-sensitive dependency parser. We introduce five different factorizations of IAA and the corresponding loss functions, and evaluate these across six different languages. We obtain robust improvements across the board using a factorization that considers dependency labels and directionality. The best method-dataset combination reaches an average overall error reduction of 6.4% in labeled attachment score.

Automatic cognate identification with gap-weighted string subsequences.

Taraka Rama

In this paper, we describe the problem of cognate identification in NLP. We introduce the idea of gap-weighted subsequences for discriminating cognates from non-cognates. We also propose a scheme to integrate phonetic features into the feature vectors for cognate identification. We show that subsequence based features perform better than state-of-the-art classifier for the purpose of cognate identification. The contribution of this paper is the use of subsequence features for cognate identification.

Short Text Understanding by Leveraging Knowledge into Topic Model

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

In this paper, we investigate the challenging task of understanding short text (STU task) by jointly considering topic modeling and knowledge incorporation. Knowledge incorporation can solve the content sparsity problem effectively for topic modeling. Specifically, the phrase topic model is proposed to leverage the auto-mined knowledge, i.e., the phrases, to guide the generative process of short text. Experimental results illustrate the effectiveness of the mechanism that utilizes knowledge to improve topic modeling's performance.

Discriminative Phrase Embedding for Paraphrase Identification

Wenpeng Yin and Hinrich Schütze

This work, concerning paraphrase identification task, on one hand contributes to expanding deep learning embeddings to include continuous and discontinuous linguistic phrases. On the other hand, it comes up with a new scheme TF-KLD-KNN to learn the discriminative weights of words and phrases specific to paraphrase task, so that a weighted sum of embeddings can represent sentences more effectively. Based on these two innovations we get competitive state-of-the-art performance on paraphrase identification.

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

Mo Yu, Matthew R. Gormley, and Mark Dredze

Compositional embedding models build a representation for a linguistic structure based on its component word embeddings. While recent work has combined these word embeddings with hand crafted features for improved performance, they are limited to using a small number of features due to model complexity, thus limiting their applicability. We propose a new model that conjoins features and word embeddings while maintaining a small number of parameters by learning feature embeddings jointly with the parameters of a compositional model. The result is a method that can scale to more features and more labels, which avoiding over-fitting. We demonstrate that our model attains state-of-the-art results on ACE and ERE fine-grain relation extraction.

LR Parsing for LCFRS

Laura Kallmeyer and Wolfgang Maier

LR parsing is a popular parsing strategy for variants of Context-Free Grammar (CFG). It has also been used for mildly context-sensitive formalisms, such as Tree-Adjoining Grammar. In this paper, we present the first LR-style parsing algorithm for Linear Context-Free Rewriting Systems (LCFRS), a mildly context-sensitive extension of CFG which has received considerable attention in the last years.

Simple task-specific bilingual word embeddings

Stephan Gouws and Anders Søgaard

We introduce a simple wrapper method that uses off-the-shelf word embedding algorithms to learn task-specific bilingual word embeddings. We use a small dictionary to produce mixed context-target pairs that we use to train embedding models. The model has the advantage that it (a) is independent of the choice of embedding algorithm, (b) does not require parallel data, and (c) can be adapted to specific tasks by re-defining the equivalence classes. We show how our method outperforms off-the-shelf bilingual embeddings on the task of unsupervised cross-language part-of-speech (POS) tagging, as well as on the task of semi-supervised cross-language super sense (SuS) tagging.

Sampling Techniques for Streaming Cross Document Coreference Resolution

Luke Shrimpton, Victor Lavrenko, and Miles Osborne

We present the first truly streaming cross document coreference resolution (CDC) system. Processing infinite streams of mentions forces us to use a constant amount of memory and so we maintain a representative, fixed sized sample at all times. For the sample to be representative it should represent a large number of entities whilst taking into account both temporal recency and distant references. We introduce new sampling techniques that take into account a notion of streaming discourse (current mentions depend on previous mentions). Using the proposed sampling techniques we are able to get a CEAFe score within 5% of a non-streaming system while using only 30% of the memory.

Clustering Sentences with Density Peaks for Multi-document Summarization

Yang Zhang, Yunqing Xia, Yi Liu, and Wenmin Wang

Multi-document Summarization (MDS) is of great value to many real world applications. Many scoring models are proposed to select appropriate sentences from documents to form the summary, in which the clusterbased methods are popular. In this work, we propose a unified sentence scoring model which measures representativeness and diversity at the same time. Experimental results on DUC04 demonstrate that our MDS method outperforms the DUC04 best method and the existing cluster-based methods, and it yields close results compared to the state-of-the-art generic MDS methods. Advantages of the proposed MDS method are two-fold: (1) The density peaks clustering algorithm is adopted, which is effective and fast. (2) No external resources such as Wordnet and Wikipedia or complex language parsing algorithms is used, making reproduction and deployment very easy in real environment.

Large-Scale Native Language Identification with Cross-Corpus Evaluation

Shervin Malmasi and Mark Dras

We present a large-scale Native Language Identification (NLI) experiment on new data, with a focus on cross-corpus evaluation to identify corpus- and genre-independent language transfer features. We test a new corpus and show it is comparable to other NLI corpora and suitable for this task. Cross-corpus evaluation on two large corpora achieves good accuracy and evidences the existence of reliable language transfer features, but lower performance also suggests that NLI models are not completely portable across corpora. Finally, we present a brief case study of features distinguishing Japanese learners' English writing, demonstrating the presence of cross-corpus and cross-genre language transfer features that are highly applicable to SLA and ESL research.

Unsupervised Sparse Vector Densification for Short Text Similarity

Yangqiu Song and Dan Roth

Sparse representations of text such as bag-of-words models or extended explicit semantic analysis (ESA) representations are commonly used in many NLP applications. However, for short texts, the similarity between two such sparse vectors is not accurate due to the small term overlap. While there have been multiple proposals for dense representations of words, measuring similarity between short texts (sentences, snippets, paragraphs) requires combining these token level similarities. In this paper, we propose to combine ESA representations and word2vec representations as a way to generate denser representations and, consequently, a better similarity measure between short texts. We study three densification mechanisms that involve aligning sparse representation via many-to-many, many-to-one, and one-to-one mappings. We then show the effectiveness of these mechanisms on measuring similarity between short texts.

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

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

In September 2014, Twitter users unequivocally reacted to the Ray Rice assault scandal by unleashing personal stories of domestic abuse via the hashtags #WhyIStayed or #WhyILeft. We explore at a macro-level firsthand accounts of domestic abuse from a substantial, balanced corpus of tweeted instances designated with these tags.

To seek insights into the reasons victims give for staying in vs. leaving abusive relationships, we analyze the corpus using linguistically motivated methods. We also report on an annotation study for corpus assessment. We perform classification, contributing a classifier that discriminates between the two hashtags exceptionally well at 82% accuracy with a substantial error reduction over its baseline.

Morphological Word-Embeddings

Ryan Cotterell and Hinrich Schütze

Linguistic similarity is multi-faceted. For instance, two words may be similar with respect to semantics, syntax, or morphology inter alia. Continuous word-embeddings have been shown to capture most of these shades of similarity to some degree. This work considers guiding word-embeddings with morphologically annotated data, a form of semi-supervised learning, encouraging the vectors to encode a word's morphology, i.e., words close in the embedded space share morphological features. We extend the log-bilinear model to this end and show that indeed our learned embeddings achieve this, using German as a case study.

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 Inclezan

In this paper we present our work on recognizing high level social constructs such as Leadership and Status from textual conversation using an approach that makes use of the background knowledge about social hierarchy and integrates statistical methods and symbolic logic based methods. We use a stratified approach in which we first detect lower level language constructs such as politeness, command and agreement that help us to infer intermediate constructs such as deference, closeness and authority that are observed between the parties engaged in conversation. These intermediate constructs in turn are used to determine the social constructs Leadership and Status. We have implemented this system successfully in both English and Korean languages and achieved considerable accuracy.

Two/Too Simple Adaptations of Word2Vec for Syntax Problems

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

We present two simple modifications to the models in the popular Word2Vec tool, in order to generate embeddings more suited to tasks involving syntax. The main issue with the original models is the fact that they are insensitive to word order. While order independence is useful for inducing semantic representations, this leads to suboptimal results when they are used to solve syntax-based problems. We show improvements in part-of-speech tagging and dependency parsing using our proposed models.

Random Walks and Neural Network Language Models on Knowledge Bases

Josu Goikoetxea, Aitor Soroa, and Eneko Agirre

Random walks over large knowledge bases like WordNet have been successfully used in word similarity, relatedness and disambiguation tasks. Unfortunately, those algorithms are relatively slow for large repositories, with significant memory footprints. In this paper we present a novel algorithm which encodes the structure of a knowledge base in a continuous vector space, combining random walks and neural net language models in order to produce novel word representations. Evaluation in word relatedness and similarity datasets yields equal or better results than those of a random walk algorithm, using a dense representation (300 dimensions instead of 117K). Furthermore, the word representations are complementary to those of the random walk algorithm and to corpus-based continuous representations, improving the state-of-the-art in the similarity dataset. Our technique opens up exciting opportunities to combine distributional and knowledge-based word representations.

Estimating Numerical Attributes by Bringing Together Fragmentary Clues

Hiroya Takamura and Jun'ichi Tsuji

This work is an attempt to automatically obtain numerical attributes of physical objects. We propose to represent each physical object as a feature vector and represent sizes as linear functions of feature vectors. We train the function in the framework of the combined regression and ranking with many types of fragmentary clues including absolute clues (e.g., A is 30cm long) and relative clues (e.g., A is larger than B).

Unsupervised POS Induction with Word Embeddings

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

Unsupervised word embeddings have been shown to be valuable as features in supervised learning problems; however, their role in unsupervised problems has been less thoroughly explored. In this paper, we show

that embeddings can likewise add value to the problem of unsupervised POS induction. In two representative models of POS induction, we replace multinomial distributions over the vocabulary with multivariate Gaussian distributions over word embeddings and observe consistent improvements in eight languages. We also analyze the effect of various choices while inducing word embeddings on “downstream” POS induction results.

Extracting Information about Medication Use from Veterinary Discussions

Haibo Ding and Ellen Riloff

Our research aims to extract information about medication use from veterinary discussion forums. We introduce the task of categorizing information about medication use to determine whether a doctor has prescribed medication, changed protocols, observed effects, or stopped use of a medication. First, we create a medication detector for informal veterinary texts and show that features derived from the Web can be very powerful. Second, we create classifiers to categorize each medication mention with respect to six categories. We demonstrate that this task benefits from a rich linguistic feature set, domain-specific semantic features produced by a weakly supervised semantic tagger, and balanced self-training.

MPQA 3.0: An Entity/Event-Level Sentiment Corpus

Lingjia Deng and Janyce Wiebe

This paper presents an annotation scheme for adding entity and event target annotations to the MPQA corpus, a rich span-annotated opinion corpus. The new corpus promises to be a valuable new resource for developing systems for entity/event-level sentiment analysis. Such systems, in turn, would be valuable in NLP applications such as Automatic Question Answering. We introduce the idea of entity and event targets (eTargets), describe the annotation scheme, and present the results of an agreement study.

GPU-Friendly Local Regression for Voice Conversion

Taylor Berg-Kirkpatrick and Dan Klein

Voice conversion is the task of transforming a source speaker’s voice so that it sounds like a target speaker’s voice. We present a GPU-friendly local regression model for voice conversion that is capable of converting speech in real-time and achieves state-of-the-art accuracy on this task. Our model uses a new approximation for computing local regression coefficients that is explicitly designed to preserve memory locality. As a result, our inference procedure is amenable to efficient implementation on the GPU. Our approach is more than 10X faster than a highly optimized CPU-based implementation, and is able to convert speech 2.7X faster than real-time.

Social Event

Tuesday, June 2, 2015, 8pm–11pm

Grand Ballroom

The NAACL 2015 Social and Networking Event will be held immediately following the Tuesday Poster Session and dinner in the Grand Ballroom at the hotel. Here you will enjoy desserts, coffee and tea, and a cash bar. Bring your boots and hats and follow along with our dance instructors to learn the latest country line dancing, then practice your moves to the sounds of a local DJ (who will play other sounds if you tire of the country twang). Enjoy networking with colleagues and have a relaxing evening!

(We are trying a second-time experiment in holding a second poster dinner and open-to-all desserts, drinks, and dancing social event combination in place of the usual Banquet).

We hope to make your conference experience not only enlightening but also entertaining and enjoyable!

Main Conference: Wednesday, June 3

Overview

7:30 – 9:00	Registration and Breakfast	<i>Plaza Exhibit All</i>	
9:00 – 10:10	Invited Talk: "A Quest for Visual Intelligence in Computers" (Fei-fei Li) <i>Plaza Ballroom A, B, & C</i>		
10:10 – 10:40	Break Session 8	<i>Plaza Exhibit All</i>	
10:40 – 11:55	NLP for Web, Social Media and Social Sciences (Long + TACL Papers) <i>Plaza Ballroom A & B</i>	Language and Vision (Long + TACL Papers) <i>Plaza Ballroom D & E</i>	Machine Translation (Long + TACL Papers) <i>Plaza Ballroom F</i>
11:55 – 1:00	Lunch		
1:00 – 2:00	Business Meeting Session 9	<i>Plaza Ballroom A & B</i>	
2:00 – 3:15	Lexical Semantics and Sentiment Analysis (Long Papers) <i>Plaza Ballroom A & B</i>	NLP-enabled Technology (Long + TACL Papers) <i>Plaza Ballroom D & E</i>	Linguistic and Psycholinguistic Aspects of CL (Long Papers) <i>Plaza Ballroom F</i>
3:15 – 3:45	Break	<i>Plaza Exhibit All</i>	
3:45 – 5:15	Best Paper Plenary Session	<i>Plaza Ballroom A, B, & C</i>	
5:15 – 5:30	Best Paper Awards and Closing Remarks	<i>Plaza Ballroom A, B, & C</i>	

Keynote Address: Fei-Fei Li

“A Quest for Visual Intelligence in Computers”

Wednesday, June 3, 2015, 9:00–10:10am

Plaza Ballroom A, B, & C

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.

Session 8 Overview – Wednesday, June 3, 2015

Track A	Track B	Track C	
<i>NLP for Web, Social Media and Social Sciences (Long + TACL Papers)</i> Plaza Ballroom A & B	<i>Language and Vision (Long + TACL Papers)</i> Plaza Ballroom D & E	<i>Machine Translation (Long + TACL Papers)</i> Plaza Ballroom F	
Testing and Comparing Computational Approaches for Identifying the Language of Framing in Political News <i>Baumer, Elovic, Qin, Polletta, and Gay</i>	Translating Videos to Natural Language Using Deep Recurrent Neural Networks <i>Venugopalan, Xu, Donahue, Rohrbach, Mooney, and Saenko</i>	A Comparison of Update Strategies for Large-Scale Maximum Expected BLEU Training <i>Wuebker, Muehr, Lehnen, Peitz, and Ney</i>	11:05
[TACL] Extracting Lexically Divergent Paraphrases from Twitter <i>Xu, Ritter, Callison-Burch, Dolan, and Ji</i>	[TACL] A Bayesian Model of Grounded Color Semantics <i>McMahan and Stone</i>	[TACL] Gappy Pattern Matching on GPUs for On-Demand Extraction of Hierarchical Translation Grammars <i>He, Lin, and Lopez</i>	11:30
Echoes of Persuasion: The Effect of Euphony in Persuasive Communication <i>Guerini, Özbal, and Strapparava</i>	Learning to Interpret and Describe Abstract Scenes <i>Gilberto Mateos Ortiz, Wolff, and Lapata</i>	Learning Translation Models from Monolingual Continuous Representations <i>Zhao, Hassan, and Auli</i>	11:55

Parallel Session 8

Session 8A: NLP for Web, Social Media and Social Sciences (Long + TACL Papers)

Plaza Ballroom A & B

Chair: Philip Resnik

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

10:40–11:05

The subconscious influence of framing on perceptions of political issues is well-documented in political science and communication research. A related line of work suggests that drawing attention to framing may help reduce such framing effects by enabling frame reflection, critical examination of the framing underlying an issue. However, definite guidance on how to identify framing does not exist. This paper presents a technique for identifying frame-invoking language. The paper first describes a human subjects pilot study that explores how individuals identify framing and informs the design of our technique. The paper then describes our data collection and annotation approach. Results show that the best performing classifiers achieve performance comparable to that of human annotators, and they indicate which aspects of language most pertain to framing. Both technical and theoretical implications are discussed.

[TACL] Extracting Lexically Divergent Paraphrases from Twitter

Wei Xu, Alan Ritter, Chris Callison-Burch, William B. Dolan, and Yangfeng Ji

11:05–11:30

We present MultiP (Multi-instance Learning Paraphrase Model), a new model suited to identify paraphrases within the short messages on Twitter. We jointly model paraphrase relations between word and sentence pairs and assume only sentence-level annotations during learning. Using this principled latent variable model alone, we achieve the performance competitive with a state-of-the-art method which combines a latent space model with a feature-based supervised classifier. Our model also captures lexically divergent paraphrases that differ from yet complement previous methods; combining our model with previous work significantly outperforms the state-of-the-art. In addition, we present a novel annotation methodology that has allowed us to crowdsource a paraphrase corpus from Twitter. We make this new dataset available to the research community.

Echoes of Persuasion: The Effect of Euphony in Persuasive Communication

Marco Guerini, Gözde Özbal, and Carlo Strapparava

11:30–11:55

While the effect of various lexical, syntactic, semantic and stylistic features have been addressed in persuasive language from a computational point of view, the persuasive effect of phonetics has received little attention. By modeling a notion of euphony and analyzing four datasets comprising persuasive and non-persuasive sentences in different domains (political speeches, movie quotes, slogans and tweets), we explore the impact of sounds on different forms of persuasiveness. We conduct a series of analyses and prediction experiments within and across datasets. Our results highlight the positive role of phonetic devices on persuasion.

Session 8B: Language and Vision (Long + TACL Papers)

Plaza Ballroom D & E

*Chair: Luke Zettlemoyer***Translating Videos to Natural Language Using Deep Recurrent Neural Networks***Subhashini Venugopalan, Huijuan Xu, Jeff Donahue, Marcus Rohrbach, Raymond Mooney, and Kate Saenko*

10:40–11:05

Solving the visual symbol grounding problem has long been a goal of artificial intelligence. The field appears to be advancing closer to this goal with recent breakthroughs in deep learning for natural language grounding in static images. In this paper, we propose to translate videos directly to sentences using a unified deep neural network with both convolutional and recurrent structure. Described video datasets are scarce, and most existing methods have been applied to toy domains with a small vocabulary of possible words. By transferring knowledge from 1.2M+ images with category labels and 100,000+ images with captions, our method is able to create sentence descriptions of open-domain videos with large vocabularies. We compare our approach with recent work using language generation metrics, subject, verb, and object prediction accuracy, and a human evaluation.

[TACL] A Bayesian Model of Grounded Color Semantics*Brian McMahan and Matthew Stone*

11:05–11:30

Natural language meanings allow speakers to encode important real-world distinctions, but corpora of grounded language use also reveal that speakers categorize the world in different ways and describe situations with different terminology. To learn meanings from data, we therefore need to link underlying representations of meaning to models of speaker judgment and speaker choice. This paper describes a new approach to this problem: we model variability through uncertainty in categorization boundaries and distributions over preferred vocabulary. We apply the approach to a large data set of color descriptions, where statistical evaluation documents its accuracy. The results are available as a Lexicon of Uncertain Color Standards (LUX), which supports future efforts in grounded language understanding and generation by probabilistically mapping 829 English color descriptions to potentially context-sensitive regions in HSV color space.

Learning to Interpret and Describe Abstract Scenes*Luis Gilberto Mateos Ortiz, Clemens Wolff, and Mirella Lapata*

11:30–11:55

Given a (static) scene, a human can effortlessly describe what is going on (who is doing what to whom, how, and why). The process requires knowledge about the world, how it is perceived, and described. In this paper we study the problem of interpreting and verbalizing visual information using abstract scenes created from collections of clip art images. We propose a model inspired by machine translation operating over a large parallel corpus of visual relations and linguistic descriptions. We demonstrate that this approach produces human-like scene descriptions which are both fluent and relevant, outperforming a number of competitive alternatives based on templates and sentence-based retrieval.

Session 8C: Machine Translation (Long + TACL Papers)

Plaza Ballroom F

Chair: Haitao Mi

A Comparison of Update Strategies for Large-Scale Maximum Expected BLEU Training *Joern Wuebker, Sebastian Muehr, Patrick Lehen, Stephan Peitz, and Hermann Ney* 10:40–11:05

This work presents a flexible and efficient discriminative training approach for statistical machine translation. We propose to use the RPROP algorithm for optimizing a maximum expected BLEU objective and experimentally compare it to several other updating schemes. It proves to be more efficient and effective than the previously proposed growth transformation technique and also yields better results than stochastic gradient descent and AdaGrad. We also report strong empirical results on two large scale tasks, namely BOLT Chinese->English and WMT German->English, where our final systems outperform results reported by Setiawan and Zhou (2013) and on matrix.statmt.org. On the WMT task, discriminative training is performed on the full training data of 4M sentence pairs, which is unsurpassed in the literature.

[TACL] Gappy Pattern Matching on GPUs for On-Demand Extraction of Hierarchical Translation Grammars

Hua He, Jimmy Lin, and Adam Lopez

11:05–11:30

Grammars for machine translation can be materialized on demand by finding source phrases in an indexed parallel corpus and extracting their translations. This approach is limited in practical applications by the computational expense of online lookup and extraction. For phrase-based models, recent work has shown that on-demand grammar extraction can be greatly accelerated by parallelization on general purpose graphics processing units (GPUs), but these algorithms do not work for hierarchical models, which require matching patterns that contain gaps. We address this limitation by presenting a novel GPU algorithm for on-demand hierarchical grammar extraction that is at least an order of magnitude faster than a comparable CPU algorithm when processing large batches of sentences. In terms of end-to-end translation, with decoding on the CPU, we increase throughput by roughly two thirds on a standard MT evaluation dataset. The GPU necessary to achieve these improvements increases the cost of a server by about a third. We believe that GPU-based extraction of hierarchical grammars is an attractive proposition, particularly for MT applications that demand high throughput.

Learning Translation Models from Monolingual Continuous Representations

Kai Zhao, Hany Hassan, and Michael Auli

11:30–11:55

Translation models often fail to generate good translations for infrequent words or phrases. Previous work attacked this problem by inducing new translation rules from monolingual data with a semi-supervised algorithm. However, this approach does not scale very well since it is very computationally expensive to generate new translation rules for only a few thousand sentences. We propose a much faster and simpler method that directly hallucinates translation rules for infrequent phrases based on phrases with similar continuous representations for which a translation is known. To speed up the retrieval of similar phrases, we investigate approximated nearest neighbor search with redundant bit vectors which we find to be three times faster and significantly more accurate than locality sensitive hashing. Our approach of learning new translation rules improves a phrase-based baseline by up to 1.6 BLEU on Arabic-English translation, it is three-orders of magnitudes faster than existing semi-supervised methods and 0.5 BLEU more accurate.

NAACL Business Meeting

Date: Wednesday, June 3, 2015

Time: 1:00–2:00 PM

Venue: Plaza Ballroom A & B

Chair: Hal Daumé III

All attendees are encouraged to participate in the business meeting.

