CONFERENCE HANDBOOK

June 9–14 The 2013 Conference of the North American Association for Computational Linguistics: Human Language Technologies

June 13–14 The Second Joint Conference on Computational and Lexical Semantics

June 14–15 SemEval

Handbook production by:
Alex Clemmer (University of Utah)
Matt Post (Johns Hopkins University)
Printing by Omnipress of Madison, Wisconsin

The 2013 Conference of the
North American Chapter of the
Association for Computational Linguistics:
Human Language Technologies

June 9–14, 2013

The Second Joint Conference on
Lexical and Computational Semantics

(with SemEval, June 14–15)

June 13–14

Conference Handbook

June 9–15, 2013
Westin Peachtree Plaza Hotel
Atlanta, Georgia

Message from the General Chair

Welcome everyone!

It is my pleasure to welcome you all to Atlanta, Georgia, for the 2013 NAACL Human Language Technologies conference. This is a great opportunity to reconnect with old friends and make new acquaintances, learn the latest in your own field and become curious about new areas, and also to experience Atlanta's warm southern hospitality. That hospitality starts with Priscilla Rasmussen! Priscilla thinks about everything that we all take for granted: the registration that just took place, the rooms in which we sit, the refreshments that keep us energized, and the social events that make this conference so fun, and many other details that you would miss if they weren't there. Please introduce yourself and say hi. Priscilla is the backbone of the NAACL organization. Thank you!

This conference started a year ago, when Hal Daumé III and Katrin Kirchhoff graciously agreed to be program co-chairs. It is no exaggeration to say how much their dedication has shaped this conference and how grateful I am for their initiative and hard work. Thank you Hal and Katrin, especially for all the fun discussion that made the work light and the year go by fast! This conference could not have happened with you.

Thanks go to the entire organizing committee. As I am writing this to be included in the proceedings, I am grateful for the fantastic detailed and proactive work by Colin Cherry and Matt Post, the publications chairs. The tutorials chairs, Katrin Erk and Jimmy Lin, selected, and solicited, 6 tutorials to present in depth material on some of the diverse topics represented in our community. Chris Dyer and Derrick Higgins considered which projects shine best when shown as a demonstration. The workshops chairs for NAACL, Sujith Ravi and Luke Zettlemoyer, worked jointly with ACL and EMNLP to select the workshops to be held at NAACL. They also worked with ICML 2013 to co-host workshops that bridge the two communities, in addition to the Joint NAACL/ICML symposium.

Posters from the student research workshop are part of the poster and demonstrations session on Monday night. This is a great opportunity for the students to be recognized in the community and to benefit from lively discussion of their presentations (attendees take note!) Annie Louis and Richard Socher are the student research workshop chairs, and Julia Hockenmaier and Eric Ringger generously share their wisdom as the faculty advisors. The student research workshop itself will be held on the first day of workshops. There are so many people who contribute their time to the behind-the-scenes organization of the conference, without which the conference cannot take place. Asking for money is probably not a natural fit for anyone, but Chris Brew worked on local sponsorship, and Dan Bikel and Patrick Pantel worked

to obtain sponsorship across the ACL conferences this year - thank you! Jacob Eisenstein had the more fun role of distributing money as the student volunteer coordinator, and we thank all of the student volunteers who will be helping to run a smooth conference. Kristy Boyer kept the communication "short and tweet" using a variety of social media (and old-fashioned media too). An important part of the behind-the-scenes efforts that enable a conference like NAACL to come together are the sponsors. We thank all of the sponsors for the contributions to the conference , both for the general funding made available as well as the specific programs that are funded through sponsorship. You can read more about these sponsors in our conference handbook.

This year there are several initiatives, and if successful, we hope they'll be part of NAACL conferences in the future. One is to make the proceedings available prior to the conference; we hope you will benefit from the extra time to read the papers beforehand. Another is for tutorials and all oral presentations to be recorded on video and made available post-conference. We are also delighted to host presentations, in both oral and poster formats, from the new Transactions of the ACL journal, to enhance the impact these will already have as journal publications. Finally, Matt Post is creating a new digital form of conference handbook to go with our digital age; thanks also go to Alex Clemmer who has prepared the paper copy that you may be reading right now. We hope you use the #NAACL2013 tag when you are tweeting about the conference or papers at the conference; together, we'll be creating a new social media corpus to explore.

Once again, we are pleased to be co-located with *SEM conference, and the SemEval workshop. We are lucky to have ICML 2013 organized so close in time and place. Several researchers who span the two communities have reconvened the Joint NAACL/ICML symposium on June 15, 2013. In addition, two workshops that address areas of interest to both NAACL and ICML members have been organized on June 16th, as part of the ICML conference.

NAACL 2013 has given me a great appreciation for the volunteering that is part of our culture. Besides the organizing committee itself, we are guided by the NAACL executive board, who think about questions with a multi-year perspective. I also want to recognize the members who first initiated and now maintain the ACL Anthology, where all of our published work will be available to all in perpetuity, a fabulous contribution and one that distinguishes our academic community.

Have a fun conference!

Lucy Vanderwende, Microsoft Research
NAACL HLT 2013 General Chair

Message from the Program Committee Co-Chairs

Welcome to NAACL HLT 2013 in Atlanta, Georgia. We have an exciting program consisting of six tutorials, 24 sessions of talks (both for long and short papers), an insane poster madness session that includes posters from the newly revamped student research workshop, ten workshops and two additional cross-pollination workshops held jointly with ICML (occurring immediately after NAACL HLT, just one block away). There are a few innovations in the conference this year, the most noticeable of which is the twitter channel #naacl2013 and the fact that we are the first conference to host papers published in the Transactions of the ACL journal – there are six such papers in our program, marked as [TACL]. We are very excited about our two invited talks, one on Monday morning and one Wednesday morning. The first is by Gina Kuperberg, who will talk about "Predicting Meaning: What the Brain tells us about the Architecture of Language Comprehension." The second presenter is our own Kathleen KcKeown, who will talk about "Natural Language Applications from Fact to Fiction."

The morning session on Tuesday includes the presentation of best paper awards to two worthy recipients. The award for Best Short Paper goes to Marta Recasens, Marie-Catherine de Marneffe and Christopher Potts for their paper "The Life and Death of Discourse Entities: Identifying Singleton Mentions" The award for Best Student Paper goes to the long paper "Automatic Generation of English Respellings" by Bradley Hauer and Greg Kondrak. We gratefully acknowledge IBM's support for the Student Best Paper Award. Finally, many thanks to the Best Paper Committee for selecting these excellent papers!

The complete program includes 95 long papers (of which six represent presentations from the journal Transactions of the ACL, a first for any ACL conference!) and 51 short papers. We are excited that the conference is able to present such a dynamic array of papers, and would like to thank the authors for their great work. We worked hard to keep the conference to three parallel sessions at any one time to hopefully maximize a participant's ability to see everything she wants! This represents an acceptance rate of 30% for long papers and 37% for short papers. More details about the distribution across areas and other statistics will be made available in the NAACL HLT Program Chair report on the ACL wiki: `aclweb.org/adminwiki/index.php?title=Reports`

The review process for the conference was double-blind, and included an author response period for clarifying reviewers' questions. We were very pleased to have the assistance of 350 reviewers, each of whom reviewed an average of 3.7 papers, in deciding the program. We are especially thankful for the reviewers who spent time reading the author responses and engaging other reviewers in the discussion board. Assigning reviewers would not have been possible without the hard work of Mark Dredze and his miracle assignment scripts. Furthermore, constructing the program would not

have been possible without 22 excellent area chairs forming the Senior Program Committee: Eugene Agichtein, Srinivas Bangalore, David Bean, Phil Blunsom, Jordan Boyd-Graber, Marine Carpuat, Joyce Chai, Vera Demberg, Bill Dolan, Doug Downey, Mark Dredze, Markus Dreyer, Sanda Harabagiu, James Henderson, Guy Lapalme, Alon Lavie, Percy Liang, Johanna Moore, Ani Nenkova, Joakim Nivre, Bo Pang, Zak Shafran, David Traum, Peter Turney, and Theresa Wilson. Area chairs were responsible for managing paper assignments, collating reviewer responses, handling papers for other area chairs or program chairs who had conflicts of interest, making recommendations for paper acceptance or rejection, and nominating best papers from their areas. We are very grateful for the time and energy that they have put into the program.

There are a number of other people that we interacted with who deserve a hearty thanks for the success of the program. Rich Gerber and the START team at Softconf have been invaluable for helping us with the mechanics of the reviewing process. Matt Post and Colin Cherry, as publications co-chairs, have been very helpful in assembling the final program and coordinating the publications of the workshop proceedings. There are several crucial parts of the overall program that were the responsibility of various contributors, including Annie Louis, Richard Socher, Julia Hockenmaier and Eric Ringger (Student Research Workshop chairs, who did an amazing job revamping the SRW); Jimmy Lin and Katrin Erk (Tutorial Chairs); Luke Zettlemoyer and Sujith Ravi (Workshop Chairs); Chris Dyer and Derrick Higgins (Demo Chairs); Jacob Eisenstein (Student Volunteer Coordinator); Chris Brew (Local Sponsorship Chair); Patrick Pantel and Dan Bikel (Sponsorship Chairs); and the new-founded Publicity chair who handled #naacl2013 tweeting among other things, Kristy Boyer.

We would also like to thank Chris Callison-Burch and the NAACL Executive Board for guidance during the process. Michael Collins was amazingly helpful in getting the inaugural TACL papers into the NAACL HLT conference. Priscilla Rasmussen deserves, as always, special mention and warmest thanks as the local arrangements chair and general business manager. Priscilla is amazing and everyone who sees her at the conference should thank her.

Finally, we would like to thank our General Chair, Lucy Vanderwende, for both her trust and guidance during this process. She helped turn the less-than-wonderful parts of this job to roses, and her ability to organize an incredibly complex event is awe inspiring. None of this would have happened without her.

We hope that you enjoy the conference!

Hal Daumé III, University of Maryland
Katrin Kirchhoff, University of Washington

NAACL HLT 2013 Organizing Committee

General Conference Chair
Lucy Vanderwende, Microsoft Research

Program Committee Chairs
Hal Daumé III, University of Maryland
Katrin Kirchhoff, University of Washington

Local Arrangements
Priscilla Rassmussen

Workshop Chairs
Luke Zettlemoyer, University of Washington
Sujith Ravi, Google

Tutorial Chairs
Jimmy Lin, University of Maryland
Katrin Erk, University of Texas at Austin

Student Research Workshop
Chairs:
Annie Louis, University of Pennsylvania
Richard Socher, Stanford University
Faculty Advisors:
Julia Hockenmaier, University of Illinois at Urbana-Champaign
Eric Ringger, Brigham Young University

Student Volunteer Coordinator
Jacob Eisenstein, School of Interactive Computing, Georgia Tech

Demonstrations Chairs
Chris Dyer, Carnegie Mellon University
Derrick Higgins, Educational Testing Service

Local Sponsorship Chair
Chris Brew, Educational Testing Service

NAACL Sponsorship Chairs
Patrick Pantel, Microsoft Research
Dan Bikel, Google

Publications Chairs
Matt Post, Johns Hopkins University
Colin Cherry, National Research Council, Canada

Publicity Chair
Kristy Boyer, North Carolina State University

Area Chairs
Phonology and Morphology, Word Segmentation
Markus Dreyer (SDL Language Weaver)

Syntax, Tagging, Chunking and Parsing
Joakim Nivre (Uppsala University)

James Henderson (Université de Genève)

Semantics
Percy Liang (Stanford University)
Peter Turney (National Research Council of Canada)

Multimodal NLP
Srinivas Bangalore (AT&T)

Discourse, Dialogue, Pragmatics
David Traum (Institute for Creative Technologies)
Joyce Chai (Michigan State University)

Linguistic Aspects of CL
Vera Demberg (Saarland University)

Summarization
Guy Lapalme (Université de Montréal)

Generation
Johanna Moore (University of Edinburgh)

ML for Language Processing
Phil Blunsom (University of Oxford)
Mark Dredze (Johns Hopkins University)

Machine Translation
Alon Lavie (Carnegie Mellon University)
Marine Carpuat (National Research Council of Canada)

Information Retrieval and QA
Eugene Agichtein (Emory University)

Information Extraction
Doug Downey (Northwestern University)
Sanda Harabagiu (University of Texas at Dallas)

Spoken Language Processing
Zak Shafran (Oregon Health and Science University)

Sentiment Analysis and Opinion Mining
Bo Pang (Cornell University)
Theresa Wilson (Johns Hopkins University)

NLP-enabled Technology
David Bean (TDW)

Document Categorization and Topic Clustering
Jordan Boyd-Graber (University of Maryland)

Social Media Analysis and Processing
Bill Dolan (Microsoft Research)

Language Resources and Evaluation Methods
Ani Nenkova (University of Pennsylvania)

Contents

1 Tutorials: Sunday, June 9 **1**

T1: Deep Learning for NLP (without Magic) (R. Socher and C.D. Manning) 2

T2: Discourse Processing (M. Stede) 4

T3: Towards Reliability-Aware Entity Analytics and Integration for Noisy Text at Scale (S. Mehta and L. V. Subramaniam) 6

T4: Semantic Role Labeling (M. Palmer, I. Titov and S. Wu) 8

T5: Spectral Learning Algorithms for Natural Language Processing (S. Cohen, M. Collins, D. P. Foster, K. Stratos and L. Ungar) 10

T6: Morphological, Syntactical and Semantic Knowledge in Statistical Machine Translation (M. R. Costa-jussà and C. Quirk) 12

2 Main Conference: Monday, June 10 **15**

Invited Talk: Gina R Kuperberg 16

Parallel Sessions 18

Poster Madness! 25

Posters and Demonstrations Session 38

3 Main Conference: Tuesday, June 11 **41**

Best Paper Awards 42

Parallel Sessions 42

Banquet 54

4 Main Conference: Wednesday, June 12 **55**

Invited Talk: Kathleen McKeown 56

Parallel Sessions 58

5 *SEM and SemEval: Thursday–Saturday, June 13–15 **69**

Organizing Committee 75

*SEM (Main conference and shared task) 77

SemEval 93

6 Workshops: Thursday–Friday, June 13–14 **103**
Student Research Workshop 104
W1: The 8th Workshop on Innovative Use of NLP for Building Educational Applications (BEA8) 105
W2: Seventh Workshop on Syntax, Semantics and Structure in Statistical Translation (SSST-7) 108
W3: First workshop on Metaphor in NLP (Meta4NLP) 109
W4: 9th workshop on Multiword Expressions (MWE) 111
W5: Language Analysis in Social Media (LASM) 113
W6: Events: Definition, Detection, Coreference, and Representation (EVENTS) . . . 114
W7: Workshop on Vision and Natural Language Processing (WVL) 115
W8: Computational Linguistics for Literature (CLFL) 117
W9: Natural Language Processing for Improving Textual Accessibility (NLP4ITA) . . 118
W10: 4th workshop on Computation Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA 2013) 119

7 Joint NAACL/ICML Symposium: Saturday, June 15 **123**

8 Venue information **125**

Restaurant Guide **126**

Hotel Maps **143**

Author Index **145**

Area map (inside back cover) **159**

Sponsors **160**

1

Tutorials: Sunday, June 9

Overview

7:30am – 6:00pm	**Registration**	(Terrace)
9:00am – 12:30pm	**Morning Session**	
	Deep Learning for NLP (without Magic) Richard Socher and Christopher D. Manning	(International B,C)
	Discourse Processing Manfred Stede	(International D)
	NLP for uncertain data at scale Sameep Mehta and L V Subramaniam	(International E,F)
12:30pm – 2:00pm	**Lunch**	
1:00pm – 6:00pm	**NAACL Board Meeting** *(Lunch provided)*	(Boardroom)
2:00pm – 5:30pm	**Afternoon Session**	
	Semantic Role Labeling Martha Palmer, Ivan Titov and Shumin Wu	(International B,C)
	Spectral Learning Algorithms for Natural Language Processing Shay Cohen, Michael Collins, Dean Foster, Karl Stratos and Lyle Ungar	(International D)
	Morphological, Syntactical and Semantic Knowledge in Statistical Machine Translation Marta Ruiz Costa-jussà and Chris Quirk	(International E,F)
6:00pm – 9:00pm	**Welcoming Reception**	(Atlanta Ballroom)

Tutorial 1

Deep Learning for NLP (without Magic)

Richard Socher and Christopher D. Manning

Sunday, June 9, 2013, 9:00am – 12:30pm

International B,C

Machine learning is everywhere in today's NLP, but by and large machine learning amounts to numerical optimization of weights for human designed representations and features. The goal of deep learning is to explore how computers can take advantage of data to develop features and representations appropriate for complex interpretation tasks. This tutorial aims to cover the basic motivation, ideas, models and learning algorithms in deep learning for natural language processing. Recently, these methods have been shown to perform very well on various NLP tasks such as language modeling, POS tagging, named entity recognition, sentiment analysis and paraphrase detection, among others. The most attractive quality of these techniques is that they can perform well without any external hand-designed resources or time-intensive feature engineering. Despite these advantages, many researchers in NLP are not familiar with these methods. Our focus is on insight and understanding, using graphical illustrations and simple, intuitive derivations. The goal of the tutorial is to make the inner workings of these techniques transparent, intuitive and their results interpretable, rather than black boxes labeled "magic here". The first part of the tutorial presents the basics of neural networks, neural word vectors, several simple models based on local windows and the math and algorithms of training via backpropagation. In this section applications include language modeling and POS tagging. In the second section we present recursive neural networks which can learn structured tree outputs as well as vector representations for phrases and sentences. We cover both equations as well as applications. We show how training can be achieved by a modified version of the backpropagation algorithm introduced before. These modifications allow the algorithm to work on tree structures. Applications include sentiment analysis and paraphrase detection. We also draw connections to recent work in semantic compositionality in vector spaces. The principle goal, again, is to make these methods appear intuitive and interpretable rather than mathematically confusing. By this point in the tutorial, the audience members should have a clear understanding of how to build a deep learning system for word-, sentence- and document-level tasks. The last part of the tutorial gives a general overview

---

Richard Socher is a PhD student at Stanford working with Chris Manning and Andrew Ng. His research interests are machine learning for NLP and vision. He is interested in developing new models that learn useful features, capture compositional and hierarchical structure in multiple modalities and perform well across different tasks. He was awarded the 2011 Yahoo! Key Scientific Challenges Award, the Distinguished Application Paper Award at ICML 2011 and a Microsoft Research PhD Fellowship in 2012.

Christopher Manning is an Associate Professor of Computer Science and Linguistics at Stanford University (PhD, Stanford, 1995). Manning has coauthored leading textbooks on statistical approaches to NLP (Manning and Schuetze 1999) and information retrieval (Manning et al. 2008). His recent work concentrates on machine learning and natural language processing, including applications such as statistical parsing and text understanding, joint probabilistic inference, clustering, and deep learning over text and images.

of the different applications of deep learning in NLP, including bag of words models. We will provide a discussion of NLP-oriented issues in modeling, interpretation, representational power, and optimization.

Tutorial 2

Discourse Processing

Manfred Stede (University of Potsdam)

Sunday, June 9, 2013, 9:00am – 12:30pm

International D

The observation that discourse is more than a mere sequence of utterances or sentences amounts to a truism. But what follows from this? In what way does the "value added" arise when segments of discourse are juxtaposed - how does hierarchical structure originate from a linearized discourse?

While many discourse phenomena apply to dialogue and monologue alike, this tutorial will center its attention on monologue written text. The perspective taken is that of practical language processing: We study methods for automatically deriving discourse information from text, and point to aspects of their implementation. The emphasis is on breadth rather than depth, so that the attendees will get an overview of the central tasks of discourse processing, with pointers to the literature for studying the individual problems in more depth. Much of the tutorial will follow the line of the recent book M. Stede: Discourse Processing. Morgan & Claypool 2011.

Specifically, we will study the most important ways of ascribing structure to discourse. This is, first, a breakdown into functional units that are characteristic for the genre of the text. A news message, for example, is conventionally structured in a different way than a scientific paper is. For grasping this level of structure, the patterns that are characteristic for the specific genre need to be modeled.

Second, an ongoing text, unless it is very short, will cover different topics and address them in a sensible linear order. This is largely independent of genre, and since the notion of topic is relatively vague, it is harder to describe and sometimes difficult to identify. The common approach is to track the distribution of content words across the text, but in addition, overt signals for topic switches can be exploited.

Third, the identification of coreference links is a central aspect of discourse processing, and has received much attention in computational linguistics. We will survey the corpus-based methods

---

Manfred Stede, University of Potsdam. After completing his dissertation on the role of lexical semantics in multilingual text generation, Manfred Stede shifted his research focus towards problems of discourse structure and its role in various applications of text understanding. For discourse structure, his work centered on coherence relations and associated structural descriptions of text, and on the linguistic signals of such relations, especially connectives. From the early 2000s on, he developed the Potsdam Commentary Corpus as an example of (German) texts analyzed simultaneously on multiple levels, including sentential syntax, coreference, and rhetorical structure; in parallel, the technical infrastructure of a database for querying and visualizing multi-layer corpora was developed. In recent years, more analysis levels have been added to the corpus (e.g., content zones, connectives and their arguments). As for applications, Manfred worked on text summarization and various tasks of information extraction; more recently, his focus has been on issues of subjectivity and sentiment analysis.

that have dominated the field in recent years, and then look at the ramifications that the set of all coreference links in a text has for its structure.

Fourth, we investigate the structure resulting from establishing coherence relations (e.g., Cause, Contrast) among adjacent text segments. The term "discourse parsing" is often used for the task of identifying such relations (by exploiting more or less explicit linguistic signals) and building tree structures that reflect the semantic or pragmatic scaffolding of a (portion of) text.

Thus emerges a picture of a text as a series of different, yet related, layers of analysis. The final part of the tutorial addresses the issue of inter-connections between these levels. As a tool for accessing such multi-layered text corpora, we will see how the (open-source) ANNIS2 database allows for querying the data across different layers, and for visualizing different structural layers in appropriate ways.

Tutorial 3

Towards Reliability-Aware Entity Analytics and Integration for Noisy Text at Scale

Sameep Mehta and L. Venkata Subramaniam

Sunday, June 9, 2013, 9:00am – 12:30pm

International E,F

Due to easy to use apps (Facebook, Twitter, etc.), higher Internet connectivity and always on facility allowed by smart phones, the key characteristics of raw data are changing. This new data can be characterized by 4V's - Volume, Velocity, Variety and Veracity. For example during a Football match, some people will Tweet about goals, penalties, etc., while others may write longer blogs and further there will be match reports filed in trusted online news media after the match. Although the sources may be varied, the data describes and provides multiple evidences for the same event. Such multiple evidences should be used to strengthen the belief in the underlying physical event as the individual data points may have inherent uncertainty. The uncertainty can arise from inconsistent, incomplete and ambiguous reports. The uncertainty is also because the trust levels of the different sources vary and affect the overall reliability. We will summarize various efforts to perform reliability aware entity integration.

The other problem in text analysis in such setting is posed by presence of noise in the text. Since the text is produced in several informal settings such as email, blogs, tweet, SMS, chat and is inherently noisy and has several veracity issues. For example, missing punctuation and the use of non-standard words can often hinder standard natural language processing techniques such as part-of-speech tagging and parsing. Further downstream applications such as entity extraction, entity resolution and entity completion have to explicitly handle noise in order to return useful results. Often, depending on the application, noise can be modeled and it may be possible to develop specific strategies to immunize the system from the effects of noise and

---

Sameep Mehta is researcher in Information Management Group at IBM Research India. He received his M.S. and Ph.D. from The Ohio State University, USA in 2006. He also holds an Adjunct Faculty position at the International Institute of Information Technology, New Delhi. Sameep regularly advises MS and PhD students at University of Delhi and IIT Delhi. He regularly delivers Tutorials at COMAD (2009, 2010 and 2011). His current research interests include Data Mining, Business Analytics, Service Science, Text Mining, and Workforce Optimization.

L Venkata Subramaniam manages the information management analytics and solutions group at IBM Research India. He received his PhD from IIT Delhi in 1999. His research focuses on unstructured information management, statistical natural language processing, noisy text analytics, text and data mining, information theory, speech and image processing. He often teaches and guides student thesis at IIT Delhi on these topics. His tutorial titled Noisy Text Analytics was the second largest at NAACL-HLT 2010. He co founded the AND (Analytics for Noisy Unstructured Text Data) workshop series and also co-chaired the first four workshops, 2007-2010. He was guest co-editor of two special issues on Noisy Text Analytics in the International Journal of Document Analysis and Recognition in 2007 and 2009.

improve performance. Also the aspect of reliability is key as a lot of this data is ambiguous, incomplete, conflicting, untrustworthy and deceptive. The key goals of this tutorial are:

- Draw the attention of researchers towards methods for doing entity analytics and integration on data with 4V characteristics.
- Differentiate between noise and uncertainty in such data.
- Provide an in-depth discussion on handling noise in NLP based methods.
- Finally, handling uncertainty through information fusion and integration.

This tutorial builds on two earlier tutorials — NAACL 2010 tutorial on Noisy Text and COMAD 2012 tutorial on Reliability Aware Data Fusion. In parallel the authors are also hosting a workshop on related topic "Reliability Aware Data Fusion" at SIAM Data Mining Conference, 2013.

Tutorial 4

Semantic Role Labeling

Martha Palmer (University of Colorado), Ivan Titov (Saarland University), and Shumin Wu (University of Colorado)

Sunday, June 9, 2013, 2:00pm – 5:30pm

International B,C

This tutorial will describe semantic role labeling, the assignment of semantic roles to eventuality participants in an attempt to approximate a semantic representation of an utterance. The linguistic background and motivation for the definition of semantic roles will be presented, as well as the basic approach to semantic role annotation of large amounts of corpora. Recent extensions to this approach that encompass light verb constructions and predicative adjectives will be included, with reference to their impact on English, Arabic, Hindi and Chinese. Current proposed extensions such as Abstract Meaning Representations and richer event representations will also be touched on.

Details of machine learning approaches will be provided, beginning with fully supervised approaches that use the annotated corpora as training material. The importance of syntactic parse information and the contributions of different feature choices, including tree kernels, will be discussed, as well as the advantages and disadvantages of particular machine learning algorithms

---

Martha Palmer is a Professor of Linguistics and Computer Science, and a Fellow of the Institute of Cognitive Science at the University of Colorado. Her current research is aimed at building domain-independent and language independent techniques for semantic interpretation based on linguistically annotated data, such as Proposition Banks. She has been the PI on NSF, NIH and DARPA projects for linguistic annotation (syntax, semantics and pragmatics) of English, Chinese, Korean, Arabic and Hindi. She has been a member of the Advisory Committee for the DARPA TIDES program, Chair of SIGLEX, Chair of SIGHAN, a past President of the Association for Computational Linguistics, and is a Co-Editor of JNLE and of LiLT and is on the CL Editorial Board. She received her Ph.D. in Artificial Intelligence from the University of Edinburgh in 1985.

Ivan Titov joined the Saarland University as a junior faculty and head of a research group in November 2009, following a postdoc at the University of Illinois at Urbana-Champaign. He received his Ph.D. in Computer Science from the University of Geneva in 2008 and his master's degree in Applied Mathematics and Informatics from the St. Petersburg State Polytechnic University (Russia) in 2003. His research interests are in statistical natural language processing (models of syntax, semantics and sentiment) and machine learning (structured prediction methods, latent variable models, Bayesian methods).

Shumin Wu is a Computer Science PhD student (advised by Dr. Martha Palmer) at the University of Colorado. His current research is aimed at developing and applying semantic mapping (aligning and jointly inferring predicate-argument structures between languages) to Chinese dropped-pronoun recovery/alignment, automatic verb class induction, and other applications relevant to machine translation.

and approaches such as joint inference. Appropriate considerations for evaluation will be presented as well as successful uses of semantic role labeling in NLP applications.

We will also cover techniques for exploiting unlabeled corpora and transferring models across languages. These include methods, which project annotations across languages using parallel data, induce representations solely from unlabeled corpora (unsupervised methods) or exploit a combination of a small amount of human annotation and a large unlabeled corpus (semi-supervised techniques). We will discuss methods based on different machine learning paradigms, including generative Bayesian models, graph-based algorithms and bootstrapping style techniques.