Session 9 Overview – Wednesday, June 3, 2015

	Track A	Track B	Track C
	<i>Lexical Semantics and Sentiment Analysis (Long Papers)</i> Plaza Ballroom A & B	<i>NLP-enabled Technology (Long + TACL Papers)</i> Plaza Ballroom D & E	<i>Linguistic and Psycholinguistic Aspects of CL (Long Papers)</i> Plaza Ballroom F
2:25	A Corpus and Model Integrating Multiword Expressions and Supersenses <i>Schneider and Smith</i>	How to Memorize a Random 60-Bit String <i>Ghazvininejad and Knight</i>	A Bayesian Model for Joint Learning of Categories and their Features <i>Ferermann and Lapata</i>
2:50	Good News or Bad News: Using Affect Control Theory to Analyze Readers' Reaction Towards News Articles <i>Alhothali and Hoey</i>	[TACL] Building a State-of-the-Art Grammatical Error Correction System <i>Rozovskaya and Roth</i>	Shared common ground influences information density in microblog texts <i>Doyle and Frank</i>
3:15	Do We Really Need Lexical Information? Towards a Top-down Approach to Sentiment Analysis of Product Reviews <i>Otmakhova and Shin</i>	[TACL] Predicting the Difficulty of Language Proficiency Tests <i>Beinborn, Zesch, and Gurevych</i>	Hierarchic syntax improves reading time prediction <i>Schijndel and Schuler</i>

Parallel Session 9

Session 9A: Lexical Semantics and Sentiment Analysis (Long Papers)

Plaza Ballroom A & B

Chair: Saif Mohammed

A Corpus and Model Integrating Multiword Expressions and Supersenses

Nathan Schneider and Noah A. Smith

14:00–14:25

This paper introduces a task of identifying and semantically classifying lexical expressions in context. We investigate the online reviews genre, adding semantic supersense annotations to a 55,000 word English corpus that was previously annotated for multiword expressions. The noun and verb supersenses apply to full lexical expressions, whether single- or multiword. We then present a sequence tagging model that jointly infers lexical expressions and their supersenses. Results show that even with our relatively small training corpus in a noisy domain, the joint task can be performed to attain 70% class labeling F1.

Good News or Bad News: Using Affect Control Theory to Analyze Readers' Reaction Towards News Articles

Areej Alhothali and Jesse Hoey

14:25–14:50

This paper proposes a novel approach to sentiment analysis that leverages work in sociology on symbolic interactionism. The proposed approach uses Affect Control Theory (ACT) to analyze readers' sentiment towards factual (objective) content and towards its entities (subject and object). ACT is a theory of affective reasoning that uses empirically derived equations to predict the sentiments and emotions that arise from events. This theory relies on several large lexicons of words with affective ratings in a three-dimensional space of evaluation, potency, and activity (EPA). The equations and lexicons of ACT were evaluated on a newly collected news-headlines corpus. ACT lexicon was expanded using a label propagation algorithm, resulting in 86,604 new words. The predicted emotions for each news headline was then computed using the augmented lexicon and ACT equations. The results had a precision of 82%, 79%, and 68% towards the event, the subject, and the object, respectively. These results are significantly higher than those of standard sentiment analysis techniques.

Do We Really Need Lexical Information? Towards a Top-down Approach to Sentiment Analysis of Product Reviews

Yulia Otmakhova and Hyopil Shin

14:50–15:15

Most of the current approaches to sentiment analysis of product reviews are dependent on lexical sentiment information and proceed in a bottom-up way, adding new layers of features to lexical data. In this paper, we maintain that a typical product review is not a bag of sentiments, but a narrative with an underlying structure and reoccurring patterns, which allows us to predict its sentiments knowing only its general polarity and discourse cues that occur in it. We hypothesize that knowing only the review's score and its discourse patterns would allow us to accurately predict the sentiments of its individual sentences. The experiments we conducted prove this hypothesis and show a substantial improvement over the lexical baseline.

Session 9B: NLP-enabled Technology (Long + TACL Papers)

Plaza Ballroom D & E

Chair: Brendan O'Connor

How to Memorize a Random 60-Bit String

Marjan Ghazvininejad and Kevin Knight

14:00–14:25

User-generated passwords tend to be memorable, but not secure. A random, computer-generated 60-bit string is much more secure. However, users cannot memorize random 60-bit strings. In this paper, we investigate methods for converting arbitrary bit strings into English word sequences (both prose and poetry), and we study their memorability and other properties.

[TACL] Building a State-of-the-Art Grammatical Error Correction System

Alla Rozovskaya and Dan Roth

14:25–14:50

This paper identifies and examines the key principles underlying building a state-of-the-art grammatical error correction system. We do this by analyzing the Illinois system that placed first among seventeen teams in the recent CoNLL-2013 shared task on grammatical error correction. The system focuses on five different types of errors common among non-native English writers. We describe four design principles that are relevant for correcting all of these errors, analyze the system along these dimensions, and show how each of these dimensions contributes to the performance.

[TACL] Predicting the Difficulty of Language Proficiency Tests

Lisa Beinborn, Torsten Zesch, and Iryna Gurevych

14:50–15:15

Language proficiency tests are used to evaluate and compare the progress of language learners. We present an approach for automatic difficulty prediction of C-tests that performs on par with human experts. On the basis of detailed analysis of newly collected data, we develop a model for C-test difficulty introducing four dimensions: solution difficulty, candidate ambiguity, inter-gap dependency, and paragraph difficulty. We show that cues from all four dimensions contribute to C-test difficulty.

Session 9C: Linguistic and Psycholinguistic Aspects of CL (Long Papers)
Plaza Ballroom F *Chair: William Schuler*

A Bayesian Model for Joint Learning of Categories and their Features

Lea Freymann and Mirella Lapata

14:00–14:25

Categories such as ANIMAL or FURNITURE are acquired at an early age and play an important role in processing, organizing, and conveying world knowledge. Theories of categorization largely agree that categories are characterized by features such as function or appearance and that feature and category acquisition go hand-in-hand, however previous work has considered these problems in isolation. We present the first model that jointly learns categories and their features. The set of features is shared across categories, and strength of association is inferred in a Bayesian framework. We approximate the learning environment with natural language text which allows us to evaluate performance on a large scale. Compared to highly engineered pattern-based approaches, our model is cognitively motivated, knowledge-lean, and learns categories and features which are perceived by humans as more meaningful.

Shared common ground influences information density in microblog texts

Gabriel Doyle and Michael Frank

14:25–14:50

If speakers use language rationally, they should structure their messages to achieve approximately uniform information density (UID), in order to optimize transmission via a noisy channel. Previous work identified a consistent increase in linguistic information across sentences in text as a signature of the UID hypothesis. This increase was derived from a predicted increase in context, but the context itself was not quantified. We use microblog texts from Twitter, tied to a single shared event (the baseball World Series), to quantify both linguistic and non-linguistic context. By tracking changes in contextual information, we predict and identify gradual and rapid changes in information content in response to in-game events. These findings lend further support to the UID hypothesis and highlights the importance of non-linguistic common ground for language production and processing.

Hierarchic syntax improves reading time prediction

Marten van Schijndel and William Schuler

14:50–15:15

Previous work has debated whether humans make use of hierarchic syntax when processing language (Frank and Bod, 2011; Fossum and Levy, 2012). This paper uses an eye-tracking corpus to demonstrate that hierarchic syntax significantly improves reading time prediction over a strong n-gram baseline. This study shows that an interpolated 5-gram baseline can be made stronger by combining n-gram statistics over entire eye-tracking regions rather than simply using the last n-gram in each region, but basic hierarchic syntactic measures are still able to achieve significant improvements over this improved baseline.

Best Paper Plenary Session

Time: 3:45–5:15

Location: Plaza Ballroom A, B, & C

Chair: Claire Cardie

Retrofitting Word Vectors to Semantic Lexicons

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

Vector space word representations are learned from distributional information of words in large corpora. Although such statistics are semantically informative, they disregard the valuable information that is contained in semantic lexicons such as WordNet, FrameNet, and the Paraphrase Database. This paper proposes a method for refining vector space representations using relational information from semantic lexicons by encouraging linked words to have similar vector representations, and it makes no assumptions about how the input vectors were constructed. Evaluated on a battery of standard lexical semantic evaluation tasks in several languages, we obtain substantial improvements starting with a variety of word vector models. Our refinement method outperforms prior techniques for incorporating semantic lexicons into the word vector training algorithms.

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

Vinodh Krishnan and Jacob Eisenstein

We present an unsupervised model for inducing signed social networks from the content exchanged across network edges. Inference in this model solves three problems simultaneously: (1) identifying the sign of each edge; (2) characterizing the distribution over content for each edge type; (3) estimating weights for triadic features that map to theoretical models such as structural balance. We apply this model to the problem of inducing the social function of address terms, such as Madame, comrade, and dude. On a dataset of movie scripts, our system obtains a coherent clustering of address terms, while at the same time making intuitively plausible judgments of the formality of social relations in each film. As an additional contribution, we provide a bootstrapping technique for identifying and tagging address terms in dialogue.

Unsupervised Morphology Induction Using Word Embeddings

Radu Soricut and Franz Och

We present a language agnostic, unsupervised method for inducing morphological transformations between words. The method relies on certain regularities manifest in high-dimensional vector spaces. We show that this method is capable of discovering a wide range of morphological rules, which in turn are used to build morphological analyzers. We evaluate this method across six different languages and nine datasets, and show significant improvements across all languages.

6

*SEM

The **Joint Conference on Lexical and Computational Semantics (*SEM)** provides a forum of exchange for the growing number of NLP researchers working on different aspects of semantic processing. After the previous editions of *SEM in Montreal (2012), Atlanta (2013), and Dublin (2014), the 2015 edition will take place in Denver on June 4 and 5 and is colocated with SemEval and NAACL. As in 2014 at COLING, also on this occasion *SEM and SemEval chose to coordinate their programs by featuring a joint invited talk. In this way, *SEM aims to bring together the ACL SIGLEX and ACL SIGSEM communities.

The acceptance rate of *SEM 2015 was quite competitive: out of 98 submissions, we accepted 36 papers for an overall acceptance of 37%. The acceptance rate of long paper that were accepted for oral presentation (18 out of 62) is 29%. The papers cover a wide range of topics including distributional semantics; lexical semantics and lexical acquisition; formal and linguistic semantics; discourse semantics; lexical resources, linked data and ontologies; semantics for applications; and extra-propositional semantics: sentiment and figurative meaning.

The *SEM 2015 program consists of oral presentations for selected long papers and a poster session for long and short papers.

Day One, June 4th:

- Joint *SEM SemEval keynote talk by **Marco Baroni**;
- Oral presentation sessions on distributional semantics, lexical semantics, and extra-propositional semantics;
- Poster session.

Day Two, June 5th:

- Keynote talk by **Preslav Nakov**;
- Oral presentation sessions on semantics for applications, lexical resources and ontologies, formal semantics, and discourse semantics;
- *SEM Best Paper Award.

We cannot finish without saying that *SEM 2015 would not have been possible without the considerable efforts of our area chairs, their reviewers, and the computational semantics community in general.

We hope you will enjoy *SEM 2015,

Martha Palmer, University of Colorado Boulder, General Chair
Gemma Boleda, University of Trento, Program Co-Chair
Paolo Rosso, Universitat Politècnica de València, Program Co-Chair

*SEM: 4th Joint Conference on Lexical and Computational Semantics

Organizers: *Martha Palmer, Gemma Boleda, and Paolo Rosso*

Venue: Governor's Square 14

Thursday, June 4, 2015

8:50–10:30 **Block 1 - Distributional semantics**

8:50–9:00 **Welcome**

9:00–10:00 **Playing ficles and running with the corbons: What (multimodal) distributional semantic models learn during their childhood (Joint *SEM and SemEval keynote by Marco Baroni)**

10:00–10:30 Neural Networks for Integrating Compositional and Non-compositional Sentiment in Sentiment Composition
Xiaodan Zhu, Hongyu Guo, and Parinaz Sobhani

10:30–11:00 **Coffee break**

11:00–12:30 **Block 2 - Distributional semantics**

11:00–11:30 Compositional Distributional Semantics with Long Short Term Memory
Phong Le and Willem Zuidema

11:30–12:00 A Hybrid Distributional and Knowledge-based Model of Lexical Semantics
Nikolaos Aletras and Mark Stevenson

12:00–12:30 Distributional semantics for ontology verification
Julien Corman, Laure Vieu, and Nathalie Aussenac-Gilles

12:30–2:00 **Lunch break**

2:00–3:30 **Block 3 - Lexical semantics**

2:00–2:30 Combining Seemingly Incompatible Corpora for Implicit Semantic Role Labeling
Parvin Sadat Feizabadi and Sebastian Padó

2:30–3:00 Identification of Caused Motion Construction
Jena D. Hwang and Martha Palmer

3:00–3:30 A Methodology for Word Sense Disambiguation at 90% based on large-scale CrowdSourcing
Oier Lopez de Lacalle and Eneko Agirre

3:30–4:00 **Coffee break**

4:00–5:00 **Block 4 - Extra-propositional semantics**

4:00–4:30 Learning Structures of Negations from Flat Annotations
Vinodkumar Prabhakaran and Branimir Boguraev

4:30–5:00 A New Dataset and Evaluation for Belief/Factuality
Vinodkumar Prabhakaran, Tomas By, Julia Hirschberg, Owen Rambow, Samira Shaikh, Tomek Strzalkowski, Jennifer Tracey, Michael Arrigo, Rupayan Basu, Micah Clark, Adam Dalton, Mona Diab, Louise Guthrie, Anna Prokofieva, Stephanie Strassel, Gregory Werner, Yorick Wilks, and Janyce Wiebe

5:00–7:00 Poster session with lightning talks intro

- Non-Orthogonal Explicit Semantic Analysis
Nitish Aggarwal, Kartik Asooja, Georgeta Bordea, and Paul Buitelaar
 - Combining Mention Context and Hyperlinks from Wikipedia for Named Entity Disambiguation
Ander Barrena, Aitor Soroa, and Eneko Agirre
 - Collective Document Classification with Implicit Inter-document Semantic Relationships
Clint Burford, Steven Bird, and Timothy Baldwin
- 11:00–11:30 SGRank: Combining Statistical and Graphical Methods to Improve the State of the Art in Unsupervised Keyphrase Extraction**
- Soheil Danesh, Tamara Sumner, and James H. Martin*
- Towards Semantic Language Classification: Inducing and Clustering Semantic Association Networks from Europarl
Steffen Eger, Niko Schenk, and Alexander Mehler
 - Ideological Perspective Detection Using Semantic Features
Heba ElFardy, Mona Diab, and Chris Callison-Burch
 - Mapping Different Rhetorical Relation Annotations: A Proposal
Farah Benamara and Maite Taboada
 - Dissecting the Practical Lexical Function Model for Compositional Distributional Semantics
Abhijeet Gupta, Jason Utt, and Sebastian Padó
 - Event Extraction as Frame-Semantic Parsing
Alex Judea and Michael Strube
 - Competence in lexical semantics
Andras Kornai, Judit Ács, Márton Makrai, Dávid Márk Nemeskey, Katalin Pajkossy, and Gábor Recski
 - Extending a Single-Document Summarizer to Multi-Document: a Hierarchical Approach
Luis Marujo, Ricardo Ribeiro, David Martins de Matos, João Neto, Anatole Gershman, and Jaime Carbonell
 - Reading Between the Lines: Overcoming Data Sparsity for Accurate Classification of Lexical Relationships
Silvia Neculescu, Sara Mendes, David Jurgens, Núria Bel, and Roberto Navigli
 - Multi-Level Alignments As An Extensible Representation Basis for Textual Entailment Algorithms
Tae-Gil Noh, Sebastian Padó, Vered Shwartz, Ido Dagan, Vivi Nastase, Kathrin Eichler, Lili Kotlerman, and Meni Adler
 - Leveraging Preposition Ambiguity to Assess Compositional Distributional Models of Semantics
Samuel Ritter, Cotie Long, Denis Paperno, Marco Baroni, Matthew Botvinick, and Adele Goldberg
 - Learning to predict script events from domain-specific text
Rachel Rudinger, Vera Demberg, Ashutosh Modi, Benjamin Van Durme, and Manfred Pinkal
 - Combining Open Source Annotators for Entity Linking through Weighted Voting
Pablo Ruiz and Thierry Poibeau
 - Automatic Generation of a Lexical Resource to support Semantic Role Labeling in Portuguese
Magali Sanches Duran and Sandra Aluísio
 - Can Selectional Preferences Help Automatic Semantic Role Labeling?
Shumin Wu and Martha Palmer

Friday, June 5, 2015

9:00–10:30 **Block 1 - Semantics for applications**

9:00–10:00 **60 Years Ago People Dreamed of Talking with a Machine. Are We Any Closer? (Keynote by Preslav Nakov)**

10:00–10:30 Implicit Entity Recognition in Clinical Documents

Sujan Perera, Pablo Mendes, Amit Sheth, Krishnaprasad Thirunarayan, Adarsh Alex, Christopher Heid, and Greg Mott

10:30–11:00 **Coffee break**

11:00–12:00 **Block 2 - Semantics for applications; Lexical resources and ontologies**

11:00–11:30 A Distant Supervision Approach to Semantic Role Labeling

Peter Exner, Marcus Klang, and Pierre Nugues

11:30–12:00 Discovering Hypernymy Relations using Text Layout

Jean-Philippe Fauconnier and Mouna Kamel

12:00–1:30 **Lunch break**

1:30–3:00 **Block 3 - Formal semantics**

1:30–2:00 The complexity of finding the maximum spanning DAG and other restrictions for DAG parsing of natural language

Natalie Schlüter

2:00–2:30 Incremental Semantic Construction Using Normal Form CCG Derivation

Yoshihide Kato and Shigeki Matsubara

2:30–3:00 Dependency-Based Semantic Role Labeling using Convolutional Neural Networks

William Foland and James Martin

3:00–3:30 **Coffee break**

3:30–5:00 **Block 4 - Discourse semantics**

3:30–4:00 A State-of-the-Art Mention-Pair Model for Coreference Resolution

Olga Uryupina and Alessandro Moschitti

4:00–4:30 Resolving Discourse-Deictic Pronouns: A Two-Stage Approach to Do It

Sujay Kumar Jauhar, Raul Guerra, Edgar González Pellicer, and

Marta Recasens

4:30–5:00 Graph-based Coherence Modeling For Assessing Readability

Mohsen Mesgar and Michael Strube

5:00 **Best Paper Award and closing**

7

Workshops

Thursday–Friday

Governor's Square 17	9th International Workshop on Semantic Evaluations	p.108
----------------------	----------------------------------------------------	-------

Thursday

Governor's Square 12	2nd Workshop on Argumentation Mining	p.118
Director's Row E	The 10th Workshop on Innovative Use of NLP for Building Educational Applications	p.120
Director's Row I	Fourth Workshop on Computational Linguistics for Literature	p.122
Tower Court A	Cognitive Modeling and Computational Linguistics	p.124
Director's Row H	The 3rd Workshop on EVENTS: Definition, Detection, Coreference, and Representation	p.125
Director's Row J	The 11th Workshop on Multiword Expressions	p.127
Governors Square 11	Ninth Workshop on Syntax, Semantics and Structure in Statistical Translation	p.129

Friday

Governors Square 17	Computational Linguistics and Clinical Psychology — From Linguistic Signal to Clinical Reality	p.131
Director's Row E	Extra-Propositional Aspects of Meaning in Computational Linguistics	p.133
Director's Row I	The 9th Linguistic Annotation Workshop	p.134
Director's Row J	The Third Workshop on Metaphor in NLP	p.136
Governor's Square 11	Third International Workshop on Natural Language Processing for Social Media	p.137
Director's Row H	Vector Space Modeling for NLP	p.138

Workshop 1: 9th International Workshop on Semantic Evaluations

Organizers: *Preslav Nakov, Torsten Zesch, Daniel Cer, and David Jurgens*

Venue: Governor's Square 17

Thursday, June 4, 2015

8:00–8:30 **Registration**

8:30–9:00 **Opening remarks**

9:00–10:00 **Joint *SEM and SemEval keynote talk by Marco Baroni, “Playing files and running with the corbons: What (multimodal) distributional semantic models learn during their childhood”**

Track I - Text Similarity and Question Answering (Part 1)

10:00–10:15 SemEval-2015 Task 1: Paraphrase and Semantic Similarity in Twitter (PIT)

Wei Xu, Chris Callison-Burch, and Bill Dolan

10:15–10:25 MITRE: Seven Systems for Semantic Similarity in Tweets
Guido Zarrella, John Henderson, Elizabeth M. Merkhofer, and Laura Strickhart

10:25–11:00 **Poster Session: Tasks 1, 2, and 3 (Part 1)**

- CICBUPnlp: Graph-Based Approach for Answer Selection in Community Question Answering Task
Helena Gomez, Darnes Vilariño, David Pinto, and Grigori Sidorov
- HLTC-HKUST: A Neural Network Paraphrase Classifier using Translation Metrics, Semantic Roles and Lexical Similarity Features
Dario Bertero and Pascale Fung
- FBK-HLT: An Effective System for Paraphrase Identification and Semantic Similarity in Twitter
Ngoc Phuoc An Vo, Simone Magnolini, and Octavian Popescu
- ECNU: Leveraging Word Embeddings to Boost Performance for Paraphrase in Twitter
Jiang Zhao and Man Lan
- ROB: Using Semantic Meaning to Recognize Paraphrases
Rob van der Goot and Gertjan van Noord
- AMRITA_CENSemEval-2015: Paraphrase Detection for Twitter using Unsupervised Feature Learning with Recursive Autoencoders
Mahalakshmi Shanumuga Sundaram, Anand Kumar Madasamy, and Soman Kotti Padannayil
- Ebiquity: Paraphrase and Semantic Similarity in Twitter using Skipgrams
Taneeya Satyapanich, Hang Gao, and Tim Finin
- RTM-DCU: Predicting Semantic Similarity with Referential Translation Machines
Ergun Bicici
- Twitter Paraphrase Identification with Simple Overlap Features and SVMs
Asli Eyecioglu and Bill Keller

- TKLBLIIR: Detecting Twitter Paraphrases with TweetingJay
Mladen Karan, Goran Glavaš, Jan Šnajder, Bojana Dalbelo Bašić, Ivan Vulić, and Marie-Francine Moens
- CDTDS: Predicting Paraphrases in Twitter via Support Vector Regression
Rafael - Michael Karampatsis
- yiGou: A Semantic Text Similarity Computing System Based on SVM
Yang Liu, Chengjie Sun, Lei Lin, and Xiaolong Wang
- USAAR-SHEFFIELD: Semantic Textual Similarity with Deep Regression and Machine Translation Evaluation Metrics
Liling Tan, Carolina Scarton, Lucia Specia, and Josef van Genabith
- TrWP: Text Relatedness using Word and Phrase Relatedness
Md Rashadul Hasan Rakib, Aminul Islam, and Evangelos Milios
- MiniExperts: An SVM Approach for Measuring Semantic Textual Similarity
Hanna Béchara, Hermani Costa, Shiva Taslimipoor, Rohit Gupta, Constantin Orasan, Gloria Corpas Pastor, and Ruslan Mitkov
- FBK-HLT: A New Framework for Semantic Textual Similarity
Ngoc Phuoc An Vo, Simone Magnolini, and Octavian Popescu
- UMDuluth-BlueTeam: SVCSTS - A Multilingual and Chunk Level Semantic Similarity System
Sakethram Karumuri, Viswanadh Kumar Reddy Vuggumudi, and Sai Charan Raj Chitrала
- SemantiKLUE: Semantic Textual Similarity with Maximum Weight Matching
Natalia Plotnikova, Gabriella Lapesa, Thomas Proisl, and Stefan Evert
- ECNU: Using Traditional Similarity Measurements and Word Embedding for Semantic Textual Similarity Estimation
Jiang Zhao, Man Lan, and Jun Feng Tian
- UQeResearch: Semantic Textual Similarity Quantification
Hamed Hassanzadeh, Tudor Groza, Anthony Nguyen, and Jane Hunter
- WSL: Sentence Similarity Using Semantic Distance Between Words
Naoko Miura and Tomohiro Takagi
- SOPA: Random Forests Regression for the Semantic Textual Similarity task
Davide Buscaldi, Jorge Garcia Flores, Ivan V. Meza, and Isaac Rodriguez
- MathLingBudapest: Concept Networks for Semantic Similarity
Gábor Recski and Judit Ács
- DCU: Using Distributional Semantics and Domain Adaptation for the Semantic Textual Similarity SemEval-2015 Task 2
Piyush Arora, Chris Hokamp, Jennifer Foster, and Gareth Jones
- DLSCU: Sentence Similarity from Word Alignment and Semantic Vector Composition
Md Arafat Sultan, Steven Bethard, and Tamara Sumner
- FCICU: The Integration between Sense-Based Kernel and Surface-Based Methods to Measure Semantic Textual Similarity
Basma Hassan, Samir AbdelRahman, and Reem Bahgat
- AZMAT: Sentence Similarity Using Associative Matrices
Evan Jaffe, Lifeng Jin, David King, and Marten van Schijndel
- NeRoSim: A System for Measuring and Interpreting Semantic Textual Similarity
Rajendra Banjade, Nobal Bikram Niraula, Nabin Maharjan, Vasile Rus, Dan Stefanescu, Mihai Lintean, and Dipesh Gautam