Tutorial 5

Towards Reliability-Aware Entity Analytics and Integration for Noisy Text at Scale

Shay Cohen (Columbia University), Michael Collins (Columbia University), Dean P. Foster (University of Pennsylvania), Karl Stratos (Columbia University), and Lyle Ungar (University of Pennsylvania)

Sunday, June 9, 2013, 2:00pm – 5:30pm

International D

Recent work in machine learning and NLP has developed spectral algorithms for many learning tasks involving latent variables. Spectral algorithms rely on singular value decomposition as a basic operation, usually followed by some simple estimation method based on the method of moments. From a theoretical point of view, these methods are appealing in that they offer consistent estimators (and PAC-style guarantees of sample complexity) for several important latent-variable models. This is in contrast to the EM algorithm, which is an extremely successful approach, but which only has guarantees of reaching a local maximum of the likelihood function.

Shay Cohen is a postdoctoral research scientist in the Department of Computer Science at Columbia University. He is a computing innovation fellow. His research interests span a range of topics in natural language processing and machine learning. He is especially interested in developing efficient and scalable parsing algorithms as well as learning algorithms for probabilistic grammars.

Michael Collins is the Vikram S. Pandit Professor of computer science at Columbia University. His research is focused on topics including statistical parsing, structured prediction problems in machine learning, and applications including machine translation, dialog systems, and speech recognition. His awards include a Sloan fellowship, an NSF career award, and best paper awards at EMNLP (2002, 2004, and 2010), UAI (2004 and 2005), and CoNLL 2008.

Dean P. Foster is currently the Marie and Joseph Melone Professor of Statistics at the Wharton School of the University of Pennsylvania. His current research interests are machine learning, stepwise regression and computational linguistics. He has been searching for new methods of finding useful features in big data sets. His current set of hammers revolve around fast matrix methods (which decompose 2nd moments) and tensor methods for decomposing 3rd moments.

Karl Stratos is a Ph.D. student in the Department of Computer Science at Columbia. His research is focused on machine learning and natural language processing. His current research efforts are focused on spectral learning of latent-variable models, or more generally, uncovering latent structure from data.

Lyle Ungar is a professor at the Computer and Information Science Department at the University of Pennsylvania. His research group develops scalable machine learning and text mining methods, including clustering, feature selection, and semi-supervised and multi-task learning for natural language, psychology, and medical research. Example projects include spectral learning of language models, multi-view learning for gene expression and MRI data, and mining social media to better understand personality and well-being.

From a practical point of view, the methods (unlike EM) have no need for careful initialization, and have recently been shown to be highly efficient (as one example, in work under submission by the authors on learning of latent-variable PCFGs, a spectral algorithm performs at identical accuracy to EM, but is around 20 times faster).

In this tutorial we will aim to give a broad overview of spectral methods, describing theoretical guarantees, as well as practical issues. We will start by covering the basics of singular value decomposition and describe efficient methods for doing singular value decomposition. The SVD operation is at the core of most spectral algorithms that have been developed.

We will then continue to cover canonical correlation analysis (CCA). CCA is an early method from statistics for dimensionality reduction. With CCA, two or more views of the data are created, and they are all projected into a lower dimensional space which maximizes the correlation between the views. We will review the basic algorithms underlying CCA, give some formal results giving guarantees for latent-variable models and also describe how they have been applied recently to learning lexical representations from large quantities of unlabeled data. This idea of learning lexical representations can be extended further, where unlabeled data is used to learn underlying representations which are subsequently used as additional information for supervised training.

We will also cover how spectral algorithms can be used for structured prediction problems with sequences and parse trees. A striking recent result by Hsu, Kakade and Zhang (2009) shows that HMMs can be learned efficiently using a spectral algorithm. HMMs are widely used in NLP and speech, and previous algorithms (typically based on EM) were guaranteed to only reach a local maximum of the likelihood function, so this is a crucial result. We will review the basic mechanics of the HMM learning algorithm, describe its formal guarantees, and also cover practical issues.

Last, we will cover work about spectral algorithms in the context of natural language parsing. We will show how spectral algorithms can be used to estimate the parameter models of latent-variable PCFGs, a model which serves as the base for state-of-the-art parsing models such as the one of Petrov et al. (2007). We will show what are the practical steps that are needed to be taken in order to make spectral algorithms for L-PCFGs (or other models in general) practical and comparable to state of the art.

Tutorial 6

Morphological, Syntactical and Semantic Knowledge in Statistical Machine Translation

Marta R. Costa-jussà (Institute for Infocomm Research) and Chris Quirk (Microsoft Research)

Sunday, June 9, 2013, 2:00pm – 5:30pm

International E,F

This tutorial focuses on how morphology, syntax and semantics may be introduced into a standard phrase-based statistical machine translation system with techniques such as machine learning, parsing and word sense disambiguation, among others.

Regarding the phrase-based system, we will describe only the key theory behind it. The main challenges of this approach are that the output contains unknown words, wrong word orders and non-adequate translated words. To solve these challenges, recent research enhances the standard system using morphology, syntax and semantics.

Morphologically-rich languages have many different surface forms, even though the stem of a word may be the same. This leads to rapid vocabulary growth, as various prefixes and suffixes can combine with stems in a large number of possible combinations. Language model probability estimation is less robust because many more word forms occur rarely in the data. This morphologically-induced sparsity can be reduced by incorporating morphological information into the SMT system. We will describe the three most common solutions to face morphology: preprocessing the data so that the input language more closely resembles the output language; using additional language models that introduce morphological information; and post-processing the output to add proper inflections.

Marta R. Costa-jussà, Institute for Infocomm Research (I2R), is a Telecommunication's Engineer by the Universitat Politècnica de Catalunya (UPC, Barcelona) and she received her PhD from the UPC in 2008. Her research experience is mainly in Automatic Speech Recognition, Machine Translation and Information Retrieval. She has worked at LIMSI-CNRS (Paris), Barcelona Media Innovation Center (Barcelona) and the Universidade de Sao Paulo (São Paulo). Since December 2012 she is working at Institute for Infocomm Research (Singapore) implementing the IMTraP project ("Integration of Machine Translation Paradigms") on Hybrid Machine Translation, funded by the European Marie Curie International Outgoing European Fellowship program. She is currently organizing the ACL Workshop HyTRA 2013 and she will be teaching a summer school course on hybrid machine translation at ESSLLI 2013.

Chris Quirk, Microsoft Research. After studying Computer Science and Mathematics at Carnegie Mellon University, Chris joined Microsoft in 2000 to work on the Intentional Programming project, an extensible compiler and development framework. He moved to the Natural Language Processing group in 2001, where his research has mostly focused on statistical machine translation powering Microsoft Translator, especially on several generations of a syntax directed translation system that powers over half of the translation systems. He is also interested in semantic parsing, paraphrase methods, and very practical problems such as spelling correction and transliteration.

Syntax differences between the source and target language may lead to significant differences in the relative word order of translated words. Standard phrase-based SMT systems surmount reordering/syntactic challenges by learning from data. Most approaches model reordering inside translation units and using statistical methodologies, which limits the performance in language pairs with different grammatical structures. We will briefly introduce some recent advances in SMT that use modeling approaches based on principles more powerful flat phrases and better suited to the hierarchical structures of language: SMT decoding with stochastic synchronous context free grammars and syntax-driven translation models.

Finally, semantics are not directly included in the SMT core algorithm, which means that challenges such as polysemy or synonymy are either learned directly from data or they are incorrectly translated. We will focus on recent attempts to introduce semantics into statistical-based systems by using source context information.

The course material will be suitable both for attendees with limited knowledge of the field, and for researchers already familiar with SMT who wish to learn about modern tendencies in hybrid SMT. The mathematical content of the course include probability and simple machine learning, so reasonable knowledge of statistics and mathematics is required. There will be a small amount of linguistics and ideas from natural language processing.

2

Main Conference: Monday, June 10

Overview

7:00am – 6:00pm	**Registration**	(Terrace)
7:30am – 8:45am	**Continental Breakfast**	(Terrace)
8:45am – 9:00am	**Welcome to NAACL 2013!**	(Peachtree Ballroom)
9:00am – 10:10am	**Invited Talk: Gina Kuperberg**	(Peachtree Ballroom)
10:10am – 10:40am	**Break**	(Terrace)
10:40am – 12:20pm	**Parallel Sessions**	
	Machine Translation	(Peachtree B,C)
	Information Extraction	(Plaza Ballroom)
	Cognitive and Psycholinguistics	(Peachtree D,E)
12:20pm – 2:00pm	**Lunch**	
12:20pm – 2:00pm	**IBM Student Lunch**	(Vinings)
2:00pm – 3:40pm	**Parallel Sessions**	
	Parsing and Syntax	(Peachtree B,C)
	Topic Modeling and Text Mining	(Plaza Ballroom)
	Spoken Language Processing	(Peachtree D,E)
3:40pm – 4:10pm	**Break**	(Terrace)
4:10pm – 6:00pm	**Poster Madness!**	(Peachtree Ballroom)
6:00pm – 6:30pm	**Break**	(Terrace)
6:30pm – 8:30pm	**Posters & Demonstrations**	(200 Building Grand Atrium)
	Buffet dinner provided	

Invited Talk: Gina R Kuperberg

"Predicting Meaning: What the Brain tells us about the Architecture of Language Comprehension"

Gina R. Kuperberg

Monday, June 10, 2013, 9:00am – 10:10am

Atlanta Ballroom

Abstract: It is well established that we draw upon our real-world knowledge to predict upcoming events and even individual words. I will discuss evidence that the neurocognitive mechanisms that we engage in retrieving conceptual information associated with incoming words are quite distinct from those engaged when these predictions are disconfirmed by the input. Drawing broad links with computational models conceptualizing language comprehension as an incremental process of belief updating, I will suggest that the engagement of these distinct neurocognitive systems allows for comprehension that is both highly efficient and highly flexible (1).

I will first discuss studies using event-related potentials (ERPs) to examine online brain activity during sentence and discourse comprehension. I will then draw some (still tentative) links between this ERP literature and some relevant fMRI and MEG studies. Finally, I will discuss the advantages of a predictive comprehension system. Predicting correctly clearly offers advantages in terms of computational efficiency. Here I will argue that the costs incurred when we predict incorrectly are also crucial for successful and flexible comprehension. Neurocognitive responses triggered by prediction errors may rescue us from interpretation errors in noisy environments, may allow us learn novel events, and may enable us to flexibly adjust our comprehension strategies in response to everchanging task and environmental demands.

Biography: Gina R Kuperberg is a Cognitive Neuroscientist and a Professor in the Department of Psychology and the Cognitive Science Center at Tufts University, Boston. She is also a Board Certified Psychiatrist and a Principal Investigator in the Psychiatry Neuroscience Program at Massachusetts General Hospital, Harvard Medical School. Her research program aims to understand the neurocognitive mechanisms by which the human brain builds meaning from language, and how these mechanisms break down in neuropsychiatric disorders, particularly schizophrenia.

Dr. Kuperberg's Lab is situated across in both the Department of Psychology at Tufts and the Martinos Center for Biomedical Imaging at Massachusetts General Hospital. The Lab uses multimodal neuroimaging methods — event-related potentials (ERPs), functional MRI (fMRI) and magneto-encephalography (MEG) — to probe both the spatial and temporal dimensions of cognition in the brain. Her research program is funded by an RO1 from the National Institute of Mental Health (NIMH), as well as awards from the Brain and Behavior Research Foundation and the Sidney Baer Trust. She, her students, postdocs and collaborators publish in a wide range of journals of Cognitive Neuroscience, Psycholinguistics, Experimental Psychology, Neuroimaging and Psychiatry.

Dr. Kuperberg has served as a standing member for the Language and Communication Study Section for the National Institute of Health, and as a committee representative for Language for the Cognitive Neuroscience society. Her research accomplishments have been recognized by several awards, including the A.E. Bennett Research Award from the Society for Biological Psychiatry, the Joseph Zubin Award for Significant Contributions to Research in Psychopathology, and an Award from Brain Research for their most highly cited article, for her review of the architecture of the language system, Neural Mechanisms of Language Comprehension: Challenges to Syntax.

Dr. Kuperberg earned her MD at St. Bartholomew's Medical School, London, and her PhD in Psychology and Cognitive Neuroscience at Kings College, University of London. She completed an internship at St. Bartholomew's Hospital and residency training in Psychiatry at the Maudsley Hospital and Institute of Psychiatry, London. In 1998, she came to Boston where she completed research fellowships in Neuroimaging and Cognitive Electrophysiology at Massachusetts General Hospital, Harvard Medical School and Tufts University, working with David Caplan, Anders Dale and Phil Holcomb.

Session M1a: Machine Translation

Model With Minimal Translation Units, But Decode With Phrases
Nadir Durrani, Alexander Fraser, and Helmut Schmid
10:40am–11:05am

N-gram-based models co-exist with their phrase-based counterparts as an alternative SMT framework. Both techniques have pros and cons. While the N-gram-based framework provides a better model that captures both source and target contexts and avoids spurious phrasal segmentation, the ability to memorize and produce larger translation units gives an edge to the phrase-based systems during decoding, in terms of better search performance and superior selection of translation units. In this paper we combine N-gram-based modeling with phrase-based decoding, and obtain the benefits of both approaches. Our experiments show that using this combination not only improves the search accuracy of the N-gram model but that it also improves the BLEU scores. Our system outperforms state-of-the-art phrase-based systems (Moses and Phrasal) and N-gram-based systems by a significant margin on German, French and Spanish to English translation tasks.

Beyond Left-to-Right: Multiple Decomposition Structures for SMT
Hui Zhang, Kristina Toutanova, Chris Quirk, and Jianfeng Gao
11:05am–11:30am

Standard phrase-based translation models do not explicitly model context dependence between translation units. As a result, they rely on large phrase pairs and target language models to recover contextual effects in translation. In this work, we explore n-gram models over Minimal Translation Units (MTUs) to explicitly capture contextual dependencies across phrase boundaries in the channel model. As there is no single best direction in which contextual information should flow, we explore multiple decomposition structures as well as dynamic bidirectional decomposition. The resulting models are evaluated in an intrinsic task of lexical selection for MT as well as a full MT system, through n-best reranking. These experiments demonstrate that additional contextual modeling does indeed benefit a phrase-based system and that the direction of conditioning is important. Integrating multiple conditioning orders provides consistent benefit, and the most important directions differ by language pair.

Improved Reordering for Phrase-Based Translation using Sparse Features
Colin Cherry
11:30am–11:55am

There have been many recent investigations into methods to tune SMT systems using large numbers of sparse features. However, there have not been nearly so many examples of helpful sparse features, especially for phrase-based systems. We use sparse features to address reordering, which is often considered a weak point of phrase-based translation. Using a hierarchical reordering model as our baseline, we show that simple features coupling phrase orientation to frequent words or word-clusters can improve translation quality, with boosts of up to 1.2 BLEU points in Chinese-English and 1.8 in Arabic-English. We compare this solution to a more traditional maximum entropy approach, where a probability model with similar features is trained on word-aligned bitext. We show that sparse decoder features outperform maximum entropy handily, indicating that there are major advantages to optimizing reordering features directly for BLEU with the decoder in the loop.

Simultaneous Word-Morpheme Alignment for Statistical Machine Translation
Elif Eyigöz, Daniel Gildea, and Kemal Oflazer
11:55am–12:20pm

Current word alignment models for statistical machine translation do not address morphology beyond merely splitting words. We present a twolevel alignment model that distinguishes between words and morphemes, in which we embed an IBM Model 1 inside an HMM based word alignment model. The model jointly induces word and morpheme alignments using an EM algorithm. We evaluated our model on Turkish-English parallel data. We obtained significant improvement of BLEU scores over IBM Model 4. Our results indicate that utilizing information from morphology improves the quality of word alignments.

Session M1b: Information Extraction

Multi-faceted Event Recognition with Bootstrapped Dictionaries

Ruihong Huang and Ellen Riloff

10:40am–11:05am

Identifying documents that describe a specific type of event is challenging due to the high complexity and variety of event descriptions. We propose a multi-faceted event recognition approach, which identifies documents about an event using event phrases as well as defining characteristics of the event. Our research focuses on civil unrest events and learns civil unrest expressions as well as phrases corresponding to potential agents and reasons for civil unrest. We present a bootstrapping algorithm that automatically acquires event phrases, agent terms, and purpose (reason) phrases from unannotated texts. We use the bootstrapped dictionaries to identify civil unrest documents and show that multi-faceted event recognition can yield high accuracy.

Named Entity Recognition with Bilingual Constraints

Wanxiang Che, Mengqiu Wang, Christopher D. Manning, and Ting Liu

11:05am–11:30am

Different languages contain complementary cues about entities, which can be used to improve Named Entity Recognition (NER) systems. We propose a method that formulates the problem of exploring such signals on unannotated bilingual text as a simple Integer Linear Program, which encourages entity tags to agree via bilingual constraints. Bilingual NER experiments on the large OntoNotes 4.0 Chinese-English corpus show that the proposed method can improve strong baselines for both Chinese and English. In particular, Chinese performance improves by over 5
We can then annotate a large amount of bilingual text (80k sentence pairs) using our method, and add it as up-training data to the original monolingual NER training corpus. The Chinese model retrained on this new combined dataset outperforms the strong baseline by over 3

Minimally Supervised Method for Multilingual Paraphrase Extraction from Definition Sentences on the Web

Yulan Yan, Chikara Hashimoto, Kentaro Torisawa, Takao Kawai, Jun'ichi Kazama, and Stijn De Saeger

11:30am–11:55am

We propose a minimally supervised method for multilingual paraphrase extraction from definition sentences on the Web. Hashimoto et al. (2011) extracted paraphrases from Japanese definition sentences on the Web, assuming that definition sentences defining the same concept tend to contain paraphrases. However, their method requires manually annotated data and is language dependent. We extend their framework and develop a minimally supervised method applicable to multiple languages. Our experiments show that our method is comparable to Hashimoto et al.'s for Japanese and outperforms previous unsupervised methods for English, Japanese, and Chinese, and that our method extracts 10,000 paraphrases with 92% precision for English, 82.5% precision for Japanese, and 82% precision for Chinese.

Relation Extraction with Matrix Factorization and Universal Schemas

Sebastian Riedel, Limin Yao, Andrew McCallum, and Benjamin M. Marlin

11:55am–12:20pm

Traditional relation extraction predicts relations within some fixed and finite target schema. Machine learning approaches to this task require either manual annotation or, in the case of distant supervision, existing structured sources of the same schema. The need for existing datasets can be avoided by using a universal schema: the union of all involved schemas (surface form predicates as in OpenIE, and relations in the schemas of pre-existing databases). This schema has an almost unlimited set of relations (due to surface forms), and supports integration with existing structured data (through the relation types of existing databases). To populate a database of such schema we present a family of matrix factorization models that predict affinity between database tuples (the users) and relations (the items). We show that this achieves substantially higher accuracy than the traditional classification approach. More importantly, by operating simultaneously on relations observed in text and in pre-existing structured DBs such as Freebase, we are able to reason about unstructured and structured data in mutually-supporting ways. By doing so our approach outperforms state-of-the-art distant supervision.

Session M1c: Cognitive and Psycholinguistics

Extracting the Native Language Signal for Second Language Acquisition
Ben Swanson and Eugene Charniak
10:40am–11:05am

We develop a method for effective extraction of linguistic patterns that are differentially expressed based on the native language of the author. This method uses multiple corpora to allow for the removal of data set specific patterns, and addresses both feature relevancy and redundancy. We evaluate different relevancy ranking metrics and show that common measures of relevancy can be inappropriate for data with many rare features. Our feature set is a broad class of syntactic patterns, and to better capture the signal we extend the Bayesian Tree Substitution Grammar induction algorithm to a supervised mixture of latent grammars. We show that this extension can be used to extract a larger set of relevant features.

An Analysis of Frequency- and Memory-Based Processing Costs
Marten van Schijndel and William Schuler
11:05am–11:30am

The frequency of words and syntactic constructions has been observed to have a substantial effect on language processing. This begs the question of what causes certain constructions to be more or less frequent. A theory of grounding (Phillips, 2010) would suggest that cognitive limitations might cause languages to develop frequent constructions in such a way as to avoid processing costs. This paper studies how current theories of working memory fit into theories of language processing and what influence memory limitations may have over reading times. Measures of such limitations are evaluated on eye-tracking data and the results are compared with predictions made by different theories of processing.

Cross-Lingual Semantic Similarity of Words as the Similarity of Their Semantic Word Responses
Ivan Vulić and Marie-Francine Moens
11:30am–11:55am

We propose a new approach to identifying semantically similar words across languages. The approach is based on an idea that two words in different languages are similar if they are likely to generate similar words (which includes both source and target language words) as their top semantic word responses. Semantic word responding is a concept from cognitive science which addresses detecting most likely words that humans output as free word associations given some cue word. The method consists of two main steps: (1) it utilizes a probabilistic multilingual topic model trained on comparable data to learn and quantify the semantic word responses, (2) it provides ranked lists of similar words according to the similarity of their semantic word response vectors. We evaluate our approach in the task of bilingual lexicon extraction (BLE) for a variety of language pairs. We show that in the cross-lingual settings without any language pair dependent knowledge the response-based method of similarity is more robust and outperforms current state-of-the art methods that directly operate in the semantic space of latent cross-lingual concepts/topics.

Combining multiple information types in Bayesian word segmentation
Gabriel Doyle and Roger Levy
11:55am–12:20pm

Humans identify word boundaries in continuous speech by combining multiple cues; existing state-of-the-art models, though, look at a single cue. We extend the generative model of Goldwater et al (2006) to segment using syllable stress as well as phonemic form. Our new model treats identification of word boundaries and prevalent stress patterns in the language as a joint inference task. We show that this model improves segmentation accuracy over purely segmental input representations, and recovers the dominant stress pattern of the data. Additionally, our model retains high performance even without single-word utterances. We also demonstrate a discrepancy in the performance of our model and human infants on an artificial-language task in which stress cues and transition-probability information are pitted against one another. We argue that this discrepancy indicates a bound on rationality in the mechanisms of human segmentation.

Session M2a: Parsing and Syntax

Training Parsers on Incompatible Treebanks

Richard Johansson

2:00pm–2:25pm

We consider the problem of training a statistical parser in the situation when there are multiple treebanks available, and these treebanks are annotated according to different linguistic conventions.

To address this problem, we present two simple adaptation methods: the first method is based on the idea of using a shared feature representation when parsing multiple treebanks, and the second method on guided parsing where the output of one parser provides features for a second one.

To evaluate and analyze the adaptation methods, we train parsers on treebank pairs in four languages: German, Swedish, Italian, and English. We see significant improvements for all eight treebanks when training on the full training sets. However, the clearest benefits are seen when we consider smaller training sets. Our experiments were carried out with unlabeled dependency parsers, but the methods can easily be generalized to other feature-based parsers.

Learning a Part-of-Speech Tagger from Two Hours of Annotation

Dan Garrette and Jason Baldridge

2:25pm–2:50pm

Most work on weakly-supervised learning for part-of-speech taggers has been based on unrealistic assumptions about the amount and quality of training data. For this paper, we attempt to create true low-resource scenarios by allowing a linguist just two hours to annotate data and evaluating on the languages Kinyarwanda and Malagasy. Given these severely limited amounts of either type supervision (tag dictionaries) or token supervision (labeled sentences), we are able to dramatically improve the learning of a hidden Markov model through our method of automatically generalizing the annotations, reducing noise, and inducing word-tag frequency information.

Experiments with Spectral Learning of Latent-Variable PCFGs

Shay B. Cohen, Karl Stratos, Michael Collins, Dean P. Foster, and Lyle Ungar

2:50pm–3:15pm

Latent-variable PCFGs (L-PCFGs) are a highly successful model for natural language parsing. Recent work (Cohen et al., 2012) has introduced a spectral algorithm for parameter estimation of L-PCFGs, which—unlike the EM algorithm—is guaranteed to give consistent parameter estimates (it has PAC-style guarantees of sample complexity). This paper describes experiments using the spectral algorithm. We show that the algorithm provides models with the same accuracy as EM, but is an order of magnitude more efficient. We describe a number of key steps used to obtain this level of performance; these should be relevant to other work on the application of spectral learning algorithms. We view our results as strong empirical evidence for the viability of spectral methods as an alternative to EM.

[TACL] Finding Optimal 1-Endpoint-Crossing Trees

Emily Pitler, Sampath Kannan, and Mitchell Marcus

3:15pm–3:40pm

Dependency parsing algorithms capable of producing the types of crossing dependencies seen in natural language sentences have traditionally been orders of magnitude slower than algorithms for projective trees. For 95.8–99.8% of dependency parses in various natural language treebanks, whenever an edge is crossed, the edges that cross it all have a common vertex. The optimal dependency tree that satisfies this 1-Endpoint-Crossing property can be found with an O(n4) parsing algorithm that recursively combines forests over intervals with one exterior point. 1-Endpoint-Crossing trees also have natural connections to linguistics and another class of graphs that has been studied in NLP.

Session M2b: Topic Modeling and Text Mining

Representing Topics Using Images

Nikolaos Aletras and Mark Stevenson

2:00pm–2:25pm

Topics generated automatically, e.g. using LDA, are now widely used in Computational Linguistics. Topics are normally represented as a set of keywords, often the n terms in a topic with the highest marginal probabilities. We introduce an alternative approach in which topics are represented using images. Candidate images for each topic are retrieved from the web by querying a search engine using the top n terms. The most suitable image is selected from this set using a graph-based algorithm which makes use of textual information from the metadata associated with each image and features extracted from the images themselves. We show that the proposed approach significantly outperforms several baselines and can provide images that are useful to represent a topic.

Drug Extraction from the Web: Summarizing Drug Experiences with Multi-Dimensional Topic Models

Michael J. Paul and Mark Dredze

2:25pm–2:50pm

Multi-dimensional latent text models, such as factorial LDA (f-LDA), capture multiple factors of corpora, creating structured output for researchers to better understand the contents of a corpus. We consider such models for clinical research of new recreational drugs and trends, an important application for mining current information for healthcare workers. We use a "three-dimensional" f-LDA variant to jointly model combinations of drug (marijuana, salvia, etc.), aspect (effects, chemistry, etc.) and route of administration (smoking, oral, etc.) Since a purely unsupervised topic model is unlikely to discover these specific factors of interest, we develop a novel method of incorporating prior knowledge by leveraging user generated tags as priors in our model. We demonstrate that this model can be used as an exploratory tool for learning about these drugs from the Web by applying it to the task of extractive summarization. In addition to providing useful output for this important public health task, our prior-enriched model provides a framework for the application of f-LDA to other tasks.

Towards Topic Labeling with Phrase Entailment and Aggregation

Yashar Mehdad, Giuseppe Carenini, Raymond T. Ng, and Shafiq Joty

2:50pm–3:15pm

We propose a novel framework for topic labeling that assigns the most representative phrases for a given set of sentences covering the same topic. We build an entailment graph over phrases that are extracted from the sentences, and use the entailment relations to identify and select the most relevant phrases. We then aggregate those selected phrases by means of phrase generalization and merging. We motivate our approach by applying over conversational data, and show that our framework improves performance significantly over baseline algorithms.

Topic Segmentation with a Structured Topic Model

Lan Du, Wray Buntine, and Mark Johnson

3:15pm–3:40pm

We present a new hierarchical Bayesian model for unsupervised topic segmentation. This new model integrates a point-wise boundary sampling algorithm used in Bayesian segmentation into a structured topic model that can capture a simple hierarchical topic structure latent in documents. We develop an MCMC inference algorithm to split/merge segment(s). Experimental results show that our model outperforms previous unsupervised segmentation methods using only lexical information on Choi's datasets and two meeting transcripts and has performance comparable to those previous methods on two written datasets.