- Samsung: Align-and-Differentiate Approach to Semantic Textual Similarity
Lushan Han, Justin Martineau, Doreen Cheng, and Christopher Thomas
- UBC: Cubes for English Semantic Textual Similarity and Supervised Approaches for Interpretable STS
Eneko Agirre, Aitor Gonzalez-Agirre, Inigo Lopez-Gazpio, Montse Maritxalar, German Rigau, and Larraitz Uria
- ASAP-II: From the Alignment of Phrases to Textual Similarity
Ana Alves, David Simões, Hugo Gonçalo Oliveira, and Adriana Ferrugento
- TATO: Leveraging on Multiple Strategies for Semantic Textual Similarity
Tu Thanh Vu, Quan Hung Tran, and Son Bao Pham
- HITSZ-ICRC: Exploiting Classification Approach for Answer Selection in Community Question Answering
Yongshuai Hou, Cong Tan, Xiaolong Wang, Yaoyun Zhang, Jun Xu, and Qingcai Chen
- QCRI: Answer Selection for Community Question Answering - Experiments for Arabic and English
Massimo Nicosia, Simone Filice, Alberto Barrón-Cedeño, Iman Saleh, Hamdy Mubarak, Wei Gao, Preslav Nakov, Giovanni Da San Martino, Alessandro Moschitti, Kareem Darwish, Lluís Márquez, Shafiq Joty, and Walid Magdy
- ICRC-HIT: A Deep Learning based Comment Sequence Labeling System for Answer Selection Challenge
Xiaoqiang Zhou, Baotian Hu, Jiaxin Lin, Yang xiang, and Xiaolong Wang
- JAIST: Combining multiple features for Answer Selection in Community Question Answering
Quan Hung Tran, Vu Tran, Tu Vu, Minh Nguyen, and Son Bao Pham
- Shiraz: A Proposed List Wise Approach to Answer Validation
Amin Heydari Alashiy, Saeed Rahmani, Meysam Roostaei, and Mostafa Fakhrahmad
- Al-Bayan: A Knowledge-based System for Arabic Answer Selection
Reham Mohamed, Maha Ragab, Heba Abdelnasser, Nagwa M. El-Makky, and Marwan Torki
- FBK-HLT: An Application of Semantic Textual Similarity for Answer Selection in Community Question Answering
Ngoc Phuoc An Vo, Simone Magnolini, and Octavian Popescu
- ECNU: Using Multiple Sources of CQA-based Information for Answers Selection and YES/NO Response Inference
Liang Yi, JianXiang Wang, and Man Lan
- Volttron: A Hybrid System For Answer Validation Based On Lexical And Distance Features
Ivan Zumanov, Marina Kraeva, Nelly Hateva, Ivana Yovcheva, Ivelina Nikolova, and Galia Angelova
- CoMiC: Adapting a Short Answer Assessment System for Answer Selection
Björn Rudzewitz and Ramon Ziai

10:30–11:00 **Coffee Break and Poster Session**

Track I - Text Similarity and Question Answering (Part 2)

11:00–11:15 SemEval-2015 Task 2: Semantic Textual Similarity, English, Spanish and Pilot on Interpretability

Eneko Agirre, Carmen Banea, Claire Cardie, Daniel Cer, Mona Diab, Aitor Gonzalez-Agirre, Weiwei Guo, Iñigo Lopez-Gazpio, Montse Marítxalar, Rada Mihalcea, German Rigau, Larraitz Uria, and Janyce Wiebe

11:15–11:25 ExB Themis: Extensive Feature Extraction from Word Alignments for Semantic Textual Similarity

Christian Häring, Robert Remus, and Xose de la Puente

11:25–11:40 SemEval-2015 Task 3: Answer Selection in Community Question Answering

Preslav Nakov, Lluís Màrquez, Walid Magdy, Alessandro Moschitti, Jim Glass, and Bilal Randeree

11:40–11:50 VectorSLU: A Continuous Word Vector Approach to Answer Selection in Community Question Answering Systems

Yonatan Belinkov, Mitra Mohtarami, Scott Cyphers, and James Glass

11:50–12:30 **Poster Session: Tasks 1, 2, and 3 (Part 2)**

12:30–1:30 **Lunch Break**

Track IV - Word Sense Disambiguation and Induction

1:30–1:45 SemEval-2015 Task 13: Multilingual All-Words Sense Disambiguation and Entity Linking

Andrea Moro and Roberto Navigli

1:45–1:55 LIMSI: Translations as Source of Indirect Supervision for Multilingual All-Words Sense Disambiguation and Entity Linking

Marianna Apidianaki and Li Gong

1:55–2:10 SemEval-2015 Task 14: Analysis of Clinical Text

Noémie Elhadad, Sameer Pradhan, Sharon Gorman, Suresh Manandhar, Wendy Chapman, and Guergana Savova

2:10–2:20 UTH-CCB: The Participation of the SemEval 2015 Challenge — Task 14

Jun Xu, Yaoyun Zhang, Jingqi Wang, Yonghui Wu, Min Jiang, Ergin Soysal, and Hua Xu

2:20–2:35 SemEval-2015 Task 15: A CPA dictionary-entry-building task

Vít Baisa, Jane Bradbury, Silvie Cinkova, Ismail El Maarouf, Adam Kilgarriff, and Octavian Popescu

2:35–2:45 BLCUNLP: Corpus Pattern Analysis for Verbs Based on Dependency Chain

Yukun Feng, Qiao Deng, and Dong Yu

2:45–4:00 **Poster Session: Tasks 13, 14, and 15**

- WSD-games: a Game-Theoretic Algorithm for Unsupervised Word Sense Disambiguation
Rocco Tripodi and Marcello Pelillo
- DFKI: Multi-objective Optimization for the Joint Disambiguation of Entities and Nouns & Deep Verb Sense Disambiguation
Dirk Weissenborn, Feiyu Xu, and Hans Uszkoreit
- EBL-Hope: Multilingual Word Sense Disambiguation Using a Hybrid Knowledge-Based Technique
Eniafe Festus Ayetiran and Guido Boella
- VUA-background : When to Use Background Information to Perform Word Sense Disambiguation
Marten Postma, Ruben Izquierdo, and Piek Vossen
- TeamUFAL: WSD+EL as Document Retrieval
Petr Fanta, Roman Sudarikov, and Ondřej Bojar

- EL92: Entity Linking Combining Open Source Annotators via Weighted Voting
Pablo Ruiz and Thierry Poibeau
- UNIBA: Combining Distributional Semantic Models and Sense Distribution for Multilingual All-Words Sense Disambiguation and Entity Linking
Pierpaolo Basile, Annalina Caputo, and Giovanni Semeraro
- SUDOKU: Treating Word Sense Disambiguation & Entity Linking as a Deterministic Problem - via an Unsupervised & Iterative Approach
Steve L. Manion
- TeamHCMUS: Analysis of Clinical Text
Nghia Huynh and Quoc Ho
- UTU: Adapting Biomedical Event Extraction System to Disorder Attribute Detection
Kai Hakala
- IHS-RD-Belarus: Identification and Normalization of Disorder Concepts in Clinical Notes
Maryna Chernyshevich and Vadim Stankevitch
- UWM: A Simple Baseline Method for Identifying Attributes of Disease and Disorder Mentions in Clinical Text
Omid Ghiasvand and Rohit Kate
- TAKELAB: Medical Information Extraction and Linking with MINERAL
Goran Glavaš
- TMUNSW: Identification of Disorders and Normalization to SNOMED-CT Terminology in Unstructured Clinical Notes
Jitendra Jonnagaddala, Siaw-Teng Liaw, Pradeep Ray, Manish Kumar, and Hong-Jie Dai
- UtahPOET: Disorder Mention Identification and Context Slot Filling with Cognitive Inspiration
Kristina Doing-Harris, Sean Igo, Jianlin Shi, and John Hurdle
- ULisboa: Recognition and Normalization of Medical Concepts
André Leal, Bruno Martins, and Francisco Couto
- ezDI: A Supervised NLP System for Clinical Narrative Analysis
Parth Pathak, Pinal Patel, Vishal Panchal, Sagar Soni, Kinjal Dani, Amrish Patel, and Narayan Choudhary
- CUAB: Supervised Learning of Disorders and their Attributes using Relations
James Gung, John Osborne, and Steven Bethard
- BioinformaticsUA: Machine Learning and Rule-Based Recognition of Disorders and Clinical Attributes from Patient Notes
Sérgio Matos, José Sequeira, and José Luís Oliveira
- LIST-LUX: Disorder Identification from Clinical Texts
Asma Ben Abacha, Aikaterini Karanasiou, Yassine Mrabet, and Julio Cesar Dos Reis
- CMILLS: Adapting Semantic Role Labeling Features to Dependency Parsing
Chad Mills and Gina-Anne Levow
- Duluth: Word Sense Discrimination in the Service of Lexicography
Ted Pedersen

3:30–4:00 **Coffee Break and Poster Session**

Track III - Sentiment

4:00–4:15 SemEval-2015 Task 9: CLIEval Implicit Polarity of Events
Irene Russo, Tommaso Caselli, and Carlo Strapparava

- 4:15–4:30 SemEval-2015 Task 10: Sentiment Analysis in Twitter
Sara Rosenthal, Preslav Nakov, Svetlana Kiritchenko, Saif Mohammad, Alan Ritter, and Veselin Stoyanov
- 4:30–4:40 UNITN: Training Deep Convolutional Neural Network for Twitter Sentiment Classification
Aliaksei Severyn and Alessandro Moschitti
- 4:40–4:55 SemEval-2015 Task 11: Sentiment Analysis of Figurative Language in Twitter
Aniruddha Ghosh, Guofu Li, Tony Veale, Paolo Rosso, Ekaterina Shutova, John Barnden, and Antonio Reyes
- 4:55–5:05 CLaC-SentiPipe: SemEval2015 Subtasks 10 B,E, and Task 11
Canberk Özdemir and Sabine Bergler
- 5:05–5:20 SemEval-2015 Task 12: Aspect Based Sentiment Analysis
Maria Pontiki, Dimitris Galanis, Haris Papageorgiou, Suresh Manandhar, and Ion Androutsopoulos
- 5:20–5:30 NLANGP: Supervised Machine Learning System for Aspect Category Classification and Opinion Target Extraction
Zhiqiang Toh and Jian Su
- 5:30–6:30 **Poster Session: Tasks 9, 10, 11, and 12**
- SHELLFBK: An Information Retrieval-based System For Multi-Domain Sentiment Analysis
Mauro Dragoni
 - DIEGOLab: An Approach for Message-level Sentiment Classification in Twitter
Abeed Sarker, Azadeh Nikfarjam, Davy Weissenbacher, and Graciela Gonzalez
 - Splusplus: A Feature-Rich Two-stage Classifier for Sentiment Analysis of Tweets
Li Dong, Furu Wei, Yichun Yin, Ming Zhou, and Ke Xu
 - IIIT-H at SemEval 2015: Twitter Sentiment Analysis – The Good, the Bad and the Neutral!
Ayushi Dalmia, Manish Gupta, and Vasudeva Varma
 - CIS-positive: A Combination of Convolutional Neural Networks and Support Vector Machines for Sentiment Analysis in Twitter
Sebastian Ebert, Ngoc Thang Vu, and Hinrich Schütze
 - GTI: An Unsupervised Approach for Sentiment Analysis in Twitter
Milagros Fernández-Gavilanes, Tamara Álvarez-López, Jonathan Juncal-Martínez, Enrique Costa-Montenegro, and Francisco Javier González-Castaño
 - Gradiant-Analytics: Training Polarity Shifters with CRFs for Message Level Polarity Detection
Héctor Cerezo-Costas and Diego Celia-Salgado
 - IOA: Improving SVM Based Sentiment Classification Through Post Processing
Peijia Li, Weiqun Xu, Chenglong Ma, Jia Sun, and Yonghong Yan
 - RoseMerry: A Baseline Message-level Sentiment Classification System
Huizhi Liang, Richard Fothergill, and Timothy Baldwin
 - UDLAP: Sentiment Analysis Using a Graph-Based Representation
Esteban Castillo, Ofelia Cervantes, Darnes Vilariño, David Báez, and Alfredo Sánchez
 - ECNU: Multi-level Sentiment Analysis on Twitter Using Traditional Linguistic Features and Word Embedding Features
Zhihua Zhang, Guoshun Wu, and Man Lan

- Lsislif: Feature Extraction and Label Weighting for Sentiment Analysis in Twitter
Hussam Hamdan, Patrice Bellot, and Frederic Bechet
 - ELiRF: A SVM Approach for SA tasks in Twitter at SemEval-2015
Mayte Giménez, Ferran Pla, and Lluís-F. Hurtado
 - Webis: An Ensemble for Twitter Sentiment Detection
Matthias Hagen, Martin Potthast, Michel Büchner, and Benno Stein
 - Sentibase: Sentiment Analysis in Twitter on a Budget
Satarupa Guha, Aditya Joshi, and Vasudeva Varma
 - UNIBA: Sentiment Analysis of English Tweets Combining Micro-blogging, Lexicon and Semantic Features
Pierpaolo Basile and Nicole Novielli
 - IITPSemEval: Sentiment Discovery from 140 Characters
Ayush Kumar, Vamsi Krishna, and Asif Ekbal
 - Swiss-Chocolate: Combining Flipout Regularization and Random Forests with Artificially Built Subsystems to Boost Text-Classification for Sentiment
Fatih Uzdilli, Martin Jaggi, Dominic Egger, Pascal Julmy, Leon Derczynski, and Mark Cieliebak
 - INESC-ID: A Regression Model for Large Scale Twitter Sentiment Lexicon Induction
Silvio Amir, Wang Ling, Ramón Astudillo, Bruno Martins, Mario J. Silva, and Isabel Trancoso
 - KLUEless: Polarity Classification and Association
Natalia Plotnikova, Micha Kohl, Kevin Volkert, Stefan Evert, Andreas Lerner, Natalie Dykes, and Heiko Ermer
 - SWASH: A Naive Bayes Classifier for Tweet Sentiment Identification
Ruth Talbot, Chloe Acheampong, and Richard Wicentowski
 - SWATCS65: Sentiment Classification Using an Ensemble of Class Projects
Richard Wicentowski
 - SWATAC: A Sentiment Analyzer using One-Vs-Rest Logistic Regression
Yousef Alhessi and Richard Wicentowski
 - TwitterHawk: A Feature Bucket Based Approach to Sentiment Analysis
William Boag, Peter Potash, and Anna Rumshisky
 - SeNTU: Sentiment Analysis of Tweets by Combining a Rule-based Classifier with Supervised Learning
Prerna Chikarsal, Soujanya Poria, and Erik Cambria
 - INESC-ID: Sentiment Analysis without Hand-Coded Features or Linguistic Resources using Embedding Subspaces
Ramón Astudillo, Silvio Amir, Wang Ling, Bruno Martins, Mario J. Silva, and Isabel Trancoso
 - WarwickDCS: From Phrase-Based to Target-Specific Sentiment Recognition
Richard Townsend, Adam Tsakalidis, Yiwei Zhou, Bo Wang, Maria Liakata, Arkaitz Zubiaga, Alexandra Cristea, and Rob Procter
 - UIR-PKU: Twitter-OpinMiner System for Sentiment Analysis in Twitter at SemEval 2015
Xu Han, Binyang Li, Jing Ma, Yuxiao Zhang, Gaoyan Ou, Tengjiao Wang, and Kam-fai Wong
 - SWAT-CMW: Classification of Twitter Emotional Polarity using a Multiple-Classifier Decision Schema and Enhanced Emotion Tagging
Riley Collins, Daniel May, Noah Weinthal, and Richard Wicentowski
 - LLT-PolyU: Identifying Sentiment Intensity in Ironic Tweets
Hongzhi Xu, Enrico Santus, Anna Laszlo, and Chu-Ren Huang
-

- KELabTeam: A Statistical Approach on Figurative Language Sentiment Analysis in Twitter
Hoang Long Nguyen, Trung Duc Nguyen, Dosam Hwang, and Jason J. Jung
- LT3: Sentiment Analysis of Figurative Tweets: piece of cake #NotReally
Cynthia Van Hee, Els Lefever, and Veronique Hoste
- PRHLT: Combination of Deep Autoencoders with Classification and Regression Techniques for SemEval-2015 Task 11
Parth Gupta and Jon Ander Gómez
- ValenTo: Sentiment Analysis of Figurative Language Tweets with Irony and Sarcasm
Delia Irazú Hernández Farías, Emilio Sulis, Viviana Patti, Giancarlo Ruffo, and Cristina Bosco
- CPH: Sentiment analysis of Figurative Language on Twitter #easypeasy #not
Sarah McGillion, Héctor Martínez Alonso, and Barbara Plank
- UPF-taln: SemEval 2015 Tasks 10 and 11. Sentiment Analysis of Literal and Figurative Language in Twitter
Francesco Barbieri, Francesco Ronzano, and Horacio Saggin
- DsUniPi: An SVM-based Approach for Sentiment Analysis of Figurative Language on Twitter
Maria Karanassou, Christos Doulkeridis, and Maria Halkidi
- V3: Unsupervised Aspect Based Sentiment Analysis for SemEval2015 Task 12
Aitor García Pablos, Montse Cuadros, and German Rigau
- LT3: Applying Hybrid Terminology Extraction to Aspect-Based Sentiment Analysis
Orphee De Clercq, Marjan Van de Kauter, Els Lefever, and Veronique Hoste
- UFRGS: Identifying Categories and Targets in Customer Reviews
Anderson Kauer and Viviane Moreira
- SINAL: Syntactic Approach for Aspect-Based Sentiment Analysis
Salud M. Jiménez-Zafra, Eugenio Martínez-Cámara, M. Teresa Martín-Valdivia, and L. Alfonso Ureña López
- ECNU: Extracting Effective Features from Multiple Sequential Sentences for Target-dependent Sentiment Analysis in Reviews
Zhihua Zhang and Man Lan
- UMDuluth-CS8761-12: A Novel Machine Learning Approach for Aspect Based Sentiment Analysis
Ravikanth Repaka, Ranga Reddy Pallelra, Akshay Reddy Koppula, and Venkata Subhash Movva
- EliXa: A Modular and Flexible ABSA Platform
Iñaki San Vicente, Xabier Saralegi, and Rodrigo Agerri
- Lisislif: CRF and Logistic Regression for Opinion Target Extraction and Sentiment Polarity Analysis
Hussam Hamdan, Patrice Bellot, and Frederic Bechet
- SIEL: Aspect Based Sentiment Analysis in Reviews
Satarupa Guha, Aditya Joshi, and Vasudeva Varma
- Sentiae: Target and Aspect based Sentiment Analysis in SemEval-2015 Task 12
José Saías
- TJUdeM: A Combination Classifier for Aspect Category Detection and Sentiment Polarity Classification
Zhifei Zhang, Jian-Yun Nie, and Hongling Wang

Friday, June 5, 2015

Track II - Time and Space (Part 1)

- 9:00–9:15 SemEval-2015 Task 4: TimeLine: Cross-Document Event Ordering
Anne-Lyse Minard, Manuela Speranza, Eneko Agirre, Itziar Aldabe, Marieke van Erp, Bernardo Magnini, German Rigau, and Ruben Urizar
- 9:15–9:25 SPINOZA_VU: An NLP Pipeline for Cross Document TimeLines
Tomaso Caselli, Antske Fokkens, Roser Morante, and Piek Vossen
- 9:25–9:40 SemEval-2015 Task 5: QA TempEval - Evaluating Temporal Information Understanding with Question Answering
Hector Llorens, Nathanael Chambers, Naushad UzZaman, Nasrin Mostafazadeh, James Allen, and James Pustejovsky
- 9:40–9:50 HLT-FBK: A Complete Temporal Processing System for QA TempEval
Paramita Mirza and Anne-Lyse Minard
- 9:50–10:05 SemEval-2015 Task 6: Clinical TempEval
Steven Bethard, Leon Derczynski, Guergana Savova, James Pustejovsky, and Marc Verhagen
- 10:05–10:15 BluLab: Temporal Information Extraction for the 2015 Clinical TempEval Challenge
Sumithra Velupillai, Danielle L Mowery, Samir Abdelrahman, Lee Christensen, and Wendy Chapman

10:15–11:00 Poster Session: Tasks 4, 5, 6, 7, and 8 (Part 1)

- GPLSIUA: Combining Temporal Information and Topic Modeling for Cross-Document Event Ordering
Borja Navarro and Estela Saquete
- HeidelToul: A Baseline Approach for Cross-document Event Ordering
Bilel Moulahi, Jannik Strötgen, Michael Gertz, and Lynda Tamine
- HITSZ-ICRC: An Integration Approach for QA TempEval Challenge
Yongshuai Hou, Cong Tan, Qingcai Chen, and Xiaolong Wang
- UFPRSheffield: Contrasting Rule-based and Support Vector Machine Approaches to Time Expression Identification in Clinical TempEval
Hegler Tissot, Genevieve Gorrell, Angus Roberts, Leon Derczynski, and Marcos Didonet Del Fabro
- IXAGroupEHUDiac: A Multiple Approach System towards the Diachronic Evaluation of Texts
Haritz Salaberri, Iker Salaberri, Olatz Arregi, and Beñat Zapirain
- USAAR-CHRONOS: Crawling the Web for Temporal Annotations
Liling Tan and Noam Ordan
- AMBRA: A Ranking Approach to Temporal Text Classification
Marcos Zampieri, Alina Maria Ciobanu, Vlad Niculae, and Liviu P. Dinu
- IXAGroupEHUSpaceEval: (X-Space) A WordNet-based approach towards the Automatic Recognition of Spatial Information following the ISO-Space Annotation Scheme
Haritz Salaberri, Olatz Arregi, and Beñat Zapirain
- UTD: Ensemble-Based Spatial Relation Extraction
Jennifer D'Souza and Vincent Ng

10:30–11:00 Coffee Break and Poster Session

Track II - Time and Space (Part 2)

- 11:00–11:15 SemEval 2015, Task 7: Diachronic Text Evaluation
Octavian Popescu and Carlo Strapparava

11:15–11:25 UCD : Diachronic Text Classification with Character, Word, and Syntactic N-grams

Terrence Szymanski and Gerard Lynch

11:25–11:40 SemEval-2015 Task 8: SpaceEval

James Pustejovsky, Parisa Kordjamshidi, Marie-Francine Moens, Aaron Levine, Seth Dworman, and Zachary Yocum

11:40–11:50 SpRL-CWW: Spatial Relation Classification with Independent Multi-class Models

Eric Nichols and Fadi Botros

11:50–12:30 **Poster Session: Tasks 4, 5, 6, 7, and 8 (Part 2)**

12:30–2:00 **Lunch Break**

Track V - Learning Semantic Relations

2:00–2:15 SemEval-2015 Task 17: Taxonomy Extraction Evaluation (TExEval)
Georgeta Bordea, Paul Buitelaar, Stefano Faralli, and Roberto Navigli

2:15–2:25 INRIASAC: Simple Hypernym Extraction Methods
Gregory Grefenstette

2:25–2:40 SemEval 2015 Task 18: Broad-Coverage Semantic Dependency Parsing
Stephan Oepen, Marco Kuhlmann, Yusuke Miyao, Daniel Zeman, Silvie Cinkova, Dan Flickinger, Jan Hajic, and Zdenka Uresova

2:40–2:50 Peking: Building Semantic Dependency Graphs with a Hybrid Parser
Yantao Du, Fan Zhang, Xun Zhang, Weivei Sun, and Xiaojun Wan

2:50–4:00 **Poster Session: Tasks 17 and 18**

- USAAR-WLV: Hypernym Generation with Deep Neural Nets
Liling Tan, Rohit Gupta, and Josef van Genabith
- NTNU: An Unsupervised Knowledge Approach for Taxonomy Extraction
Bamfa Ceesay and Wen Juan Hou
- LT3: A Multi-modular Approach to Automatic Taxonomy Construction
Els Lefever
- TALN-UPF: Taxonomy Learning Exploiting CRF-Based Hypernym Extraction on Encyclopedic Definitions
Luis Espinosa Anke, Horacio Saggion, and Francesco Ronzano
- QASSIT: A Pretopological Framework for the Automatic Construction of Lexical Taxonomies from Raw Texts
Guillaume Cleuziou, Davide Buscaldi, Gaël Dias, Vincent Levorato, and Christine Largeron
- Riga: from FrameNet to Semantic Frames with C6.0 Rules
Guntis Barzdins, Peteris Paikens, and Didzis Gosko
- Turku: Semantic Dependency Parsing as a Sequence Classification
Jenna Kanerva, Juhani Luotolahti, and Filip Ginter
- Lisbon: Evaluating TurboSemanticParser on Multiple Languages and Out-of-Domain Data
Mariana S. C. Almeida and André F. T. Martins

3:30–4:00 **Coffee Break and Poster Session**

4:00–4:40 **SemEval-2016 Task Announcements**

4:40–5:40 **Closing Session (statistics, polls, questions)**

Workshop 2: 2nd Workshop on Argumentation Mining

Organizers: *Claire Cardie, Nancy Green, Iryna Gurevych, Graeme Hirst, Diane Litman, Smaranda Muresan, Georgios Petasis, Manfred Stede, Marilyn Walker, and Janyce Wiebe*

Venue: Governor's Square 12

Thursday, June 4, 2015

7:30–8:45 **Breakfast**

8:45–9:00 **Introductions**

9:00–9:40 **Setting the Stage: Overview on Argumentation Mining by Manfred Stede, Nancy Green and Ivan Habernal**

9:40–10:30 **Paper Session 1**

9:40–10:05 Linking the Thoughts: Analysis of Argumentation Structures in Scientific Publications
Christian Kirschner, Judith Eckle-Kohler, and Iryna Gurevych

10:05–10:30 Identifying Argumentation Schemes in Genetics Research Articles
Nancy Green

10:30–11:00 **Break**

11:00–12:30 **Paper Session 2**

11:00–11:20 Extracting Argument and Domain Words for Identifying Argument Components in Texts
Huy Nguyen and Diane Litman

11:20–11:30 **Poster Madness: 1-minute presentation for each poster**

11:30–12:30 **Poster Session**

- Towards relation based Argumentation Mining
Lucas Carstens and Francesca Toni
- A Shared Task on Argumentation Mining in Newspaper Editorials
Johannes Kiesel, Khalid Al Khatib, Matthias Hagen, and Benno Stein
- Conditional Random Fields for Identifying Appropriate Types of Support for Propositions in Online User Comments
Joonsuk Park, Arzoo Katiyar, and Bishan Yang
- A Computational Approach for Generating Toulmin Model Argumentation
Paul Reisert, Naoya Inoue, Naoki Okazaki, and Kentaro Inui
- Argument Extraction from News
Christos Sardianos, Ioannis Manousos Katakos, Georgios Petasis, and Vangelis Karkaletsis
- From Argumentation Mining to Stance Classification
Parinaz Sobhani, Diana Inkpen, and Stan Matwin
- Argument Discovery and Extraction with the Argument Workbench
Adam Wyner, Wim Peters, and David Price

12:30–2:00 **Lunch**

2:00–3:30 **Paper Session 3**

2:00–2:25 Automatic Claim Negation: Why, How and When
Yonatan Bilu, Daniel Hershcovitch, and Noam Slonim

- 2:25–2:50 Learning Sentence Ordering for Opinion Generation of Debate
Toshihiko Yanase, Toshinori Miyoshi, Kohsuke Yanai, Misa Sato, Makoto Iwayama, Yoshiki Niwa, Paul Reisert, and Kentaro Inui
- 2:50–3:10 Towards Detecting Counter-considerations in Text
Andreas Peldszus and Manfred Stede
- 3:10–3:30 Identifying Prominent Arguments in Online Debates Using Semantic Textual Similarity
Filip Boltužić and Jan Šnajder
- 3:30–4:00 **Break**
- 4:00–4:50 **Paper Session 4**
- 4:00–4:25 And That's A Fact: Distinguishing Factual and Emotional Argumentation in Online Dialogue
Shereen Oraby, Lena Reed, Ryan Compton, Ellen Riloff, Marilyn Walker, and Steve Whittaker
- 4:25–4:50 Combining Argument Mining Techniques
John Lawrence and Chris Reed
- 4:50–5:30 **Wrap-up Discussion**

Workshop 3: The 10th Workshop on Innovative Use of NLP for Building Educational Applications

Organizers: *Joel Tetreault, Jill Burstein, and Claudia Leacock*

Venue: Director's Row E

Thursday, June 4, 2015

8:45–9:00 **Lead Presentations**

9:00–9:15 **Opening Remarks**

9:15–9:40 Candidate evaluation strategies for improved difficulty prediction of language tests
Lisa Beinborn, Torsten Zesch, and Iryna Gurevych

9:40–10:05 Feature selection for automated speech scoring
Anastassia Loukina, Klaus Zechner, Lei Chen, and Michael Heilman

10:05–10:30 Incorporating Coherence of Topics as a Criterion in Automatic Response-to-Text Assessment of the Organization of Writing
Zahra Rahimi, Diane Litman, Elaine Wang, and Richard Correnti

10:30–11:00 **Break**

11:00–11:25 Automatic morphological analysis of learner Hungarian
Scott Ledbetter and Markus Dickinson

11:25–11:45 Automated Scoring of Picture-based Story Narration
Swapna Somasundaran, Chong Min Lee, Martin Chodorow, and Xinhao Wang

11:45–12:05 Measuring Feature Diversity in Native Language Identification
Shervin Malmasi and Aoife Cahill

12:05–12:25 Automated Evaluation of Scientific Writing: AESW Shared Task Proposal
Vidas Daudaravicius

12:30–2:00 **Lunch**

2:00–3:30 **Poster Sessions**

2:00–2:45 **Poster Session A**

- Scoring Persuasive Essays Using Opinions and their Targets
Noura Farra, Swapna Somasundaran, and Jill Burstein
- Towards Automatic Description of Knowledge Components
Cyril Goutte, Guillaume Durand, and Serge Leger
- The Impact of Training Data on Automated Short Answer Scoring Performance
Michael Heilman and Nitin Madnani
- Interpreting Questions with a Log-Linear Ranking Model in a Virtual Patient Dialogue System
Evan Jaffe, Michael White, William Schuler, Eric Fosler-Lussier, Alex Rosenfeld, and Douglas Danforth
- Identifying Patterns For Short Answer Scoring Using Graph-based Lexico-Semantic Text Matching
Lakshmi Ramachandran, Jian Cheng, and Peter Foltz
- Lark Trills for Language Drills: Text-to-speech technology for language learners
Elena Volodina and Dijana Pijetlovic

- The Jinan Chinese Learner Corpus
Maolin Wang, Shervin Malmasi, and Mingxuan Huang
- Reducing Annotation Efforts in Supervised Short Answer Scoring
Torsten Zesch, Michael Heilman, and Aoife Cahill
- Annotation and Classification of Argumentative Writing Revisions
Fan Zhang and Diane Litman

2:45–3:30 **Poster Session B**

- Embarrassed or Awkward? Ranking Emotion Synonyms for ESL Learners' Appropriate Wording
Wei-Fan Chen, MeiHua Chen, and Lun-Wei Ku
- RevUP: Automatic Gap-Fill Question Generation from Educational Texts
Girish Kumar, Rafael Banchs, and Luis Fernando D'Haro
- Preliminary Experiments on Crowdsourced Evaluation of Feedback Granularity
Nitin Madnani, Martin Chodorow, Aoife Cahill, Melissa Lopez, Yoko Futagi, and Yigal Attali
- Oracle and Human Baselines for Native Language Identification
Shervin Malmasi, Joel Tetreault, and Mark Dras
- Using PEGWriting® to Support the Writing Motivation and Writing Quality of Eighth-Grade Students: A Quasi-Experimental Study
Joshua Wilson and Trish Martin
- Towards Creating Pedagogic Views from Encyclopedic Resources
Ditty Mathew, Dhivya Eswaran, and Sutanu Chakraborti
- Judging the Quality of Automatically Generated Gap-fill Question using Active Learning
Nobal Bikram Niraula and Vasile Rus
- Generating Reference Texts for Short Answer Scoring Using Graph-based Summarization
Lakshmi Ramachandran and Peter Foltz
- Evaluating the performance of Automated Text Scoring systems
Helen Yannakoudakis and Ronan Cummins
- Task-Independent Features for Automated Essay Grading
Torsten Zesch, Michael Wojatzki, and Dirk Scholten-Akoun

3:30–4:00 **Break**

4:00–4:25 Using Learner Data to Improve Error Correction in Adjective—Noun Combinations
Ekaterina Kochmar and Ted Briscoe

4:25–4:50 Using NLP to Support Scalable Assessment of Short Free Text Responses
Alistair Willis

4:50–5:15 Automatically Scoring Freshman Writing: A Preliminary Investigation
Courtney Napoles and Chris Callison-Burch

5:15–5:30 **Closing Remarks**

Workshop 4: Fourth Workshop on Computational Linguistics for Literature

Organizers: *Anna Feldman, Anna Kazantseva, Stan Szpakowicz, and Corina Koole*

Venue: Director's Row I

Thursday, June 4, 2015

Session 1

8:57–9:00 **Welcome**

9:00–10:00 **Exploratory Programming for Literary Work: invited talk, Nick Montfort, MIT**

- Tools for Digital Humanities: Enabling Access to the Old Occitan Romance of Flamenca
Olga Scrivner and Sandra Kübler

Coffee break

Session 2

- RhymeDesign: A Tool for Analyzing Sonic Devices in Poetry
Nina McCurdy, Vivek Srikumar, and Miriah Meyer
- Rhetorical Figure Detection: the Case of Chiasmus
Marie Dubremetz and Joakim Nivre
- Validating Literary Theories Using Automatic Social Network Extraction
Prashant Jayannavar, Apoorv Agarwal, Melody Ju, and Owen Rambow

Lunch break

Session 3

2:00–3:00 **The (not so) Simple Shape of Stories: invited talk, Matthew Jockers, UNL**

3:00–3:30 **Poster teaser talks**

Coffee break

Session 4

4:00–4:30 **Poster session**

- GutenTag: an NLP-driven Tool for Digital Humanities Research in the Project Gutenberg Corpus
Julian Brooke, Adam Hammond, and Graeme Hirst
- A Pilot Experiment on Exploiting Translations for Literary Studies on Kafka's "Verwandlung"
Fabienne Cap, Ina Rösiger, and Jonas Kuhn
- Identifying Literary Texts with Bigrams
Andreas van Cranenburgh and Corina Koole
- Visualizing Poetry with SPARSAR — Visual Maps from Poetic Content
Rodolfo Delmonte

- Towards a better understanding of Burrows's Delta in literary authorship attribution
Stefan Evert, Thomas Proisl, Thorsten Vitt, Christof Schöch, Fotis Jannidis, and Steffen Pielström
- Gender-Based Vocation Identification in Swedish 19th Century Prose Fiction using Linguistic Patterns, NER and CRF Learning
Dimitrios Kokkinakis, Ann Ighe, and Mats Malm
- Rule-based Coreference Resolution in German Historic Novels
Markus Krug, Frank Puppe, Fotis Jannidis, Luisa Macharowsky, Isabella Reger, and Lukas Weimar
- A computational linguistic approach to Spanish Golden Age Sonnets: metrical and semantic aspects
Borja Navarro

Session 5

- AUTOMATED TRANSLATION OF A LITERARY WORK: A PILOT STUDY
Laurent Besacier and Lane Schwartz
- Translating Literary Text between Related Languages using SMT
Antonio Toral and Andy Way

5:30–5:33 **Farewell**

Workshop 5: Cognitive Modeling and Computational Linguistics

Organizers: *Timothy J. O'Donnell and Marten van Schijndel*

Venue: Tower Court A

Thursday, June 4, 2015

8:55–9:00 Opening Remarks

9:00–10:00 Invited Talk by Andrew Kehler

10:00–10:30 Predictions for self-priming from incremental updating models unifying comprehension and production
Cassandra L. Jacobs

10:30–11:00 Coffee Break

11:00–11:30 Pragmatic Alignment on Social Support Type in Health Forum Conversations
Yafei Wang, John Yen, and David Reitter

11:30–12:00 Audience size and contextual effects on information density in Twitter conversations
Gabriel Doyle and Michael Frank

12:00–12:30 Centre Stage: How Social Network Position Shapes Linguistic Coordination
Bill Noble and Raquel Fernandez

12:30–1:30 Lunch Break

1:30–2:00 Fusion of Compositional Network-based and Lexical Function Distributional Semantic Models
Spiros Georgilidakis, Elias Iosif, and Alexandros Potamianos

2:00–2:30 Verb polysemy and frequency effects in thematic fit modeling
Clayton Greenberg, Vera Demberg, and Asad Sayeed

2:30–3:00 An Evaluation and Comparison of Linguistic Alignment Measures
Yang Xu and David Reitter

3:00–3:30 Utility-based evaluation metrics for models of language acquisition: A look at speech segmentation
Lawrence Phillips and Lisa Pearl

3:30–4:00 Coffee Break

4:00–4:30 Evidence of syntactic working memory usage in MEG data
Marten van Schijndel, Brian Murphy, and William Schuler

4:30–5:00 Modeling fMRI time courses with linguistic structure at various grain sizes
John Hale, David Lutz, Wen-Ming Luh, and Jonathan Brennan

5:00–6:00 Invited Talk by Mark Steedman

Workshop 6: The 3rd Workshop on EVENTS: Definition, Detection, Coreference, and Representation

Organizers: *Eduard Hovy, Teruko Mitamura, and Martha Palmer*

Venue: Director's Row H

Thursday, June 4, 2015

7:30–9:00 Breakfast

9:00–9:15 Welcome

9:15–10:30 Invited Talk: **Bernardo Magnini**

10:30–10:50 Coffee break

10:50–12:00 Session I: Subevent structure and shared task

12:00–1:30 Lunch

1:30–2:40 Session II: Event mention (nugget) detection

2:40–3:30 Poster Session

- Translating Granularity of Event Slots into Features for Event Coreference Resolution.
Agata Cybulska and Piek Vossen
- Word Sense Disambiguation via PropStore and OntoNotes for Event Mention Detection
Nicolas R Fauciglia, Yiu-Chang Lin, Xuezhe Ma, and Eduard Hovy
- Opposition Relations among Verb Frames
Anna Feltracco, Elisabetta Jezek, and Bernardo Magnini
- Encoding event structure in Urdu/Hindi VerbNet
Annette Hautli-Janisz, Tracy Holloway King, and Gilian Ramchand
- Using Topic Modeling and Similarity Thresholds to Detect Events
Nathan Keane, Connie Yee, and Liang Zhou
- Detecting Causally Embedded Structures Using an Evolutionary Algorithm
Chen Li and Roxana Girju
- Evaluation Algorithms for Event Nugget Detection : A Pilot Study
Zhengzhong Liu, Teruko Mitamura, and Eduard Hovy
- Event analysis for information extraction from business-based technical documents
Bell Manrique-Losada and Carlos Mario Zapata Jaramillo
- Event Nugget Annotation: Processes and Issues
Teruko Mitamura, Yukari Yamakawa, Susan Holm, Zhiyi Song, Ann Bies, Seth Kulick, and Stephanie Strassel
- Game-Changing Event Definition and Detection in an eSports Corpus
Emily Olshefski
- Identifying Various Kinds of Event Mentions in K-Parser Output
Arpit Sharma, Nguyen Vo, Somak Aditya, and Chitta Baral
- From Light to Rich ERE: Annotation of Entities, Relations, and Events
Zhiyi Song, Ann Bies, Stephanie Strassel, Tom Riese, Justin Mott, Joe Ellis, Jonathan Wright, Seth Kulick, Neville Ryant, and Xiaoyi Ma

- Classification and Acquisition of Contradictory Event Pairs using
Crowdsourcing
*Yu Takabatake, Hajime Morita, Daisuke Kawahara, Sadao Kurohashi,
Ryuichiro Higashinaka, and Yoshihiro Matsuo*
- Semantic Interoperability for Cross-lingual and cross-document Event
Detection
Piek Vossen, Egoitz Laparra, German Rigau, and Itziar Aldabe
- Modeling and Characterizing Social Media Topics Using the Gamma
Distribution
Connie Yee, Nathan Keane, and Liang Zhou

3:30–3:50 **Coffee Break**

3:50–5:00 **Session III: Inter-event relation annotation**

5:00–5:30 **Discussion: Future event-related directions**

5:30–5:45 **Closing**

Workshop 7: The 11th Workshop on Multiword Expressions

Organizers: *Valia Kordoni, Markus Egg, Kostadin Cholakov, Stella Markantonatou, and Shuly Wintner*

Venue: Director's Row J

Thursday, June 4, 2014

Oral Session 1

9:00–9:30 A Method of Accounting Bigrams in Topic Models
Michael Nokel and Natalia Loukachevitch

9:30–10:00 Multiword Expression Identification with Recurring Tree Fragments and Association Measures
Federico Sangati and Andreas van Cranenburgh

10:00–10:30 How to Account for Idiomatic German Support Verb Constructions in Statistical Machine Translation
Fabienne Cap, Manju Nirmal, Marion Weller, and Sabine Schulte im Walde

10:30–11:00 Coffee Break

Oral Session 2

11:00–11:20 A Multiword Expression Data Set: Annotating Non-Compositionality and Conventionalization for English Noun Compounds
Meghdad Farahmand, Aaron Smith, and Joakim Nivre

11:20–11:40 Modeling the Statistical Idiosyncrasy of Multiword Expressions
Meghdad Farahmand and Joakim Nivre

11:40–12:00 Clustering-based Approach to Multiword Expression Extraction and Ranking
Elena Tutubalina

Invited Talk by Laura A. Michaelis

12:00–1:00 How Constructions Mean
Paul Kay and Laura A. Michaelis

1:00–2:00 Lunch

2:00–2:30 Poster Booster Session (5 minutes per poster)

- Never-Ending Multiword Expressions Learning
Alexandre Rondon, Helena Caseli, and Carlos Ramisch
- The Impact of Multiword Expression Compositionality on Machine Translation Evaluation
Bahar Salehi, Nitika Mathur, Paul Cook, and Timothy Baldwin
- The Bare Necessities: Increasing Lexical Coverage for Multi-Word Domain Terms with Less Lexical Data
Branimir Boguraev, Esme Manandise, and Benjamin Segal
- Phrase translation using a bilingual dictionary and n-gram data: A case study from Vietnamese to English
Khang Nhut Lam, Feras Al Tarouti, and Jugal Kalita
- Annotation and Extraction of Multiword Expressions in Turkish Treebanks
Gülşen Eryiğit, Kübra ADALI, Dilara Torunoglu-Selamet, Umut Sulubacak, and Tuğba Pamay

- Event Categorization beyond Verb Senses
Aron Marvel and Jean-Pierre Koenig

2:30–3:30 **Poster Session**

3:30–4:00 **Coffee Break**

Oral Session 3

4:00–4:30 Muddying The Multiword Expression Waters: How Cognitive Demand Affects Multiword Expression Production
Adam Goodkind and Andrew Rosenberg

4:30–5:00 Building a Lexicon of Formulaic Language for Language Learners
Julian Brooke, Adam Hammond, David Jacob, Vivian Tsang, Graeme Hirst, and Fraser Shein

Workshop 8: Ninth Workshop on Syntax, Semantics and Structure in Statistical Translation

Organizers: *Dekai Wu, Marine Carpuat, Eneko Agirre, and Nora Aranberri*

Venue: Governors Square 11

June 4, 2015

8:55–9:00 Opening Remarks

9:00–10:30 Session 1

9:00–10:00 Invited Talk (Philipp Koehn)

10:00–10:30 Harmonizing word alignments and syntactic structures for extracting phrasal translation equivalents
Dun Deng, Nianwen Xue, and Shiman Guo

10:30–11:00 Coffee Break

11:00–12:30 Session 2

11:00–11:30 Non-projective Dependency-based Pre-Reordering with Recurrent Neural Network for Machine Translation
Antonio Valerio Miceli Barone and Giuseppe Attardi

11:30–12:00 Translating Negation: Induction, Search And Model Errors
Federico Fancellu and Bonnie Webber

12:00–12:30 SMT error analysis and mapping to syntactic, semantic and structural fixes
Nora Aranberri

12:30–1:55 Lunch Break

1:55–2:30 Session 3

1:55–2:00 QTLeap Best Paper Award

2:00–2:30 Unsupervised False Friend Disambiguation Using Contextual Word Clusters and Parallel Word Alignments
Maryam Aminian, Mahmoud Ghoneim, and Mona Diab

2:30–3:30 Session 4: Posters

2:30–2:35 METEOR-WSD: Improved Sense Matching in MT Evaluation
Marianna Apidianaki and Benjamin Marie

2:35–2:40 Analyzing English-Spanish Named-Entity enhanced Machine Translation
Mikel Artetxe, Eneko Agirre, Iñaki Alegria, and Gorka Labaka

2:40–2:45 Predicting Prepositions for SMT
Marion Weller, Alexander Fraser, and Sabine Schulte im Walde

2:45–2:50 Translation reranking using source phrase dependency features
Antonio Valerio Miceli Barone

2:50–2:55 Semantics-based pretranslation for SMT using fuzzy matches
Tom Vanallemeersch and Vincent Vandeghinste

3:30–4:00 Coffee Break

4:00–5:00 **Session 5**

4:00–4:30 What Matters Most in Morphologically Segmented SMT Models?