Session M2c: Spoken Language Processing

Text Alignment for Real-Time Crowd Captioning

Iftekhar Naim, Daniel Gildea, Walter Lasecki, and Jeffrey P. Bigham

2:00pm–2:25pm

The primary way of providing real-time captioning for deaf and hard of hearing people is to employ expensive professional stenographers who can type as fast as natural speaking rates. Recent work has shown that a feasible alternative is to combine the partial captions of ordinary typists, each of whom types part of what they hear. In this paper, we describe an improved method for combining partial captions into a final output based on weighted A* search and multiple sequence alignment (MSA). In contrast to prior work, our method allows the tradeoff between accuracy and speed to be tuned, and provides formal error bounds. Our method outperforms the current state-of-the-art on Word Error Rate (WER) (29.6%), BLEU Score (41.4%), and F-measure (36.9%). The end goal is for these captions to be used by people, and so we also compare how these metrics correlate with the judgments of 50 study participants, which may assist others looking to make further progress on this problem.

Discriminative Joint Modeling of Lexical Variation and Acoustic Confusion for Automated Narrative Retelling Assessment

Maider Lehr, Izhak Shafran, Emily Prud'hommeaux, and Brian Roark

2:25pm–2:50pm

Automatically assessing the fidelity of a retelling to the original narrative – a task of growing clinical importance – is challenging, given extensive paraphrasing during retelling along with cascading automatic speech recognition (ASR) errors. We present a word tagging approach using conditional random fields (CRFs) that allows a diversity of features to be considered during inference, including some capturing acoustic confusions encoded in word confusion networks. We evaluate the approach under several scenarios, including both supervised and unsupervised training, the latter achieved by training on the output of a baseline automatic word-alignment model. We also adapt the ASR models to the domain, and evaluate the impact of error rate on performance. We find strong robustness to ASR errors, even using just the 1-best system output. A hybrid approach making use of both automatic alignment and CRFs trained tagging models achieves the best performance, yielding strong improvements over using either approach alone.

Using Out-of-Domain Data for Lexical Addressee Detection in Human-Human-Computer Dialog

Heeyoung Lee, Andreas Stolcke, and Elizabeth Shriberg

2:50pm–3:15pm

Addressee detection (AD) is an important problem for dialog systems in human-human-computer scenarios (contexts involving multiple people and a system) because system-directed speech must be distinguished from human-directed speech. Recent work on AD (Shriberg et al., 2012) showed good results using prosodic and lexical features trained on in-domain data. In-domain data, however, is expensive to collect for each new domain. In this study we focus on lexical models and investigate how well out-of-domain data (either outside the domain, or from single-user scenarios) can fill in for matched in-domain data. We find that human-addressed speech can be modeled using out-of-domain conversational speech transcripts, and that human-computer utterances can be modeled using single-user data: the resulting AD system outperforms a system trained only on matched in-domain data. Further gains (up to a 4% reduction in equal error rate) are obtained when in-domain and out-of-domain models are interpolated. Finally, we examine which parts of an utterance are most useful. We find that the first 1.5 seconds of an utterance contain most of the lexical information for AD, and analyze which lexical items convey this. Overall, we conclude that the H-H-C scenario can be approximated by combining data from H-C and H-H scenarios only.

Segmentation Strategies for Streaming Speech Translation

Vivek Kumar Rangarajan Sridhar, John Chen, Srinivas Bangalore, Andrej Ljolje, and Rathinavelu Chengalvarayan

3:15pm–3:40pm

The study presented in this work is a first effort at real-time speech translation of TED talks, a compendium of public talks with different speakers addressing a variety of topics. We address the goal of achieving a system that balances translation accuracy and latency. In order to improve ASR performance for our diverse

data set, adaptation techniques such as constrained model adaptation and vocal tract length normalization are found to be useful. In order to improve machine translation (MT) performance, techniques that could be employed in real-time such as monotonic and partial translation retention are found to be of use. We also experiment with inserting text segmenters of various types between ASR and MT in a series of real-time translation experiments. Among other results, our experiments demonstrate that a good segmentation is useful, and a novel conjunction-based segmentation strategy improves translation quality nearly as much as other strategies such as comma-based segmentation. It was also found to be important to synchronize various pipeline components in order to minimize latency.

Poster Madness!

Prior to the poster session, poster presenters are given one minute each to pitch their paper. Posters from the Student Research Workshop are included. Following the Poster Madness! session, there will be a buffet dinner and a combined Posters and Demonstrations session.

Main Conference Posters

Enforcing Subcategorization Constraints in a Parser Using Sub-parses Recombining
Seyed Abolghasem Mirroshandel, Alexis Nasr, and Benoît Sagot

Treebanks are not large enough to adequately model subcategorization frames of predicative lexemes, which is an important source of lexico-syntactic constraints for parsing. As a consequence, parsers trained on such treebanks usually make mistakes when selecting the arguments of predicative lexemes. In this paper, we propose an original way to correct subcategorization errors by combining sub-parses of a sentence S that appear in the list of the n-best parses of S. The subcategorization information comes from three different resources, the first one is extracted from a treebank, the second one is computed on a large corpora and the third one is an existing syntactic lexicon. Experiments on the French Treebank showed a 15.24% reduction of erroneous subcategorization frames (SF) selections for verbs as well as a relative decrease of the error rate of 4% Labeled Accuracy Score on the state of the art parser on this treebank.

Large-Scale Discriminative Training for Statistical Machine Translation Using Held-Out Line Search
Jeffrey Flanigan, Chris Dyer, and Jaime Carbonell

We introduce a new large-scale discriminative learning algorithm for machine translation that is capable of learning parameters in models with extremely sparse features. To ensure their reliable estimation and to prevent overfitting, we use a two-phase learning algorithm. First, the contribution of individual sparse features is estimated using large amounts of parallel data. Second, a small development corpus is used to determine the relative contributions of the sparse features and standard dense features. Not only does this two-phase learning approach prevent overfitting, the second pass optimizes corpus-level BLEU of the Viterbi translation of the decoder. We demonstrate significant improvements using sparse rule indicator features in three different translation tasks. To our knowledge, this is the first large-scale discriminative training algorithm capable of showing improvements over the MERT baseline with only rule indicator features in addition to the standard MERT features.

Measuring Term Informativeness in Context
Zhaohui Wu and C. Lee Giles

Measuring term informativeness is a fundamental NLP task. Existing methods, mostly based on statistical information in corpora, do not actually measure informativeness of a term with regard to its semantic context. This paper proposes a new lightweight feature-free approach to encode term informativeness in context by leveraging web knowledge. Given a term and its context, we model context-aware term informativeness based on semantic similarity between the context and the term's most featured context in a knowledge base, Wikipedia. We apply our method to three applications: core term extraction from snippets (text segment), scientific keywords extraction (paper), and back-of-the-book index generation (book). The performance is state-of-the-art or close to it for each application, demonstrating its effectiveness and generality.

Unsupervised Learning Summarization Templates from Concise Summaries
Horacio Saggion

We here present and compare two unsupervised approaches for inducing the main conceptual information in rather stereotypical summaries in two different languages. We evaluate the two approaches in two different information extraction settings: monolingual and cross-lingual information extraction. The extraction systems are trained on auto-annotated summaries (containing the induced concepts) and evaluated on human-annotated documents. Extraction results are promising, being close in performance to those achieved when the system is trained on human-annotated summaries.

Classification of South African languages using text and acoustic based methods: A case of six selected languages
Peleira Nicholas Zulu

Language variations are generally known to have a severe impact on the performance of Human Language Technology Systems. In order to predict or improve system performance, a thorough investigation into these variations, similarities and dissimilarities, is required. Distance measures have been used in several applications of speech processing to analyze different varying speech attributes. However, not much work has been done on language distance measures, and even less work has been done involving South African languages. This study explores two methods for measuring the linguistic distance of six South African languages. It concerns a text based method, (the Levenshtein Distance), and an acoustic approach using extracted mean pitch values. The Levenshtein distance uses parallel word transcriptions from all six languages with as little as 144 words, whereas the pitch method is text-independent and compares mean language pitch differences. Cluster analysis resulting from the distance matrices from both methods correlates closely with human perceptual distances and existing literature about the six languages.

Improving Syntax-Augmented Machine Translation by Coarsening the Label Set
Greg Hanneman and Alon Lavie

We present a new variant of the Syntax-Augmented Machine Translation (SAMT) formalism with a category-coarsening algorithm originally developed for tree-to-tree grammars. We induce bilingual labels into the SAMT grammar, use them for category coarsening, then project back to monolingual labeling as in standard SAMT. The result is a "collapsed" grammar with the same expressive power and format as the original, but many fewer nonterminal labels. We show that the smaller label set provides improved translation scores by 1.14 BLEU on two Chinese–English test sets while reducing the occurrence of sparsity and ambiguity problems common to large label sets.

Keyphrase Extraction for N-best Reranking in Multi-Sentence Compression
Florian Boudin and Emmanuel Morin

Multi-Sentence Compression (MSC) is the task of generating a short single sentence summary from a cluster of related sentences. This paper presents an N-best reranking method based on keyphrase extraction. Compression candidates generated by a word graph-based MSC approach are reranked according to the number and relevance of keyphrases they contain. Both manual and automatic evaluations were performed using a dataset made of clusters of newswire sentences. Results show that the proposed method significantly improves the informativity of the generated compressions.

Development of a Persian Syntactic Dependency Treebank
Mohammad Sadegh Rasooli, Manouchehr Kouhestani, and Amirsaeid Moloodi

This paper describes the annotation process and linguistic properties of the Persian syntactic dependency treebank. The treebank consists of approximately 30,000 sentences annotated with syntactic roles in addition to morpho-syntactic features. One of the unique features of this treebank is that there are almost 4800 distinct verb lemmas in its sentences making it a valuable resource for educational goals. The treebank is constructed with a bootstrapping approach by means of available tagging and parsing tools and manually correcting the annotations. The data is splitted into standard train, development and test set in the CoNLL dependency format and is freely available to researchers.

Improving reordering performance using higher order and structural features
Mitesh M. Khapra, Ananthakrishnan Ramanathan, and Karthik Visweswariah

Recent work has shown that word aligned data can be used to learn a model for reordering source sentences to match the target order. This model learns the cost of putting a word immediately before another word and finds the best reordering by solving an instance of the Traveling Salesman Problem (TSP). However, for efficiently solving the TSP, the model is restricted to pairwise features which examine only a pair of words and their neighborhood. In this work, we go beyond these pairwise features and learn a model to rerank the n-best reorderings produced by the TSP model using higher order and structural features which help in capturing longer range dependencies. In addition to using a more informative set of source side features, we also capture target side features indirectly by using the translation score assigned to a reordering. Our experiments, involving Urdu-English, show that the proposed approach outperforms a state-of-the-art PBSMT system which uses the TSP model for reordering by 1.3 BLEU points, and a publicly available state-of-the-art

MT system, Hiero, by 3 BLEU points.

Massively Parallel Suffix Array Queries and On-Demand Phrase Extraction for Statistical Machine Translation Using GPUs
Hua He, Jimmy Lin, and Adam Lopez

Translation models in statistical machine translation can be scaled to large corpora and arbitrarily-long phrases by looking up translations of source phrases "on the fly" in an indexed parallel corpus using suffix arrays. However, this can be slow because on-demand extraction of phrase tables is computationally expensive. We address this problem by developing novel algorithms for general purpose graphics processing units (GPUs), which enable suffix array queries for phrase lookup and phrase extraction to be massively parallelized. Compared to a highly-optimized, state-of-the-art serial CPU-based implementation, our techniques achieve at least an order of magnitude improvement in terms of throughput. This work demonstrates the promise of massively parallel architectures and the potential of GPUs for tackling computationally-demanding problems in statistical machine translation and language processing.

[TACL] Branch and Bound Algorithm for Dependency Parsing with Non-local Features
Xian Qian and Yang Liu

Graph based dependency parsing is inefficient when handling non-local features due to high computational complexity of inference. In this paper, we proposed an exact and efficient decoding algorithm based on the Branch and Bound (B&B) framework where non-local features are bounded by a linear combination of local features. Dynamic programming is used to search the upper bound. Experiments are conducted on English PTB and Chinese CTB datasets. We achieved competitive Unlabeled Attachment Score (UAS) when no additional resources are available: 93.17% for English and 87.25% for Chinese. Parsing speed is 177 words per second for English and 97 words per second for Chinese. Our algorithm is general and can be adapted to non-projective dependency parsing or other graphical models.

Discriminative Training of 150 Million Translation Parameters and Its Application to Pruning
Hendra Setiawan and Bowen Zhou

Until recently, the application of discriminative training to log linear-based statistical machine translation has been limited to tuning the weights of a handful of features or training features with a limited number of parameters. In this paper, we propose to scale up discriminative training to train features with 150 million parameters, which is one order of magnitude higher than previously published system, and to apply discriminative training to redistribute probability mass that is lost due to model pruning. The experimental results confirm the effectiveness of our proposals on NIST MT06 test set over a strong hierarchical phrase-based baseline.

Applying Pairwise Ranked Optimisation to Improve the Interpolation of Translation Models
Barry Haddow

In Statistical Machine Translation we often have to combine different sources of parallel training data to build a good system. One way of doing this is to build separate translation models from each data set and linearly interpolate them, and to date the main method for optimising the interpolation weights is to minimise the model perplexity on a heldout set. In this work, rather than optimising for this indirect measure, we directly optimise for BLEU on the tuning set and show improvements in average performance over two data sets and 8 language pairs.

Dialectal Arabic to English Machine Translation: Pivoting through Modern Standard Arabic
Wael Salloum and Nizar Habash

Modern Standard Arabic (MSA) has a wealth of natural language processing (NLP) tools and resources. In comparison, resources for dialectal Arabic (DA), the unstandardized spoken varieties of Arabic, are still lacking. We present ELISSA, a machine translation (MT) system for DA to MSA. ELISSA employs a rule-based

approach that relies on morphological analysis, transfer rules and dictionaries in addition to language models to produce MSA paraphrases of DA sentences. ELISSA can be employed as a general preprocessor for DA when using MSA NLP tools. A manual error analysis of ELISSA's output shows that it produces correct MSA translations over 93% of the time. Using ELISSA to produce MSA versions of DA sentences as part of an MSA-pivoting DA-to-English MT solution, improves BLEU scores on multiple blind test sets between 0.6% and 1.4%.

What to do about bad language on the internet
Jacob Eisenstein

The rise of social media has brought computational linguistics in ever-closer contact with bad language: text that defies our expectations about vocabulary, spelling, and syntax. This paper surveys the landscape of bad language, and offers a critical review of the NLP community's response, which has largely followed two paths: normalization and domain adaptation. Each approach is evaluated in the context of theoretical and empirical work on computer-mediated communication. In addition, the paper presents a quantitative analysis of the lexical diversity of social media text, and its relationship to other corpora.

Minibatch and Parallelization for Online Large Margin Structured Learning
Kai Zhao and Liang Huang

Online learning algorithms such as perceptron and MIRA have become popular for many NLP tasks thanks to their simpler architect-ure and faster convergence over batch learning methods. However, while batch learning such as CRF is easily parallelizable, online learning is much harder to parallelize: previous efforts often witness a decrease in the converged accuracy, and the speedup is typically very small (~3) even with many (10+) processors. We instead present a much simpler architecture based on "mini-batches", which is trivially parallelizable. We show that, unlike previous methods, minibatch learning (in serial mode) actually improves the converged accuracy for both perceptron and MIRA learning, and when combined with simple parallelization, minibatch leads to very significant speedups (up to 9x on 12 processors) on state-of-the-art parsing and tagging systems.

Improved Part-of-Speech Tagging for Online Conversational Text with Word Clusters
Olutobi Owoputi, Brendan O'Connor, Chris Dyer, Kevin Gimpel, Nathan Schneider, and Noah A. Smith

We consider the problem of part-of-speech tagging for informal, online conversational text. We systematically evaluate the use of large-scale unsupervised word clustering and new lexical features to improve tagging accuracy. With these features, our system achieves state-of-the-art tagging results on both Twitter and IRC POS tagging tasks; Twitter tagging is improved from 90% to 93% accuracy (more than 3% absolute). Qualitative analysis of these word clusters yields insights about NLP and linguistic phenomena in this genre. Additionally, we contribute the first POS annotation guidelines for such text and release a new dataset of English language tweets annotated using these guidelines. Tagging software, annotation guidelines, and large-scale word clusters are available at: http://www.ark.cs.cmu.edu/TweetNLP This paper describes release 0.3 of the "CMU Twitter Part-of-Speech Tagger" and annotated data.

Parser lexicalisation through self-learning
Marek Rei and Ted Briscoe

We describe a new self-learning framework for parser lexicalisation that requires only a plain-text corpus of in-domain text. The method first creates augmented versions of dependency graphs by applying a series of modifications designed to directly capture higher-order lexical path dependencies. Scores are assigned to each edge in the graph using statistics from an automatically parsed background corpus. As bilexical dependencies are sparse, a novel directed distributional word similarity measure is used to smooth edge score estimates. Edge scores are then combined into graph scores and used for reranking the top-n analyses found by the unlexicalised parser. The approach achieves significant improvements on WSJ and biomedical text over the unlexicalised baseline parser, which is originally trained on a subset of the Brown corpus.

Mining User Relations from Online Discussions using Sentiment Analysis and Probabilistic Matrix Factorization
Minghui Qiu, Liu Yang, and Jing Jiang

Advances in sentiment analysis have enabled extraction of user relations implied in online textual exchanges such as forum posts. However, recent studies in this direction only consider direct relation extraction from text. As user interactions can be sparse in online discussions, we propose to apply collaborative filtering through probabilistic matrix factorization to generalize and improve the opinion matrices extracted from forum posts. Experiments with two tasks show that the learned latent factor representation can give good performance on a relation polarity prediction task and improve the performance of a subgroup detection task.

Focused training sets to reduce noise in NER feature models
Amber McKenzie

Feature and context aggregation play a large role in current NER systems, allowing significant opportunities for research into op-timizing these features to cater to different domains. This work strives to reduce the noise introduced into aggregated features from dis-parate and generic training data in order to al-low for contextual features that more closely model the entities in the target data. The pro-posed approach trains models based on only a part of the training set that is more similar to the target domain. To this end, models are trained for an existing NER system using the top documents from the training set that are similar to the target document in order to demonstrate that this technique can be applied to improve any pre-built NER system. Initial results show an improvement over the Illinois tagger with a weighted average F1 score of 91.67 compared to the University of Illinois NE tagger's score of 91.32. This research serves as a proof-of-concept for future planned work to cluster the training documents to produce a number of more focused models from a given training set, thereby reducing noise and extracting a more representative feature set.

Learning to Relate Literal and Sentimental Descriptions of Visual Properties
Mark Yatskar, Svitlana Volkova, Asli Celikyilmaz, Bill Dolan, and Luke Zettlemoyer

Language can describe our visual world at many levels, including not only what is literally there but also the sentiment that it invokes. In this paper, we study visual language, both literal and sentimental, that describes the overall appearance and style of virtual characters. Sentimental properties, including labels such as "youthful" or "country western," must be inferred from descriptions of the more literal properties, such as facial features and clothing selection. We present a new dataset, collected to describe Xbox avatars, as well as models for learning the relationships between these avatars and their literal and sentimental descriptions. In a series of experiments, we demonstrate that such learned models can be used for a range of tasks, including predicting sentimental words and using them to rank and build avatars. Together, these results demonstrate that sentimental language provides a concise (though noisy) means of specifying low-level visual properties.

Morphological Analysis and Disambiguation for Dialectal Arabic
Nizar Habash, Ryan Roth, Owen Rambow, Ramy Eskander, and Nadi Tomeh

The many differences between Dialectal Arabic and Modern Standard Arabic (MSA) pose a challenge to the majority of Arabic natural language processing tools, which are designed for MSA. In this paper, we retarget an existing state-of-the-art MSA morphological tagger to Egyptian Arabic (ARZ). Our evaluation demonstrates that our ARZ morphology tagger outperforms its MSA variant on ARZ input in terms of accuracy in part-of-speech tagging, diacritization, lemmatization and tokenization; and in terms of utility for ARZ-to-English statistical machine translation.

Using a Supertagged Dependency Language Model to Select a Good Translation in System Combination
Wei-Yun Ma and Kathleen McKeown

We present a novel, structured language model - Supertagged Dependency Language Model to model the syntactic dependencies between words. The goal is to identify ungrammatical hypotheses from a set of candidate translations in a MT system combination framework and help select the best translation candidates using a variety of sentence-level features. We use a two-step mechanism based on constituent parsing and elementary tree extraction to obtain supertags and their dependency relations. Our experiments show that the structured language model provides significant improvement in the framework of sentence-level system combination.

Dudley North visits North London: Learning When to Transliterate to Arabic
Mahmoud Azab, Houda Bouamor, Behrang Mohit, and Kemal Oflazer

We report the results of our work on automating the transliteration decision of named entities for English to Arabic machine translation. We construct a classification-based framework to automate this decision, evaluate our classifier both in the limited news and the diverse Wikipedia domains, and achieve promising accuracy. Moreover, we demonstrate a reduction of translation error and an improvement in the performance of an English-to-Arabic machine translation system.

Better Twitter Summaries?
Joel Judd and Jugal Kalita

This paper describes an approach to improve summaries for a collection of Twitter posts created using the Phrase Reinforcement (PR) Algorithm (Sharifi et al. 2010). The PR algorithm often generates summaries with excess text and noisy speech. We parse these summaries using a dependency parser and use the dependencies to eliminate some of the excess text and build better-formed summaries. We compare the results to those obtained using the PR Algorithm.

Training MRF-Based Phrase Translation Models using Gradient Ascent
Jianfeng Gao and Xiaodong He

This paper presents a general, statistical framework for modeling phrase translation via Markov random fields. The model allows for arbituary features extracted from a phrase pair to be incorporated as evidence. The parameters of the model are estimated using a large-scale discriminative training approach that is based on stochastic gradient ascent and an N-best list based expected BLEU as the objective function. The model is easy to be incoporated into a standard phrase-based statistical machine translation system, requiring no code change in the runtime engine. Evaluation is performed on two Europarl translation tasks, German-English and French-English. Results show that incoporating the Markov random field model significantly improves the performance of a state-of-the-art phrase-based machine translation system, leading to a gain of 0.8-1.3 BLEU points.

Automatic Morphological Enrichment of a Morphologically Underspecified Treebank
Sarah Alkuhlani, Nizar Habash, and Ryan Roth

In this paper, we study the problem of automatic enrichment of a morphologically underspecified treebank for Arabic, a morphologically rich language. We show that we can map from a tagset of size six to one with 485 tags at an accuracy rate of 94%-95%. We can also identify the unspecified lemmas in the treebank with an accuracy over 97%. Furthermore, we demonstrate that using our automatic annotations improves the performance of a state-of-the-art Arabic morphological tagger. Our approach combines a variety of techniques from corpus-based statistical models to linguistic rules that target specific phenomena. These results suggest that the cost of treebanking can be reduced by designing underspecified treebanks that can be subsequently enriched automatically.

A Beam-Search Decoder for Normalization of Social Media Text with Application to Machine Translation
Pidong Wang and Hwee Tou Ng

Social media texts are written in an informal style, which hinders other natural language processing (NLP) applications such as machine translation. Text normalization is thus important for processing of social media text. Previous work mostly focused on normalizing words by replacing an informal word with its formal form. In this paper, to further improve other downstream NLP applications, we argue that other normalization operations should also be performed, e.g., missing word recovery and punctuation correction. A novel beam-search decoder is proposed to effectively integrate various normalization operations. Empirical results show that our system obtains statistically significant improvements over two strong baselines in both normalization and translation tasks, for both Chinese and English.

Parameter Estimation for LDA-Frames
Jiří Materna

LDA-frames is an unsupervised approach for identifying semantic frames from semantically unlabeled text corpora, and seems to be a useful competitor for manually created databases of selectional preferences. The most limiting property of the algorithm is such that the number of frames and roles must be predefined. In this paper we present a modification of the LDA-frames algorithm allowing the number of frames and roles to be determined automatically, based on the character and size of training data.

Approximate PCFG Parsing Using Tensor Decomposition
Shay B. Cohen, Giorgio Satta, and Michael Collins

We provide an approximation algorithm for PCFG parsing, which asymptotically improves time complexity with respect to the input grammar size, and prove upper bounds on the approximation quality. We test our algorithm on two treebanks, and get significant improvements in parsing speed.

Negative Deceptive Opinion Spam
Myle Ott, Claire Cardie, and Jeffrey T. Hancock

The rising influence of user-generated online reviews has led to growing incentive for businesses to solicit and manufacture deceptive opinion spam—fictitious reviews that have been deliberately written to sound authentic and deceive the reader. Recently, Ott et al. (2011) have introduced an opinion spam dataset containing gold standard deceptive positive hotel reviews. However, the complementary problem of negative deceptive opinion spam, intended to slander competitive offerings, remains largely unstudied. Following an approach similar to Ott et al. (2011), in this work we create and study the first dataset of deceptive opinion spam with negative sentiment reviews. Based on this dataset, we find that standard n-gram text categorization techniques can detect negative deceptive opinion spam with performance far surpassing that of human judges. Finally, in conjunction with the aforementioned positive review dataset, we consider the possible interactions between sentiment and deception, and present initial results that encourage further exploration of this relationship.

Improving speech synthesis quality by reducing pitch peaks in the source recordings
Luisina Violante, Pablo Rodríguez Zivic, and Agustín Gravano

We present a method for improving the perceived naturalness of corpus-based speech synthesizers. It consists in removing pronounced pitch peaks in the original recordings, which typically lead to noticeable discontinuities in the synthesized speech. We perceptually evaluated this method using two concatenative and two HMM-based synthesis systems, and found that using it on the source recordings managed to improve the naturalness of the synthesizers and had no effect on their intelligibility.

Robust Systems for Preposition Error Correction Using Wikipedia Revisions
Aoife Cahill, Nitin Madnani, Joel Tetreault, and Diane Napolitano

We show that existing methods for training preposition error correction systems, whether using well-edited text or error-annotated corpora, do not generalize across very different test sets. We present a new, large errorannotated corpus and use it to train systems that generalize across three different test sets, each from a different domain and with different error characteristics. This new corpus is automatically extracted from Wikipedia revisions and contains over one million instances of preposition corrections.

Supervised Bilingual Lexicon Induction with Multiple Monolingual Signals
Ann Irvine and Chris Callison-Burch

Prior research into learning translations from source and target language monolingual texts has treated the task as an unsupervised learning problem. Although many techniques take advantage of a seed bilingual lexicon, this work is the first to use that data for supervised learning to combine a diverse set of signals derived from a pair of monolingual corpora into a single discriminative model. Even in a low resource machine translation setting, where induced translations have the potential to improve performance substantially, it is reasonable to assume access to some amount of data to perform this kind of optimization. Our work shows that only a few hundred translation pairs are needed to achieve strong performance on the bilingual lexicon induction task, and our approach yields an average relative gain in accuracy of nearly 50% over an unsupervised baseline. Large

gains in accuracy hold for all 22 languages (low and high resource) that we investigate.

Creating Reverse Bilingual Dictionaries
Khang Nhut Lam and Jugal Kalita

Bilingual dictionaries are expensive resources and not many are available when one of the languages is resource-poor. In this paper, we propose algorithms for creation of new reverse bilingual dictionaries from existing bilingual dictionaries in which English is one of the two languages. The main idea of our algorithms is to exploit the similarity between word-concept pairs using the English Wordnet to produce reverse dictionary entries. Since our algorithms rely on available bilingual dictionaries, they are applicable to any bilingual dictionary as long as one of the two languages has Wordnet type lexical ontology.

Identification of Temporal Event Relationships in Biographical Accounts
Lucian Silcox and Emmett Tomai

This paper examines the efficacy of the application of a pre-existing technique in the area of event-event temporal relationship identification. We attempt to both reproduce the results of said technique, as well as extend the previous work with application to a newly-created domain of biographical data. We find that initially the simpler feature sets perform as expected, but that the final improvement to the feature set underperforms. In response, we provide an analysis of the individual features and identify differences existing between the two corpora.