Mohammad Salameh, Colin Cherry, and Grzegorz Kondrak

4:30–5:00 Improving Chinese-English PropBank Alignment

Shumin Wu and Martha Palmer

Workshop 9: Computational Linguistics and Clinical Psychology — From Linguistic Signal to Clinical Reality

Organizers: *Margaret Mitchell, Glen Coppersmith, and Kristy Hollingshead*

Venue: Governors Square 17

Friday, June 5, 2015

9:00–9:15 **Opening Remarks (Margaret Mitchell, Glen Coppersmith, Kristy Hollingshead)**

9:15–11:00 **Oral Presentations, Session 1**

- From ADHD to SAD: Analyzing the Language of Mental Health on Twitter through Self-Reported Diagnoses
Glen Coppersmith, Mark Dredze, Craig Harman, and Kristy Hollingshead
- Quantifying the Language of Schizophrenia in Social Media
Margaret Mitchell, Kristy Hollingshead, and Glen Coppersmith
- The role of personality, age, and gender in tweeting about mental illness
Daniel Preoțiuc-Pietro, Johannes Eichstaedt, Gregory Park, Maarten Sap, Laura Smith, Victoria Tobolsky, H. Andrew Schwartz, and Lyle Ungar

11:00–11:15 **Break**

11:15–11:45 **Shared Task**

- CLPsych 2015 Shared Task: Depression and PTSD on Twitter
Glen Coppersmith, Mark Dredze, Craig Harman, Kristy Hollingshead, and Margaret Mitchell
- Mental Illness Detection at the World Well-Being Project for the CLPsych 2015 Shared Task
Daniel Preoțiuc-Pietro, Maarten Sap, H. Andrew Schwartz, and Lyle Ungar
- Screening Twitter Users for Depression and PTSD with Lexical Decision Lists
Ted Pedersen
- The University of Maryland CLPsych 2015 Shared Task System
Philip Resnik, William Armstrong, Leonardo Claudino, and Thang Nguyen

11:35–11:45 **Discussion (Philip Resnik)**

11:45–12:45 **Poster Presentations**

- Computational cognitive modeling of inflectional verb morphology in Spanish-speakers for the characterization and diagnosis of Alzheimer's Disease
M. Dolores del Castillo, J. Ignacio Serrano, and Jesús Oliva
- Recursive Neural Networks for Coding Therapist and Patient Behavior in Motivational Interviewing
Michael Tanana, Kevin Hallgren, Zac Imel, David Atkins, Padhraic Smyth, and Vivek Srikumar
- Putting Feelings into Words: Cross-Linguistic Markers of the Referential Process
Sean Murphy, Bernard Maskit, and Wilma Bucci

- Towards Developing an Annotation Scheme for Depressive Disorder Symptoms: A Preliminary Study using Twitter Data
Danielle Mowery, Craig Bryan, and Mike Conway
- Beyond LDA: Exploring Supervised Topic Modeling for Depression-Related Language in Twitter
Philip Resnik, William Armstrong, Leonardo Claudino, Thang Nguyen, Viet-An Nguyen, and Jordan Boyd-Graber
- Automated morphological analysis of clinical language samples
Kyle Gorman, Steven Bedrick, Geza Kiss, Eric Morley, Rosemary Ingham, Metrah Mohammed, Katina Papadakis, and Jan van Santen

12:45–2:00 **Lunch**

Invited Talk

2:00–2:45 **Invited Talk (Munmun De Choudhury)**

2:45–3:00 **Break**

3:00–4:45 **Oral Presentations, Session 2**

- Similarity Measures for Quantifying Restrictive and Repetitive Behavior in Conversations of Autistic Children
Masoud Rouhizadeh, Richard Sproat, and Jan van Santen
- Practical issues in developing semantic frameworks for the analysis of verbal fluency data: A Norwegian data case study
Mark Rosenstein, Peter Foltz, Anja Vaskinn, and Brita Elvevåg
- A Computer Program for Tracking the Evolution of a Psychotherapy Treatment
Bernard Maskit, Wilma Bucci, and Sean Murphy

4:45–5:00 **Closing Remarks**

Workshop 10: Extra-Propositional Aspects of Meaning in Computational Linguistics

Organizers: *Eduardo Blanco, Roser Morante, and Caroline Sporleder*

Venue: Director's Row E

Friday, June 5, 2015

8:00–9:00 **Breakfast**

9:15–9:30 **Opening remarks**

9:30–10:30 **Invited Talk: Lauri Karttunen**

9:30–10:30 Limits of Natural Logic
Lauri Karttunen

10:30–11:00 **Coffee break**

11:00–12:30 **Session 1**

11:00–11:30 Translating Negation: A Manual Error Analysis
Federico Fancellu and Bonnie Webber

11:30–12:00 Filled Pauses in User-generated Content are Words with
Extra-propositional Meaning
Ines Rehbein

12:00–12:30 A Compositional Interpretation of Biomedical Event Factuality
Halil Kilicoglu, Graciela Rosemblat, Michael Cairelli, and Thomas Rindflesch

12:30–2:00 **Lunch break**

2:00–3:00 **Session 2**

2:10–2:30 Committed Belief Tagging on the Factbank and LU Corpora: A Comparative Study
Gregory Werner, Vinodkumar Prabhakaran, Mona Diab, and Owen Rambow

2:30–3:00 Extending NegEx with Kernel Methods for Negation Detection in Clinical Text
Chaitanya Shivade, Marie-Catherine de Marneffe, Eric Fosler-Lussier, and Albert M. Lai

3:00–3:30 **Discussion and closing remarks**

Workshop 11: The 9th Linguistic Annotation Workshop

Organizers: *Adam Meyers, Ines Rehbein, and Heike Zinsmeister*

Venue: Director's Row I

Friday, June 5, 2015

8:45–9:00 **Opening Remarks**

9:00–10:30 **Session 1**

Oral Presentations

9:00–9:30 Scaling Semantic Frame Annotation

Nancy Chang, Praveen Paritosh, David Huynh, and Collin Baker

9:30–10:00 An Analytic and Empirical Evaluation of Return-on-Investment-Based Active Learning

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

10:00–10:30 Annotating genericity: a survey, a scheme, and a corpus

Annemarie Friedrich, Alexis Palmer, Melissa Peate Sørensen, and Manfred Pinkal

10:30–11:00 **Coffee break**

11:00–12:30 **Session 2**

Poster presentations

- Design and Annotation of the First Italian Corpus for Text Simplification
Dominique Brunato, Felice Dell'Orletta, Giulia Venturi, and Simonetta Montemagni

- On the Discursive Structure of Computer Graphics Research Papers
Beatrix Fisas, Horacio Saggion, and Francesco Ronzano

- Semantic Annotation of Japanese Functional Expressions and its Impact on Factuality Analysis
Yudai Kamioka, Kazuya Narita, Junta Mizuno, Miwa Kanno, and Kentaro Inui

- A Qualitative Analysis of a Corpus of Opinion Summaries based on Aspects

Roque Lopez, Thiago Pardo, Lucas Avanço, Pedro Filho, Alessandro Bokan, Paula Cardoso, Márcio Dias, Fernando Nóbrega, Marco Cabezudo, Jackson Souza, Andressa Zacarias, Eloize Seno, and Ariani Di Felippo

- Developing Language-tagged Corpora for Code-switching Tweets
Suraj Maharjan, Elizabeth Blair, Steven Bethard, and Thamar Solorio

- Annotating Geographical Entities on Microblog Text
Koji Matsuda, Akira Sasaki, Naoaki Okazaki, and Kentaro Inui

- The Annotation Process of the ITU Web Treebank
Tuğba Pamay, Umut Sulubacak, Dilara Torunoğlu-Selamet, and Gülşen Eryiğit

- Part of Speech Annotation of Intermediate Versions in the Keystroke Logged Translation Corpus
Tatiana Šerbina, Paula Niemietz, Matthias Fricke, Philipp Meisen, and Stella Neumann

- A Hierarchy with, of, and for Preposition Supersenses
Nathan Schneider, Vivek Srikumar, Jena D. Hwang, and Martha Palmer
- Bilingual English-Czech Valency Lexicon Linked to a Parallel Corpus
Zdenka Uresova, Ondřej Dušek, Eva Fucikova, Jan Hajic, and Jana Sindlerova
- Correction Annotation for Non-Native Arabic Texts: Guidelines and Corpus
Wajdi Zaghouani, Nizar Habash, Houda Bouamor, Alla Rozovskaya, Behrang Mohit, Abeer Heider, and Kemal Oflazer

12:30–2:00 **Lunch break**

2:00–3:30 **Session 3**

Panel

- Balancing the Existing and the New in the Context of Annotating Non-Canonical Language
Ann Bies
- Parsing Learner Text: to Shoehorn or not to Shoehorn
Aoife Cahill
- Non-canonical language is not harder to annotate than canonical language
Barbara Plank, Héctor Martínez Alonso, and Anders Søgaard
- What I've learned about annotating informal text (and why you shouldn't take my word for it)
Nathan Schneider

4:00–6:00 **Session 4**

Oral presentations

- 4:00–4:30 On Grammaticality in the Syntactic Annotation of Learner Language
Markus Dickinson and Marwa Ragheb
- 4:30–5:00 Across Languages and Genres: Creating a Universal Annotation Scheme for Textual Relations
Ekaterina Lapshinova-Koltunski, Anna Nedoluzhko, and Kerstin Anna Kunz
- 5:00–5:30 Annotating the Implicit Content of Sluices
Pranav Anand and Jim McCloskey
- 5:30–6:00 Annotating Causal Language Using Corpus Lexicography of Constructions
Jesse Dunietz, Lori Levin, and Jaime Carbonell

6:00–6:10 **Closing**

Workshop 12: The Third Workshop on Metaphor in NLP

Organizers: *Ekaterina Shutova, Beata Beigman Klebanov, and Patricia Lichtenstein*

Venue: Director's Row J

Friday, June 5, 2015

9:00–9:05 **Opening remarks**

9:05–10:05 **Invited talk: "Metaphor in lexical resources" (Susan Windisch Brown and James H. Martin and Martha Palmer)**

Session 1

10:05–10:30 Effects of Situational Factors on Metaphor Detection in an Online Discussion Forum

Hyeju Jang, MiaoMiao Wen, and Carolyn Rose

10:30–11:00 **Coffee break**

Session 2

11:00–11:25 Supervised Word-Level Metaphor Detection: Experiments with Concreteness and Reweighting of Examples

Beata Beigman Klebanov, Chee Wee Leong, and Michael Flor

11:25–11:50 Modeling the interaction between sensory and affective meanings for detecting metaphor

Andrew Gargett and John Barnden

11:50–12:15 Exploring Sensorial Features for Metaphor Identification

Serra Sinem Tekiroglu, Gözde Özbal, and Carlo Strapparava

12:15–12:40 MetaNet: Deep semantic automatic metaphor analysis

Ellen Dodge, Jisup Hong, and Elise Stickles

12:40–2:15 **Lunch**

Session 3

2:15–2:40 High-Precision Abductive Mapping of Multilingual Metaphors

Jonathan Gordon, Jerry Hobbs, Jonathan May, and Fabrizio Morbini

2:40–3:05 A Corpus of Rich Metaphor Annotation

Jonathan Gordon, Jerry Hobbs, Jonathan May, Michael Mohler,

Fabrizio Morbini, Bryan Rink, Marc Tomlinson, and Suzanne Wertheim

3:05–3:30 Understanding Cultural Conflicts using Metaphors and Sociolinguistic Measures of Influence

Samira Shaikh, Tomek Strzalkowski, Sarah Taylor, John Lien, Ting Liu, George Aaron Broadwell, Laurie Feldman, Boris Yamrom, Kit Cho, and Yuliya Peshkova

3:30–4:00 **Coffee break**

Session 4

4:00–4:25 Chinese CogBank: Where to See the Cognitive Features of Chinese Words

Bin Li, Xiaopeng Bai, Siqi Yin, and Jie Xu

4:25–4:50 Fighting Words and Antagonistic Worlds

Tony Veale

Workshop 13: Third International Workshop on Natural Language Processing for Social Media

Organizers: *Shou-De Lin, Lun-Wei Ku, Cheng-Te Li, and Erik Cambria*

Venue: Governor's Square 11

Friday, June 5, 2015

9:20–9:30 **Opening**

9:30–10:30 **Keynote Speech: Variation and Change in Social Media Language (Jacob Eisenstein (Georgia Institute of Technology))**

10:30–11:00 **Coffee Break (I)**

11:00–12:00 **Natural Language Processing Session**

11:00–11:30 Location Name Disambiguation Exploiting Spatial Proximity and Temporal Consistency

Takashi Awamura, Daisuke Kawahara, Eiji Aramaki, Tomohide Shibata, and Sadao Kurohashi

11:30–12:00 Paraphrase Identification and Semantic Similarity in Twitter with Simple Features

Ngoc Phuoc An Vo, Simone Magnolini, and Octavian Popescu

12:00–1:30 **Lunch**

1:30–2:00 **Panel Discussion**

2:00–3:00 **Keynote Speech: Predicting Pragmatic Reasoning about Language Use in Context (Michael C. Frank (Stanford University))**

3:00–3:30 **Coffee Break (II)**

3:30–5:00 **Social Media Session**

3:30–4:00 A Language Detection System for Short Chats in Mobile Games
Pidong Wang, Nikhil Bojja, and Shivasankari Kannan

4:00–4:30 Long Nights, Rainy Days, and Misspent Youth: Automatically Extracting and Categorizing Occasions Associated with Consumer Products.
David Bracewell

4:30–5:00 A Deep Learning and Knowledge Transfer Based Architecture for Social Media User Characteristic Determination
Matthew Riemer, Sophia Krasikov, and Harini Srinivasan

5:00–5:20 **Best Paper Award/Closing**

Workshop 14: Vector Space Modeling for NLP

Organizers: *Phil Blunsom, Shay Cohen, Paramveer Dhillon, and Percy Liang*

Venue: Director's Row H

9:00–9:15 **Opening Remarks**

9:15–10:15 **Invited Talk (Chris Manning)**

10:15–12:15 **Poster session 1**

- A Simple Word Embedding Model for Lexical Substitution
Oren Melamud, Omer Levy, and Ido Dagan
- Unsupervised Text Normalization Using Distributed Representations of Words and Phrases
Vivek Kumar Rangarajan Sridhar
- A Multi-classifier Approach to support Coreference Resolution in a Vector Space Model
Ana Zelaia, Olatz Arregi, and Basilio Sierra
- Neural context embeddings for automatic discovery of word senses
Mikael Kägebäck, Fredrik Johansson, Richard Johansson, and Devdatt Dubhashi
- Distributional Representations of Words for Short Text Classification
Chenglong Ma, Weiqun Xu, Peijia Li, and Yonghong Yan
- Relation Extraction: Perspective from Convolutional Neural Networks
Thien Huu Nguyen and Ralph Grishman
- Distributional Semantic Concept Models for Entity Relation Discovery
Jay Urbain, Glenn Bushee, and George Kowalski
- A Deep Architecture for Non-Projective Dependency Parsing
Erick Fonseca and Sandra Aluísio
- Short Text Clustering via Convolutional Neural Networks
Jiaming Xu, Peng Wang, Guanhua Tian, Bo Xu, Jun Zhao, Fangyuan Wang, and Hongwei Hao
- A Word-Embedding-based Sense Index for Regular Polysemy Representation
Marco Del Tredici and Núria Bel
- Simple Semi-Supervised POS Tagging
Karl Stratos and Michael Collins
- Learning Distributed Representations for Multilingual Text Sequences
Hieu Pham, Thang Luong, and Christopher Manning
- Combining Distributed Vector Representations for Words
Justin Garten, Kenji Sagae, Volkan Ustun, and Morteza Dehghani

12:15–1:30 **Lunch**

1:30–2:30 **Invited Talk (Xavier Carreras)**

2:30–4:30 **Poster session 2**

- Dependency Link Embeddings: Continuous Representations of Syntactic Substructures
Mohit Bansal
- DeepNL: a Deep Learning NLP pipeline
Giuseppe Attardi
- A Vector Space Approach for Aspect Based Sentiment Analysis
Abdulaziz Alghunaim, Mitra Mohtarami, Scott Cyphers, and Jim Glass

- Word Embeddings vs Word Types for Sequence Labeling: the Curious Case of CV Parsing
Melanie Tosik, Carsten Lygteskov Hansen, Gerard Goossen, and Mihai Rotaru
- Morpho-syntactic Regularities in Continuous Word Representations: A multilingual study.
Garrett Nicolai, Colin Cherry, and Grzegorz Kondrak
- Towards Combined Matrix and Tensor Factorization for Universal Schema Relation Extraction
Sameer Singh, Tim Rocktäschel, and Sebastian Riedel
- Neural word embeddings with multiplicative feature interactions for tensor-based compositions
Joo-Kyung Kim, Marie-Catherine de Marneffe, and Eric Fosler-Lussier
- Bilingual Word Representations with Monolingual Quality in Mind
Thang Luong, Hieu Pham, and Christopher D. Manning
- Distributed Word Representations Improve NER for e-Commerce
Mahesh Joshi, Ethan Hart, Mirko Vogel, and Jean-David Ruvini
- Semantic Information Extraction for Improved Word Embeddings
Jiaqiang Chen and Gerard de Melo
- Named Entity Recognition for Arabic Social Media
Ayah Zirikly and Mona Diab
- Vector Space Models for Scientific Document Summarization
John Conroy and Sashka Davis
- Unsupervised Topic Modeling for Short Texts Using Distributed Representations of Words
Vivek Kumar Rangarajan Sridhar
- Estimating User Location in Social Media with Stacked Denoising Auto-encoders
Ji Liu and Diana Inkpen

4:30–5:30 **Invited Talk (Marco Baroni)**

5:30–5:45 **Concluding Remarks and Prizes**

5:45–7:00 **Farewell Reception**

8

Anti-harassment policy

The open exchange of ideas, the freedom of thought and expression, and respectful scientific debate are central to the aims and goals of a NAACL conference. These require a community and an environment that recognizes the inherent worth of every person and group, that fosters dignity, understanding, and mutual respect, and that embraces diversity. For these reasons, NAACL is dedicated to providing a harassment-free experience for participants at our events and in our programs.

Harassment and hostile behavior are unwelcome at any NAACL conference. This includes: speech or behavior (including in public presentations and on-line discourse) that intimidates, creates discomfort, or interferes with a person's participation or opportunity for participation in the conference. We aim for NAACL conferences to be an environment where harassment in any form does not happen, including but not limited to: harassment based on race, gender, religion, age, color, national origin, ancestry, disability, sexual orientation, or gender identity. Harassment includes degrading verbal comments, deliberate intimidation, stalking, harassing photography or recording, inappropriate physical contact, and unwelcome sexual attention.

It is the responsibility of the community as a whole to promote an inclusive and positive environment for our scholarly activities. In addition, any participant who experiences harassment or hostile behavior may contact any current member of the NAACL Board or contact Priscilla Rasmussen, who is usually available at the registration desk of the conference. Please be assured that if you approach us, your concerns will be kept in strict confidence, and we will consult with you on any actions taken.

The NAACL board members are listed at:

<http://naacl.org/officers/officers-2015.html>

The full policy and its implementation is defined at:

<http://naacl.org/policies/anti-harassment.html>

Index

- Abdelnasser, Heba, 110
AbdelRahman, Samir, 109
Abdelrahman, Samir, 116
Abend, Omri, 73
Acheampong, Chloe, 114
Ács, Judit, 104, 109
ADALI, Kübra, 127
Aditya, Somak, 86, 125
Adler, Meni, 104
Agarwal, Apoorv, 43, 122
Agerri, Rodrigo, 115
Aggarwal, Nitish, 104
Agichtein, Eugene, 54
Agirre, Eneko, 47, 86, 103, 104, 110, 111, 116, 129
Agrawal, Nishant, 52
Ahlberg, Malin, 66
Aizawa, Akiko, 78
Al Khatib, Khalid, 118
Al Tarouti, Feras, 127
Aldabe, Itziar, 116, 126
Alegria, Iñaki, 129
Aletras, Nikolaos, 103
Alex, Adarsh, 105
Alfonseca, Enrique, 70
Alghunaim, Abdulaziz, 138
Alhessi, Yousef, 114
Alhothali, Areej, 97
Allen, James, 116
Almeida, Mariana S. C., 117
Alpert-Abrams, Hannah, 66
Aluísio, Sandra, 104, 138
Álvarez-López, Tamara, 113
Alves, Ana, 110
Ambati, Bharat Ram, 27
Aminian, Maryam, 129
Amir, Silvio, 114
Ammar, Waleed, 39, 86
Anand, Pranav, 41, 135
Andrade, Daniel, 82
Andreas, Jacob, 37
Andrews, Nicholas, 78
Androutsopoulos, Ion, 113
Angelova, Galia, 110
Ann Dey, Shirin, 43
Apidianaki, Marianna, 111, 129
Apostolova, Emilia, 38
Aramaki, Eiji, 137
Aranberri, Nora, 129
Armstrong, William, 131, 132
Arora, Piyush, 109
Arora, Raman, 38
Arregi, Olatz, 116, 138
Arrigo, Michael, 103
Arroyo-Fernández, Ignacio, 53
Artetxe, Mikel, 129
Arthur, Philip, 46
Asooja, Kartik, 104
Astudillo, Ramón, 114
Atkins, David, 131
Attali, Yigal, 121
Attardi, Giuseppe, 129, 138
Auli, Michael, 33, 94
Aussenac-Gilles, Nathalie, 103
Avanço, Lucas, 134
Awamura, Takashi, 137
Ayetiran, Eniafe Festus, 111
Azab, Mahmoud, 79

- Bachrach, Yoram, 11, 17
Báez, David, 113
Baeza-Yates, Ricardo, 80
Bahgat, Reem, 109
Bai, Xiaopeng, 136
Baisa, Vít, 111
Baker, Collin, 11, 21, 134
Balasubramanian, Sriramkumar, 43
Baldwin, Timothy, 64, 74, 79, 104, 113, 127
Baldwin, Tyler, 49
Ballesteros, Miguel, 40, 76
Baltescu, Paul, 43
Bamman, David, 44
Banchs, Rafael, 121
Banea, Carmen, 111
Banjade, Rajendra, 109
Bansal, Mohit, 37, 45, 138
Bao Pham, Son, 110
Baral, Chitta, 86, 125
Barbieri, Francesco, 115
Barnden, John, 113, 136
Baroni, Marco, 32, 64, 104
Barrault, Loïc, 65
Barrena, Ander, 104
Barrón-Cedeño, Alberto, 110
Barzdins, Guntis, 117
Barzilay, Regina, 27, 72
Basile, Pierpaolo, 112, 114
Basu, Rupayan, 103
Baumer, Eric, 92
Béchara, Hanna, 109
Bechet, Frederic, 114, 115
Bedrick, Steven, 132
Beigman Klebanov, Beata, 136
Beinborn, Lisa, 98, 120
Bel, Núria, 104, 138
Belinkov, Yonatan, 27, 111
Bellot, Patrice, 114, 115
Ben Abacha, Asma, 112
Ben-David, Naama, 60
Benamara, Farah, 104
Benton, Adrian, 36
Berg-Kirkpatrick, Taylor, 66, 87
Berger, Jonah, 49
Bergler, Sabine, 113
Bertero, Dario, 108
Besacier, Laurent, 123
Bethard, Steven, 28, 44, 109, 112, 116, 134
Bhattacharyya, Pushpak, 74, 77, 80
Bhingardive, Sudha, 80
Bianchi, Francesca, 81
Bicici, Ergun, 108
Biemann, Chris, 64
Bies, Ann, 125, 135
Bilgic, Mustafa, 41
Bilu, Yonatan, 118
Bird, Steven, 104
Biswas, Kanad K., 55
Black, Alan W, 86
Black, Kevin, 61
Blain, Frédéric, 65
Blair, Elizabeth, 134
Blanco, Eduardo, 41, 133
Blunsom, Phil, 43, 138
Boag, William, 114
Boella, Guido, 111
Boguraev, Branimir, 103, 127
Bohnet, Bernd, 40, 76
Bojar, Ondrej, 111
Bojja, Nikhil, 137
Bokan, Alessandro, 134
Boleda, Gemma, 103
Boltužić, Filip, 119
Bordea, Georgeta, 104, 117
Bosco, Cristina, 115
Botros, Fadi, 117
Botvinick, Matthew, 104
Bouamor, Houda, 135
Bougares, Fethi, 65
Bouman, Danny, 69
Boyd-Graber, Jordan, 28, 41, 54, 69, 132
Boyer, Kristy, 60
Bracewell, David, 137
Bradbury, Jane, 111
Brennan, Jonathan, 124
Brew, Chris, 78
Briscoe, Ted, 39, 121
Broadwell, George Aaron, 136
Brockett, Chris, 33
Brooke, Julian, 122, 128
Brunato, Dominique, 134
Bryan, Craig, 132
Bucci, Wilma, 131, 132
Büchner, Michel, 114
Buitelaar, Paul, 104, 117
Bunescu, Aurelia, 3
Burford, Clint, 104
Burkhart, Clinton, 82
Burstein, Jill, 120
Buscaldi, Davide, 109, 117
Bushee, Glenn, 138
By, Tomas, 103
Byrne, Bill, 48, 65
- C. de Souza, José G., 49
Cabezudo, Marco, 134
Cahill, Aoife, 120, 121, 135
Cairelli, Michael, 133
Callison-Burch, Chris, 11, 14, 49, 92, 104, 108, 121
Camacho-Collados, José, 46
Cambria, Erik, 114, 137