Predicative Adjectives: An Unsupervised Criterion to Extract Subjective Adjectives
Michael Wiegand, Josef Ruppenhofer, and Dietrich Klakow

We examine predicative adjectives as an unsupervised criterion to extract subjective adjectives. We do not only compare this criterion with a weakly supervised extraction method but also with gradable adjectives, i.e. another highly subjective subset of adjectives that can be extracted in an unsupervised fashion. In order to prove the robustness of this extraction method, we will evaluate the extraction with the help of two different state-of-the-art sentiment lexicons (as a gold standard).

Modeling Syntactic and Semantic Structures in Hierarchical Phrase-based Translation
Junhui Li, Philip Resnik, and Hal Daumé III

Incorporating semantic structure into a linguistics-free translation model is challenging, since semantic structures are closely tied to syntax. In this paper, we propose a two-level approach to exploiting predicate-argument structure reordering in a hierarchical phrase-based translation model. First, we introduce linguistically motivated constraints into a hierarchical model, guiding translation phrase choices in favor of those that respect syntactic boundaries. Second, based on such translation phrases, we propose a predicate-argument structure reordering model that predicts reordering not only between an argument and its predicate, but also between two arguments. Experiments on Chinese-to-English translation demonstrate that both advances significantly improve translation accuracy.

Using Derivation Trees for Informative Treebank Inter-Annotator Agreement Evaluation
Seth Kulick, Ann Bies, Justin Mott, Mohamed Maamouri, Beatrice Santorini, and Anthony Kroch

This paper discusses the extension of a system developed for automatic discovery of treebank annotation inconsistencies over an entire corpus to the particular case of evaluation of inter-annotator agreement. This system makes for a more informative IAA evaluation than other systems because it pinpoints the inconsistencies and groups them by their structural types. We evaluate the system on two corpora - (1) a corpus of English web text, and (2) a corpus of Modern British English.

Embracing Ambiguity: A Comparison of Annotation Methodologies for Crowdsourcing Word Sense Labels
David Jurgens

Word sense disambiguation aims to identify which meaning of a word is present in a given usage. Gathering word sense annotations is a laborious and difficult task. Several methods have been proposed to gather sense annotations using large numbers of untrained annotators, with mixed results. We propose three new annotation methodologies for gathering word senses where untrained annotators are allowed to use multiple labels and weight the senses. Our findings show that given the appropriate annotation task, untrained workers can obtain at least as high agreement as annotators in a controlled setting, and in aggregate generate equally as good of a sense labeling.

Compound Embedding Features for Semi-supervised Learning
Mo Yu, Tiejun Zhao, Daxiang Dong, Hao Tian, and Dianhai Yu

To solve data sparsity problem, recently there has been a trend in discriminative methods of NLP to use representations of lexical items learned from unlabeled data as features. In this paper, we investigated the usage of word representations learned by neural language models, i.e. word embeddings. The direct usage has disadvantages such as large amount of computation, inadequacy with dealing word ambiguity and rare-words, and the problem of linear non-separability. To overcome these problems, we instead built compound features from continuous word embeddings based on clustering. Experiments showed that the compound features not only improved the performances on several NLP tasks, but also ran faster, suggesting the potential of embeddings.

On Quality Ratings for Spoken Dialogue Systems – Experts vs. Users
Stefan Ultes, Alexander Schmitt, and Wolfgang Minker

In the field of Intelligent User Interfaces, Spoken Dialogue Systems (SDSs) play a key role as speech represents a true intuitive means of human communication. Deriving information about its quality can help rendering SDSs more user-adaptive. Work on automatic estimation of subjective quality usually relies on statistical models. To create those, manual data annotation is required, which may be performed by actual users or by experts. Here, both variants have their advantages and drawbacks. In this paper, we analyze the relationship between user and expert ratings by investigating models which combine the advantages of both types of ratings. We explore two novel approaches using statistical classification methods and evaluate those with a preexisting corpus providing user and expert ratings. After analyzing the results, we eventually recommend to use expert ratings instead of user ratings in general.

Overcoming the Memory Bottleneck in Distributed Training of Latent Variable Models of Text
Yi Yang, Alexander Yates, and Doug Downey

Large unsupervised latent variable models (LVMs) of text, such as Latent Dirichlet Allocation models or Hidden Markov Models (HMMs), are constructed using parallel training algorithms on computational clusters. The memory required to hold LVM parameters forms a bottleneck in training more powerful models. In this paper, we show how the memory required for parallel LVM training can be reduced by partitioning the training corpus to minimize the number of unique words on any computational node. We present a greedy document partitioning technique for the task. For large corpora, our approach reduces memory consumption by over 50%, and trains the same models up to three times faster, when compared with existing approaches for parallel LVM training.

Processing Spontaneous Orthography
Ramy Eskander, Nizar Habash, Owen Rambow, and Nadi Tomeh

In cases in which there is no standard orthography for a language or language variant, written texts will display a variety of orthographic choices. This is problematic for natural language processing (NLP) because it creates spurious data sparseness. We study the transformation of spontaneously spelled Egyptian Arabic into a conventionalized orthography which we have previously proposed for NLP purposes. We show that a two-stage process can reduce divergences from this standard by 69%, making subsequent processing of Egyptian Arabic easier.

Purpose and Polarity of Citation: Towards NLP-based Bibliometrics
Amjad Abu-Jbara, Jefferson Ezra, and Dragomir Radev

Bibliometric measures are commonly used to estimate the popularity and the impact of published research. Existing bibliometric measures provide "quantitative" indicators of how good a published paper is. This does not necessarily reflect the "quality" of the work presented in the paper. For example, when h-index is computed for a researcher, all incoming citations are treated equally, ignoring the fact that some of these citations might be negative. In this paper, we propose using NLP to add a "qualitative" aspect to biblometrics. We analyze the text that accompanies citations in scientific articles (which we term citation context). We propose supervised methods for identifying citation text and analyzing it to determine the purpose (i.e. author intention) and the polarity (i.e. author sentiment) of citation.

Estimating effect size across datasets
Anders Søgaard

Most NLP tools are applied to text that is different from the kind of text they were evaluated on. Common evaluation practice prescribes significance testing across data points in available test data, but typically we only have a single test sample. This short paper argues that in order to assess the robustness of NLP tools we need to evaluate them on diverse samples, and we consider the problem of finding the most appropriate way to estimate the true effect size across datasets of our systems over their baselines. We apply meta-analysis and show experimentally - by comparing estimated error reduction over observed error reduction on held-out datasets - that this method is significantly more predictive of success than the usual practice of using macro- or micro-averages. Finally, we present a new parametric meta-analysis based on non-standard assumptions that seems superior to standard parametric meta-analysis.

Systematic Comparison of Professional and Crowdsourced Reference Translations for Machine Translation
Rabih Zbib, Gretchen Markiewicz, Spyros Matsoukas, Richard Schwartz, and John Makhoul

We present a systematic study of the effect of crowdsourced translations on Machine Translation performance. We compare Machine Translation systems trained on the same data but with translations obtained using Amazon's Mechanical Turk vs. professional translations, and show that the same performance is obtained from Mechanical Turk translations at 1/5th the cost. We also show that adding a Mechanical Turk reference translation of the de- velopment set improves parameter tuning and output evaluation.

Down-stream effects of tree-to-dependency conversions
Jakob Elming, Anders Johannsen, Sigrid Klerke, Emanuele Lapponi, Hector Martinez Alonso, and Anders Søgaard

Dependency analysis relies on morphosyntactic evidence, as well as semantic evidence. In some cases, however, morphosyntactic evidence seems to be in conflict with semantic evidence. For this reason dependency grammar theories, annotation guidelines and tree-to-dependency conversion schemes often differ in how they analyze various syntactic constructions. Most experiments for which constituent-based treebanks such as the Penn Treebank are converted into dependency treebanks rely blindly on one of four-five widely used tree-to-dependency conversion schemes. This paper evaluates the down-stream effect of choice of conversion scheme, showing that it has dramatic impact on end results.

Student Research Workshop Posters

From Language to Family and Back: Native Language and Language Family Identification from English Text
Ariel Stolerman, Aylin Caliskan, and Rachel Greenstadt

Revealing an anonymous author's traits from text is a well-researched area. In this paper we aim to identify the native language and language family of a non-native English author, given his/her English writings. We extract features from the text based on prior work, and extend or modify it to construct different feature sets, and use support vector machines for classification. We show that native language identification accuracy can

be improved by up to 6.43% for a 9-class task, depending on the feature set, by introducing a novel method to incorporate language family information. In addition we show that introducing grammar-based features improves accuracy of both native language and language family identification.

Critical Reflections on Evaluation Practices in Coreference Resolution
Gordana Ilic Holen

In this paper we revisit the task of quantitative evaluation of coreference resolution systems. We review the most used metrics (MUC, B-CUBED, CEAF and BLANC) on basis of their evaluation of coreference resolution in five texts from the OntoNotes corpus. We examine both correlation between the metrics and the degree to which our human judgement of coreference resolution agrees with the metrics. In conclusion we claim that loss of information value is an essential factor in human perception of the degree of success or failure of coreference resolution and that including a layer of mention information weight could improve both the evaluation and coreference resolution itself.

A Rule-based Approach for Karmina Generation
Franky Franky

We present our work in generating Karmina, an old Malay poetic form for Indonesian language. Karmina is a poem with two lines that consists of a hook (sampiran) on the first line and a message on the second line. One of the unique aspects of Karmina is in the absence of discourse relation between its hook and message. We approached the problem by generating the hooks and the messages in separate processes using predefined schemas and a manually built knowledge base. The Karminas were produced by randomly pairing the messages with the hooks, subject to the constraints imposed on the rhymes and on the structure similarity. Syllabifications were performed on the cue words of the hooks and messages to ensure the generated pairs have matching rhymes. We were able to generate a number of positive examples while still leaving room for improvement, particularly in the generation of the messages, which currently are still limited, and in filtering the negative results.

A Machine Learning Approach to Automatic Term Extraction using a Rich Feature Set
Merley Conrado, Thiago Pardo, and Solange Rezende

In this paper we propose an automatic term extraction approach that uses machine learning incorporating varied and rich features of candidate terms. In our preliminary experiments, we also tested different attribute selection methods to verify which features are more relevant for automatic term extraction. We achieved state of the art results for unigram extraction in Brazilian Portuguese.

Ontology Label Translation
Mihael Arcan and Paul Buitelaar

Our research investigates the translation of ontology labels, which has applications in multilingual knowledge access. Ontologies are often defined only in one language, mostly English. To enable knowledge access across languages, such monolingual ontologies need to be translated into other languages. The primary challenge in ontology label translation is the lack of context, which makes this task rather different than document translation. The core objective therefore, is to provide statistical machine translation (SMT) systems with additional context information. In our approach, we first extend standard SMT by enhancing a translation model with context information that keeps track of surrounding words for each translation. We then match these context words with the ontology vocabulary. The ontology label placement will now inform an approach that computes a semantic similarity between the phrase pair context vector from the parallel corpus and a vector of noun phrases that occur in surrounding ontology labels, i.e. siblings and parents. We applied our approach to the translation of a financial ontology, translating from English to German, using Europarl as parallel corpus. This experiment showed that our approach can provide a slight improvement over standard SMT for this task, without exploiting any additional domain-specific resources. The core benefit of the approach is in disambiguation of translations rather than translation generation.

Entrainment in Spoken Dialogue Systems: Adopting, Predicting and Influencing User Behavior
Rivka Levitan

Entrainment is the phenomenon of conversational partners becoming more similar to each other as they speak. This thesis proposal presents a comprehensive look at entrainment in human conversations and how entrainment may be incorporated into the design of spoken dialogue systems in order to improve system performance and user satisfaction. We compare different kinds of entrainment in both classic and novel dimensions, provide experimental results on the utility of entrainment, and show that entrainment can be used to improve a system's ASR performance and turn-taking decisions.

Helpfulness-Guided Review Summarization
Wenting Xiong

Review mining and summarization has been a hot topic for the past decade. A lot of effort has been devoted to aspect detection and sentiment analysis under the assumption that every review has the same utility for related tasks. However, reviews are not equally helpful as indicated by user-provided helpfulness assessment associated with the reviews. In this thesis, we propose a novel review summarization framework which summarizes review content under the supervision of automated assessment of review helpfulness. This helpfulness-guided framework can be easily adapted to traditional review summarization tasks, for a wide range of domains.

Statistical Machine Translation in Low Resource Settings
Ann Irvine

My thesis will explore ways to improve the performance of statistical machine translation (SMT) in low resource conditions. Specifically, it aims to reduce the dependence of modern SMT systems on expensive parallel data. We define low resource settings as having only small amounts of parallel data available, which is the case for many language pairs. All current SMT models use parallel data during training for extracting translation rules and estimating translation probabilities. The theme of our approach is the integration of information from alternate data sources, other than parallel corpora, into the statistical model. In particular, we focus on making use of large monolingual and comparable corpora. By augmenting components of the SMT framework, we hope to extend its applicability beyond the small handful of language pairs with large amounts of available parallel text.

Domain-Independent Captioning of Domain-Specific Images
Rebecca Mason

Automatically describing visual content is an extremely difficult task, with hard AI problems in Computer Vision (CV) and Natural Language Processing (NLP) at its core. Previous work has relied on supervised visual recognition systems to determine the content of images. Because these systems require massive amounts of hand-labeled data for training, the number of visual classes that can be recognized is typically very small. We argue that these approaches place unrealistic limits on the kinds of images that can be captioned, and are unlikely to produce captions which reflect human interpretations.

We present a framework for image caption generation that does not rely on visual recognition systems, which we have implemented on a dataset of online shopping images and product descriptions. We propose future work to improve this method, and extensions for other domains of images and natural text.

Reversing Morphological Tokenization in English-to-Arabic SMT
Mohammad Salameh, Colin Cherry, and Grzegorz Kondrak

Morphological tokenization has been used in machine translation for morphologically complex languages to reduce lexical sparsity. Unfortunately, when translating into a morphologically complex language, recombining segmented tokens to generate original word forms is not a trivial task, due to morphological, phonological and orthographic adjustments that occur during tokenization. We review a number of detokenization schemes for Arabic, such as rule-based and table-based approaches and show their limitations. We then propose a novel detokenization scheme that uses a character-level discriminative string transducer to predict the original form of a segmented word. In a comparison to a state-of-the-art approach, we demonstrate slightly better detokenization error rates, without the need for any hand-crafted rules. We also demonstrate the effectiveness of our approach in an English-to-Arabic translation task.

Reducing Annotation Effort on Unbalanced Corpus based on Cost Matrix
Wencan Luo, Diane Litman, and Joel Chan

Annotated corpora play a significant role in many NLP applications. However, annotation by humans is time-consuming and costly. In this paper, a high recall predictor based on a cost-sensitive learner is proposed as a method to semi-automate the annotation of unbalanced classes. We demonstrate the effectiveness of our approach in the context of one form of unbalanced task: annotation of transcribed human-human dialogues for presence/ absence of uncertainty. In two data sets, our cost-matrix based method of uncertainty annotation achieved high levels of recall while maintaining acceptable levels of accuracy. The method is able to reduce human annotation effort by about 80% without a significant loss in data quality, as demonstrated by an extrinsic evaluation showing that results originally achieved using manually-obtained uncertainty annotations can be replicated using semi-automatically obtained uncertainty annotations.

User Goal Change Model for Spoken Dialog State Tracking
Yi Ma

In this paper, a Maximum Entropy Markov Model (MEMM) for dialog state tracking is proposed to efficiently handle user goal evolvement in two steps. The system first predicts the occurrence of a user goal change based on linguistic features and dialog context for each dialog turn, and then the proposed model could utilize this user goal change information to infer the most probable dialog state sequence which underlies the evolvement of user goal during the dialog. It is believed that with the suggested various domain independent feature functions, the proposed model could better exploit not only the intra-dependencies within long ASR N-best lists but also the inter-dependencies of the observations across dialog turns, which leads to more efficient and accurate dialog state inference.

Large-Scale Paraphrasing for Natural Language Understanding
Juri Ganitkevitch

We examine the application of data-driven paraphrasing to natural language understanding. We leverage bilingual parallel corpora to extract a large collection of syntactic paraphrase pairs, and introduce an adaptation scheme that allows us to tackle a variety of text transformation tasks via paraphrasing. We evaluate our system on the sentence compression task. Further, we use distributional similarity measures based on context vectors derived from large monolingual corpora to annotate our paraphrases with an orthogonal source of information. This yields significant improvements in our compression system's output quality, achieving state-of-the-art performance. Finally, we propose a refinement of our paraphrases by classifying them into natural logic entailment relations. By extending the synchronous parsing paradigm towards these entailment relations, we will enable our system to perform recognition of textual entailment.

Posters and Demonstrations Session

The Poster and Demonstrations Session will be held in the Grand Atrium of the 200 Building from 6:30–8:30pm, and will include a buffet dinner. The Demonstrations abstracts are printed below, since they were not a part of the Poster Madness! session.

DALE: A Word Sense Disambiguation System for Biomedical Documents Trained using Automatically Labeled Examples
Judita Preiss and Mark Stevenson

Automatic interpretation of documents is hampered by the fact that language contains terms which have multiple meanings. These ambiguities can still be found when language is restricted to a particular domain, such as biomedicine. Word Sense Disambiguation (WSD) systems attempt to resolve these ambiguities but are often only able to identify the meanings for a small set of ambiguous terms. DALE (Disambiguation using Automatically Labeled Examples) is a supervised WSD system that can disambiguate a wide range of ambiguities found in biomedical documents. DALE uses the UMLS Metathesaurus as both a sense inventory and as a source of information for automatically generating labeled training examples. DALE is able to disambiguate biomedical documents with the coverage of unsupervised approaches and accuracy of supervised methods.

Topic Models and Metadata for Visualizing Text Corpora
Justin Snyder, Rebecca Knowles, Mark Dredze, Matthew Gormley, and Travis Wolfe

Effectively exploring and analyzing large text corpora requires visualizations that provide a high level summary. Past work has relied on faceted browsing of document metadata or on natural language processing of document text. In this paper, we present a new web-based tool that integrates topics learned from an unsupervised topic model in a faceted browsing experience. The user can manage topics, filter documents by topic and summarize views with metadata and topic graphs. We report a user study of the usefulness of topics in our tool.

TMTprime: A Recommender System for MT and TM Integration
Aswarth Abhilash Dara, Sandipan Dandapat, Declan Groves, and Josef van Genabith

TMTprime is a recommender system that facilitates the effective use of both translation memory (TM) and machine translation (MT) technology within industrial language service providers (LSPs) localization workflows. LSPs have long used Translation Memory (TM) technology to assist the translation process. Recent research shows how MT systems can be combined with TMs in Computer Aided Translation (CAT) systems, selecting either TM or MT output based on sophisticated translation quality estimation without access to a reference. However, to date there are no commercially available frameworks for this. TMTprime takes confidence estimation out of the lab and provides a commercially viable platform that allows for the seamless integration of MT with legacy TM systems to provide the most effective (least effort/cost) translation options to human translators, based on the TMTprime confidence score.

Anafora: A Web-based General Purpose Annotation Tool
Wei-Te Chen and Will Styler

Anafora is a newly-developed open source web-based text annotation tool built to be lightweight, flexible, easy to use and capable of annotating with a variety of schemas, simple and complex. Anafora allows secure web-based annotation of any plaintext file with both spanned (e.g. named entity or markable) and relation annotations, as well as adjudication for both types of annotation. Anafora offers automatic set assignment and progress-tracking, centralized and human-editable XML annotation schemas, and file-based storage and organization of data in a human-readable single-file XML format.

A Web Application for the Diagnostic Evaluation of Machine Translation over Specific Linguistic Phenomena
Antonio Toral, Sudip Kumar Naskar, Joris Vreeke, Federico Gaspari, and Declan Groves

This paper presents a web application and a web service for the diagnostic evaluation of Machine Translation (MT). These web-based tools are built on top of DELiC4MT, an open-source software package that assesses the performance of MT systems over user-defined linguistic phenomena (lexical, morphological, syntactic and

semantic). The advantage of the web-based scenario is clear; compared to the standalone tool, the user does not need to carry out any installation, configuration or maintenance of the tool.

KooSHO: Japanese Text Input Environment based on Aerial Hand Writing
Masato Hagiwara and Soh Masuko

Hand gesture-based input systems have been in active research, yet most of them focus only on single character recognition. We propose KooSHO: an environment for Japanese text input based on aerial hand gestures. The system provides an integrated experience of character input, Kana-Kanji conversion, and search result visualization. To achieve faster input, users only have to input consonant, which is then converted directly to Kanji sequences by direct consonant decoding. The system also shows suggestions to complete the user input. The comparison with a voice recognition system and a screen keyboard showed that KooSHO can be a more practical solution compared to the existing system.

UMLS::Similarity: Measuring the Relatedness and Similarity of Biomedical Concepts
Bridget McInnes, Ted Pedersen, Serguei Pakhomov, Ying Liu, and Genevieve Melton-Meaux

UMLS::Similarity is freely available open source software that allows a user to measure the semantic similarity or relatedness of biomedical terms found in the Unified Medical Language System (UMLS). It is written in Perl and can be used via a command line interface, an API, or a Web interface.

KELVIN: a tool for automated knowledge base construction
Paul McNamee, James Mayfield, Tim Finin, Tim Oates, Dawn Lawrie, Tan Xu, and Douglas Oard

We present KELVIN, an automated system for processing a large text corpus and distilling a knowledge base about persons, organizations, and locations. We have tested the KELVIN system on several corpora, including: (a) the TAC KBP 2012 Cold Start corpus which consists of public Web pages from the University of Pennsylvania, and (b) a subset of 26k news articles taken from English Gigaword 5th edition.

The system essentially creates a Wikipedia automatically, but one without the narrative text. Any text corpus can be used, so the method is not restricted to well-known entities.

Our NAACL HLT 2013 demonstration permits a user to interact with a set of searchable HTML pages, which are automatically generated from the knowledge base. Each page contains information analogous to the semi-structured details about an entity that are present in Wikipedia Infoboxes, along with hyperlink citations to supporting text.

Argviz: Interactive Visualization of Topic Dynamics in Multi-party Conversations
Viet-An Nguyen, Yuening Hu, Jordan Boyd-Graber, and Philip Resnik

We introduce an efficient, interactive framework—Argviz—for experts to analyze the dynamic topical structure of multi-party conversations. Users inject their needs, expertise, and insights into models via iterative topic refinement. The refined topics feed into a segmentation model, whose outputs are shown to users via multiple coordinated views.

3

Main Conference: Tuesday, June 11

Overview

7:00am – 6:00pm	**Registration**	(Terrace)
7:30am – 9:00am	**Continental Breakfast**	(Terrace)
9:15am – 10:15am	**Best Paper Awards Session**	(Peachtree Ballroom)
10:15am – 10:45am	**Break**	**(Terrace)**
10:45am – 12:00pm	**Parallel Sessions: Short papers**	
	Machine Translation and Multilinguality	(Peachtree B,C)
	Sentiment Analysis and Topic Modeling	(Plaza Ballroom)
	Spoken Language Processing	(Peachtree D,E)
12:00pm – 2:00pm	**Lunch**	
1:00pm – 2:00pm	**Business Meeting**	(Peachtree B,C)
	All are welcome!	
2:00pm – 3:15pm	**Parallel Sessions: Short papers**	
	Semantics	(Peachtree B,C)
	Information Extraction	(Plaza Ballroom)
	Discourse and Dialog	(Peachtree D,E)
3:15pm – 3:45pm	**Break**	**(Terrace)**
3:45pm – 5:25pm	**Parallel Sessions**	
	Semantics	(Peachtree B,C)
	Information Extraction	(Plaza Ballroom)
	Discourse	(Peachtree D,E)
7:00pm – 11:00pm	**Banquet at World of Coca-Cola Museum**	

Best Paper Awards

Best Short Paper

The Life and Death of Discourse Entities: Identifying Singleton Mentions
Marta Recasens, Marie-Catherine de Marneffe, and Christopher Potts
9:25am–9:45am

A discourse typically involves numerous entities, but few are mentioned more than once. Distinguishing discourse entities that die out after just one mention (singletons) from those that lead longer lives (coreferent) would benefit NLP applications such as coreference resolution, protagonist identification, topic modeling, and discourse coherence. We build a logistic regression model for predicting the singleton/coreferent distinction, drawing on linguistic insights about how discourse entity lifespans are affected by syntactic and semantic features. The model is effective in its own right (78% accuracy), and incorporating it into a state-of-the-art coreference resolution system yields a significant improvement.

IBM Best Student Paper

Automatic Generation of English Respellings
Bradley Hauer and Grzegorz Kondrak
9:45am–10:15am

A respelling is an alternative spelling of a word in the same writing system, intended to clarify pronunciation. We introduce the task of automatic generation of a respelling from the word's phonemic representation. Our approach combines machine learning with linguistic constraints and electronic resources. We evaluate our system both intrinsically through a human judgment experiment, and extrinsically by passing its output to a letter-to-phoneme converter. The results show that the forms generated by our system are better on average than those found on the Web, and approach the quality of respellings designed by an expert.

Session T1a: Machine Translation and Multilinguality

A Simple, Fast, and Effective Reparameterization of IBM Model 2

Chris Dyer, Victor Chahuneau, and Noah A. Smith

10:45am–11:00am

We present a simple log-linear reparameterization of IBM Model 2 that overcomes problems arising from Model 1's strong assumptions and Model 2's overparameterization. Efficient inference, likelihood evaluation, and parameter estimation algorithms are provided. Training the model is consistently ten times faster than Model 4. On three large-scale translation tasks, systems built using our alignment model outperform IBM Model 4.

Phrase Training Based Adaptation for Statistical Machine Translation

Saab Mansour and Hermann Ney

11:00am–11:15am

We present a novel approach for translation model (TM) adaptation using phrase training. The proposed adaptation procedure is initialized with a standard general-domain TM, which is then used to perform phrase training on a smaller in-domain set. This way, we bias the probabilities of the general TM towards the in-domain distribution. Experimental results on two different lectures translation tasks show significant improvements of the adapted systems over the general ones. Additionally, we compare our results to mixture modeling, where we report gains when using the suggested phrase training adaptation method.

Translation Acquisition Using Synonym Sets

Daniel Andrade, Masaki Tsuchida, Takashi Onishi, and Kai Ishikawa

11:15am–11:30am

We propose a new method for translation acquisition which uses a set of synonyms to acquire translations from comparable corpora. The motivation is that, given a certain query term, it is often possible for a user to specify one or more synonyms. Using the resulting set of query terms has the advantage that we can overcome the problem that a single query term's context vector does not always reliably represent a terms meaning due to the context vector's sparsity. Our proposed method uses a weighted average of the synonyms' context vectors, that is derived by inferring the mean vector of the von Mises-Fisher distribution. We evaluate our method, using the synsets from the cross-lingually aligned Japanese and English WordNet. The experiments show that our proposed method significantly improves translation accuracy when compared to a previous method for smoothing context vectors.

Supersense Tagging for Arabic: the MT-in-the-Middle Attack

Nathan Schneider, Behrang Mohit, Chris Dyer, Kemal Oflazer, and Noah A. Smith

11:30am–11:45am

We consider the task of tagging Arabic nouns with WordNet supersenses. Three approaches are evaluated. The first uses an expert-crafted but limited-coverage lexicon, Arabic WordNet, and heuristics. The second uses unsupervised sequence modeling. The third and most successful approach uses machine translation to translate the Arabic into English, which is automatically tagged with English supersenses, the results of which are then projected back into Arabic. Analysis shows gains and remaining obstacles in four Wikipedia topical domains.

Zipfian corruptions for robust POS tagging

Anders Søgaard

11:45am–12:00pm

Inspired by robust generalization and adversarial learning we describe a novel approach to learning structured perceptrons for part-of-speech (POS) tagging that is less sensitive to domain shifts. The objective of our method is to minimize average loss under random distribution shifts. We restrict the possible target distributions to mixtures of the source distribution and random Zipfian distributions. Our algorithm is used for POS tagging and evaluated on the English Web Treebank and the Danish Dependency Treebank with an average 4.4% error reduction in tagging accuracy.