- Cap, Fabienne, 122, 127
Caputo, Annalina, 112
Carbonell, Jaime, 104, 135
Cardie, Claire, 68, 100, 111, 118
Cardoso, Paula, 134
Carpuat, Marine, 129
Carstens, Lucas, 118
Caseli, Helena, 127
Caselli, Tommaso, 112, 116
Cassidy, Taylor, 44, 70
Castillo, Esteban, 113
Castillo, M. Dolores del, 131
Ceesay, Bamfa, 117
Celix-Salgado, Diego, 113
Cer, Daniel, 108, 111
Cerezo-Costas, Héctor, 113
Cervantes, Ofelia, 113
Chakraborti, Sutanu, 121
Chambers, Nathanael, 44, 116
Chandlee, Jane, 44
Chang, Jason, 78
Chang, Jim, 78
Chang, Kai-Wei, 11, 12
Chang, Ming-Wei, 43
Chang, Nancy, 134
Chapman, Wendy, 111, 116
Charniak, Eugene, 28
Che, Wanxiang, 31
Chen, Chen, 69
Chen, Jianfu, 42
Chen, Jiaqiang, 139
Chen, Lei, 120
Chen, MeiHua, 121
Chen, Qingcai, 110, 116
Chen, Wei-Fan, 121
Chen, Yun-Nung, 39
Chen, Zhiyuan, 54
Cheng, Doreen, 110
Cheng, Jian, 120
Chernyshevich, Maryna, 112
Cherry, Colin, 41, 62, 130, 139
Chiang, David, 47, 65, 83
Chikersal, Prerna, 114
Chisholm, Andrew, 73
Chitirala, Sai Charan Raj, 109
Cho, Kit, 136
Chodorow, Martin, 120, 121
Choi, Jinho D., 55
Choi, Yejin, 32, 36, 42
Cholakov, Kostadin, 127
Choudhary, Narayan, 112
Chow, Kam-Pui, 45
Christensen, Lee, 116
Chuang, Jason, 33
Cieliebak, Mark, 114
Cimiano, Philipp, 60
Cinkova, Silvie, 111, 117
Ciobanu, Alina Maria, 116
Clark, Jonathan H., 44
Clark, Micah, 103
Clark, Peter, 36, 60
Claudino, Leonardo, 131, 132
Cleuziou, Guillaume, 117
Cohan, Arman, 66
Cohen, Kevin B., 159
Cohen, Shay B., 73
Cohen, Shay, 138
Cohn, Trevor, 79
Collins, Michael, 138
Collins, Riley, 114
Colmenares, Carlos A., 31
Compton, Ryan, 119
Conroy, John, 139
Conway, Mike, 132
Cook, Paul, 64, 127
Coppersmith, Glen, 131
Corman, Julien, 103
Corpas Pastor, Gloria, 109
Correnti, Richard, 120
Costa, Hernani, 109
Costa-Montenegro, Enrique, 113
Cotterell, Ryan, 62, 86
Couto, Francisco, 112
Craneburgh, Andreas van, 122, 127
Cristea, Alexandra, 114
Cromieres, Fabien, 74
Cuadros, Montse, 115
Culotta, Aron, 33
Cummins, Ronan, 121
Cuong, Hoang, 49
Cybulska, Agata, 125
Cyphers, Scott, 111, 138
Da San Martino, Giovanni, 110
Dabre, Raj, 74
Dagan, Ido, 45, 64, 104, 138
Dai, Hong-Jie, 112
Dalbelo Bašić, Bojana, 109
Dalmia, Ayushi, 113
Dalton, Adam, 103
Dalton, Jeff, 54
Danesh, Soheil, 104
Danforth, Douglas, 120
Dani, Kinjal, 112
Daniele, Falavigna, 49
Darwish, Kareem, 27, 110
Daudaravicius, Vidas, 120
Daumé III, Hal, 11, 12, 50, 62, 95
Davis, Sashka, 139
Dayrell, Carmen, 81
De Clercq, Orphee, 115
de Gispert, Adrià, 65

- D'Egidio, Angela, 81
Dehghani, Morteza, 138
Del Tredici, Marco, 138
Dell'Orletta, Felice, 134
Delmonte, Rodolfo, 122
Demberg, Vera, 26, 104, 124
Dembowski, Julia, 80
Deng, Dun, 129
Deng, Li, 61
Deng, Lingjia, 52, 87
Deng, Qiao, 111
Deoskar, Tejaswini, 27
Derczynski, Leon, 114, 116
Deri, Aliya, 35
DeYoung, Jay, 78
D'Haro, Luis Fernando, 121
Dhillon, Paramveer, 138
Di Felippo, Ariani, 134
Diab, Mona, 103, 104, 111, 129, 133, 139
Dias, Gaël, 117
Dias, Márcio, 134
Dickinson, Markus, 120, 135
Ding, Haibo, 87
Dinu, Liviu P., 116
Dodge, Ellen, 136
Dodge, Jesse, 100
Doing-Harris, Kristina, 112
Dolan, Bill, 33, 108
Dolan, William B., 92
Donahue, Jeff, 93
Dong, Li, 113
Dong, Yuanzhe, 65
Dos Reis, Julio Cesar, 112
Dou, Dejing, 80
Doulkeridis, Christos, 115
Doyle, Gabriel, 99, 124
Dragoni, Mauro, 113
Dras, Mark, 85, 121
Dredze, Mark, 26, 36, 44, 78, 84, 131
Dreyer, Markus, 65, 75
D'Souza, Jennifer, 116
Du, Yantao, 117
Dubhashi, Devdatt, 138
Dubremetz, Marie, 122
Duh, Kevin, 46, 61
Dunietz, Jesse, 135
Dupoux, Emmanuel, 46, 62
Durand, Guillaume, 120
Durrett, Greg, 37, 73
Dušek, Ondřej, 135
Dworman, Seth, 117
Dyer, Chris, 39, 40, 44, 86, 100
Dykes, Natalie, 114
Ebert, Sebastian, 113
Ebrahimi, Javid, 80
Eckle-Kohler, Judith, 118
Etemadi, Sauleh, 54
Eger, Steffen, 104
Egg, Markus, 127
Egger, Dominic, 114
Eichler, Kathrin, 104
Eichstaedt, Johannes, 49, 131
Eisenstein, Jacob, 48, 100
Eisner, Jason, 62
Ekbal, Asif, 114
El Maarouf, Ismail, 111
Elfandy, Heba, 104
Elhadad, Noémie, 111
Ellis, Joe, 125
Ellsworth, Michael, 11, 21
Elovic, Elisha, 92
Elvevåg, Brita, 132
Ermer, Heiko, 114
Erp, Marieke van, 116
Eryigit, Gülşen, 127, 134
Espinosa Anke, Luis, 117
Eswaran, Dhivya, 121
Ettinger, Allyson, 50
Evert, Stefan, 72, 109, 114, 123
Exner, Peter, 105
Eyecioglu, Asli, 108
Eyraud, Remi, 44
Fabro, Marcos Didonet Del, 116
Fakhrahmad, Mostafa, 110
Fan, Xiangmin, 75
Fancellu, Federico, 129, 133
Fanta, Petr, 111
Farahmand, Meghdad, 127
Faralli, Stefano, 117
Farra, Noura, 120
Faruqui, Manaal, 79, 100
Fauceglia, Nicolas R., 125
Fauconnier, Jean-Philippe, 105
Feizabadi, Parvin Sadat, 103
Feldman, Anna, 122
Feldman, Laurie, 136
Felice, Mariano, 39
Felt, Paul, 61, 134
Feltracco, Anna, 125
Feng, Vanessa Wei, 69
Feng, Yukun, 111
Ferguson, James, 37
Fernandez, Raquel, 124
Fernández-Gavilanes, Milagros, 113
Ferraro, Francis, 78
Ferrugento, Adriana, 110
Filho, Pedro, 134
Filice, Simone, 110
Filippova, Katja, 70
Finin, Tim, 108

- Fisas, Beatriz, 134
Flanigan, Jeffrey, 11, 16, 68
Flickinger, Dan, 117
Flor, Michael, 136
Fokkens, Antske, 116
Foland, William, 105
Foltz, Peter, 120, 121, 132
Fonseca, Erick, 138
Forsberg, Markus, 66
Fosler-Lussier, Eric, 50, 120, 133, 139
Foster, Jennifer, 109
Fothergill, Richard, 113
Fox Tree, Jean E., 41
Frank, Michael, 99, 124
Fraser, Alexander, 129
Fraser, Kathleen C., 60
Frermann, Lea, 99
Fricke, Matthias, 134
Friedrich, Annemarie, 134
Fucikova, Eva, 135
Fujita, Atsushi, 39
Fung, Pascale, 108
Futagi, Yoko, 121
Fyshe, Alona, 26
- Galanis, Dimitris, 113
Galley, Michel, 33
Gao, Hang, 108
Gao, Jianfeng, 11, 19, 33, 61
Gao, Mingkun, 49
Gao, Wei, 110
Garcia Flores, Jorge, 109
García Pablos, Aitor, 115
Gargett, Andrew, 136
Garrette, Dan, 66
Garten, Justin, 138
Gaspers, Judith, 60
Gautam, Dipesh, 109
Gay, Geri, 92
Genabith, Josef van, 109, 117
Georgilidakis, Spiros, 124
Gershman, Anatole, 104
Gertz, Michael, 116
Ghazvininejad, Marjan, 98
Ghiasvand, Omid, 112
Ghoneim, Mahmoud, 129
Ghosh, Aniruddha, 113
Gilberto Mateos Ortiz, Luis, 93
Gildea, Daniel, 32, 64
Giménez, Mayte, 114
Gimpel, Kevin, 37, 45
Ginter, Filip, 76, 117
Girju, Roxana, 125
Glass, James, 111
Glass, Jim, 111, 138
Glavaš, Goran, 109, 112
- Globerson, Amir, 27, 81
Göckel, Thierry, 75
Goharian, Nazli, 66
Goikoetxea, Josu, 86
Goldberg, Adele, 104
Goldberg, Yoav, 81
Goldberger, Jacob, 45
Gomez, Helena, 108
Gómez, Jon Ander, 115
Gong, Li, 111
González Pellicer, Edgar, 105
Gonzalez, Graciela, 113
González-Agirre, Aitor, 110, 111
González-Castaño, Francisco Javier, 113
Goodkind, Adam, 128
Goossen, Gerard, 139
Goot, Rob van der, 108
Gordon, Jonathan, 136
Gorinski, Philip John, 68
Gorman, Kyle, 132
Gorman, Sharon, 111
Gormley, Matthew R., 78, 84
Gorrell, Genevieve, 116
Gosko, Didzis, 117
Goutte, Cyril, 120
Gouws, Stephan, 84
Graehl, Jonathan, 75
Graham, Naida, 60
Graham, Yvette, 74
Green, Nancy, 118
Greenberg, Clayton, 26, 124
Grefenstette, Gregory, 117
Grimmer, Justin, 33
Grishman, Ralph, 36, 76, 138
Groza, Tudor, 109
Guerini, Marco, 92
Guerra, Raul, 105
Guha, Anupam, 69
Guha, Satarupa, 114, 115
Gung, James, 112
Guo, Hongyu, 41, 103
Guo, Shiman, 129
Guo, Weiwei, 111
Guo, Yufan, 51
Gupta, Abhijeet, 104
Gupta, Manish, 113
Gupta, Parth, 115
Gupta, Rohit, 109, 117
Gupta, Sonal, 79
Gurevych, Iryna, 98, 118, 120
Guthrie, Louise, 103
Gutiérrez-Vasques, Ximena, 55
- Ha Vo, Nguyen, 86
Haas, Carolin, 83
Haas, Michael, 40

- Habash, Nizar, 135
Hachey, Ben, 73
Haertel, Robbie, 61, 134
Haffari, Gholamreza, 61
Hagen, Matthias, 114, 118
Hajic, Jan, 117, 135
Hajishirzi, Hannaneh, 35, 60, 81
Hakala, Kai, 112
Hale, John, 124
Halkidi, Maria, 115
Hallgren, Kevin, 131
Hamdan, Hussam, 114, 115
Hammond, Adam, 122, 128
Han, Lushan, 110
Han, Sangdo, 79
Han, Xu, 114
Hänig, Christian, 111
Hao, Hongwei, 138
Harman, Craig, 78, 131
Hart, Ethan, 139
Hassan, Basma, 109
Hassan, Hany, 74, 94
Hassanzadeh, Hamed, 109
Hastings, Jenny, 86
Hateva, Nelly, 110
Hautli-Janisz, Annette, 125
Hazem, Amir, 65
He, He, 11, 12
He, Hua, 94
He, Xiaodong, 11, 19, 61, 72
He, Yifan, 36, 76
Hearne, James, 82
Heer, Jeffrey, 33
Heid, Christopher, 105
Heider, Abeer, 135
Heilman, Michael, 66, 120, 121
Heinz, Jeffrey, 44
Henderson, John, 108
Hermjakob, Ulf, 70
Hernández Farías, Delia Irazú, 115
Hershovich, Daniel, 118
Hewitt, Luke, 77
Heydari Alashty, Amin, 110
Hicks, Tom, 75
Hieber, Felix, 74
Higashinaka, Ryuichiro, 126
Hirao, Tsutomu, 50
Hirschberg, Julia, 103
Hirst, Graeme, 60, 69, 118, 122, 128
Hixon, Ben, 60
Hladka, Barbora, 75
Ho, Quoc, 112
Hobbs, Jerry, 136
Hoenen, Armin, 83
Hoey, Jesse, 97
Hokamp, Chris, 79, 109
Hollingshead, Kristy, 131
Holm, Susan, 125
Homan, Christopher, 85
Hong, Jisup, 136
Hoste, Veronique, 115
Hou, Yongshuai, 110, 116
Hovy, Dirk, 80
Hovy, Eduard, 40, 100, 125
Hu, Baotian, 110
Huang, Chu-Ren, 114
Huang, Jingwen, 69
Huang, Jonathan, 32
Huang, Liang, 32, 66, 81
Huang, Mingxuan, 121
Hulden, Mans, 66
Hunter, Jane, 109
Hurdle, John, 112
Hurtado, Lluís-F., 114
Huynh, David, 134
Huynh, Nghia, 112
Hwang, Dosam, 115
Hwang, Jena D., 103, 135
Hwang, William, 35

Ighe, Ann, 123
Iglesias, Gonzalo, 65
Igo, Sean, 112
Iida, Ryu, 38
Imel, Zac, 131
Inclezan, Daniela, 86
Ingham, Rosemary, 132
Inkpen, Diana, 118, 139
Inoue, Naoya, 118
Intxaurreondo, Ander, 47
Inui, Kentaro, 118, 119, 134
Iosif, Elias, 124
Isabelle, Pierre, 39
Islam, Aminul, 109
Iwayama, Makoto, 119
Iyyer, Mohit, 69
Izquierdo, Ruben, 111

Jacob, David, 128
Jacobs, Cassandra L., 124
Jaffe, Evan, 109, 120
Jagfeld, Glorianna, 52
Jaggi, Martin, 114
Jang, Hyeju, 136
Jannidis, Fotis, 123
Jansen, Aren, 46
Jansen, Peter, 36
Jauhar, Sujay Kumar, 40, 100, 105
Jayannavar, Prashant, 122
Jeong, Minwoo, 28
Jezek, Elisabetta, 125
Ji, Heng, 70, 83

- Ji, Yangfeng, 33, 92
Jiang, Min, 111
Jiménez-Zafra, Salud M., 115
Jin, Lifeng, 64, 109
Jing, Wang, 45
Johansson, Fredrik, 138
Johansson, Richard, 81, 138
Johnson, Mark, 27, 46
Johnson, Rie, 28
Johnston, Nicholas, 32
Jones, Gareth, 109
Jonnagaddala, Jitendra, 112
Joshi, Aditya, 114, 115
Joshi, Mahesh, 139
Joty, Shafiq, 110
Ju, Melody, 122
Juan Hou, Wen, 117
Judea, Alex, 104
Julmy, Pascal, 114
Juncal-Martínez, Jonathan, 113
Jung, Jason J., 115
Jurafsky, Dan, 40
Jurczyk, Tomasz, 55
Jurgens, David, 82, 104, 108
- Kalita, Jugal, 127
Kallmeyer, Laura, 84
Kamath, Shruti, 43
Kamel, Mouna, 105
Kamioka, Yudai, 134
Kanerva, Jenna, 76, 117
Kannan, Shivasankari, 137
Kanno, Miwa, 134
Karampats, Rafael - Michael, 109
Karan, Mladen, 109
Karanašiou, Aikaterini, 112
Karanasou, Maria, 115
Karkaletsis, Vangelis, 118
Karttunen, Lauri, 133
Karumuri, Sakethram, 109
Katakis, Ioannis Manousos, 118
Kate, Rohit, 112
Katiyar, Arzoo, 118
Kato, Yoshihide, 105
Kauer, Anderson, 115
Kautz, Henry, 32
Kawahara, Daisuke, 126, 137
Kay, Paul, 127
Kazantseva, Anna, 122
Keane, Nathan, 125, 126
Keller, Bill, 108
Kennington, Casey, 38
Kern, Margaret, 49
Khashabi, Daniel, 50
Khoddam, Ehsan, 26
Kiesel, Johannes, 118
- Kilgarriff, Adam, 111
Kilicoglu, Halil, 133
Kim, Byungsoo, 79
Kim, Joo-Kyung, 139
Kim, Young-Bum, 28
King, David, 109
King, Tracy Holloway, 125
Kirchhoff, Katrin, 65
Kiritchenko, Svetlana, 50, 113
Kirschner, Christian, 118
Kiss, Geza, 132
Klang, Marcus, 105
Klein, Dan, 37, 66, 73, 87
Knight, Kevin, 35, 70, 98
Kochmar, Ekaterina, 121
Koenig, Jean-Pierre, 128
Kohl, Micha, 114
Kokkinakis, Dimitrios, 123
Kondrak, Grzegorz, 43, 62, 130, 139
Kong, Lingpeng, 42
Koolen, Corina, 122
Koppula, Akshay Reddy, 115
Kordjamshidi, Parisa, 117
Kordoni, Valia, 127
Korhonen, Anna, 51
Kornai, Andras, 104
Kosinski, Michal, 49
Kotlerman, Lili, 104
Kotti Padannayil, Soman, 108
Kowalski, George, 138
Kozareva, Zornitsa, 82
Kågebäck, Mikael, 138
Kraeva, Marina, 110
Krasikov, Sophia, 137
Krause, Sebastian, 70
Krishna, Vamsi, 114
Krishnan, Vinodh, 100
Kříž, Vincent, 75
Krug, Markus, 123
Kruszewski, Germán, 64
Ku, Lun-Wei, 121, 137
Kübler, Sandra, 122
Kuhlmann, Marco, 117
Kuhn, Jonas, 122
Kulick, Seth, 125
Kulshreshtha, Rajat, 52
Kumar, Akshay, 52
Kumar, Ayush, 114
Kumar, Girish, 121
Kumar, Manish, 112
Kumar, Shankar, 79
Kunchukuttan, Anoop, 77
Kunz, Kerstin Anna, 135
Kurohashi, Sadao, 74, 126, 137
Kuznetsova, Polina, 42
Kwon, Soonchoul, 79

- Labaka, Gorka, 129
Lai, Albert M., 50, 133
Lam, Khang Nhut, 127
Lambrou-Latreille, Konstantinos, 52
Lan, Man, 108–110, 113, 115
Langford, John, 11, 12
Laparra, Egoitz, 126
Lapata, Mirella, 68, 72, 93, 99
Lapesa, Gabriella, 72, 109
Lapshinova-Koltunski, Ekaterina, 135
Largeron, Christine, 117
Laszlo, Anna, 114
Lavie, Alon, 44
Lavrenko, Victor, 85
Lawrence, John, 119
Lazaridou, Angeliki, 32
Le, Phong, 48, 103
Leacock, Claudia, 120
Leal, André, 112
Ledbetter, Scott, 120
Lee, Chong Min, 120
Lee, Gary Geunbae, 79
Lee, Jackson, 55
Lee, Joohyung, 86
Lee, Lillian, 24
Lefever, Els, 115, 117
Leger, Serge, 120
Lehnens, Patrick, 94
Lei, Tao, 27, 72
Leong, Chee Wee, 136
Lerner, Andreas, 114
Levin, Lori, 86, 135
Levinboim, Tomer, 47, 83
Levine, Aaron, 117
Levorato, Vincent, 117
Levow, Gina-Anne, 112
Levy, Omer, 64, 138
Lewis, William, 66
Li, Bin, 136
Li, Binyang, 114
Li, Cheng-Te, 137
Li, Chengtao, 27
Li, Chen, 42, 82, 125
Li, Fei-Fei, 90
Li, Guofu, 113
Li, Peijia, 113, 138
Li, Sujian, 83
Li, Xiaolong, 60
Li, Yunyao, 49
Liakata, Maria, 114
Liang, Huizhi, 113
Liang, Percy, 138
Liaw, Siaw-Teng, 112
Lichtenstein, Patricia, 136
Lien, John, 136
Lin, Chin-Yew, 83
Lin, Chu-Cheng, 86
Lin, Jiaxin, 110
Lin, Jimmy, 94
Lin, Lei, 109
Lin, Shou-De, 137
Lin, Yiu-Chang, 125
Lin, Yiye, 65
Ling, Wang, 86, 114
Lintean, Mihai, 109
Litman, Diane, 75, 118, 120, 121
Litvak, Marina, 31
Liu, Chao, 65
Liu, Fei, 31, 35, 68
Liu, Ji, 139
Liu, Qiguang, 32
Liu, Ting, 69, 136
Liu, Xiaodong, 61
Liu, Yang, 42, 44, 82, 109
Liu, Yi, 85
Liu, Yijia, 31
Liu, Yudong, 82
Liu, Zhengzhong, 125
Livescu, Karen, 37
Llorens, Hector, 116
Long, Cotie, 104
Lopez de Lacalle, Oier, 47, 103
Lopez, Adam, 76, 94
Lopez, Melissa, 121
Lopez, Roque, 134
Lopez-Gazpio, Inigo, 110, 111
Loukachevitch, Natalia, 127
Loukina, Anastassia, 120
Lu, Ang, 37
Lu, Wei-Lwun, 54
Lu, Weiming, 84
Lu, Ziyu, 45
Ludusan, Bogdan, 62
Luh, Wen-Ming, 124
Lumpkin, Barry, 86
Lund, Jeffrey, 41
Luo, Jiebo, 32
Luo, Liang, 82
Luo, Wencan, 75
Luong, Thang, 138, 139
Luotolahti, Juhani, 76, 117
Lutz, David, 124
Lygteskov Hansen, Carsten, 139
Lynch, Gerard, 117

M. El-Makky, Nagwa, 110
Ma, Chenglong, 113, 138
Ma, Jing, 114
Ma, Xiaoyi, 125
Ma, Xuezhe, 125
Maas, Andrew, 40
Macharowsky, Luisa, 123
-