Session T1b: Sentiment Analysis and Topic Modeling

A Multi-Dimensional Bayesian Approach to Lexical Style
Julian Brooke and Graeme Hirst
10:45am–11:00am

We adapt the popular LDA topic model (Blei et al., 2003) to the representation of stylistic lexical information, evaluating our model based on human-interpretability at the word and text level. We show, in particular, that this model can be applied to the task of inducing stylistic lexicons, and that a multi-dimensional approach is warranted given the correlations among stylistic dimensions.

Unsupervised Domain Tuning to Improve Word Sense Disambiguation
Judita Preiss and Mark Stevenson
11:00am–11:15am

The topic of a document can prove to be useful information for Word Sense Disambiguation (WSD) since certain meanings tend to be associated with particular topics. This paper presents an LDA-based approach for WSD, which is trained using any available WSD system to establish a sense per (Latent Dirichlet allocation based) topic. The technique is tested using three unsupervised and one supervised WSD algorithms within the sport and finance domains giving a performance increase each time, suggesting that the technique may be useful to improve the performance of any available WSD system.

What's in a Domain? Multi-Domain Learning for Multi-Attribute Data
Mahesh Joshi, Mark Dredze, William W. Cohen, and Carolyn P. Rosé
11:15am–11:30am

Multi-Domain learning assumes that a single metadata attribute is used in order to divide the data into so-called domains. However, real-world datasets often have multiple metadata attributes that can divide the data into domains. It is not always apparent which single attribute will lead to the best domains, and more than one attribute might impact classification. We propose extensions to two multi-domain learning techniques for our multi-attribute setting, enabling them to simultaneously learn from several metadata attributes. Experimentally, they outperform the multi-domain learning baseline, even when it selects the single "best" attribute.

An opinion about opinions about opinions: subjectivity and the aggregate reader
Asad Sayeed
11:30am–11:45am

This opinion piece proposes that recent advances in opinion detection are limited in the extent to which they can detect important categories of opinion because they are not designed to capture some of the pragmatic aspects of opinion. A component of these is the perspective of the user of an opinion-mining system as to what an opinion really is, which is in itself a matter of opinion (metasubjectivity). We propose a way to define this component of opinion and describe the challenges it poses for corpus development and sentence-level detection technologies. Finally, we suggest that investment in techniques to handle metasubjectivity will likely bear costs but bring benefits in the longer term.

An Examination of Regret in Bullying Tweets
Jun-Ming Xu, Benjamin Burchfiel, Xiaojin Zhu, and Amy Bellmore
11:45am–12:00pm

Social media users who post bullying related tweets may later experience regret, potentially causing them to delete their posts. In this paper, we construct a corpus of bullying tweets and periodically check the existence of each tweet in order to infer if and when it becomes deleted. We then conduct exploratory analysis in order to isolate factors associated with deleted posts. Finally, we propose the construction of a regrettable posts predictor to warn users if a tweet might cause regret.

Session T1c: Spoken Language Processing

A Cross-language Study on Automatic Speech Disfluency Detection

Wen Wang, Andreas Stolcke, Jiahong Yuan, and Mark Liberman

10:45am–11:00am

We investigate two systems for automatic disfluency detection on English and Mandarin conversational speech data. The first system combines various lexical and prosodic features in a Conditional Random Field model for detecting edit disfluencies. The second system combines acoustic and language model scores for detecting filled pauses through constrained speech recognition. We compare the contributions of different knowledge sources to detection performance between these two languages.

Distributional semantic models for the evaluation of disordered language

Masoud Rouhizadeh, Emily Prud'hommeaux, Brian Roark, and Jan van Santen

11:00am–11:15am

Atypical semantic and pragmatic expression is frequently reported in the language of children with autism. Although this atypicality often manifests itself in the use of unusual or unexpected words and phrases, the rate of use of such unexpected words is rarely directly measured or quantified. In this paper, we use distributional semantic models to automatically identify unexpected words in narrative retellings by children with autism. The classification of unexpected words is sufficiently accurate to distinguish the retellings of children with autism from those with typical development. These techniques demonstrate the potential of applying automated language analysis techniques to clinically elicited language data for diagnostic purposes.

Atypical Prosodic Structure as an Indicator of Reading Level and Text Difficulty

Julie Medero and Mari Ostendorf

11:15am–11:30am

Automatic assessment of reading ability builds on applying speech recognition tools to oral reading, measuring words correct per minute. This work looks at more fine-grained analysis that accounts for effects of prosodic context using a large corpus of read speech from a literacy study. Experiments show that lower-level readers tend to produce relatively more lengthening on words that are not likely to be final in a prosodic phrase, i.e. in less appropriate locations. The results have impli- cations for automatic assessment of text dif- ficulty in that locations of atypical prosodic lengthening are indicative of difficult lexical items and syntactic constructions.

Using Document Summarization Techniques for Speech Data Subset Selection

Kai Wei, Yuzong Liu, Katrin Kirchhoff, and Jeff Bilmes

11:30am–11:45am

In this paper we leverage methods from submodular function optimization developed for document summarization and apply them to the problem of subselecting acoustic data. We evaluate our results on data subset selection for a phone recognition task. Our framework shows significant improvements over random selection and previously proposed methods using a similar amount of resources.

Semi-Supervised Discriminative Language Modeling with Out-of-Domain Text Data

Arda Çelebi and Murat Saraçlar

11:45am–12:00pm

One way to improve the accuracy of automatic speech recognition (ASR) is to use discriminative language modeling (DLM), which enhances discrimination by learning where the ASR hypotheses deviate from the uttered sentences. However, DLM requires large amounts of ASR output to train. Instead, we can simulate the output of an ASR system, in which case the training becomes semi-supervised. The advantage of using simulated hypotheses is that we can generate as many hypotheses as we want provided that we have enough text material. In typical scenarios, transcribed in-domain data is limited but large amounts of out-of-domain (OOD) data is available. In this study, we investigate how semi-supervised training performs with OOD data. We find out that OOD data can yield improvements comparable to in-domain data.

Session T2a: Semantics

More than meets the eye: Study of Human Cognition in Sense Annotation

Salil Joshi, Diptesh Kanojia, and Pushpak Bhattacharyya

2:00pm–2:15pm

Word Sense Disambiguation (WSD) approaches have reported good accuracies in recent years. However, these approaches can be classified as weak AI systems. According to the classical definition, a strong AI based WSD system should perform the task of sense disambiguation in the same manner and with similar accuracy as human beings. In order to accomplish this, a detailed understanding of the human techniques employed for sense disambiguation is necessary. Instead of building yet another WSD system that uses contextual evidence for sense disambiguation, as has been done before, we have taken a step back - we have endeavored to discover the cognitive faculties that lie at the very core of the human sense disambiguation technique.

In this paper, we present a hypothesis regarding the cognitive sub-processes involved in the task of WSD. We support our hypothesis using the experiments conducted through the means of an eye-tracking device. We also strive to find the levels of difficulties in annotating various classes of words, with senses. We believe, once such an in-depth analysis is performed, numerous insights can be gained to develop a robust WSD system that conforms to the principle of strong AI.

Improving Lexical Semantics for Sentential Semantics: Modeling Selectional Preference and Similar Words in a Latent Variable Model

Weiwei Guo and Mona Diab

2:15pm–2:30pm

Sentence Similarity [SS] computes a similarity score between two sentences. The SS task differs from document level semantics tasks in that it features the sparsity of words in a data unit, i.e. a sentence. Accordingly it is crucial to robustly model each word in a sentence to capture the complete semantic picture of the sentence. In this paper, we hypothesize that by better modeling lexical semantics we can obtain better sentential semantics. We incorporate both corpus-based (selectional preference information) and knowledge-based (similar words extracted in a dictionary) lexical semantics into a latent variable model. The experiments show state-of-the-art performance among unsupervised systems on two SS datasets.

Linguistic Regularities in Continuous Space Word Representations

Tomas Mikolov, Wen-tau Yih, and Geoffrey Zweig

2:30pm–2:45pm

Continuous space language models have recently demonstrated outstanding results across a variety of tasks. In this paper, we examine the vector-space word representations that are implicitly learned by the input-layer weights. We find that these representations are surprisingly good at capturing syntactic and semantic regularities in language, and that each relationship is characterized by a relation-specific vector offset. This allows vector-oriented reasoning based on the offsets between words. For example, the male/female relationship is automatically learned, and with the induced vector representations, "King - Man + Woman" results in a vector very close to "Queen." We demonstrate that the word vectors capture syntactic regularities by means of syntactic analogy questions (provided with this paper), and are able to correctly answer almost 40% of the questions. We demonstrate that the word vectors capture semantic regularities by using the vector offset method to answer SemEval-2012 Task 2 questions. Remarkably, this method outperforms the best previous systems.

TruthTeller: Annotating Predicate Truth

Amnon Lotan, Asher Stern, and Ido Dagan

2:45pm–3:00pm

We propose a novel semantic annotation type of assigning truth values to predicate occurrences, and present TruthTeller, a standalone publicly-available tool that produces such annotations. TruthTeller integrates a range of linguistic phenomena, such as negation, modality, presupposition, implicativity, and more, which were dealt only partly in previous works. Empirical evaluations against human annotations show satisfactory results and suggest the usefulness of this new type of tool for NLP.

PPDB: The Paraphrase Database
Juri Ganitkevitch, Benjamin Van Durme, and Chris Callison-Burch
3:00pm–3:15pm

We present the 1.0 release of our paraphrase database, PPDB. Its English portion, PPDB:Eng, contains over 220 million paraphrase pairs, consisting of 73 million phrasal and 8 million lexical paraphrases, as well as 140 million paraphrase patterns, which capture many meaning-preserving syntactic transformations. The paraphrases are extracted from bilingual parallel corpora totaling over 100 million sentence pairs and over 2 billion English words. We also release PPDB:Spa, a collection of 196 million Spanish paraphrases. Each paraphrase pair in PPDB contains a set of associated scores, including paraphrase probabilities derived from the bitext data and a variety of monolingual distributional similarity scores computed from the Google n-grams and the Annotated Gigaword corpus. Our release includes pruning tools that allow users to determine their own precision/recall tradeoff.

Session T2b: Information Extraction

Exploiting the Scope of Negations and Heterogeneous Features for Relation Extraction: A Case Study for Drug-Drug Interaction Extraction
Md. Faisal Mahbub Chowdhury and Alberto Lavelli
2:00pm–2:15pm

This paper presents an approach that exploits the scope of negation cues for relation extraction (RE) without the need of using any specifically annotated dataset for building a separate negation scope detection classifier. New features are proposed which are used in two different stages. These also include non-target entity specific features. The proposed RE approach outperforms the previous state of the art for drug-drug interaction (DDI) extraction

Graph-Based Seed Set Expansion for Relation Extraction Using Random Walk Hitting Times
Joel Lang and James Henderson
2:15pm–2:30pm

Iterative bootstrapping methods are widely employed for relation extraction, especially because they require only a small amount of human supervision Unfortunately, a phenomenon known as semantic drift can affect the accuracy of iterative bootstrapping and lead to poor extractions. This paper proposes an alternative bootstrapping method, which ranks relation tuples by measuring their distance to the seed tuples in a bipartite tuple-pattern graph. In contrast to previous bootstrapping methods, our method is not susceptible to semantic drift, and it empirically results in better extractions than iterative methods.

Distant Supervision for Relation Extraction with an Incomplete Knowledge Base
Bonan Min, Ralph Grishman, Li Wan, Chang Wang, and David Gondek
2:30pm–2:45pm

Distant supervision, heuristically labeling a corpus using a knowledge base, has emerged as a popular choice for training relation extractors. In this paper, we show that a significant number of "negative" examples generated by the labeling process are false negatives because the knowledge base is incomplete. Therefore the heuristic for generating negative examples has a serious flaw. Building on a state-of-the-art distantly-supervised extraction algorithm, we proposed an algorithm that learns from only positive and unlabeled labels at the pair-of-entity level. Experimental results demonstrate its advantage over existing algorithms.

Measuring the Structural Importance through Rhetorical Structure Index
Narine Kokhlikyan, Alex Waibel, Yuqi Zhang, and Joy Ying Zhang
2:45pm–3:00pm

In this paper, we propose a novel Rhetorical Structure Index (RSI) to measure the structural importance of a word or a phrase. Unlike TF-IDF and other content-driven measurements, RSI identifies words or phrases that are structural cues in an unstructured document. We show structurally motivated features with high RSI values are more useful than content-driven features for applications such as segmenting the unstructured lecture transcription into meaningful segments. Experiments show that using RSI significantly improves the segmentation accuracy compared to the traditional content-based feature weighting scheme such as TF-IDF.

Separating Fact from Fear: Tracking Flu Infections on Twitter
Alex Lamb, Michael J. Paul, and Mark Dredze
3:00pm–3:15pm

Twitter has been shown to be a fast and reliable method for disease surveillance of common illnesses like influenza. However, previous work has relied on simple content analysis, which conflates flu tweets that report infection with those that express concerned awareness of the flu. By discriminating these categories, as well as tweets about the authors versus about others, we demonstrate significant improvements on influenza surveillance using Twitter.

Session T2c: Discourse and Dialog

Differences in User Responses to a Wizard-of-Oz versus Automated System

Jesse Thomason and Diane Litman

2:00pm–2:15pm

Wizard-of-Oz experimental setup in a dialogue system is commonly used to gather data for informing an automated version of that system. Previous work has exposed dependencies between user behavior towards systems and user belief about whether the system is automated or human-controlled. This work examines whether user behavior changes when user belief is held constant and the system's operator is varied. We perform a post-hoc experiment using generalizable prosodic and lexical features of user responses to a dialogue system backed with and without a human wizard. Our results suggest that user responses are different when communicating with a wizarded and an automated system, indicating that wizard data may be less reliable for informing automated systems than generally assumed.

Improving the Quality of Minority Class Identification in Dialog Act Tagging

Adinoyi Omuya, Vinodkumar Prabhakaran, and Owen Rambow

2:15pm–2:30pm

We present a method of improving the performance of dialog act tagging in identifying minority classes by using per-class feature optimization and a method of choosing the class based not on confidence, but on a cascade of classifiers. We show that it gives a minority class F-measure error reduction of 22.8%, while also reducing the error for other classes and the overall error by about 10%.

Discourse Connectors for Latent Subjectivity in Sentiment Analysis

Rakshit Trivedi and Jacob Eisenstein

2:30pm–2:45pm

Document-level sentiment analysis can benefit from fine-grained subjectivity, so that sentiment polarity judgments are based on the relevant parts of the document. While fine-grained subjectivity annotations are rarely available, encouraging results have been obtained by modeling subjectivity as a latent variable. However, latent variable models fail to capitalize on our linguistic knowledge about discourse structure. We present a new method for injecting linguistic knowledge into latent variable subjectivity modeling, using discourse connectors. Connector-augmented transition features allow the latent variable model to learn the relevance of discourse connectors for subjectivity transitions, without subjectivity annotations. This yields significantly improved performance on document-level sentiment analysis in English and Spanish. We also describe a simple heuristic for automatically identifying connectors when no predefined list is available.

Coherence Modeling for the Automated Assessment of Spontaneous Spoken Responses

Xinhao Wang, Keelan Evanini, and Klaus Zechner

2:45pm–3:00pm

This study focuses on modeling discourse coherence in the context of automated assessment of spontaneous speech from non-native speakers. Discourse coherence has always been used as a key metric in human scoring rubrics for various assessments of spoken language. However, very little research has been done to assess a speaker's coherence in automated speech scoring systems. To address this, we present a corpus of spoken responses that has been annotated for discourse coherence quality. Then, we investigate the use of several features originally developed for essays to model coherence in spoken responses. An analysis on the annotated corpus shows that the prediction accuracy for human holistic scores of an automated speech scoring system can be improved by around 10% relative after the addition of the coherence features. Further experiments indicate that a weighted F-Measure of 73% can be achieved for the automated prediction of the coherence scores.

Disfluency Detection Using Multi-step Stacked Learning

Xian Qian and Yang Liu

3:00pm–3:15pm

In this paper, we propose a multi-step stacked learning model for disfluency detection. Our method incorpo-

rates refined n-gram features step by step from different word sequences. First, we detect filler words. Second, edited words are detected using n-gram features extracted from both the original text and filler filtered text. In the third step, additional n-gram features are extracted from edit removed texts together with our newly induced in-between features to improve edited word detection. We use Max-Margin Markov Networks (M3Ns) as the classifier with the weighted hamming loss to balance precision and recall. Experiments on the Switchboard corpus show that the refined n-gram features from multiple steps and M3Ns with weighted hamming loss can significantly improve the performance. Our method for disfluency detection achieves the best reported F-score 0:841 without the use of additional resources.

Session T3a: Semantics

[TACL] Weakly Supervised Learning of Semantic Parsers for Mapping Instructions to Actions

Yoav Artzi and Luke Zettlemoyer
3:45pm–4:10pm

The context in which language is used provides a strong signal for learning to recover its meaning. In this paper, we show it can be used within a grounded CCG semantic parsing approach that learns a joint model of mean- ing and context for interpreting and executing natural language instructions, using various types of weak supervision. The joint nature provides crucial benefits by allowing situated cues, such as the set of visible objects, to directly influence learning. It also enables algorithms that learn while executing instructions, for example by trying to replicate human actions. Experiments on a benchmark naviga- tional dataset demonstrate strong performance under differing forms of supervision, including correctly executing 60% more instruction sets relative to the previous state of the art.

Using Semantic Unification to Generate Regular Expressions from Natural Language

Nate Kushman and Regina Barzilay
4:10pm–4:35pm

We consider the problem of translating natural language text queries into regular expressions which represent their meaning. The mismatch in the level of abstraction between the natural language representation and the regular expression representation make this a novel and challenging problem. However, a given regular expression can be written in many semantically equivalent forms, and we exploit this flexibility to facilitate translation by finding a form which more directly corresponds to the natural language. We evaluate our technique on a set of natural language queries and their associated regular expressions which we gathered from Amazon Mechanical Turk. Our model substantially outperforms a state-of-the-art semantic parsing baseline, yielding a 29% absolute improvement in accuracy.

Probabilistic Frame Induction

Jackie Chi Kit Cheung, Hoifung Poon, and Lucy Vanderwende
4:35pm–5:00pm

In natural-language discourse, related events tend to appear near each other to describe a larger scenario. Such structures can be formalized by the notion of a frame (a.k.a. template), which comprises a set of related events and prototypical participants and event transitions. Identifying frames is a prerequisite for information extraction and natural language generation, and is usually done manually. Methods for inducing frames have been proposed recently, but they typically use ad hoc procedures and are difficult to diagnose or extend. In this paper, we propose the first probabilistic approach to frame induction, which incorporates frames, events, and participants as latent topics and learns those frame and event transitions that best explain the text. The number of frame components is inferred by a novel application of a split-merge method from syntactic parsing. In end-to-end evaluations from text to induced frames and extracted facts, our method produces state-of-the-art results while substantially reducing engineering effort.

A Quantum-Theoretic Approach to Distributional Semantics

William Blacoe, Elham Kashefi, and Mirella Lapata
5:00pm–5:25pm

In this paper we explore the potential of quantum theory as a formal framework for capturing lexical meaning. We present a novel semantic space model that is syntactically aware, takes word order into account, and features key quantum aspects such as superposition and entanglement. We define a dependency-based Hilbert space and show how to represent the meaning of words by density matrices that encode dependency neighborhoods. Experiments on word similarity and association reveal that our model achieves results competitive with a variety of classical models.

Session T3b: Information Extraction

Answer Extraction as Sequence Tagging with Tree Edit Distance

Xuchen Yao, Benjamin Van Durme, Chris Callison-Burch, and Peter Clark

3:45pm–4:10pm

Our goal is to extract answers from pre-retrieved sentences for Question Answering (QA).

We construct a linear-chain Conditional Random Field based on pairs of questions and their possible answer sentences, learning the association between questions and answer types. This casts answer extraction as an answer sequence tagging problem for the first time, where knowledge of shared structure between question and source sentence is incorporated through features based on Tree Edit Distance (TED). Our model is free of manually created question and answer templates, fast to run (processing 200 QA pairs per second excluding parsing time), and yields an F1 of 63.3
prior TREC QA evaluations. The developed system is open-source, and includes an implementation of the TED model that is state of the art in the task of ranking QA pairs.

Open Information Extraction with Tree Kernels

Ying Xu, Mi-Young Kim, Kevin Quinn, Randy Goebel, and Denilson Barbosa

4:10pm–4:35pm

Traditional relation extraction seeks to identify pre-specified semantic relations within natural language text, while open Information Extraction (Open IE) takes a more general approach, and looks for a variety of relations without restriction to a fixed relation set. With this generalization comes the question, what is a relation? For example, should the more general task be restricted to relations mediated by verbs, nouns, or both? To help answer this question, we propose two levels of subtasks for Open IE. One task determines if a sentence potentially contains a relation between two entities. The other task looks to confirm explicit relation words for two entities. We propose multiple SVM models with dependency tree kernels for both tasks. For explicit relation extraction, our system can extract both noun and verb relations. Our results on three datasets show that our system is superior when compared to state-of-the-art systems like REVERB and OLLIE for both tasks. For example, in some experiments our system achieves 33% improvement on nominal relation extraction over OLLIE. In addition we propose an unsupervised rule-based approach which can serve as a strong baseline for Open IE systems.

Finding What Matters in Questions

Xiaoqiang Luo, Hema Raghavan, Vittorio Castelli, Sameer Maskey, and Radu Florian

4:35pm–5:00pm

In natural language question answering (QA) systems, questions often contain terms and phrases that are critically important for retrieving or finding anwers from documents. We present a learnable system that can extract and rank these terms and phrases (dubbed "mandatory matching phrases" or MMPs), and demonstrate their utility in a QA system on Internet discussion forum data sets. The system relies on deep syntactic and semantic analysis of questions only and is independent of relevant documents. Our proposed model can predict MMPs with high accuracy. When used in a QA system features derived from the MMP model improve performance significantly over a state-of-the-art baseline. The final QA system was the best performing system in the DARPA BOLT-IR evaluation.

A Just-In-Time Keyword Extraction from Meeting Transcripts

Hyun-Je Song, Junho Go, Seong-Bae Park, and Se-Young Park

5:00pm–5:25pm

In a meeting, it is often desirable to extract keywords from each utterance as soon as it is spoken. Thus, this paper proposes a just-in-time keyword extraction from meeting transcripts. The proposed method considers two major factors that make it different from keyword extraction from normal texts. The first factor is the temporal history of preceding utterances that grants higher importance to recent utterances than old ones, and the second is topic-relevance that forces only the preceding utterances relevant to the current utterance to be considered in keyword extraction. Our experiments on two data sets in English and Korean show that the consideration of the factors results in performance improvement in keyword extraction from meeting transcripts.

Session T3c: Discourse

Same Referent, Different Words: Unsupervised Mining of Opaque Coreferent Mentions
Marta Recasens, Matthew Can, and Daniel Jurafsky
3:45pm–4:10pm

Coreference resolution systems rely heavily on string overlap (e.g., "Google Inc." and "Google"), performing badly on mentions with very different words ("opaque" mentions) like "Google" and "the search giant". Yet prior attempts to resolve opaque pairs using ontologies or distributional semantics hurt precision more than improved recall. We present a new unsupervised method for mining opaque pairs. Our intuition is to "restrict" distributional semantics to articles about the same event, thus promoting referential match. Using an English comparable corpus of tech news, we built a dictionary of opaque coreferent mentions (only 3% are in WordNet). Our dictionary can be integrated into any coreference system (it increases the performance of a state-of-the-art system by 1% F1 on all measures) and is easily extendable by using news aggregators.

Global Inference for Bridging Anaphora Resolution
Yufang Hou, Katja Markert, and Michael Strube
4:10pm–4:35pm

We present the first work on antecedent selection for bridging resolution without restrictions on anaphor or relation types. Our model integrates global constraints on top of a rich local feature set in the framework of Markov logic networks. The global model improves over the local one and both strongly outperform a reimplementation of prior work.

Classifying Temporal Relations with Rich Linguistic Knowledge
Jennifer D'Souza and Vincent Ng
4:35pm–5:00pm

We examine the task of temporal relation classification. Unlike existing approaches to this task, we (1) classify an event-event or event-time pair as one of the 14 temporal relations defined in the TimeBank corpus, rather than as one of the six relations collapsed from the original 14; (2) employ sophisticated linguistic knowledge derived from a variety of semantic and discourse relations, rather than focusing on morpho-syntactic knowledge; and (3) leverage a novel combination of rule-based and learning-based approaches, rather than relying solely on one or the other. Experiments with the TimeBank corpus show that our knowledge-rich, hybrid approach yields a 15–16% relative reduction in error over a state-of-the-art learning-based baseline system.

Improved Information Structure Analysis of Scientific Documents Through Discourse and Lexical Constraints
Yufan Guo, Roi Reichart, and Anna Korhonen
5:00pm–5:25pm

Inferring the information structure of scientific documents is useful for many down-stream applications. Existing feature-based machine learning approaches to this task require substantial training data and suffer from limited performance. Our idea is to guide feature-based models with declarative domain knowledge encoded as posterior distribution constraints. We explore a rich set of discourse and lexical constraints which we incorporate through the Generalized Expectation (GE) criterion. Our constrained model improves the performance of existing fully and lightly supervised models. Even a fully unsupervised version of this model outperforms lightly supervised feature-based models, showing that our approach can be useful even when no labeled data is available.

Banquet

NAACL HLT 2013 Banquet at World of Coca Cola

©Wikimedia foundation, released under GPL

Tuesday, June 11, 2013, 7:00pm – 11:00pm

The World of Coca-Cola
121 Baker St. NW
Atlanta, GA 30313-1807

This year's NAACL HLT banquet promises to be a unique and interactive experience! Taking place in the intriguing World of Coca Cola Museum, attendees will have access to a multi-sensory 4D theater, an extraordinary 1880s soda fountain, the smallest bottling line in the world, plus an opportunity to sample nearly 70 different beverages from around the globe. Multiple buffet stations and seating areas will allow banquet-goers to enjoy the food and explore the museum at their leisure. And, of course there will be a DJ for your dancing pleasure. The banquet will take place at the World of Coca Cola Museum , Tuesday evening, June 11th. So, save yourselves the standard $15-16 entrance fee and join us on Tuesday.

4

Main Conference: Wednesday, June 12

Overview

7:00am – 6:00pm	**Registration**	(Terrace)
7:30am – 9:00am	**Continental Breakfast**	(Terrace)
9:00am – 10:10am	**Invited Talk: Kathleen McKeown**	(Atlanta Ballroom)
10:10am – 10:40am	**Break**	(Terrace)
10:40am – 12:20pm	**Parallel Sessions**	
	Machine Translation	(Peachtree B,C)
	Semantics	(Plaza Ballroom)
	Social Media Processing	(Peachtree D,E)
12:20pm – 2:00pm	**Lunch**	
2:00pm – 3:15pm	**Parallel Sessions**	
	Parsing and Syntax	(Peachtree B,C)
	Dialog	(Plaza Ballroom)
	Annotation and Language Resources	(Peachtree D,E)
3:15pm – 3:45pm	**Break**	(Terrace)
3:45pm – 5:00pm	**Parallel Sessions**	
	Semantics and Syntax	(Peachtree B,C)
	Summarization and Generation	(Plaza Ballroom)
	Morphology and Phonology	(Peachtree D,E)
5:00pm	**Conference adjourns!**	

Invited Talk: Kathleen McKeown

"Natural Language Applications from Fact to Fiction"

Kathleen McKeown

Wednesday, June 12, 2013, 9:00am – 10:10am

Atlanta Ballroom

Abstract: Much research in the natural language field has been carried out on news, given the large amount of annotated data now available for this genre.

Certainly, the ability to analyze the facts of real world events, enabling systems to answer questions and summarize the events of the day is an important application. As research has moved to analysis of new genres, whether fact, opinion or fiction, new approaches to existing applications have arisen and opportunities for new applications have emerged. In this talk, I will present research in my group at Columbia as it has moved from news to scientific articles to online discussion forums to novels. I will touch on summarization, open-ended question answering, social influence and social networks.

Biography: A leading scholar and researcher in the field of natural language processing, McKeown focuses her research on big data; her interests include text summarization, question answering, natural language generation, multimedia explanation, digital libraries, and multilingual applications. Her research group's Columbia Newsblaster, which has been live since 2001, is an online system that automatically tracks the day's news, and demonstrates the group's new technologies for multi-document summarization, clustering, and text categorization, among others. Currently, she leads a large research project involving prediction of technology emergence from a large collection of journal articles.

McKeown joined Columbia in 1982, immediately after earning her Ph.D. from University of Pennsylvania. In 1989, she became the first woman professor in the school to receive tenure, and later the first woman to serve as a department chair (1998-2003). McKeown has received numerous honors and awards for her research and teaching. She received the National Science Foundation Presidential Young Investigator Award in 1985, and also is the recipient of a National Science Foundation Faculty Award for Women, was selected as an AAAI Fellow, a Founding Fellow of the Association for Computational Linguistics and an ACM Fellow. In 2010, she won

both the Columbia Great Teacher Award—an honor bestowed by the students—and the Anita Borg Woman of Vision Award for Innovation.

McKeown served as a board member of the Computing Research Association and as secretary of the board. She was president of the Association of Computational Linguistics in 1992, vice president in 1991, and secretary treasurer for 1995-1997. She was also a member of the Executive Council of the Association for Artificial Intelligence and the co-program chair of their annual conference in 1991.

Session W1a: Machine Translation

Adaptation of Reordering Models for Statistical Machine Translation

Boxing Chen, George Foster, and Roland Kuhn

10:40am–11:05am

Previous research on domain adaptation (DA) for statistical machine translation (SMT) has mainly focused on the translation model (TM) and the language model (LM). To the best of our knowledge, there is no previous work on reordering model (RM) adaptation for phrase-based SMT. In this paper, we demonstrate that mixture model adaptation of a lexicalized RM can significantly improve SMT performance, even when the system already contains a domain-adapted TM and LM. We find that, surprisingly, different training corpora can vary widely in their reordering characteristics for particular phrase pairs. Furthermore, particular training corpora may be highly suitable for training the TM or the LM, but unsuitable for training the RM, or vice versa, so mixture weights for these models should be estimated separately. An additional contribution of the paper is to propose two improvements to mixture model adaptation: smoothing the in-domain sample, and weighting instances by document frequency. Applied to mixture RMs in our experiments, these techniques (especially smoothing) yield significant performance improvements.

Multi-Metric Optimization Using Ensemble Tuning

Baskaran Sankaran, Anoop Sarkar, and Kevin Duh

11:05am–11:30am

This paper examines tuning for statistical machine translation (SMT) with respect to multiple evaluation metrics. We propose several novel methods for tuning towards multiple objectives, including some based on 'ensemble decoding' methods. Pareto-optimality is a natural way to think about multi-metric optimization (MMO) and our methods can effectively combine several Pareto-optimal solutions, obviating the need to choose one. Our best performing 'ensemble tuning' method is a new algorithm for multi-metric optimization that searches for Pareto-optimal ensemble models. We study the effectiveness of our methods through experiments on multiple as well as single reference(s) datasets. Our experiments show simultaneous gains across several metrics (BLEU, RIBES), without any significant reduction in other metrics. This contrasts the traditional tuning where gains are usually limited to a single metric. Our human evaluation results confirm that in order to produce better MT output, optimizing multiple metrics is better than optimizing only one.

Grouping Language Model Boundary Words to Speed K–Best Extraction from Hypergraphs

Kenneth Heafield, Philipp Koehn, and Alon Lavie

11:30am–11:55am

We propose a new algorithm to approximately extract top-scoring hypotheses from a hypergraph when the score includes an N-gram language model.

In the popular cube pruning algorithm, every hypothesis is annotated with boundary words and permitted to recombine only if all boundary words are equal. However, many hypotheses share some, but not all, boundary words. We use these common boundary words to group hypotheses and do so recursively, resulting in a tree of hypotheses. This tree forms the basis for our new search algorithm that iteratively refines groups of boundary words on demand. Machine translation experiments show our algorithm makes translation 1.50 to 3.51 times as fast as with cube pruning in common cases.

A Systematic Bayesian Treatment of the IBM Alignment Models

Yarin Gal and Phil Blunsom

11:55am–12:20pm

The dominant yet ageing IBM and HMM word alignment models underpin most popular Statistical Machine Translation implementations in use today. Though beset by the limitations of implausible independence assumptions, intractable optimisation problems, and an excess of tunable parameters, these models provide a scalable and reliable starting point for inducing translation systems. In this paper we build upon this venerable base by recasting these models in the non-parametric Bayesian framework. By replacing the categorical distributions at their core with hierarchical Pitman-Yor processes, and through the use of collapsed Gibbs sampling, we provide a more flexible formulation and sidestep the original heuristic optimisation techniques. The resulting models are highly extendible, naturally permitting the introduction of phrasal dependencies. We present extensive experimental results showing improvements in both AER and BLEU when benchmarked against

Giza++, including significant improvements over IBM model 4.

Session W1b: Semantics

Unsupervised Metaphor Identification Using Hierarchical Graph Factorization Clustering
Ekaterina Shutova and Lin Sun
10:40am–11:05am

We present a novel approach to automatic metaphor identification, that discovers both metaphorical associations and metaphorical expressions in unrestricted text. Our system first performs hierarchical graph factorization clustering of nouns and then searches the resulting graph for metaphorical connections between concepts. It then makes use of the salient features of the metaphorically connected clusters to identify the actual metaphorical expressions. In contrast to previous work, our method is fully unsupervised. Despite this fact, it operates with an encouraging precision (0.69) and recall (0.61). Our approach is also the first one in NLP to exploit the cognitive findings on the differences in organisation of abstract and concrete concepts in the human brain.

Three Knowledge-Free Methods for Automatic Lexical Chain Extraction
Steffen Remus and Chris Biemann
11:05am–11:30am

We present three approaches to lexical chaining based on the LDA topic model and evaluate them intrinsically on a manually annotated set of German documents. After motivating the choice of statistical methods for lexical chaining with their adaptability to different languages and subject domains, we describe our new two-level chain annotation scheme, which rooted in the concept of cohesive harmony. Also, we propose a new measure for direct evaluation of lexical chains. Our three LDA-based approaches outperform two knowledge-based state-of-the art methods to lexical chaining by a large margin, which can be attributed to lacking coverage of the knowledge resource. Subsequent analysis shows that the three methods yield a different chaining behavior, which could be utilized in tasks that use lexical chaining as a component within NLP applications.

Combining Heterogeneous Models for Measuring Relational Similarity
Alisa Zhila, Wen-tau Yih, Christopher Meek, Geoffrey Zweig, and Tomas Mikolov
11:30am–11:55am

In this work, we study the problem of measuring relational similarity between two word pairs (e.g., silverware:fork and clothing:shirt). Due to the large number of possible relations, we argue that it is important to combine multiple models based on heterogeneous information sources. Our overall system consists of two novel general-purpose relational similarity models and three specific word relation models. When evaluated in the setting of a recently proposed SemEval-2012 task, our approach outperforms the previous best system substantially, achieving a 54.1% relative increase in Spearman's rank correlation.

[TACL] Modeling Semantic Relations Expressed by Prepositions
Vivek Srikumar and Dan Roth
11:55am–12:20pm

This paper introduces the problem of predicting semantic relations expressed by prepositions and develops statistical learning models for predicting the relations, their arguments and the semantic types of the arguments. First, building on the word sense disambiguation task for prepositions, we define an inventory of 32 relations by collapsing related senses across prepositions. Then, given a preposition in a sentence, we jointly model the preposition relation and its arguments, along with their semantic types to support predicting the semantic relation. The annotated data, though, only provides labels for the relation label, and not the arguments and types. We address this by presenting two models for preposition relation labeling. Our generalization of latent structure SVM gives an accuracy of 89.43% on relation labeling. Further, by jointly predicting the relation, arguments, and their types along with preposition sense, we show that we can not only improve the relation accuracy to 90.26%, but also significantly improve sense prediction accuracy.

Session W1c: Social Media Processing

Broadly Improving User Classification via Communication-Based Name and Location Clustering on Twitter

Shane Bergsma, Mark Dredze, Benjamin Van Durme, Theresa Wilson, and David Yarowsky

10:40am–11:05am

Hidden properties of social media users, such as their ethnicity, gender, and location, are often reflected in their observed attributes, such as their first and last names. Furthermore, users who communicate with each other often have similar hidden properties. We propose an algorithm that exploits these insights to cluster the observed attributes of hundreds of millions of Twitter users. Attributes such as user names are grouped together if users with those names communicate with other similar users. We separately cluster millions of unique first names, last names, and user-provided locations. The efficacy of these clusters is then evaluated on a diverse set of classification tasks that predict hidden users properties such as ethnicity, geographic location, gender, language, and race, using only profile names and locations when appropriate. Our readily-replicable approach and publicly-released clusters are shown to be remarkably effective and versatile, substantially outperforming state-of-the-art approaches and human accuracy on each of the tasks studied.

To Link or Not to Link? A Study on End-to-End Tweet Entity Linking

Stephen Guo, Ming-Wei Chang, and Emre Kiciman

11:05am–11:30am

Information extraction from microblog posts is an important task, as today microblogs capture an unprecedented amount of information and provide a view into the pulse of the world. As the core component of information extraction, we consider the task of Twitter entity linking in this paper.

In the current entity linking literature, mention detection and entity disambiguation are frequently cast as equally important but distinct problems. However, in our task, we find that mention detection is often the performance bottleneck. The reason is that messages on microblogs are short, noisy and informal texts with little context, and often contain phrases with ambiguous meanings.

To rigorously address the Twitter entity linking problem, we propose a structural SVM algorithm for entity linking that jointly optimizes mention detection and entity disambiguation as a single end-to-end task. By combining structural learning and a variety of first-order, second-order, and context-sensitive features, our system is able to outperform existing state-of-the art entity linking systems by 15% F1.

A Latent Variable Model for Viewpoint Discovery from Threaded Forum Posts

Minghui Qiu and Jing Jiang

11:30am–11:55am

Threaded discussion forums provide an important social media platform. Its rich user generated content has served as an important source of public feedback. To automatically discover the viewpoints or stances on hot issues from forum threads is an important and useful task. In this paper, we propose a novel latent variable model for viewpoint discovery from threaded forum posts. Our model is a principled generative latent variable model which captures three important factors: viewpoint specific topic preference, user identity and user interactions. Evaluation results show that our model clearly outperforms a number of baseline models in terms of both clustering posts based on viewpoints and clustering users with different viewpoints.

Identifying Intention Posts in Discussion Forums

Zhiyuan Chen, Bing Liu, Meichun Hsu, Malu Castellanos, and Riddhiman Ghosh

11:55am–12:20pm

This paper proposes to study the problem of identifying intention posts in online discussion forums. For example, in a discussion forum, a user wrote "I plan to buy a camera," which indicates a buying intention. This intention can be easily exploited by advertisers. To the best of our knowledge, there is still no reported study of this problem. Our research found that this problem is particularly suited to transfer learning because in different domains, people express the same intention in similar ways. We then propose a new transfer learning method which, unlike a general transfer learning algorithm, exploits several special characteristics of the problem. Experimental results show that the proposed method outperforms several strong baselines, including supervised learning in the target domain and a recent transfer learning method.

Session W2a: Parsing and Syntax

[TACL] Unsupervised Dependency Parsing with Acoustic Cues

John K Pate and Sharon Goldwater

2:00pm–2:25pm

Unsupervised parsing is a difficult task that infants readily perform. Progress has been made on this task using text-based models, but few computational approaches have considered how infants might benefit from acoustic cues. This paper explores the hypothesis that word duration can help with learning syntax. We describe how duration information can be incorporated into an unsupervised Bayesian depen- dency parser whose only other source of information is the words themselves (without punctuation or parts of speech). Our results, evaluated on both adult-directed and child-directed utterances, show that using word duration can improve parse quality relative to words-only baselines. These results support the idea that acoustic cues provide useful evidence about syntactic structure for language-learning infants, and motivate the use of word duration cues in NLP tasks with speech.

Dependency-based empty category detection via phrase structure trees

Nianwen Xue and Yaqin Yang

2:25pm–2:50pm

We describe a novel approach to detecting empty categories (EC) as represented in dependency trees as well as a new metric for measuring EC detection accuracy. The new metric takes into account not only the position and type of an EC, but also the head it is a dependent of in a dependency tree. We also introduce a variety of new features that are more suited for this approach. Tested on a subset of the Chinese Treebank, our system improved significantly over the best previously reported results even when evaluated with this new, arguably more stringent metric.

Target Language Adaptation of Discriminative Transfer Parsers

Oscar Täckström, Ryan McDonald, and Joakim Nivre

2:50pm–3:15pm

We study multi-source transfer parsing for resource-poor target languages; specifically methods for target language adaptation of delexicalized discriminative graph-based dependency parsers. We first show how recent insights on selective parameter sharing, based on typological and language-family features, can be applied to a discriminative parser by carefully decomposing its model features. We then show how the parser can be relexicalized and adapted using unlabeled target language data and a learning method that can incorporate diverse knowledge sources through ambiguous labelings. In the latter scenario, we exploit two sources of knowledge: arc marginals derived from the base parser in a self-training algorithm, and arc predictions from multiple transfer parsers in an ensemble-training algorithm. Our final model outperforms the state of the art in multi-source transfer parsing on 15 out of 16 evaluated languages.

Session W2b: Dialog

Emergence of Gricean Maxims from Multi-Agent Decision Theory

Adam Vogel, Max Bodoia, Christopher Potts, and Daniel Jurafsky

2:00pm–2:25pm

Grice characterized communication in terms of the cooperative principle, which enjoins speakers to make only contributions that serve the evolving conversational goals. We show that the cooperative principle and the associated maxims of relevance, quality, and quantity emerge from multi-agent decision theory. We utilize the Decentralized Partially Observable Markov Decision Process (Dec-POMDP) model of multi-agent decision making which relies only on basic definitions of rationality and the ability of agents to reason about each other's beliefs in maximizing joint utility. Our model uses cognitively-inspired heuristics to simplify the otherwise intractable task of reasoning jointly about actions, the environment, and the nested beliefs of other actors. Our experiments on a cooperative language task show that reasoning about others' belief states, and the resulting emergent Gricean communicative behavior, leads to significantly improved task performance.

Open Dialogue Management for Relational Databases

Ben Hixon and Rebecca J. Passonneau

2:25pm–2:50pm

We present open dialogue management and its application to relational databases. An open dialogue manager generates dialogue states, actions, and default strategies from the semantics of its application domain. We define three open dialogue management tasks. First, vocabulary selection finds the intelligible attributes in each database table. Second, focus discovery selects candidate dialogue foci, tables that have the most potential to address basic user goals. Third, a focus agent is instantiated for each dialogue focus with a default dialogue strategy governed by efficiency. We demonstrate the portability of open dialogue management on three very different databases. Evaluation of our system with simulated users shows that users with realistically limited domain knowledge have dialogues nearly as efficient as those of users with complete domain knowledge.

A method for the approximation of incremental understanding of explicit utterance meaning using predictive models in finite domains

David DeVault and David Traum

2:50pm–3:15pm

This paper explores the relationship between explicit and predictive models of incremental speech understanding in a dialogue system that supports a finite set of user utterance meanings. We present a method that enables the approximation of explicit understanding using information implicit in a predictive understanding model for the same domain. We show promising performance for this method in a corpus evaluation, and discuss its practical application and annotation costs in relation to some alternative approaches.

Session W2c: Annotation and Language Resources

Paving the Way to a Large-scale Pseudosense-annotated Dataset
Mohammad Taher Pilehvar and Roberto Navigli
2:00pm–2:25pm

In this paper we propose a new approach to the generation of pseudowords, i.e., artificial words which model real polysemous words. Our approach simultaneously addresses the two important issues that hamper the generation of large pseudosense-annotated datasets: semantic awareness and coverage. We evaluate these pseudowords from three different perspectives showing that they can be used as reliable substitutes for their real counterparts.

Labeling the Languages of Words in Mixed-Language Documents using Weakly Supervised Methods
Ben King and Steven Abney
2:25pm–2:50pm

In this paper we consider the problem of labeling the languages of words in mixed-language documents. This problem is approached in a weakly supervised fashion, as a sequence labeling problem with monolingual text samples for training data. Among the approaches evaluated, a conditional random field model trained with generalized expectation criteria was the most accurate and performed consistently as the amount of training data was varied.

Learning Whom to Trust with MACE
Dirk Hovy, Taylor Berg-Kirkpatrick, Ashish Vaswani, and Eduard Hovy
2:50pm–3:15pm

Non-expert annotation services like Amazon's Mechanical Turk (AMT) are cheap and fast ways to evaluate systems and provide categorical annotations for training data. Unfortunately, some annotators choose bad labels in order to maximize their pay. Manual identification is tedious, so we experiment with an item-response model. It learns in an unsupervised fashion to a) identify which annotators are trustworthy and b) predict the correct underlying labels. We match performance of more complex state-of-the-art systems and perform well even under adversarial conditions. We show considerable improvements over standard baselines, both for predicted label accuracy and trustworthiness estimates. The latter can be further improved by introducing a prior on model parameters and using Variational Bayes inference. Additionally, we can achieve even higher accuracy by focusing on the instances our model is most confident in (trading in some recall), and by incorporating annotated control instances. Our system, MACE (Multi-Annotator Competence Estimation), is available for download.

Session W3a: Semantics and Syntax

Supervised All-Words Lexical Substitution using Delexicalized Features

György Szarvas, Chris Biemann, and Iryna Gurevych

3:45pm–4:10pm

We propose a supervised lexical substitution system that does not use separate classifiers per word and is therefore applicable to any word in the vocabulary. Instead of learning word-specific substitution patterns, a global model for lexical substitution is trained on delexicalized (i.e., non lexical) features, which allows to exploit the power of supervised methods while being able to generalize beyond target words in the training set. This way, our approach remains technically straightforward, provides better performance and similar coverage in comparison to unsupervised approaches. Using features from lexical resources, as well as a variety of features computed from large corpora (n-gram counts, distributional similarity) and a ranking method based on the posterior probabilities obtained from a Maximum Entropy classifier, we improve over the state of the art in the LexSub Best-Precision metric and the Generalized Average Precision measure. Robustness of our approach is demonstrated by evaluating it successfully on two different datasets.

A Tensor-based Factorization Model of Semantic Compositionality

Tim Van de Cruys, Thierry Poibeau, and Anna Korhonen

4:10pm–4:35pm

In this paper, we present a novel method for the computation of compositionality within a distributional framework. The key idea is that compositionality is modeled as a multi-way interaction between latent factors, which are automatically constructed from corpus data. We use our method to model the composition of subject verb object triples. The method consists of two steps. First, we compute a latent factor model for nouns from standard co-occurrence data. Next, the latent factors are used to induce a latent model of three-way subject verb object interactions. Our model has been evaluated on a similarity task for transitive phrases, in which it exceeds the state of the art.

[TACL] An HDP Model for Inducing Combinatory Categorial Grammars

Yonatan Bisk and Julia Hockenmaier

4:35pm–5:00pm

We introduce a novel nonparametric Bayesian model for the induction of Combinatory Categorial Grammars from POS-tagged text. It achieves state of the art performance on a number of languages, and induces linguistically plausible lexicons.

Session W3b: Summarization and Generation

A Participant-based Approach for Event Summarization Using Twitter Streams

Chao Shen, Fei Liu, Fuliang Weng, and Tao Li

3:45pm–4:10pm

Twitter offers an unprecedented advantage on live reporting of the events happening around the world. However, summarizing the Twitter event has been a challenging task that was not fully explored in the past. In this paper, we propose a participant-based event summarization approach that "zooms-in" the Twitter event streams to the participant level, detects the important sub-events associated with each participant using a novel mixture model that combines the "burstiness" and "cohesiveness" properties of the event tweets, and generates the event summaries progressively. We evaluate the proposed approach on different event types. Results show that the participant-based approach can effectively capture the sub-events that have otherwise been shadowed by the long-tail of other dominant sub-events, yielding summaries with considerably better coverage than the state-of-the-art.

Towards Coherent Multi-Document Summarization

Janara Christensen, Mausam, Stephen Soderland, and Oren Etzioni

4:10pm–4:35pm

This paper presents G-FLOW, a novel system for coherent extractive multi-document summarization (MDS). Where previous work on MDS considered sentence selection and ordering separately, G-FLOW introduces a joint model for selection and ordering that balances coherence and salience. G-FLOW's core representation is a graph that approximates the discourse relations across sentences based on indicators including discourse cues, deverbal nouns, co-reference, and more. This graph enables G-FLOW to estimate the coherence of a candidate summary. We evaluate G-FLOW on Mechanical Turk, and find that it generates dramatically better summaries than an extractive summarizer based on a pipeline of state-of-the-art sentence selection and reordering components, underscoring the value of our joint model.

Generating Expressions that Refer to Visible Objects

Margaret Mitchell, Kees van Deemter, and Ehud Reiter

4:35pm–5:00pm

We introduce a novel algorithm for generating referring expressions, informed by human and computer vision and designed to refer to visible objects. Our method separates absolute properties like color from relative properties like size to stochastically generate a diverse set of outputs. Expressions generated using this method are often overspecified and may be underspecified, akin to expressions produced by people. We call such expressions identifying descriptions. The algorithm out-performs the well-known Incremental Algorithm (Dale and Reiter, 1995) and the Graph-Based Algorithm (Krahmer et al., 2003; Viethen et al., 2008) across a variety of images in two domains. We additionally motivate an evaluation method for referring expression generation that takes the proposed algorithm's non-determinism into account.

Session W3c: Morphology and Phonology

Supervised Learning of Complete Morphological Paradigms

Greg Durrett and John DeNero

3:45pm–4:10pm

We describe a supervised approach to predicting the set of all inflected forms of a lexical item. Our system automatically acquires the orthographic transformation rules of morphological paradigms from labeled examples, and then learns the contexts in which those transformations apply using a discriminative sequence model. Because our approach is completely data-driven and the model is trained on examples extracted from Wiktionary, our method can extend to new languages without change. Our end-to-end system is able to predict complete paradigms with 86.1% accuracy and individual inflected forms with 94.9% accuracy, averaged across three languages and two parts of speech.

Optimal Data Set Selection: An Application to Grapheme-to-Phoneme Conversion

Young-Bum Kim and Benjamin Snyder

4:10pm–4:35pm

In this paper we introduce the task of unlabeled, optimal, data set selection. Given a large pool of unlabeled examples, our goal is to select a small subset to label, which will yield a high performance supervised model over the entire data set. Our first proposed method, based on the rank-revealing QR matrix factorization, selects a subset of words which span the entire word-space effectively. For our second method, we develop the concept of feature coverage which we optimize with a greedy algorithm. We apply these methods to the task of grapheme-to-phoneme prediction. Experiments over a data-set of 8 languages show that in all scenarios, our selection methods are effective at yielding a small, but optimal set of labelled examples. When fed into a state-of-the-art supervised model for grapheme-to-phoneme prediction, our methods yield average error reductions of 20

Knowledge-Rich Morphological Priors for Bayesian Language Models

Victor Chahuneau, Noah A. Smith, and Chris Dyer

4:35pm–5:00pm

We present a morphology-aware nonparametric Bayesian model of language whose prior distribution uses manually constructed finite-state transducers to capture the word formation processes of particular languages. This relaxes the word independence assumption and enables sharing of statistical strength across, for example, stems or inflectional paradigms in different contexts. Our model can be used in virtually any generative model of text where multinomial distributions over words would be used. We obtain state-of-the-art results in language modeling, word alignment, and unsupervised morphological disambiguation in a variety of morphologically rich languages.

5

Second Joint Conference on Lexical and Computational Semantics (*SEM)

Introduction to *SEM 2013

Building on the momentum generated by the spectacular success of the Joint Conference on Lexical and Computational Semantics (*SEM) in 2012, bringing together the ACL SIGLEX and ACL SIGSEM communities, we are delighted to bring to you the second edition of the conference, as a top-tier showcase of the latest research in computational semantics. We accepted 14 papers (11 long and 3 short) for publication at the conference, out of a possible 45 submissions (a 31% acceptance rate). This is on par with some of the most competitive conferences in computational linguistics, and we are confident will set the stage for a scintillating conference.

This year, we started a tradition that we intend to maintain in all future iterations of the conference in integrating a shared task into the conference. The shared task was selected by an independent committee comprising members from SIGLEX and SIGSEM, based on an open call for proposals, and revolved around Semantic Textual Similarity (STS). The task turned out to be a huge success with 34 teams participating, submitting a total of 103 system runs.

*SEM 2013 features a number of highlight events:

Day One, June 13th:

- A timely and impressive panel on *Towards Deep Natural Language Understanding*, featuring the following panelists:
  - Kevin Knight (USC/Information Sciences Institute)
  - Chris Manning (Stanford University)
  - Martha Palmer (University of Colorado at Boulder)
  - Owen Rambow (Columbia University)
  - Dan Roth (University of Illinois at Urbana-Champaign)
- A Reception and Shared Task Poster Session in the evening, thanks to the generous sponsorship of the DARPA Deft program.

Day Two, June 14th:

- In the morning, a keynote address by David Forsyth from the Computer Science Department at the University of Illinois at Urbana Champagne on issues of Vision and Language. It promises to be an extremely stimulating speech, and is not to be missed.
- In the early afternoon, a panel on the relation between and future of *SEM, the *SEM Shared Task, SemEval and other events on computational semantics. In this panel, we will attempt to clarify and explain as well as devise plans for these different entities.
- Finally, at the end of the day, an award ceremony for the Best Long Paper and Best Short Paper.

As always, *SEM 2013 would not have been possible without the considerable efforts of our area chairs and an impressive assortment of reviewers, drawn from the ranks of SIGLEX and SIGSEM, and the computational semantics community at large. We would also like to acknowledge the generous support for the STS Task from the DARPA Deft Program.

We hope you enjoy *SEM 2013, and look forward to engaging with all of you,

Mona Diab (The George Washington University, General Chair)
Timothy Baldwin (The University of Mebourne, Program Committee Co-Chair)
Marco Baroni (University of Trento, Program Committee Co-Chair)

Introduction to SemEval

The Semantic Evaluation (SemEval) series of workshops focus on the evaluation and comparison of systems that can analyse diverse semantic phenomena in text with the aim of extending the current state-of-the-art in semantic analysis and creating high quality annotated datasets in a range of increasingly challenging problems in natural language semantics. SemEval provides an exciting forum for researchers to propose challenging research problems in semantics and to build systems/techniques to address such research problems.

SemEval-2013 is the seventh workshop in the series. The first three workshops, SensEval-1 (1998), SensEval-2 (2001), and SensEval-3 (2004), were focused on word sense disambiguation, each time growing in the number of languages offered in the tasks and in the number of participating teams. In 2007 the workshop was renamed SemEval and in the next three workshops SemEval-2007, SemEval-2010 and SemEval-2012 the nature of the tasks evolved to include semantic analysis tasks outside of word sense disambiguation. Starting in 2012 SemEval turned into a yearly event associated with *SEM.

This volume contains papers accepted for presentation at the SemEval-2013 International Workshop on Semantic Evaluation Exercises. SemEval-2013 is co-organized with the *SEM-2013 The Second Joint Conference on Lexical and Computational Semantics and co-located with The 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT).

SemEval-2013 included the following 12 tasks for evaluation:

- TempEval-3 Temporal Annotation
- Sentiment Analysis in Twitter
- Spatial Role Labeling
- Free Paraphrases of Noun Compounds
- Evaluating Phrasal Semantics
- The Joint Student Response Analysis and 8th Recognizing Textual Entailment Challenge
- Cross-lingual Textual Entailment for Content Synchronization
- Extraction of Drug-Drug Interactions from BioMedical Texts
- Cross-lingual Word Sense Disambiguation
- Evaluating Word Sense Induction & Disambiguation within An End-User Application
- Multilingual Word Sense Disambiguation
- Word Sense Induction for Graded and Non-Graded Senses

About 100 teams submitted more than 300 systems for the 12 tasks of SemEval-2013. This volume contains both Task Description papers that describe each of the above tasks and System Description papers that describe the systems that participated in the above tasks. A total of 12 task description papers and 101 system description papers are included in this volume.