- Madasamy, Anand Kumar, 108
Madnani, Nitin, 66, 120, 121
Magdy, Walid, 110, 111
Magnini, Bernardo, 116, 125
Magnolini, Simone, 108–110, 137
Maharjan, Nabin, 109
Maharjan, Suraj, 134
Mahidadia, Ashesh, 31
Maier, Wolfgang, 84
Makrai, Márton, 104
Malm, Mats, 123
Malmasi, Shervin, 85, 120, 121
Malmaud, Jonathan, 32
Manandhar, Suresh, 111, 113
Manandise, Esme, 127
Manion, Steve L., 112
Manning, Christopher D., 79, 139
Manning, Christopher, 138
Manrique-Losada, Bell, 125
Mantrach, Amin, 31
Marchetti, Galen, 68
Marcu, Daniel, 4
Marie, Benjamin, 129
Maritxalar, Montse, 110, 111
Markantonatou, Stella, 127
Marneffe, Marie-Catherine de, 50, 133, 139
Márquez, Lluís, 72, 110, 111
Martin, Eric, 31
Martin, James H., 104
Martin, James, 105
Martin, Trish, 121
Martín-Valdivia, M. Teresa, 115
Martineau, Justin, 110
Martínez Alonso, Héctor, 80, 84, 115, 135
Martínez-Cámara, Eugenio, 115
Martins de Matos, David, 104
Martins, André F. T., 117
Martins, Bruno, 112, 114
Martschat, Sebastian, 75
Marujo, Luis, 104
Marvel, Aron, 128
Maskit, Bernard, 131, 132
Mathew, Ditty, 121
Mathur, Nitika, 74, 127
Matos, Sérgio, 112
Matsubara, Shigeki, 105
Matsuda, Koji, 134
Matsuо, Yoshihiro, 126
Matwin, Stan, 118
May, Daniel, 114
May, Jonathan, 136
McCloskey, Jim, 135
McCurdy, Nina, 122
McDonald, Ryan, 40
McDowell, Bill, 44
McGillion, Sarah, 115
McMahan, Brian, 93
Mehler, Alexander, 104
Meisen, Philipp, 134
Melamud, Oren, 45, 138
Mélanie, Frédérique, 76
Melo, Gerard de, 139
Mendes, Pablo, 105
Mendes, Sara, 104
Menekse, Muhsin, 75
Menezes, Arul, 75
Merkhofer, Elizabeth M., 108
Mesgar, Mohsen, 105
Meyer, Miriah, 122
Meyers, Adam, 134
Meza, Ivan V., 109
Mi, Haitao, 66, 94
Miceli Barone, Antonio Valerio, 129
Michaelis, Laura A., 127
Mihalcea, Rada, 79, 111
Milios, Evangelos, 109
Mille, Simon, 40, 76
Mills, Chad, 112
Minard, Anne-Lyse, 116
Minato, Shin-ichi, 50
Mirza, Paramita, 116
Misra, Amita, 41
Mitamura, Teruko, 125
Mitchell, Margaret, 33, 131
Mitchell, Tom M., 26
Mitkov, Ruslan, 109
Miura, Naoko, 109
Miwa, Makoto, 64
Miyao, Yusuke, 117
Miyoshi, Toshinori, 119
Mizuno, Junta, 134
Modi, Ashutosh, 104
Moens, Marie-Francine, 109, 117
Mohamed, Reham, 110
Mohammad, Saif, 50, 113
Mohammady Ardehaly, Ehsan, 33
Mohammed, Metrah, 132
Mohammed, Saif, 97
Mohit, Behrang, 135
Mohler, Michael, 136
Mohtarami, Mitra, 111, 138
Montemagni, Simonetta, 134
Montes, Manuel, 28
Mooney, Raymond, 93
Morante, Roser, 116, 133
Morbini, Fabrizio, 136
Moreira, Viviane, 115
Morita, Hajime, 126
Morley, Eric, 132
Moro, Andrea, 111
Moschitti, Alessandro, 72, 80, 105, 110, 111,
113

- Mostafazadeh, Nasrin, 116
Mott, Greg, 105
Mott, Justin, 125
Moulahi, Bilel, 116
Movva, Venkata Subhash, 115
Mowery, Danielle L, 116
Mowery, Danielle, 132
Mrabet, Yassine, 112
Mubarak, Hamdy, 110
Muehr, Sebastian, 94
Müller, Thomas, 38
Mundkowsky, Robert, 78
Murawaki, Yugo, 47
Muresan, Smaranda, 118
Murphy, Brian, 26, 124
Murphy, Kevin, 32
Murphy, Sean, 131, 132

Nagata, Masaaki, 45, 50
Nagesh, Ajay, 52, 61
Naim, Iftekhar, 32
Nakamura, Satoshi, 76
Nakov, Preslav, 108, 110, 111, 113
Napoles, Courtney, 121
Napolitano, Diane, 78
Naradowsky, Jason, 77
Narita, Kazuya, 134
Nastase, Vivi, 104
Naung, Zaw, 86
Navarro, Borja, 116, 123
Navigli, Roberto, 46, 77, 104, 111, 117
Neculescu, Silvia, 104
Nedoluzhko, Anna, 135
Neelakantan, Arvind, 43
Negri, Matteo, 49
Nemeskey, Dávid Márk, 104
Nenkova, Ani, 35, 69, 82
Neto, João, 104
Neubig, Graham, 46, 76
Neumann, Stella, 134
Ney, Hermann, 94
Ng, Andrew, 40
Ng, Hwee Tou, 46
Ng, Vincent, 69, 73, 116
Nguyen, Anthony, 109
Nguyen, Hoang Long, 115
Nguyen, Huy, 118
Nguyen, Minh, 110
Nguyen, Thang, 41, 131, 132
Nguyen, Thien Huu, 138
Nguyen, Trung Duc, 115
Nguyen, Viet-An, 132
Ni, Aiming, 55
Nichols, Eric, 117
Nicolai, Garrett, 43, 62, 139
Nicosia, Massimo, 110

Niculae, Vlad, 116
Nie, Jian-Yun, 33, 115
Niemietz, Paula, 134
Nieto Piña, Luis, 81
Nikfarjam, Azadeh, 113
Nikolova, Ivelina, 110
Nirala, Nobal Bikram, 109, 121
Nirmal, Manju, 127
Nishino, Masaaki, 50
Nivre, Joakim, 122, 127
Niwa, Yoshiaki, 119
Noble, Bill, 124
Nóbrega, Fernando, 134
Noh, Tae-Gil, 104
Nokel, Michael, 127
Noord, Gertjan van, 108
Novielli, Nicole, 114
Nugues, Pierre, 105
Nye, Benjamin, 82

Oard, Douglas, 46
Och, Franz, 100
O'Connor, Brendan, 98
Oda, Yusuke, 76
Oepen, Stephan, 117
Oflazer, Kemal, 135
O'Gorman, Tim, 11, 16
Okazaki, Naoaki, 118, 134
Okumura, Manabu, 79
Oliva, Jesús, 131
Oliveira, Hugo Gonçalo, 110
Oliveira, José Luis, 112
Olshefski, Emily, 125
Ono, Masataka, 64
Oraby, Shereen, 119
Orasan, Constantin, 109
Ordan, Noam, 116
Osborne, John, 112
Osborne, Miles, 85
Ostendorf, Mari, 35, 81
O'Donnell, Timothy J., 124
Otmakhova, Yulia, 97
Ou, Gaoyan, 114
Ovesdotter Alm, Cecilia, 85
Özbal, Gözde, 92, 136
Özdemir, Canberk, 113

Padó, Sebastian, 103, 104
Paetzold, Gustavo, 51
Paik, Jiaul, 46
Paikens, Peteris, 117
Pajkossy, Katalin, 104
Pallerla, Ranga Reddy, 115
Palmer, Alexis, 134
Palmer, Martha, 103, 104, 125, 130, 135
Pamay, Tugba, 127, 134

- Pan, Xiaoman, 70
Panchal, Vishal, 112
Papadakis, Katina, 132
Papageorgiou, Haris, 113
Paperno, Denis, 104
Pardo, Thiago, 134
Parikh, Ankur P., 42
Paritosh, Praveen, 134
Park, Gregory, 49, 131
Park, Joon Suk, 118
Park, Seonyeong, 79
Pasca, Marius, 39
Patel, Amrish, 112
Patel, Pinal, 112
Pater, Joe, 46
Pathak, Parth, 112
Patti, Viviana, 115
Paul, Michael J., 44
Pavlick, Ellie, 11, 14, 35
Pearl, Lisa, 124
Peate Sørensen, Melissa, 134
Pedersen, Ted, 112, 131
Peitz, Stephan, 94
Peldszus, Andreas, 119
Pelillo, Marcello, 111
Peng, Haoruo, 50
Peng, Nanyun, 78
Perera, Sujan, 105
Pershina, Maria, 36
Peshkova, Yuliya, 136
Petasis, Georgios, 118
Peters, Wim, 118
Petruck, Miriam R L, 11, 21
Pham, Hieu, 138, 139
Pham, Nghia The, 32
Pham, Son Bao, 110
Phillips, Lawrence, 124
Piao, Scott, 81
Pielström, Steffen, 123
Pighin, Daniele, 70
Pijetlovic, Dijana, 120
Pilehvar, Mohammad Taher, 46, 77, 82
Pinkal, Manfred, 104, 134
Pinto, David, 108
Pitler, Emily, 40
Pla, Ferran, 114
Plank, Barbara, 80, 84, 115, 135
Plas, Lonneke van der, 52
Plotnikova, Nataliia, 109, 114
Poibeau, Thierry, 76, 104, 112
Polletta, Francesca, 92
Pontiki, Maria, 113
Poon, Hoifung, 26, 42
Popescu, Octavian, 53, 108–111, 116, 137
Poria, Soujanya, 114
Post, Matt, 3
Postma, Marten, 111
Potamianos, Alexandros, 124
Potash, Peter, 114
Potthast, Martin, 114
Pourashraf, Payam, 38
Pourabzi-Sangdeh, Forough, 54
Prabhakaran, Vinod Kumar, 103, 133
Pradhan, Sameer, 48, 111
Preoțiuc-Pietro, Daniel, 131
Price, David, 118
Procter, Rob, 114
Proisl, Thomas, 109, 123
Prokofieva, Anna, 103
Ptucha, Raymond, 85
Puduppully, Ratish, 77
Puente, Xose de la, 111
Puppe, Frank, 123
Pustejovsky, James, 116, 117
Pyysalo, Sampo, 76
Qian, Xian, 44
Qin, Bing, 31, 69
Qin, Ying, 92
Quirk, Chris, 75
Rabinovich, Andrew, 32
Ragab, Maha, 110
Ragheb, Marwa, 135
Rahimi, Afshin, 79
Rahimi, Zahra, 120
Rahmani, Saeed, 110
Rakib, Md Rashadul Hasan, 109
Rama, Taraka, 84
Ramachandran, Lakshmi, 120, 121
Ramakrishnan, Ganesh, 61
Rambow, Owen, 103, 122, 133
Ramchand, Gilian, 125
Ramisch, Carlos, 127
Randeree, Bilal, 111
Rangarajan Sridhar, Vivek Kumar, 138, 139
Rao, Sudha, 11, 12, 50
Rastogi, Pushpendre, 38
Rathod, Vivek, 32
Ray, Pradeep, 112
Rayson, Paul, 81
Recasens, Marta, 105
Recski, Gábor, 104, 109
Reddy, Siva, 72
Reddy, Sravana, 77
Redkar, Hanumant, 80
Reed, Chris, 119
Reed, Lena, 119
Refaee, Eshrag, 53
Reger, Isabella, 123
Rehbein, Ines, 133, 134
Reichart, Roi, 51

- Reisert, Paul, 118, 119
Reitter, David, 124
Rello, Luz, 80
Remus, Robert, 111
Remus, Steffen, 64
Repaka, Ravikanth, 115
Resnik, Philip, 50, 92, 131, 132
Reyes, Antonio, 113
Ribeiro, Ricardo, 104
Ribeyre, Corentin, 27
Richey, Colleen, 65
Riedel, Sebastian, 70, 77, 139
Riemer, Matthew, 137
Riese, Tom, 125
Rieser, Verena, 53
Riezler, Stefan, 74, 83
Rigau, German, 110, 111, 115, 116, 126
Riloff, Ellen, 87, 119
Rindflesch, Thomas, 133
Ringger, Eric, 41, 61, 134
Rink, Bryan, 136
Ritter, Alan, 92, 113
Ritter, Samuel, 104
Roberts, Angus, 116
Roberts, Margaret E., 33
Rochon, Elizabeth, 60
Rocktäschel, Tim, 70, 77, 139
Rodriguez, Isaac, 109
Rohrbach, Marcus, 93
Rondon, Alexandre, 127
Ronzano, Francesco, 115, 117, 134
Roostaei, Meysam, 110
Rose, Carolyn, 136
Rosemblat, Graciela, 133
Rosenberg, Andrew, 128
Rosenfeld, Alex, 120
Rosenstein, Mark, 132
Rosenthal, Sara, 113
Rösiger, Ina, 122
Rosso, Paolo, 103, 113
Rotaru, Mihai, 139
Roth, Dan, 50, 51, 85, 98
Rouhizadeh, Masoud, 132
Roy, Subhro, 51
Rozovskaya, Alla, 98, 135
Rudinger, Rachel, 104
Rudnický, Alexander, 39
Rudzewitz, Björn, 110
Ruffo, Giancarlo, 115
Ruiz, Pablo, 76, 104, 112
Rumshisky, Anna, 114
Rus, Vasile, 109, 121
Rush, Alexander M., 42
Russo, Irene, 112
Rutherford, Attapol, 43
Ruvini, Jean-David, 139
Ryant, Neville, 125
Sack, Jeffrey, 38
Sadamasu, Kunihiko, 82
Sadeh, Norman, 68
Saenko, Kate, 93
Sagae, Kenji, 138
Saggion, Horacio, 115, 117, 134
Saias, José, 115
Sakaguchi, Keisuke, 66
Sakti, Sakriani, 76
Salaberri, Haritz, 116
Salaberri, Iker, 116
Salameh, Mohammad, 50, 130
Saleh, Iman, 110
Salehi, Bahar, 64, 127
San Vicente, Iñaki, 115
Sanchez Duran, Magali, 104
Sánchez, Alfredo, 113
Sangati, Federico, 127
Sankepally, Rashmi, 46
Santen, Jan van, 132
Santus, Enrico, 114
Sap, Maarten, 49, 131
Saphra, Naomi, 76
Sapkota, Upendra, 28
Saquete, Estela, 116
Saralegi, Xabier, 115
Sardianos, Christos, 118
Sarikaya, Ruhi, 28
Sarker, Abeed, 113
Sasaki, Akira, 134
Sasaki, Yutaka, 64
Sasano, Ryohei, 79
Sato, Misa, 119
Satyapanich, Taneeya, 108
Savenkov, Denis, 54
Savova, Guergana, 111, 116
Sayeed, Asad, 26, 124
Scarton, Carolina, 54, 109
Schenk, Niko, 104
Scherl, Richard, 86
Schijndel, Marten van, 99, 109, 124
Schilder, Frank, 78
Schlangen, David, 38
Schluter, Natalie, 105
Schneider, Nathan, 11, 16, 21, 97, 135
Schöch, Christof, 123
Scholten-Akoun, Dirk, 121
Schrading, Nicolas, 85
Schuetze, Hinrich, 38
Schuler, William, 64, 99, 120, 124
Schulte im Walde, Sabine, 127, 129
Schütze, Hinrich, 61, 84, 86, 113
Schwartz, H. Andrew, 49, 131
Schwartz, Lane, 123

- Schwenk, Holger, 65
Scrivner, Olga, 122
Seddah, Djamé, 27
Segal, Benjamin, 127
Seligman, Martin, 49
Semeraro, Giovanni, 112
Seno, Eloize, 134
Seppi, Kevin, 41, 61, 134
Sequeira, José, 112
Serbina, Tatiana, 134
Serrano, J. Ignacio, 131
Severyn, Aliaksei, 80, 113
Shaikh, Samira, 103, 136
Shanumuga Sundaram, Mahalakshmi, 108
Sharma, Arpit, 125
Sharma, Manali, 41
Sharma, Nidhi, 55
Sharma, Richa, 55
Sharma, Vasu, 52
Sharp, Rebecca, 36
Sheehan, Kathleen, 78
Shein, Fraser, 128
Shepard, Jason, 55
Sheth, Amit, 105
Shi, Jianlin, 112
Shibata, Tomohide, 137
Shim, Hyosup, 79
Shin, Andrew, 79
Shin, Hyopil, 97
Shivade, Chaitanya, 50, 133
Shrimpton, Luke, 85
Shutova, Ekaterina, 113, 136
Shwartz, Vered, 104
Sidorov, Grigori, 108
Sierra, Basilio, 138
Silva, Mario J., 114
Silvestri, Fabrizio, 31
Sima'an, Khalil, 49
Simões, David, 110
Sindlerova, Jana, 135
Singh, Dhirendra, 80
Singh, Puneet, 52
Singh, Sameer, 70, 77, 139
Skjærholt, Arne, 84
Slonim, Noam, 118
Smiley, Charese, 78
Smith, Aaron, 127
Smith, Laura, 131
Smith, Noah A., 27, 42, 44, 68, 97, 100
Smyth, Padhraic, 131
Sobhani, Parinaz, 103, 118
Søgaard, Anders, 80, 84, 135
Soldaini, Luca, 66
Soler-Company, Juan, 76
Solorio, Thamar, 28, 134
Somasundaran, Swapna, 120
Song, Dezhao, 78
Song, Yangqiu, 85
Song, Young C., 32
Song, Zhiyi, 125
Soni, Sagar, 112
Sordoni, Alessandro, 33
Soricut, Radu, 100
Soroa, Aitor, 86, 104
Souza, Jackson, 134
Soyer, Hubert, 78
Soysal, Ergin, 111
Specia, Lucia, 109
Speranza, Manuela, 116
Sporleder, Caroline, 133
Sproat, Richard, 132
Srikumar, Vivek, 122, 131, 135
Srinivasan, Harini, 137
Stanford, James, 77
Stankevitch, Vadim, 112
Staubs, Robert, 46
Stede, Manfred, 118, 119
Steedman, Mark, 27, 72, 73
Steele, David, 54
Stefanescu, Dan, 109
Stein, Benno, 114, 118
Stenetrop, Pontus, 78
Stevenson, Mark, 103
Stewart, Brandon M., 33
Stickles, Elise, 136
Stillwell, David, 49
Stone, Matthew, 93
Stoyanov, Veselin, 113
Strapparava, Carlo, 92, 112, 116, 136
Strassel, Stephanie, 103, 125
Stratos, Karl, 28, 138
Strickhart, Laura, 108
Strötgen, Jannik, 116
Strube, Michael, 75, 104, 105
Strzalkowski, Tomek, 103, 136
Su, Jian, 113
Subramanya, Amarnag, 61
Sudarikov, Roman, 111
Sudooh, Katsuhito, 45
Sulis, Emilio, 115
Sultan, Md Arafat, 109
Sulubacak, Umut, 127, 134
Sumner, Tamara, 104, 109
Sun, Chengjie, 109
Sun, Jia, 113
Sun, Weiwei, 117
Surdeanu, Mihai, 36, 47, 75
Sweet, Dawn M., 86
Synnaeve, Gabriel, 62
Szpakowicz, Stan, 122
Szymanski, Terrence, 117

- Taboada, Maite, 104
Taghipour, Kaveh, 46
Takabatake, Yu, 126
Takagi, Tomohiro, 109
Takamura, Hiroya, 79, 86
Talbot, Ruth, 114
Talukdar, Partha P., 26
Tam, Yik-Cheung, 65
Tamine, Lynda, 116
Tamura, Akihiro, 82
Tan, Cong, 110, 116
Tan, Liling, 109, 116, 117
Tanana, Michael, 131
Tanner, Chris, 28
Taslimipoor, Shiva, 109
Taub-Tabib, Hillel, 81
Taylor, Sarah, 136
Tekiroglu, Serra Sinem, 136
Tetreault, Joel, 38, 120, 121
Thirunarayanan, Krishnaprasad, 105
Thomas, Christopher, 110
Thomas, Max, 78
Thomson, Sam, 68
Tian, Guanhua, 138
Tian, Jun Feng, 109
Tingley, Dustin, 33
Tissot, Hegler, 116
Titov, Ivan, 26
Tobolsky, Victoria, 131
Toda, Tomoki, 76
Toh, Zhiqiang, 113
Tokunaga, Takenobu, 38
Tomlinson, Marc, 136
Toni, Francesca, 118
Topić, Goran, 78
Toral, Antonio, 123
Torki, Marwan, 110
Torunoğlu-Selamet, Dilara, 127, 134
Tosik, Melanie, 139
Toutanova, Kristina, 42, 54
Townsend, Richard, 114
Tracey, Jennifer, 103
Tran, Quan Hung, 110
Tran, Vu, 110
Trancoso, Isabel, 86, 114
Tripodi, Rocco, 111
Tsakalidis, Adam, 114
Tsang, Vivian, 128
Tsuchida, Masaaki, 82
Tsujii, Jun'ichi, 86
Tsvetkov, Yulia, 39
Tu, Wenting, 45
Turchi, Marco, 49
Tutubalina, Elena, 127

Urbain, Jay, 138
Ureña López, L. Alfonso, 115
Uresova, Zdenka, 117, 135
Uria, Larraitz, 110, 111
Urizar, Ruben, 116
Uryupina, Olga, 105
Ustun, Volkan, 138
Uszkoreit, Hans, 111
Utt, Jason, 104
Uzdilli, Fatih, 114
UzZaman, Naushad, 116

Šnajder, Jan, 109, 119
V, Rudramurthy, 80
Valenzuela-Escarcega, Marco Antonio, 75
Van de Kauter, Marjan, 115
Van Durme, Benjamin, 11, 17, 26, 38, 78, 104
Van Hee, Cynthia, 115
Vanallemeersch, Tom, 129
Vandeghinste, Vincent, 129
Vanderwende, Lucy, 70, 75
Varma, Vasudeva, 113–115
Vaskinn, Anja, 132
Vaswani, Ashish, 47
Veale, Tony, 113, 136
Velupillai, Sumithra, 116
Vempala, Alakananda, 41
Venturi, Giulia, 134
Venugopalan, Subhashini, 93
Verhagen, Marc, 116
Versley, Yannick, 40
Vieira, Tim, 51
Vieu, Laure, 103
Vilarinho, Darnes, 108, 113
Villemonte de la Clergerie, Eric, 27
Vitt, Thorsten, 123
Vo, Ngoc Phuoc An, 53, 108–110, 137
Vo, Nguyen, 125
Vogel, Mirko, 139
Volkert, Kevin, 114
Volkova, Svitlana, 11, 17
Volodina, Elena, 120
Vossen, Piek, 111, 116, 125, 126
Vu, Duy, 79
Vu, Ngoc Thang, 113
Vu, Tu Thanh, 110
Vu, Tu, 110
Vuggumudi, Viswanadh Kumar Reddy, 109
Vulić, Ivan, 109

Waite, Aurelien, 48
Walker, Marilyn, 41, 60, 118, 119
Wan, Xiaojun, 117
Wang, Bo, 114
Wang, Chuan, 48
Wang, Dong, 65
-

- Wang, Elaine, 120
Wang, Fangyuan, 138
Wang, Hongling, 115
Wang, JianXiang, 110
Wang, Jingqi, 111
Wang, Jingtao, 75
Wang, Lu, 68
Wang, Maolin, 121
Wang, Peng, 138
Wang, Pidong, 137
Wang, Tengjiao, 114
Wang, Weiran, 37
Wang, Wen, 65
Wang, Wenmin, 85
Wang, William Yang, 39, 47
Wang, Xiaolong, 109, 110, 116
Wang, Xinhao, 120
Wang, Xun, 45
Wang, Yafei, 124
Wang, Ye-Yi, 61
Wanner, Leo, 40, 76
Warren, David, 42
Way, Andy, 123
Webber, Bonnie, 129, 133
Wehbe, Leila, 26
Wei, Baogang, 84
Wei, Furu, 113
Weimar, Lukas, 123
Weingarten, Evan, 49
Weinthal, Noah, 114
Weiss, Rebecca, 33
Weissenbacher, Davy, 113
Weissenborn, Dirk, 111
Weller, Marion, 127, 129
Wen, Miaomiao, 47, 136
Werner, Gregory, 103, 133
Wertheim, Suzanne, 136
White, Jerome, 46
White, Michael, 120
Whittaker, Steve, 119
Wicentowski, Richard, 114
Wiebe, Janyce, 87, 103, 111, 118
Wilkinson, Bryan, 53
Wilks, Yorick, 103
Willis, Alistair, 121
Wilson, Joshua, 121
Wintner, Shuly, 127
Wintrode, Jonathan, 51
Wojatzki, Michael, 121
Wolfe, Travis, 26, 78
Wolff, Clemens, 93
Wolff, Phillip, 55
Wong, Kam-fai, 114
Wrede, Britta, 60
Wright, Jonathan, 125
Wu, Dekai, 129
Wu, Guoshun, 113
Wu, Shumin, 104, 130
Wu, Wei, 35
Wu, Yonghui, 111
Wuebker, Joern, 94
Wyner, Adam, 118
Xia, Yunqing, 85
xiang, Yang, 110
Xie, Pengtao, 41
Xie, Ziang, 40
Xing, Chao, 65
Xing, Eric, 41
Xu, Bo, 138
Xu, Feiyu, 111
Xu, Han, 31
Xu, Hongzhi, 114
Xu, Hua, 111
Xu, Huijuan, 93
Xu, Jiaming, 138
Xu, Jie, 136
Xu, Jun, 110, 111
Xu, Ke, 113
Xu, Wei, 49, 68, 92, 108
Xu, Weiqun, 113, 138
Xu, Yang, 124
Xue, Nianwen, 43, 48, 129
Yamakawa, Yukari, 125
Yamrom, Boris, 136
Yan, Yonghong, 113, 138
Yanai, Kohsuke, 119
Yanase, Toshihiko, 119
Yang, Bishan, 118
Yang, Dezhi, 84
Yang, Diyi, 41
Yang, Min, 45
Yang, Shansong, 84
Yang, Yi, 48
Yannakoudakis, Helen, 121
Yao, Lei, 62
Yao, Liang, 84
Yao, Xuchen, 77
Yarowsky, David, 11, 17
Yasuda, Norihito, 50
Ye, Jieping, 86
Yee, Connie, 125, 126
Yen, John, 124
Yi, Liang, 110
Yih, Wen-tau, 11, 19
Yimam, Seid Muhie, 53
Yin, Siqi, 136
Yin, Wenpeng, 45, 61, 84
Yin, Yichun, 113
Yocum, Zachary, 117
Yovcheva, Ivana, 110

- Yu, Clement T., 45
Yu, Dian, 83
Yu, Dong, 111
Yu, Mo, 78, 84
- Zacarias, Andressa, 134
Zaghouni, Wajdi, 135
Zamani, Hamed, 49
Zamanov, Ivan, 110
Zampieri, Marcos, 116
Zapata Jaramillo, Carlos Mario, 125
Zapirain, Benat, 116
Zarrella, Guido, 108
Zayats, Victoria, 81
Zechner, Klaus, 120
Zelaia, Ana, 138
Zeman, Daniel, 117
Zesch, Torsten, 98, 108, 120, 121
Zettlemoyer, Luke, 93
Zhang, Fan, 117, 121
Zhang, Muyu, 69
Zhang, Tong, 28, 33, 37
Zhang, Xun, 117
Zhang, Yang, 85
Zhang, Yaoyun, 110, 111
Zhang, Yuan, 27, 72
Zhang, Yue, 31
Zhang, Yuxiao, 114
Zhang, Zhifei, 115
Zhang, Zhihua, 113, 115
Zhao, Jiang, 108, 109
Zhao, Jun, 138
Zhao, Kai, 81, 94
Zhao, Lin, 42, 82
Zheng, Jiehan, 43
Zhou, Liang, 125, 126
Zhou, Ming, 113
Zhou, Xiaoqiang, 110
Zhou, Yiwei, 114
Zhu, Xiaodan, 103
Zhuang, Di, 41
Ziai, Ramon, 110
Ziebart, Brian D., 45
Zinsmeister, Heike, 134
Zirkly, Ayah, 139
Zubiaga, Arkaitz, 114
Zuidema, Willem, 48, 103

Local Guide

This guide was written by Kevin Cohen. For the most up-to-date version, please visit <http://bretonnel.com/2015/04/29/visiting-denver-for-naacl2015/>.