We are indebted to all program committee members for their high quality, elaborate and thoughtful reviews. The papers in this proceedings have surely benefited from this feedback. We are grateful to *SEM 2013 and NAACL-HLT 2013 conference organizers for local organization and the forum. We most gratefully acknowledge the support of our sponsors, the ACL Special Interest Group on the Lexicon (SIGLEX) and the ACL Special Interest Group on Computational Semantics (SIGSEM).

Welcome to SemEval-2013!

Suresh Manandhar and Deniz Yuret

*SEM Organizing Committee

Organizers:

General Chair:
Mona Diab (George Washington University)

Program Committee Chairs:
Tim Baldwin (University of Melbourne)
Marco Baroni (University of Trento)

STS Shared Task Committee Chairs:
Anna Korhonen (University of Cambridge)
Malvina Nissim (University of Bologna)

SemEval Chairs:
Suresh Manandhar (University of York, UK)
Deniz Yuret (Koc University, Turkey)

Publications Chair:
Yuval Marton (Microsoft)

Sponsorship Chair:
Bernardo Magnini (Fondazione Bruno Kessler)

Panel organizer:
Martha Palmer (University of Colorado, Boulder)

Area Chairs:
Shane Bergsma (Johns Hopkins University)
Chris Biemann (TU Darmstadt)
Eduardo Blanco (Lymba Corporation)
Gemma Boleda (University of Texas, Austin)
Francis Bond (Nanyang Technological University)
Paul Cook (University of Melbourne)
Amaç Herdagdelen (Facebook)
Lauri Karttunen (Stanford University)
Diana McCarthy (University of Cambridge)
Roser Morante (University of Antwerp)
Smara Muresan (Rutgers University)
Preslav Nakov (Qatar Computing Research Institute)
Roberto Navigli (Sapienza University of Rome)
Hwee Tou Ng (National University of Singapore)
Becky Passonneau (Columbia University)
Laura Rimell (University of Cambridge)
Caroline Sporleder (University of Trier)
Fabio Massimo Zanzotto (University of Rome Tor Vergata)

Invited Speaker:

David Forsyth (University of Illinois, Urbana-Champaign)

Panelists for *SEM panel:

Kevin Knight (USC Information Sciences Institute)

Chris Manning (Stanford University)

Martha Palmer (University of Colorado at Boulder)

Owen Rambow (Columbia University)

Dan Roth (University of Illinois at Urbana-Champaign)

Panelists for Shared *SEM/SemEval panel:

*SEM and SemEval organizers

*SEM (Main conference and shared task)
Thursday, June 13–Friday, June 14

Venue: International D

Thursday, June 13, 2013

8:00–08:45 **Registration**

(08:45–10:30) **Session *SEM1: Opening Remarks and *SEM Long Papers (1)**

09:00–09:30 Towards a Formal Distributional Semantics: Simulating Logical Calculi with Tensors
Edward Grefenstette

09:30–10:00 Montague Meets Markov: Deep Semantics with Probabilistic Logical Form
Islam Beltagy, Cuong Chau, Gemma Boleda, Dan Garrette, Katrin Erk, and Raymond Mooney

10:00–10:30 Coarse to Fine Grained Sense Disambiguation in Wikipedia
Hui Shen, Razvan Bunescu, and Rada Mihalcea

10:30–11:00 **Coffee Break**

(11:00–12:30) **Session ST1: STS Shared Task (1)**

11:00–11:30 *SEM 2013 shared task: Semantic Textual Similarity
Eneko Agirre, Daniel Cer, Mona Diab, Aitor Gonzalez-Agirre, and Weiwei Guo

11:30–11:50 UMBC_EBIQUITY-CORE: Semantic Textual Similarity Systems
Lushan Han, Abhay L. Kashyap, Tim Finin, James Mayfield, and Jonathan Weese

11:50–12:10 iKernels-Core: Tree Kernel Learning for Textual Similarity
Aliaksei Severyn, Massimo Nicosia, and Alessandro Moschitti

12:10–12:30 UNITOR-CORE_TYPED: Combining Text Similarity and Semantic Filters through SV Regression
Danilo Croce, Valerio Storch, and Roberto Basili

12:30–2:00 **Lunch**

(2:00–2:40) **Session ST2a: STS Shared Task (2)**

2:00–2:20 NTNU-CORE: Combining strong features for semantic similarity
Erwin Marsi, Hans Moen, Lars Bungum, Gleb Sizov, Björn Gambäck, and André Lynum

2:20–2:40 SXUCFN-Core: STS Models Integrating FrameNet Parsing Information
Sai Wang, Ru Li, Ruibo Wang, Zhiqiang Wang, and Xia Zhang

2:40–3:30 **Session ST2b: STS Poster boosters**

3:30–4:00 **Coffee Break**

(4:00–4:25) **Session *SEM2a: *SEM Short Papers (1)**

4:00–4:25 Distinguishing Common and Proper Nouns
Judita Preiss and Mark Stevenson

4:30–6:00 **Session *SEM2b: *SEM Panel: Toward Deep Natural Language Understanding: Kevin Knight (USC/ISI), Christopher Manning (Stanford University), Martha Palmer (University of Colorado, Boulder), Owen Rambow (Columbia University), Dan Roth (University of Illinois Urbana Champagne)**

6:30–8:30 **Session PLN1: *SEM Opening Reception and STS Poster Session**

- UCAM-CORE: Incorporating structured distributional similarity into STS
  Tamara Polajnar, Laura Rimell, and Douwe Kiela
- PolyUCOMP-CORE_TYPED: Computing Semantic Textual Similarity using Overlapped Senses
  Jian Xu and Qin Lu
- HENRY-CORE: Domain Adaptation and Stacking for Text Similarity
  Michael Heilman and Nitin Madnani
- DeepPurple: Lexical, String and Affective Feature Fusion for Sentence-Level Semantic Similarity Estimation
  Nikolaos Malandrakis, Elias Iosif, Vassiliki Prokopi, Alexandros Potamianos, and Shrikanth Narayanan
- UMCC_DLSI: Textual Similarity based on Lexical-Semantic features
  Alexander Chávez, Héctor Dávila, Yoan Gutiérrez, Armando Collazo, José I. Abreu, Antonio Fernández Orquín, Andrés Montoyo, and Rafael Muñoz
- BUT-TYPED: Using domain knowledge for computing typed similarity
  Lubomir Otrusina and Pavel Smrz
- ECNUCS: Measuring Short Text Semantic Equivalence Using Multiple Similarity Measurements
  Zhu Tiantian and Man Lan
- UBC_UOS-TYPED: Regression for typed-similarity
  Eneko Agirre, Nikolaos Aletras, Aitor Gonzalez-Agirre, German Rigau, and Mark Stevenson
- KnCe2013-CORE:Semantic Text Similarity by use of Knowledge Bases
  Hermann Ziak and Roman Kern
- UPC-CORE: What Can Machine Translation Evaluation Metrics and Wikipedia Do for Estimating Semantic Textual Similarity?
  Alberto Barrón-Cedeño, Lluís Màrquez, Maria Fuentes, Horacio Rodriguez, and Jordi Turmo
- MayoClinicNLP–CORE: Semantic representations for textual similarity
  Stephen Wu, Dongqing Zhu, Ben Carterette, and Hongfang Liu
- SRIUBC-Core: Multiword Soft Similarity Models for Textual Similarity
  Eric Yeh
- LIPN-CORE: Semantic Text Similarity using n-grams, WordNet, Syntactic Analysis, ESA and Information Retrieval based Features
  Davide Buscaldi, Joseph Le Roux, Jorge J. Garcia Flores, and Adrian Popescu
- UNIBA-CORE: Combining Strategies for Semantic Textual Similarity
  Annalina Caputo, Pierpaolo Basile, and Giovanni Semeraro
- DLS@CU-CORE: A Simple Machine Learning Model of Semantic Textual Similarity
  Md. Sultan, Steven Bethard, and Tamara Sumner
- KLUE-CORE: A regression model of semantic textual similarity
  Paul Greiner, Thomas Proisl, Stefan Evert, and Besim Kabashi
- IBM_EG-CORE: Comparing multiple Lexical and NE matching features in measuring Semantic Textual similarity
  Sara Noeman

- SOFTCARDINALITY-CORE: Improving Text Overlap with Distributional Measures for Semantic Textual Similarity
  Sergio Jimenez, Claudia Becerra, and Alexander Gelbukh
- CLaC-CORE: Exhaustive Feature Combination for Measuring Textual Similarity
  Ehsan Shareghi and Sabine Bergler
- UniMelb_NLP-CORE: Integrating predictions from multiple domains and feature sets for estimating semantic textual similarity
  Spandana Gella, Bahar Salehi, Marco Lui, Karl Grieser, Paul Cook, and Timothy Baldwin
- CFILT-CORE: Semantic Textual Similarity using Universal Networking Language
  Avishek Dan and Pushpak Bhattacharyya
- CPN-CORE: A Text Semantic Similarity System Infused with Opinion Knowledge
  Carmen Banea, Yoonjung Choi, Lingjia Deng, Samer Hassan, Michael Mohler, Bishan Yang, Claire Cardie, Rada Mihalcea, and Jan Wiebe
- INAOE_UPV-CORE: Extracting Word Associations from Document Corpora to estimate Semantic Textual Similarity
  Fernando Sánchez-Vega, Manuel Montes-y Gómez, Paolo Rosso, and Luis Villaseñor-Pineda
- CNGL-CORE: Referential Translation Machines for Measuring Semantic Similarity
  Ergun Bicici and Josef Van Genabith

Friday, June 14, 2013

08:00–08:30 **Registration**

(08:30–09:30) **Session *SEM3: *SEM Short Papers (2)**

08:30–08:55 A Dataset of Syntactic-Ngrams over Time from a Very Large Corpus of English Books
Yoav Goldberg and Jon Orwant

08:55–09:20 Unsupervised Word Usage Similarity in Social Media Texts
Spandana Gella, Paul Cook, and Bo Han

9:20–9:30 **Short break for re-organization**

9:30–10:30 **Session PLN2: Keynote address: David Forsyth (University of Illinois Urbana Champagne)**

09:30–10:30 More Words and Bigger Pictures
David Forsyth

10:30–11:00 **Coffee Break**

(11:00–12:30) **Session *SEM4: *SEM Long Papers (2)**

11:00–11:30 Exploring Vector Space Models to Predict the Compositionality of German Noun-Noun Compounds
Sabine Schulte im Walde, Stefan Müller, and Stefan Roller

11:30–12:00 Predicting the Compositionality of Multiword Expressions Using Translations in Multiple Languages
Bahar Salehi and Paul Cook

12:00–12:30 Metaphor Identification as Interpretation
Ekaterina Shutova

12:30–1:30 **Lunch**

1:30–2:30 **Session PLN3: *SEM / STS Shared Task / SemEval Panel**

(2:30–3:30) **Session *SEM5: *SEM Long Papers (3)**

2:30–3:00 Using the text to evaluate short answers for reading comprehension exercises
Andrea Horbach, Alexis Palmer, and Manfred Pinkal

3:00–3:30 Choosing the Right Words: Characterizing and Reducing Error of the Word Count Approach
Hansen Andrew Schwartz, Johannes Eichstaedt, Eduardo Blanco, Lukasz Dziurzyński, Margaret L. Kern, Stephanie Ramones, Martin Seligman, and Lyle Ungar

3:30–4:00 **Coffee Break**

(4:00–5:30) **Session *SEM6: *SEM Long Papers (4)**

4:00–4:30 Automatically Identifying Implicit Arguments to Improve Argument Linking and Coherence Modeling
Michael Roth and Anette Frank

4:30–5:00 Bootstrapping Semantic Role Labelers from Parallel Data
Mikhail Kozhevnikov and Ivan Titov

5:00–5:30 Semantic Parsing Freebase: Towards Open-domain Semantic Parsing
Qingqing Cai and Alexander Yates

5:30–6:00 **Session *SEM7: Best Papers Awards and Closing Remarks**

Session *SEM1: (08:45–10:30) Opening Remarks and *SEM Long Papers (1)

Towards a Formal Distributional Semantics: Simulating Logical Calculi with Tensors

Edward Grefenstette
09:00am–09:30am

The development of compositional distributional models of semantics reconciling the empirical aspects of distributional semantics with the compositional aspects of formal semantics is a popular topic in the contemporary literature. This paper seeks to bring this reconciliation one step further, by showing how the mathematical constructs commonly used in compositional distributional models, such as tensors and matrices, can be used to simulate different aspects of predicate logic.

This paper discusses how the canonical isomorphism between tensors and multilinear maps can be exploited to simulate a full-blown quantifier free predicate calculus using tensors. It provides tensor interpretations of the set of logical connectives required to model propositional calculi. It suggests a variant of these tensor calculi capable of modelling quantifiers, using few non-linear operations. It finally discusses the relation between these variants, and how this relation should constitute the subject of future work.

Montague Meets Markov: Deep Semantics with Probabilistic Logical Form

Islam Beltagy, Cuong Chau, Gemma Boleda, Dan Garrette, Katrin Erk, and Raymond Mooney
09:30am–10:00am

We combine logical and distributional representations of natural language meaning by transforming distributional similarity judgments into weighted inference rules using Markov Logic Networks (MLNs). We show that this framework supports both judging sentence similarity and recognizing textual entailment by appropriately adapting the MLN implementation of logical connectives. We also show that distributional phrase similarity, used as textual inference rules created on the fly improves its performance.

Coarse to Fine Grained Sense Disambiguation in Wikipedia

Hui Shen, Razvan Bunescu, and Rada Mihalcea
10:00am–10:30am

Wikipedia articles are annotated by volunteer contributors with numerous links that connect words and phrases to relevant titles. Links to general senses of a word are used concurrently with links to more specific senses, without being distinguished explicitly. We present an approach to training coarse to fine grained sense disambiguation systems in the presence of such annotation inconsistencies. Experimental results show that accounting for annotation ambiguity in Wikipedia links leads to significant improvements in disambiguation.

Session ST1: (11:00–12:30) STS Shared Task (1)

***SEM 2013 shared task: Semantic Textual Similarity**
Eneko Agirre, Daniel Cer, Mona Diab, Aitor Gonzalez-Agirre, and Weiwei Guo
11:00am–11:30am

Semantic Textual Similarity (STS) measures the degree of semantic equivalence, on a scale from 0 to 5. This year we set up two tasks. The core task is similar to STS 2012, with pairs of sentences from news headlines, ma- chine translation evaluation datasets and lexi- cal resource glosses. The typed-similarity task is novel and involves pairs of cultural heritage items which are described with metadata like title, author or description. Several types of similarity have been defined, including similar author, similar time period or similar location. The annotation leverages crowdsourcing, with relative high inter-annotator correlation, rang- ing from 62% to 87%. The core task attracted 34 participants corresponding to 89 runs, and the typed-similarity task attracted 6 teams cor- responding to 14 runs.

UMBC_EBIQUITY-CORE: Semantic Textual Similarity Systems
Lushan Han, Abhay L. Kashyap, Tim Finin, James Mayfield, and Jonathan Weese
11:30am–11:50am

We describe three semantic text similarity systems developed for the *SEM 2013 STS shared task and the results of the corresponding three runs. All of them shared a word similarity feature that combined LSA word similarity and WordNet knowledge. The first, which achieved the best mean score of the 89 submitted runs, used a simple term alignment algorithm augmented with penalty terms. The other two runs, ranked second and fourth, used support vector regression models to combine larger sets of features.

iKernels-Core: Tree Kernel Learning for Textual Similarity
Aliaksei Severyn, Massimo Nicosia, and Alessandro Moschitti
11:50am–12:10pm

This paper describes the participation of iKernels system in the Semantic Textual Similarity (STS) shared task at *SEM 2013. Different from the majority of approaches, where a large number of pairwise similarity features are used to learn a regression model, our model directly encodes the input texts into syntactic/semantic structures. Our systems rely on tree kernels to automatically extract a rich set of syntactic patterns to learn a similarity score correlated with human judgements. We experiment with different structural representations derived from constituency and dependency trees. While showing large improvements over the top results from the previous year task (STS-2012), our best system ranks 21st out of total 88 participated in the STS-2013 task. Nevertheless, a slight refinement to our model makes it rank 4th.

UNITOR-CORE_TYPED: Combining Text Similarity and Semantic Filters through SV Regression
Danilo Croce, Valerio Storch, and Roberto Basili
12:10pm–12:30pm

This paper presents the UNITOR system that participated in the *SEM 2013 shared task on Semantic Textual Similarity (STS). The task is modeled as a Support Vector (SV) regression problem, where a similarity scoring function between text pairs is acquired from examples. The proposed approach has been implemented in a system that aims at providing high applicability and robustness, in order to reduce the risk of over-fitting over a specific datasets. Moreover, the approach does not require any manually coded resource (e.g. WordNet), but mainly exploits distributional analysis of unlabeled corpora. A good level of accuracy is achieved over the shared task: in the Typed STS task the proposed system ranks in 1st and 2nd position.

Session ST2a: (2:00–2:40) STS Shared Task (2)

NTNU-CORE: Combining strong features for semantic similarity
Erwin Marsi, Hans Moen, Lars Bungum, Gleb Sizov, Björn Gambäck, and André Lynum
2:00pm–2:20pm

The paper outlines the work carried out at NTNU as part of the *SEM'13 shared task on Semantic Textual Similarity, using an approach which combines shallow textual, distributional and knowledge-based features by a support vector regression model. Feature sets include (1) aggregated similarity based on named entity recognition with WordNet and Levenshtein distance through the calculation of maximum weighted bipartite graphs; (2) higher order word co-occurrence similarity using a novel method called "Multi-sense Random Indexing"; (3) deeper semantic relations based on the RelEx semantic dependency relationship extraction system; (4) graph edit-distance on dependency trees; (5) reused features of the TakeLab and DKPro systems from the STS'12 shared task. The NTNU systems obtained 9th place overall (5th best team) and 1st place on the SMT data set.

SXUCFN-Core: STS Models Integrating FrameNet Parsing Information
Sai Wang, Ru Li, Ruibo Wang, Zhiqiang Wang, and Xia Zhang
2:20pm–2:40pm

This paper describes our system submitted to *SEM 2013 Semantic Textual Similarity (STS) core task which aims to measure semantic similarity of two given text snippets. In this shared task, we propose an interpolation STS model named Model_LIM integrating FrameNet parsing information, which has a good performance with low time complexity compared with former submissions.

Session *SEM2a: (4:00–4:25) *SEM Short Papers (1)

Distinguishing Common and Proper Nouns
Judita Preiss and Mark Stevenson
4:00pm–4:25pm

We describe a number of techniques for automatically deriving lists of common and proper nouns, and show that the distinction between the two can be made automatically using a vector space model learning algorithm. We present a direct evaluation on the British National Corpus, and application based evaluations on Twitter messages and on automatic speech recognition (where the system could be employed to restore case).

Session *SEM3: (08:30–09:30) *SEM Short Papers (2)

A Dataset of Syntactic-Ngrams over Time from a Very Large Corpus of English Books
Yoav Goldberg and Jon Orwant
08:30pm–08:55pm

We created a dataset of syntactic-ngrams (counted dependency-tree fragments) based on a corpus of 3.5 million English books. The dataset includes over 10 billion distinct items covering a wide range of syntactic configurations. It also includes temporal information, facilitating new kinds of research into lexical semantics over time. This paper describes the dataset, the syntactic representation, and the kinds of information provided.

Unsupervised Word Usage Similarity in Social Media Texts
Spandana Gella, Paul Cook, and Bo Han
08:55pm–09:20am

We propose an unsupervised method for automatically calculating word usage similarity in social media data based on topic modelling, which we contrast with a baseline distributional method and Weighted Textual Matrix Factorisation. We evaluate these methods against a novel dataset made up of human ratings over 550 Twitter message pairs annotated for usage similarity for a set of 10 nouns. The results show that our topic modelling approach outperforms the other two methods.

More Words and Bigger Pictures
David Forsyth
09:30am–10:30am

Object recognition is a little like translation: a picture (text in a source language) goes in, and a description (text in a target language) comes out. I will use this analogy, which has proven fertile, to describe recent progress in object recognition.

We have very good methods to spot some objects in images, but extending these methods to produce descriptions of images remains very difficult. The description might come in the form of a set of words, indicating objects, and boxes or regions spanned by the object. This representation is difficult to work with, because some objects seem to be much more important than others, and because objects interact. An alternative is a sentence or a paragraph describing the picture, and recent work indicates how one might generate rich structures like this. Furthermore, recent work suggests that it is easier and more effective to generate descriptions of images in terms of chunks of meaning (”person on a horse”) rather than just objects (”person”; “horse”).

Finally, if the picture contains objects that are unfamiliar, then we need to generate useful descriptions that will make it possible to interact with them, even though we don’t know what they are.

Session *SEM4: (11:00–12:30) *SEM Long Papers (2)

Exploring Vector Space Models to Predict the Compositionality of German Noun-Noun Compounds

Sabine Schulte im Walde, Stefan Müller, and Stefan Roller

11:00am–11:30am

This paper explores two hypotheses regarding vector space models that predict the compositionality of German noun-noun compounds: (1) Against our intuition, we demonstrate that window-based rather than syntax-based distributional features perform better predictions, and that not adjectives or verbs but nouns represent the most salient part-of-speech. Our overall best result is state-of-the-art, reaching Spearman's rho=0.65 with a word-space model of nominal features from a 20-word window of a 1.5 billion word web corpus. (2) While there are no significant differences in predicting compound–modifier vs. compound–head ratings on compositionality, we show that the modifier (rather than the head) properties predominantly influence the degree of compositionality of the compound.

Predicting the Compositionality of Multiword Expressions Using Translations in Multiple Languages

Bahar Salehi and Paul Cook

11:30am–12:00pm

In this paper, we propose a simple, language-independent and highly effective method for predicting the degree of compositionality of multiword expressions (MWEs). We compare the translations of an MWE with the translations of its components, using a range of different languages and string similarity measures. We demonstrate the effectiveness of the method on two types of English MWEs: noun compounds and verb particle constructions. The results show that our approach is competitive with or superior to state-of-the-art methods over standard datasets.

Metaphor Identification as Interpretation

Ekaterina Shutova

12:00pm–12:30pm

Automatic metaphor identification and interpretation in text have been traditionally considered as two separate tasks in natural language processing (NLP) and addressed individually within computational frameworks. However, cognitive evidence suggests that humans are likely to perform these two tasks simultaneously, as part of a holistic metaphor comprehension process. We present a novel method that performs metaphor identification through its interpretation, being the first one in NLP to combine the two tasks in one step. It outperforms the previous approaches to metaphor identification both in terms of accuracy and coverage, as well as providing an interpretation for each identified expression.

Session *SEM5: (2:30–3:30) *SEM Long Papers (3)

Using the text to evaluate short answers for reading comprehension exercises

Andrea Horbach, Alexis Palmer, and Manfred Pinkal

2:30pm–3:00pm

Short answer questions for reading comprehension are a common task in foreign language learning. Automatic short answer scoring is the task of automatically assessing the semantic content of a student's answer, marking it e.g. as correct or incorrect. While previous approaches mainly focused on comparing a learner answer to some reference answer provided by the teacher, we explore the use of the underlying reading texts as additional evidence for the classification. First, we conduct a corpus study targeting the relation between answers and sentences in reading texts for learners of German. Second, we use the reading text directly for classification, considering three different models: an answer-based classifier extended with textual features, a simple text-based classifier, and a model that combines the two according to confidence of the text-based classification. The most promising approach is the first one, results for which show that textual features improve classification accuracy. While the other two models do not improve classification accuracy, they do investigate the role of the text and suggest possibilities for developing automatic answer scoring systems with less supervision needed from instructors.

Choosing the Right Words: Characterizing and Reducing Error of the Word Count Approach

Hansen Andrew Schwartz, Johannes Eichstaedt, Eduardo Blanco, Lukasz Dziurzyński, Margaret L. Kern, Stephanie Ramones, Martin Seligman, and Lyle Ungar

3:00pm–3:30pm

Social scientists are increasingly using the vast amount of text available on social media to measure variation in happiness and other psychological states. Such studies count words deemed to be indicators of happiness and track how the word frequencies change across locations or time. This "word count" approach is simple and scalable, yet often picks up false signals, as words can appear in different contexts and take on different meanings. We characterize the types of errors that occur using the word count approach, and find lexical ambiguity to be the most prevalent. We then show that one can reduce error with a simple refinement to such lexica by automatically eliminating highly ambiguous words. The resulting refined lexica improve precision as measured by human judgments of word occurrences in Facebook posts.

Session *SEM6: (4:00–5:30) *SEM Long Papers (4)

Automatically Identifying Implicit Arguments to Improve Argument Linking and Coherence Modeling

Michael Roth and Anette Frank

4:00pm–4:30pm

Implicit arguments are a discourse-level phenomenon that has not been extensively studied in semantic processing. One reason for this lies in the scarce amount of annotated data sets available. We argue that more data of this kind would be helpful to improve existing approaches to linking implicit arguments in discourse and to enable more in-depth studies of the phenomenon itself. In this paper, we present a range of studies that empirically validate this claim. Our contributions are threefold: we present a heuristic approach to automatically identify implicit arguments and their antecedents by exploiting comparable texts; we show how the induced data can be used as training data for improving existing argument linking models; finally, we present a novel approach to modeling local coherence that extends previous approaches by taking into account non-explicit entity references.

Bootstrapping Semantic Role Labelers from Parallel Data

Mikhail Kozhevnikov and Ivan Titov

4:30pm–5:00pm

We present an approach which uses the similarity in semantic structure of bilingual parallel sentences to bootstrap a pair of semantic role labeling (SRL) models. The setting is similar to co-training, except for the intermediate model required to convert the SRL structure between the two annotation schemes used for different languages. Our approach can facilitate the construction of SRL models for resource-poor languages, while preserving the annotation schemes designed for the target language and making use of the limited resources available for it. We evaluate the model on four language pairs, English vs German, Spanish, Czech and Chinese. Consistent improvements are observed over the self-training baseline.

Semantic Parsing Freebase: Towards Open-domain Semantic Parsing

Qingqing Cai and Alexander Yates

5:00pm–5:30pm

Existing semantic parsing research has steadily improved accuracy on a few domains and their corresponding databases. This paper introduces FreeParser, a system that trains on one domain and one set of predicate and constant symbols, and then can parse sentences for any new domain, including sentences that refer to symbols never seen during training. FreeParser uses a domain-independent architecture to automatically identify sentences relevant to each new database symbol, which it uses to supplement its manually-annotated training data from the training domain. In cross-domain experiments involving 23 domains, FreeParser can parse sentences for which it has seen comparable unannotated sentences with an F1 of 0.71.

Posters

UCAM-CORE: Incorporating structured distributional similarity into STS

Tamara Polajnar, Laura Rimell, and Douwe Kiela

This paper describes methods that were submitted as part of the *SEM shared task on Semantic Textual Similarity.

Multiple kernels provide different views of syntactic structure, from both tree and dependency parses. The kernels are then combined with simple lexical features using Gaussian process regression, which is trained on different subsets of training data for each run. We found that the simplest combination has the highest consistency across the different data sets, while introduction of more training data and models requires training and test data with matching qualities.

PolyUCOMP-CORE_TYPED: Computing Semantic Textual Similarity using Overlapped Senses

Jian Xu and Qin Lu

The Semantic Textual Similarity (STS) task aims to exam the degree of semantic equivalence between sentences (Agirre et al., 2012). This paper presents the work of the Hong Kong Polytechnic University (PolyUCOMP) team which has participated in the STS core and typed tasks of SemEval-2013. For the STS core task, the PolyUCOMP system disambiguates words senses using contexts and then determine sentence similarity by counting the number of senses they shared. For the STS typed task, the string kernel (Lodhi et al., 2002) is used to compute similarity between two entities to avoid string variations in entities.