The North American Association for Computational Linguistics annual meeting (NAACL 2015) will be held in Denver, Colorado this year. Here are some things that I think might be useful or enjoyable for visiting computational linguists, natural language processing people, and the like. I'm not going to talk about the mountains, Red Rocks, or any of that kind of stuff—you can find that in tourist guides, hotel propaganda, and pretty much anywhere else. These are some of the things that make life in Denver bearable, and that I don't think you'll hear about elsewhere.

Airport The Denver International Airport looks quite distinctive. Opinions differ as to whether it is meant to look like teepees, the Rocky Mountains, or what. It features in a number of conspiracy theories, which mainly claim that it is built over an underground complex that will be the seat of the government of the New World Order. On the way from the airport to Denver, be sure to look for the Demon Horse statue. We call it the Demon Horse for two reasons: (1) it has glowing red eyes, and (2) during its construction, the head fell off and crushed the artist, killing him.

Bookstores Here's a link to a web page listing ten good or great independent Denver bookstores:

www.westword.com/arts/the-ten-best-bookstores-in-denver-6659360

Bars The classic Denver bar that no one else knows about is El Chapultepec. Either arrive early, or be prepared to stand all evening. You can reach it from the conference hotel with a ride down the 16th Street Mall free shuttle and a short walk through a lively neighborhood. The LoDo area has many bars that are quite busy on weekend nights. Use caution around the time that the bars close. Again, you can reach LoDo quite easily from the conference hotel via the free 16th Street Mall shuttle.

Restaurants A rare hippie restaurant treat in Denver is the Mercury Cafe, known to us locally as "The Merc." It's a step back into the 1960s, sorta. Take a cab there, or the light rail—don't try to walk from the conference hotel, as the neighborhood is not always safe.

Marijuana Marijuana is legal here under state law. You can find it easily; the stores (usually known as “dispensaries”) are typically marked with a green cross or a green marijuana leaf. However, it is NOT legal under federal law—if you are not an American citizen, don’t take a chance with this. It’s a legal gray zone, and people do occasionally get burnt. Also, like alcohol, it is not legal to consume it in public. (And, no: I don’t indulge!)

Mexican food The Denver population is 30% Hispanic, and we have fantastic Mexican food here—also Salvadoran and some Peruvian, if you don’t mind leaving the area of the conference hotel to find it. Mexican food is an integral part of American food in this part of the country. A good place to get it is Real de Minas, on Colfax. Avoid the cheese-smothered burrito platters and have something that’s actually Mexican, like tacos de carne asada, or ribs (costillas) in green chile sauce. You can get there on the number 15 bus—more on that below.

Decadent snacks Voodoo Doughnuts is an import from Portland, Oregon—some of you may remember it from ACL 2011. Truly amazing doughnuts—be prepared to stand in line. You can get there on the number 15 bus from the conference hotel.

Local beers Denver has a lot of microbreweries, and many good local beers. One of the main favorites is Fat Tire (which is now nationally distributed, so you may have had it before).

The Number 15 bus The number 15 bus goes up and down Colfax Avenue, allegedly the longest street in America, and probably one of the sleaziest. (Colfax runs quite close to the conference hotel.) Everyone in Denver has a story about the number 15 bus, typically involving a drunk, a drug addict, or vomit. It’s actually pretty safe, although you should be careful on Colfax at night, as you would in any big city anywhere in the world. The last stop on the eastbound leg of the route (away from the mountains) is the Anschutz/Fitzsimons medical campus. Stop by the Biomedical Text Mining Group in the Center for Computational Bioscience for one of the best views of downtown Denver and the mountains that you’ll find.

Safety Denver is a pretty safe city. Aurora is not quite as safe, particularly in the older parts of town. In general, you should be aware of your surroundings in the evening, as you would be in any big city.

Exterior of Denver Art Museum by Jeff Wells, courtesy Denver Art Museum. Other photos by Steve Trice.

Art at Every Angle



"The Mile High City is remaking itself as a world capital of art and architecture." **Sunset Magazine**

The spectacular **Denver Art Museum** (DAM) is the largest art museum between Kansas City and the West Coast. Designed by world-renowned architect Daniel Libeskind, the Hamilton Building is a work of art in itself. Since it was completed in 2006, the DAM has become a bona fide arts-world icon. "The building is smashing," raved *Newsweek*. The addition doubled the size of the museum and allows the DAM to host major touring exhibits, including "King Tut," through January 9, 2011.

Inside, the 356,000-square-foot DAM is filled with classics by Monet, Picasso and Matisse, alongside more modern works by Warhol and O'Keefe. All in all, the museum contains a collection numbering more than 68,000 works from around the world, including intriguing pieces from Africa and pre-Columbian America. The DAM is also home to one of the greatest collections of Western American art in the world.

Denver Performing Arts Complex (DPAC)



Tony Award®-winning theater, Broadway touring productions, contemporary dance and ballet, magnificent chorales, a major symphony orchestra, internationally acclaimed opera – it's all happening at this 12-acre, four-square-block complex in the heart of downtown Denver. **DPAC** is the second-largest arts complex in the nation, with 10 performance spaces seating 10,000 people, all connected by an 80-foot-tall glass ceiling.

What's GOING ON?

To find everything happening at Denver's cultural facilities and art galleries, visit: www.Denver365.com. Here you'll find the latest on theater, dance, symphony, opera, art show openings, film, museums, kid-friendly activities and annual cultural events, from the Cherry Creek Arts Festival to Denver Arts Week.

Art Districts



Hip galleries, world-class museums and First Friday artwalks – Denver's art districts are the pulse of the Mile High City's creative community. **Downtown Denver** has a variety of art galleries, many of which are found in LoDo with beautifully restored brick warehouses, and **Cherry Creek North** features galleries amidst its boutiques and restaurants. More Denver neighborhoods, including **Highlands**, **Navajo Street**, **Old South Pearl** and **Old South Gaylord Street**, are home to eclectic galleries and studios. The **ArtDistrict on Santa Fe** boasts more than 40 galleries, shops, and restaurants.

The **Golden Triangle Museum District** is home to eight museums, in addition to more than 50 galleries, fine art studios and specialty stores. Just north of downtown, the **River North Art District**, which goes by the catchy nickname **RiNo**, is where art is made. The District is also earning a name for itself amongst in-the-know creative types, thanks to a see-and-be-seen First Friday celebration, with galleries and shops staying open late.

Museum of Contemporary Art Denver (MCA)



Denver's cutting-edge arts world was given a 27,000-square-foot home in 2007 with the completion of the David Adjaye-designed **Museum of Contemporary Art Denver** in Lower Downtown (LoDo). This lovely, modern building – with five intimate gallery spaces and a panoramic rooftop café – houses a constantly refreshed set of exhibits. No visit here is ever the same.



Climate Chart

Month	Temperature Max/Min	% of Humidity AM/PM	Precipitation Inches	Sunshine Noon %
JANUARY	43.2/16.1 F	6.21/-8.8 C	63/49	.50
FEBRUARY	46.6/20.2 F	8.1/-6.6 C	67/44	.57
MARCH	52.2/25.8 F	11.2/-3.4 C	68/40	1.28
APRIL	61.8/34.5 F	16.6/-1.4 C	67/35	1.71
MAY	70.8/43.6 F	21.6/6.4 C	70/38	2.40
JUNE	81.4/52.4 F	21.6/6.4 C	65/35	1.79
JULY	88.2/58.6 F	31.2/14.8 C	68/34	1.91
AUGUST	85.5/56.9 F	29.7/13.8 C	69/35	1.51
SEPTEMBER	76.9/47.6 F	24.9/8.7 C	69/34	1.24
OCTOBER	66.3/36.4 F	19.1/2.4 C	65/35	.98
NOVEMBER	52.5/25.4 F	11.4/-3.7 C	68/48	.87
DECEMBER	44.5/17.4 F	6.9/-8.1 C	65/51	.64
ANNUAL AVERAGE	64.2/36.2 F	17.9/2.3 C	67/40	15.4
				70

Source: U.S. National Oceanic and Atmospheric Administration

Denver International Airport

Denver International Airport is the largest airport in the nation, covering 53 square miles – large enough to hold both Dallas-Fort Worth and Chicago O'Hare airports. The airport is a short 24-mile drive from downtown Denver, making it easy to access the Mile High City.



High Altitude Tips



Drink lots of water – even if you're not thirsty.

Use sunscreen: There is 25 percent less protection from the sun's rays at Denver's 5,280-foot elevation.

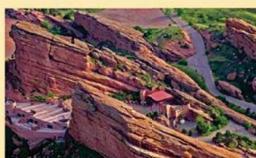
In mile-high Denver's rarefied air, a golf ball goes 10 percent farther. So does a cocktail. Drink alcohol in moderation.

Take some time to look at the sky. There is less water vapor in the air a mile above sea level, so the sky really is bluer.

The Napa Valley of Beer



Denver is the "Napa Valley of Beer," brewing more than 90 different beers daily. Coors Brewery is the largest single brewing site in the world and the Great American Beer Festival, held each September, is the nation's largest celebration of suds with 1,900 beers to sample. Denver Beer Fest overlaps from September 10-19. Go to DenverBeerFest.com for more information.



5 FREE things to do in Denver

1 Stand a "mile high" on the steps of the State Capitol – 5,280 feet above sea level.



at Red Rocks Amphitheatre & Park.

2 Ride the FREE hybrid shuttle bus on the mile-long 16th Street Mall.

4 Hike the spectacular trails

5 Drink a free beer (for those over 21) after touring the world's largest brewing site, Coors Brewery, in Golden.

Downtown - 16th Street Mall



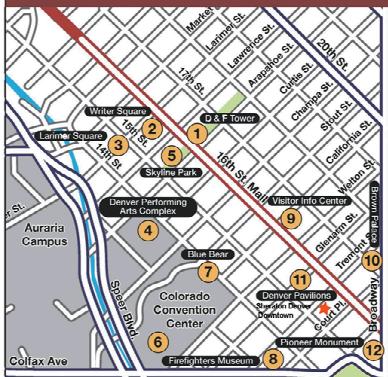
DOWNTOWN DENVER

The 16th Street Mall



The heart of downtown Denver is the 16th Street Mall, a mile-long pedestrian promenade lined with 200 trees, 50,000 flowers, outdoor cafés and inviting shops. Designed by famed architect I.M. Pei, the mall features a pattern in colored granite that resembles the design on a diamondback rattlesnake when viewed from above.

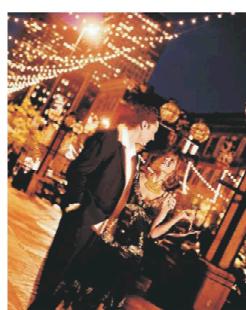
Free hybrid buses leave either end of the mall as often as every 90 seconds, stopping on every corner. The pedestrian portion of the 16th Street Mall continues over two footbridges to Commons Park on the South Platte River and to the Highlands neighborhood.



1 DANIELS & FISHER TOWER. Known by locals as the D&F Tower, this was the highest building west of the Mississippi River when it opened in 1909 and the third highest structure in the U.S. It is a $\frac{3}{4}$ -scale model of the celebrated bell tower of San Marco in Venice.

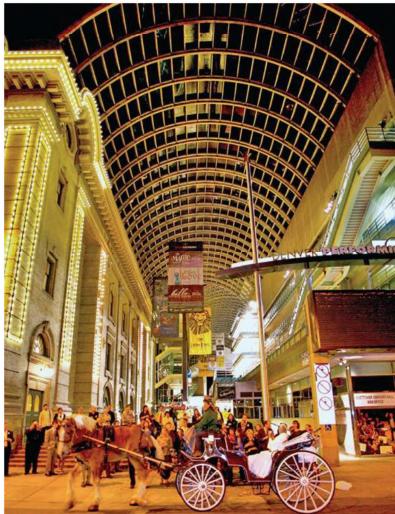


2 WRITER SQUARE. This pleasant plaza named after a local Denver family, is lined with flower baskets and outdoor sculptures, offering quiet cafés and shops.



3 LARIMER SQUARE. Downtown Denver's trendiest block is also one of its most historic. Beautiful Victorian brick-and-stone buildings now house chic eateries, one-of-a-kind galleries and clothing boutiques, nightspots, a comedy club, and some of Denver's top restaurants. This is the perfect place to watch a sunset from a sidewalk cafe.

Downtown Map: 1. D&F Tower by Riccardo Caviglia; 2. Writer Square by Steve Crandall; 3. Blue Bear by Rich Grana; 4. Denver Pavilions by Steve Crandall; 5. Colorado Convention Center by Steve Crandall; 6. Firefighters Museum by Steve Crandall; 7. Blue Bear by Rich Grana; 8. Pioneer Monument by Steve Crandall.



4 DENVER PERFORMING ARTS COMPLEX. The second-largest performing arts complex in the nation offers 10 venues seating more than 10,000 people for Tony® Award-winning theater, opera, symphony and ballet. The new \$92-million Ellie Caulkins Opera House opened in 2005 and Boettcher Concert Hall is undergoing a massive refurbishing. Ninety-minute backstage tours are available, visiting everything from the dressing rooms to costume and scenery shops.



8 FIREFIGHTERS MUSEUM. The hottest place in town tells the story of Denver's firefighters, from 1866 to present, featuring hands-on fun for children and adults with a unique gift shop.

9 DENVER VISITOR INFORMATION CENTER. Pop into our one-stop-shop for everything Denver – hundreds of free brochures and maps, personal assistance in making all your plans, a TicketMaster outlet and a gift shop.



10 BROWN PALACE HOTEL & SPA. Built in 1892, Denver's grand dame hotel features a dramatic eight-story atrium topped by a stained-glass ceiling. Enjoy afternoon tea in the lobby or take a historic walking tour of the hotel every Wednesday and Saturday at 3 p.m. Private tours can also be scheduled for other days or groups.



5 SKYLINE PARK. This three-block-long oasis in the heart of the city offers a pleasant green space to relax. Concerts and special events take place here throughout the year.



7 BLUE BEAR. One of the most popular public art statues in Denver is Lawrence Argent's 40-foot high Blue Bear, "I See What You Mean," which appears to be peeking into the convention center. The 10,000-pound bear is made up of 4,000 interlocking triangles.



11 DENVER PAVILIONS. Check out two square blocks and three levels of shopping, dining and entertainment with Niketown, Barnes & Noble bookstore, Lucky Strike Lanes, Hard Rock Café and more.



12 PIONEER MONUMENT & FOUNTAIN. The bronze figure on the rearing horse pointing the way west is Christopher "Kit" Carson, the legendary explorer and scout who played an important role in developing the Rocky Mountain West.

Climate Chart

Month	Temperature Max/Min	% of Humidity AM/PM	Precipitation Inches	Sunshine Noon %
JANUARY	43.2/16.1 F	6.21/-8.8 C	63/49	.50
FEBRUARY	46.6/20.2 F	8.1/-6.6 C	67/44	.57
MARCH	52.2/25.8 F	11.2/-3.4 C	68/40	1.28
APRIL	61.8/34.5 F	16.6/-1.4 C	67/35	1.71
MAY	70.8/43.6 F	21.6/6.4 C	70/38	2.40
JUNE	81.4/52.4 F	21.6/6.4 C	65/35	1.79
JULY	88.2/58.6 F	31.2/14.8 C	68/34	1.91
AUGUST	85.5/56.9 F	29.7/13.8 C	69/35	1.51
SEPTEMBER	76.9/47.6 F	24.9/8.7 C	69/34	1.24
OCTOBER	66.3/36.4 F	19.1/2.4 C	65/35	.98
NOVEMBER	52.5/25.4 F	11.4/-3.7 C	68/48	.87
DECEMBER	44.5/17.4 F	6.9/-8.1 C	65/51	.64
ANNUAL AVERAGE	64.2/36.2 F	17.9/2.3 C	67/40	15.4
				70

Source: U.S. National Oceanic and Atmospheric Administration

Denver International Airport

Denver International Airport is the largest airport in the nation, covering 53 square miles – large enough to hold both Dallas-Fort Worth and Chicago O'Hare airports. The airport is a short 24-mile drive from downtown Denver, making it easy to access the Mile High City.



Approximate Flight Times
to/from Denver International Airport

5
FREE

things to do
in Denver

1 Stand a "mile high" on the steps of the State Capitol – 5,280 feet above sea level.

2 Ride the FREE hybrid shuttle bus on the mile-long 16th Street Mall.

3 Tour the U.S. Mint, where more than 50 million coins are made each year.

4 Hike the spectacular trails



at Red Rocks Amphitheatre & Park.

5 Drink a free beer (for those over 21) after touring the world's largest brewing site, Coors Brewery, in Golden.

High Altitude Tips



Drink lots of water – even if you're not thirsty.

Use sunscreen: There is 25 percent less protection from the sun's rays at Denver's 5,280-foot elevation.

In mile-high Denver's rarefied air, a golf ball goes 10 percent farther. So does a cocktail. Drink alcohol in moderation.

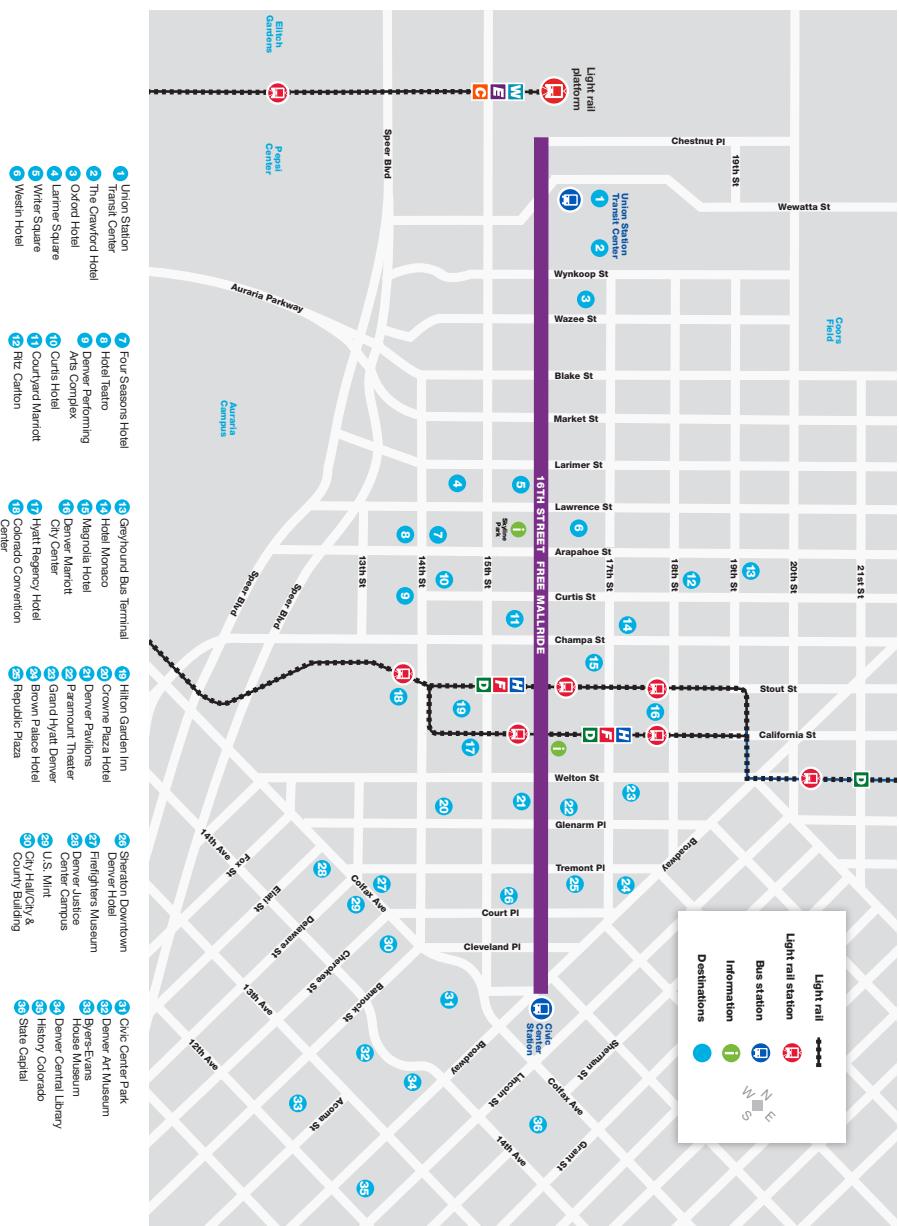
Take some time to look at the sky. There is less water vapor in the air a mile above sea level, so the sky really is bluer.

The Napa Valley of Beer



Denver is the "Napa Valley of Beer," brewing more than 90 different beers daily. Coors Brewery is the largest single brewing site in the world and the Great American Beer Festival, held each September, is the nation's largest celebration of suds with 1,900 beers to sample. Denver Beer Fest overlaps from September 10-19. Go to DenverBeerFest.com for more information.





NAACL gratefully acknowledges the following sponsors for their support:

Platinum

Bloomberg

Gold

Google™

**Goldman
Sachs**

A9™

 **IBM Watson**

Silver


NUANCE


**SDL
Research**

Bronze

YAHOO! LABS

Supporter


Digital roots


**UNIVERSITY of
WASHINGTON**
Computational Linguistics

Student Research Workshop


NSF

Google™


BABAB

fuse | machines

Wireless


MZ

MACHINEZONE