HENRY-CORE: Domain Adaptation and Stacking for Text Similarity

Michael Heilman and Nitin Madnani

This paper describes a system for automatically measuring the semantic similarity between two texts, which was the aim of the 2013 Semantic Textual Similarity (STS) task (Agirre et al., 2013). For the 2012 STS task, Heilman and Madnani (2012) submitted the PERP system, which performed competitively in relation to other submissions. However, approaches including word and n-gram features also performed well (Baer et al., 2012; Saric et al., 2012), and the 2013 STS task focused more on predicting similarity for text pairs from new domains. Therefore, for the three variations of our system that we were allowed to submit, we used stacking (Wolpert, 1992) to combine PERP with word and n-gram features and applied the domain adaptation approach outlined by Daume III (2007) to facilitate generalization to new domains. Our submissions performed well at most subtasks, particularly at measuring the similarity of news headlines, where one of our submissions ranked 2nd among 90 from 36 teams, but there is still room for improvement.

DeepPurple: Lexical, String and Affective Feature Fusion for Sentence-Level Semantic Similarity Estimation

Nikolaos Malandrakis, Elias Iosif, Vassiliki Prokopi, Alexandros Potamianos, and Shrikanth Narayanan

This paper describes our submission for SemEval2013 Task 6: Semantic Textual Similarity. We estimate the semantic similarity between two sentences using regression models with features: 1) n-gram hit rates (lexical matches) between sentences, 2) lexical semantic similarity between non-matching words, 3) string similarity metrics, 4) affective content similarity and 5) sentence length. Domain adaptation is applied in the form of independent models and a model selection strategy achieving a mean correlation of 0.47.

UMCC_DLSI: Textual Similarity based on Lexical-Semantic features

Alexander Chávez, Héctor Dávila, Yoan Gutiérrez, Armando Collazo, José I. Abreu, Antonio Fernández Orquín, Andrés Montoyo, and Rafael Muñoz

This paper describes the specifications and results of UMCC_DLSI system, which participated in the Semantic Textual Similarity task (STS) of SemEval-2013. Our supervised system uses different types of lexical and semantic features to train a Bagging classifier used to decide the correct option. Related to the different features we can highlight the resource ISR-WN used to extract semantic relations among words and the use of different algorithms to establish semantic and lexical similarities. In order to establish which features are the

most appropriate to improve STS results we participated with three runs using different set of features. Our best run reached the position 44 in the official ranking, obtaining a general correlation coefficient of 0.61.

BUT-TYPED: Using domain knowledge for computing typed similarity
Lubomir Otrusina and Pavel Smrz

This paper deals with knowledge-based text processing which aims at an intuitive notion of textual similarity. Entities and relations relevant for a particular domain are identified and disambiguated by means of semi-supervised machine learning techniques and resulting annotations are applied for computing typed-similarity of individual texts.

The work described in this paper particularly shows effects of the mentioned processes in the context of the *SEM 2013 pilot task on typed-similarity, a part of the Semantic Textual Similarity shared task. The goal is to evaluate the degree of semantic similarity between semi-structured records. As the evaluation dataset has been taken from Europeana - a collection of records on European cultural heritage objects - we focus on computing a semantic distance on field author which has the highest potential to benefit from the domain knowledge.

Specific features that are employed in our system BUT-TYPED are briefly introduced together with a discussion on their efficient acquisition. Support Vector Regression is then used to combine the features and to provide a final similarity score. The system ranked third on the attribute author among 15 submitted runs in the typed-similarity task.

ECNUCS: Measuring Short Text Semantic Equivalence Using Multiple Similarity Measurements
Zhu Tiantian and Man Lan

This paper reports our submissions to the Semantic Textual Similarity (STS) task in SemEval 2013 (Task 6). We submitted three Support Vector Regression (SVR) systems in core task, using 6 types of similarity measures, i.e., string similarity, number similarity, knowledge-based similarity, corpus-based similarity, syntactic dependency similarity and machine translation similarity. Our third system with different training data and different feature sets for each test data set performs the best and ranks 35 out of 90 runs. We also submitted two systems in typed task using string based measure and Named Entity based measure. Our best system ranks 5 out of 15 runs.

UBC_UOS-TYPED: Regression for typed-similarity
Eneko Agirre, Nikolaos Aletras, Aitor Gonzalez-Agirre, German Rigau, and Mark Stevenson

We approach the typed-similarity task using several kinds of heuristic similarity techniques for each type of similarity, based on the information of appropriated metadata fields. In addition we train a linear regressor for each type of similarity. The results indicate that the linear regression is key for good performance. Our best system ranked third in the task.

KnCe2013-CORE:Semantic Text Similarity by use of Knowledge Bases
Hermann Ziak and Roman Kern

In this paper we describe KnCe2013-CORE, a system to compute the semantic similarity of two short text snippets. The system computes a number of features which are gathered from different knowledge bases, namely WordNet, Wikipedia and Wiktionary. The similarity scores derived from these features are then fed into a multilayer perceptron neuronal networks. Depending on the size of the text snippets different parameters for the neural networks are used. The final output of the neural network is compared to human judgements. In the evaluation our system performed sufficiently well for text snippets of equal length, but the performance dropped considerably once the pairs of text snippets differ in size.

UPC-CORE: What Can Machine Translation Evaluation Metrics and Wikipedia Do for Estimating Semantic Textual Similarity?
Alberto Barrón-Cedeño, Lluís Màrquez, Maria Fuentes, Horacio Rodriguez, and Jordi Turmo

In this paper we discuss our participation to the 2013 Semeval Semantic Textual Similarity task. Our core fea-

tures include (i) a set of metrics borrowed from automatic machine translation, originally intended to evaluate automatic against reference translations and (ii) an instance of explicit semantic analysis, built upon opening paragraphs of Wikipedia 2010 articles. Our similarity estimator relies on a support vector regressor with RBF kernel. Our best approach required 13 machine translation metrics + explicit semantic analysis and ranked 65 in the competition. Our post-competition analysis shows that the features have a good expression level, but overfitting and —mainly— normalization issues caused our correlation values to decrease.

MayoClinicNLP–CORE: Semantic representations for textual similarity
Stephen Wu, Dongqing Zhu, Ben Carterette, and Hongfang Liu

The Semantic Textual Similarity (STS) task examines semantic similarity at a sentence-level. We explored three representations of semantics (implicit or explicit): named entities, semantic vectors, and structured vectorial semantics. From a DKPro baseline, we also performed feature selection and used source-specific linear regression models to combine our features. Our systems placed 5th, 6th, and 8th among 90 submitted systems.

SRIUBC-Core: Multiword Soft Similarity Models for Textual Similarity
Eric Yeh

In this year's Semantic Textual Similarity evaluation, we explore the contribution of models that provide soft similarity scores across spans of multiple words, over the previous year's system. To this end, we explored the use of neural probabilistic language models and a TF-IDF weighted variant of Explicit Semantic Analysis. The neural language model systems used vector representations of individual words, where these vectors were derived by training them against the context of words encountered, and thus reflect the distributional characteristics of their usage. To generate a similarity score between spans, we experimented with using tiled vectors and Restricted Boltzmann Machines to identify similar encodings. We find that these soft similarity methods generally outperformed our previous year's systems, albeit they did not perform as well in the overall rankings. A simple analysis of the soft similarity resources over two word phrases is provided, and future areas of improvement are described.

LIPN-CORE: Semantic Text Similarity using n-grams, WordNet, Syntactic Analysis, ESA and Information Retrieval based Features
Davide Buscaldi, Joseph Le Roux, Jorge J. Garcia Flores, and Adrian Popescu

This paper describes the system used by the LIPN team in the Semantic Textual Similarity task at SemEval 2013. It uses a support vector regression model, combining different text similarity measures that constitute the features. These measures include simple distances like Levenshtein edit distance, cosine, Named Entities overlap and more complex distances like Explicit Semantic Analysis, WordNet-based similarity, IR-based similarity, and a similarity measure based on syntactic dependencies.

UNIBA-CORE: Combining Strategies for Semantic Textual Similarity
Annalina Caputo, Pierpaolo Basile, and Giovanni Semeraro

This paper describes the UNIBA participation in the Semantic Textual Similarity (STS) core task 2013. We exploited three different systems for computing the similarity between two texts. A system is used as baseline, which represents the best model emerged from our previous participation in STS 2012. Such system is based on a distributional model of semantics capable of taking into account also syntactic structures that glue words together. In addition, we investigated the use of two different learning strategies exploiting both syntactic and semantic features. The former uses a combination strategy in order to combine the best machine learning techniques trained on 2012 training and test sets. The latter tries to overcame the limit of working with different datasets with varying characteristics by selecting only the more suitable dataset for the training purpose.

DLS@CU-CORE: A Simple Machine Learning Model of Semantic Textual Similarity
Md. Sultan, Steven Bethard, and Tamara Sumner

We present a system submitted in the Semantic Textual Similarity (STS) task at the Second Joint Conference on Lexical and Computational Semantics (*SEM 2013). Given two short text fragments, the goal of the system is to determine their semantic similarity. Our system makes use of three different measures of text similarity: word n-gram overlap, character n-gram overlap and semantic overlap. Using these measures as features, it

trains a support vector regression model on SemEval STS 2012 data. This model is then applied on the STS 2013 data to compute textual similarities. Two different selections of training data result in very different performance levels: while a correlation of 0.4135 with gold standards was observed in the official evaluation (ranked 63rd among all systems) for one selection, the other resulted in a correlation of 0.5352 (that would rank 21st).

KLUE-CORE: A regression model of semantic textual similarity
Paul Greiner, Thomas Proisl, Stefan Evert, and Besim Kabashi

This paper describes our system entered for the *SEM 2013 shared task on Semantic Textual Similarity (STS). We focus on the core task of predicting the semantic textual similarity of sentence pairs.

The current system utilizes machine learning techniques trained on semantic similarity ratings from the *SEM 2012 shared task; it achieved rank 20 out of 90 submissions from 35 different teams. Given the simple nature of our approach, which uses only WordNet and unannotated corpus data as external resources, we consider this a remarkably good result, making the system an interesting tool for a wide range of practical applications.

IBM_EG-CORE: Comparing multiple Lexical and NE matching features in measuring Semantic Textual similarity
Sara Noeman

We present in this paper the systems we participated with in the Semantic Textual Similarity task at SEM 2013. The Semantic Textual Similarity Core task (STS) computes the degree of semantic equivalence between two sentences where the participant systems will be compared to the manual scores, which range from 5 (semantic equivalence) to 0 (no relation). We combined multiple text similarity measures of varying complexity. The experiments illustrate the different effect of four feature types including direct lexical matching, idf-weighted lexical matching, modified BLEU N-gram matching and named entities matching. Our team submitted three runs during the task evaluation period and they ranked number 11, 15 and 19 among the 90 participating systems according to the official Mean Pearson correlation metric for the task. We also report an unofficial run with mean Pearson correlation of 0.59221 on STS2013 test dataset, ranking as the 3rd best system among the 90 participating systems.

SOFTCARDINALITY-CORE: Improving Text Overlap with Distributional Measures for Semantic Textual Similarity
Sergio Jimenez, Claudia Becerra, and Alexander Gelbukh

Soft cardinality has been shown to be a very strong text-overlapping baseline for the task of measuring semantic textual similarity (STS), obtaining 3rd place in SemEval-2012. At *SEM-2013 shared task, beside the plain text-overlapping approach, we tested within soft cardinality two distributional word-similarity functions derived from the ukWack corpus. Unfortunately, we combined these measures with other features using regression, obtaining positions 18th, 22nd and 23rd among the 90 participants systems in the official ranking. Already after the release of the gold standard annotations of the test data, we observed that using only the similarity measures without combining them with other features would have obtained positions 6th, 7th and 8th; moreover, an arithmetic average of these similarity measures would have been 4th(mean=0.5747). This paper describes both the 3 systems as they were submitted and the similarity measures that would obtained those better results.

CLaC-CORE: Exhaustive Feature Combination for Measuring Textual Similarity
Ehsan Shareghi and Sabine Bergler

CLaC-CORE, an exhaustive feature combination system ranked 4th among 34 teams in the Semantic Textual Similarity shared task STS 2013. Using a core set of 11 lexical features of the most basic kind, it uses a support vector regressor which uses a combination of these lexical features to train a model for predicting similarity between sentences in a two phase method, which in turn uses all combinations of the features in the feature space and trains separate models based on each combination. Then it creates a meta-feature space and trains a final model based on that. This two step process improves the results achieved by single-layer standard learning methodology over the same simple features. We analyze the correlation of feature combinations with the data sets over which they are effective.

UniMelb_NLP-CORE: Integrating predictions from multiple domains and feature sets for estimating semantic textual similarity
Spandana Gella, Bahar Salehi, Marco Lui, Karl Grieser, Paul Cook, and Timothy Baldwin

In this paper we present our systems for calculating the degree of semantic similarity between two texts that we submitted to the Semantic Textual Similarity task at SemEval-2013. Our systems predict similarity using a regression over features based on the following sources of information: string similarity, topic distributions of the texts based on latent Dirichlet allocation, and similarity between the documents returned by an information retrieval engine when the target texts are used as queries. We also explore methods for integrating predictions using different training datasets and feature sets. Our best system was ranked 17th out of 89 participating systems. In our post-task analysis, we identify simple changes to our system that further improve our results.

CFILT-CORE: Semantic Textual Similarity using Universal Networking Language
Avishek Dan and Pushpak Bhattacharyya

This paper describes the system that was submitted in the *SEM 2013 Semantic Textual Similarity shared task. The task aims to find the similarity score between a pair of sentences. We describe a Universal Networking Language (UNL) based semantic extraction system for measuring the semantic similarity. Our approach combines syntactic and word level similarity measures along with the UNL based semantic similarity measures for finding similarity scores between sentences.

CPN-CORE: A Text Semantic Similarity System Infused with Opinion Knowledge
Carmen Banea, Yoonjung Choi, Lingjia Deng, Samer Hassan, Michael Mohler, Bishan Yang, Claire Cardie, Rada Mihalcea, and Jan Wiebe

This article provides a detailed overview of the CPN text-to-text similarity system that we participated with in the Semantic Textual Similarity task evaluations hosted at *Sem 2013. In addition to more traditional components, such as knowledge-based and corpus-based metrics leveraged in a machine learning framework, we also use opinion analysis features to achieve a stronger semantic representation of textual units. While the evaluation datasets are not designed to test the similarity of opinions, as a component of textual similarity, nonetheless, our system variations ranked number 38, 39 and 45 among the 88 participating systems.

INAOE_UPV-CORE: Extracting Word Associations from Document Corpora to estimate Semantic Textual Similarity
Fernando Sánchez-Vega, Manuel Montes-y Gómez, Paolo Rosso, and Luis Villaseñor-Pineda

This paper presents three methods to evaluate the Semantic Textual Similarity (STS). The first two methods do not require labeled training data; instead, they automatically extract semantic knowledge in the form of word associations from a given reference corpus. Two kinds of word associations are considered: cooccurrence statistics and the similarity of word contexts. The third method was done in collaboration with LIPN and UMCC_DLSI. It uses several word similarity measures as features in order to construct an accurate prediction model for the STS.

CNGL-CORE: Referential Translation Machines for Measuring Semantic Similarity
Ergun Bicici and Josef Van Genabith

We invent referential translation machines (RTMs), a computational model for identifying the translation acts between any two data sets with respect to a reference corpus selected in the same domain, which can be used for judging the semantic similarity between text. RTMs make quality and semantic similarity judgments possible by using retrieved relevant training data as interpretants for reaching shared semantics. An MTPP (machine translation performance predictor) model derives features measuring the closeness of the test sentences to the training data, the difficulty of translating them, and the presence of acts of translation involved. We view semantic similarity as paraphrasing between any two given texts. Each view is modeled by an RTM model, giving us a new perspective on the binary relationship between the two. Our prediction model is the 15th on some tasks and 30th overall according to the official results of the Semantic Textual Similarity (STS 2013) challenge.

SemEval
Friday, June 14–Saturday, June 15

Venue: International E

Friday, June 14, 2013

8:00–08:30 **Registration**

08:30–08:40 **Opening remarks**

(08:30–09:30) **Session SE1: Session 1**

08:40–09:00 SemEval-2013 Task 1: TempEval-3: Evaluating Time Expressions, Events, and Temporal Relations
Naushad UzZaman, Hector Llorens, Leon Derczynski, James Allen, Marc Verhagen, and James Pustejovsky

09:00–09:20 ClearTK-TimeML: A minimalist approach to TempEval 2013
Steven Bethard

09:20–09:30 HeidelTime: Tuning English and Developing Spanish Resources for TempEval-3
Jannik Strötgen, Julian Zell, and Michael Gertz

09:30–10:30 **Session PLN2: Keynote address: David Forsyth**

10:30–11:00 **Coffee Break**

(11:00–12:30) **Session SE2: Session 2**

11:00–11:10 ATT1: Temporal Annotation Using Big Windows and Rich Syntactic and Semantic Features
Hyuckchul Jung and Amanda Stent

11:10–11:30 Semeval-2013 Task 8: Cross-lingual Textual Entailment for Content Synchronization
Matteo Negri, Alessandro Marchetti, Yashar Mehdad, Luisa Bentivogli, and Danilo Giampiccolo

11:30–11:50 SOFTCARDINALITY: Learning to Identify Directional Cross-Lingual Entailment from Cardinalities and SMT
Sergio Jimenez, Claudia Becerra, and Alexander Gelbukh

11:50–12:10 SemEval-2013 Task 5: Evaluating Phrasal Semantics
Ioannis Korkontzelos, Torsten Zesch, Fabio Massimo Zanzotto, and Chris Biemann

12:10–12:30 HsH: Estimating Semantic Similarity of Words and Short Phrases with Frequency Normalized Distance Measures
Christian Wartena

12:30–13:30 **Session SP1: Lunch Break + Poster Session 1 for Tasks 1, 5, 8**

- SemEval-2013 Task 1: TempEval-3: Evaluating Time Expressions, Events, and Temporal Relations
Naushad UzZaman, Hector Llorens, Leon Derczynski, James Allen, Marc Verhagen, and James Pustejovsky
- ClearTK-TimeML: A minimalist approach to TempEval 2013
Steven Bethard

- ManTIME: Temporal expression identification and normalization in the TempEval-3 challenge
  Michele Filannino, Gavin Brown, and Goran Nenadic
- HeidelTime: Tuning English and Developing Spanish Resources for TempEval-3
  Jannik Strötgen, Julian Zell, and Michael Gertz
- FSS-TimEx for TempEval-3: Extracting Temporal Information from Text
  Vanni Zavarella and Hristo Tanev
- ATT1: Temporal Annotation Using Big Windows and Rich Syntactic and Semantic Features
  Hyuckchul Jung and Amanda Stent
- JU_CSE: A CRF Based Approach to Annotation of Temporal Expression, Event and Temporal Relations
  Anup Kumar Kolya, Amitava Kundu, Rajdeep Gupta, Asif Ekbal, and Sivaji Bandyopadhyay
- NavyTime: Event and Time Ordering from Raw Text
  Nate Chambers
- SUTime: Evaluation in TempEval-3
  Angel Chang and Christopher D. Manning
- KUL: Data-driven Approach to Temporal Parsing of Newswire Articles
  Oleksandr Kolomiyets and Marie-Francine Moens
- UTTime: Temporal Relation Classification using Deep Syntactic Features
  Natsuda Laokulrat, Makoto Miwa, Yoshimasa Tsuruoka, and Takashi Chikayama
- SemEval-2013 Task 5: Evaluating Phrasal Semantics
  Ioannis Korkontzelos, Torsten Zesch, Fabio Massimo Zanzotto, and Chris Biemann
- HsH: Estimating Semantic Similarity of Words and Short Phrases with Frequency Normalized Distance Measures
  Christian Wartena
- UMCC_DLSI-(EPS): Paraphrases Detection Based on Semantic Distance
  Héctor Dávila, Antonio Fernández Orquín, Alexander Chávez, Yoan Gutiérrez, Armando Collazo, José I. Abreu, Andrés Montoyo, and Rafael Muñoz
- MELODI: Semantic Similarity of Words and Compositional Phrases using Latent Vector Weighting
  Tim Van de Cruys, Stergos Afantenos, and Philippe Muller
- IIRG: A Naive Approach to Evaluating Phrasal Semantics
  Lorna Byrne, Caroline Fenlon, and John Dunnion
- ClaC: Semantic Relatedness of Words and Phrases
  Reda Siblini and Leila Kosseim
- UNAL: Discriminating between Literal and Figurative Phrasal Usage Using Distributional Statistics and POS tags
  Sergio Jimenez, Claudia Becerra, and Alexander Gelbukh
- Semeval-2013 Task 8: Cross-lingual Textual Entailment for Content Synchronization
  Matteo Negri, Alessandro Marchetti, Yashar Mehdad, Luisa Bentivogli, and Danilo Giampiccolo
- ECNUCS: Recognizing Cross-lingual Textual Entailment Using Multiple Text Similarity and Text Difference Measures
  Jiang Zhao, Man Lan, and Zheng-Yu Niu
- BUAP: N-gram based Feature Evaluation for the Cross-Lingual Textual Entailment Task
  Darnes Vilariño, David Pinto, Saul León, Yuridiana Aleman, and Helena Gómez

- ALTN: Word Alignment Features for Cross-lingual Textual Entailment
  Marco Turchi and Matteo Negri
- SOFTCARDINALITY: Learning to Identify Directional Cross-Lingual Entailment from Cardinalities and SMT
  Sergio Jimenez, Claudia Becerra, and Alexander Gelbukh
- Umelb: Cross-lingual Textual Entailment with Word Alignment and String Similarity Features
  Yvette Graham, Bahar Salehi, and Timothy Baldwin

13:30–14:30 **Session PLN3: Joint Panel: Future of *SEM / STS Shared Task / SemEval**

(14:30–15:30) **Session SE3: Session 3**

14:30–14:50 UNAL: Discriminating between Literal and Figurative Phrasal Usage Using Distributional Statistics and POS tags
Sergio Jimenez, Claudia Becerra, and Alexander Gelbukh

14:50–15:10 SemEval-2013 Task 4: Free Paraphrases of Noun Compounds
Iris Hendrickx, Zornitsa Kozareva, Preslav Nakov, Diarmuid Ó Séaghdha, Stan Szpakowicz, and Tony Veale

15:10–15:30 MELODI: A Supervised Distributional Approach for Free Paraphrasing of Noun Compounds
Tim Van de Cruys, Stergos Afantenos, and Philippe Muller

(15:30–16:30) **Session SP2: Coffee Break + Poster Session 2 for Tasks 4, 10, 11, 12**

- SemEval-2013 Task 4: Free Paraphrases of Noun Compounds
  Iris Hendrickx, Zornitsa Kozareva, Preslav Nakov, Diarmuid Ó Séaghdha, Stan Szpakowicz, and Tony Veale
- SFS-TUE: Compound Paraphrasing with a Language Model and Discriminative Reranking
  Yannick Versley
- IIIT-H: A Corpus-Driven Co-occurrence Based Probabilistic Model for Noun Compound Paraphrasing
  Nitesh Surtani, Arpita Batra, Urmi Ghosh, and Soma Paul
- MELODI: A Supervised Distributional Approach for Free Paraphrasing of Noun Compounds
  Tim Van de Cruys, Stergos Afantenos, and Philippe Muller
- SemEval-2013 Task 10: Cross-lingual Word Sense Disambiguation
  Els Lefever and Véronique Hoste
- XLING: Matching Query Sentences to a Parallel Corpus using Topic Models for WSD
  Liling Tan and Francis Bond
- HLTDI: CL-WSD Using Markov Random Fields for SemEval-2013 Task 10
  Alex Rudnick, Can Liu, and Michael Gasser
- LIMSI : Cross-lingual Word Sense Disambiguation using Translation Sense Clustering
  Marianna Apidianaki
- WSD2: Parameter optimisation for Memory-based Cross-Lingual Word-Sense Disambiguation
  Maarten van Gompel and Antal van den Bosch
- NRC: A Machine Translation Approach to Cross-Lingual Word Sense Disambiguation (SemEval-2013 Task 10)
  Marine Carpuat
- SemEval-2013 Task 11: Word Sense Induction and Disambiguation within an End-User Application
  Roberto Navigli and Daniele Vannella

- Duluth : Word Sense Induction Applied to Web Page Clustering
  Ted Pedersen
- SATTY : Word Sense Induction Application in Web Search Clustering
  Satyabrata Behera, Upasana Gaikwad, Ramakrishna Bairi, and Ganesh Ramakrishnan
- UKP-WSI: UKP Lab Semeval-2013 Task 11 System Description
  Hans-Peter Zorn and Iryna Gurevych
- unimelb: Topic Modelling-based Word Sense Induction for Web Snippet Clustering
  Jey Han Lau, Paul Cook, and Timothy Baldwin
- SemEval-2013 Task 12: Multilingual Word Sense Disambiguation
  Roberto Navigli, David Jurgens, and Daniele Vannella
- GETALP System : Propagation of a Lesk Measure through an Ant Colony Algorithm
  Didier Schwab, Andon Tchechmedjiev, Jérôme Goulian, Mohammad Nasiruddin, Gilles Sérasset, and Hervé Blanchon
- UMCC_DLSI: Reinforcing a Ranking Algorithm with Sense Frequencies and Multidimensional Semantic Resources to solve Multilingual Word Sense Disambiguation
  Yoan Gutiérrez, Yenier Castañeda, Andy González, Rainel Estrada, Dennys D. Piug, Jose I. Abreu, Roger Pérez, Antonio Fernández Orquín, Andrés Montoyo, Rafael Muñoz, and Franc Camara
- DAEBAK!: Peripheral Diversity for Multilingual Word Sense Disambiguation
  Steve L. Manion and Raazesh Sainudiin

(16:30–18:30) **Session SE4: Session 4**

16:30–16:50 SemEval-2013 Task 10: Cross-lingual Word Sense Disambiguation
Els Lefever and Véronique Hoste

16:50–17:10 HLTDI: CL-WSD Using Markov Random Fields for SemEval-2013 Task 10
Alex Rudnick, Can Liu, and Michael Gasser

17:10–17:30 SemEval-2013 Task 11: Word Sense Induction and Disambiguation within an End-User Application
Roberto Navigli and Daniele Vannella

17:30–17:50 unimelb: Topic Modelling-based Word Sense Induction for Web Snippet Clustering
Jey Han Lau, Paul Cook, and Timothy Baldwin

17:50–18:10 SemEval-2013 Task 12: Multilingual Word Sense Disambiguation
Roberto Navigli, David Jurgens, and Daniele Vannella

18:10–18:20 UMCC_DLSI: Reinforcing a Ranking Algorithm with Sense Frequencies and Multidimensional Semantic Resources to solve Multilingual Word Sense Disambiguation
Yoan Gutiérrez, Yenier Castañeda, Andy González, Rainel Estrada, Dennys D. Piug, Jose I. Abreu, Roger Pérez, Antonio Fernández Orquín, Andrés Montoyo, Rafael Muñoz, and Franc Camara

18:20–18:30 DAEBAK!: Peripheral Diversity for Multilingual Word Sense Disambiguation
Steve L. Manion and Raazesh Sainudiin

Saturday, June 15, 2013

(08:40–10:30) **Session SE5: Session 5**

08:40–09:00 SemEval-2013 Task 3: Spatial Role Labeling
Oleksandr Kolomiyets, Parisa Kordjamshidi, Marie-Francine Moens, and Steven Bethard

09:00–09:20 SemEval-2013 Task 7: The Joint Student Response Analysis and 8th Recognizing Textual Entailment Challenge
Myroslava Dzikovska, Rodney Nielsen, Chris Brew, Claudia Leacock, Danilo Giampiccolo, Luisa Bentivogli, Peter Clark, Ido Dagan, and Hoa Trang Dang

09:20–09:35 ETS: Domain Adaptation and Stacking for Short Answer Scoring
Michael Heilman and Nitin Madnani

09:35–09:50 SOFTCARDINALITY: Hierarchical Text Overlap for Student Response Analysis
Sergio Jimenez, Claudia Becerra, and Alexander Gelbukh

09:50–10:00 UKP-BIU: Similarity and Entailment Metrics for Student Response Analysis
Omer Levy, Torsten Zesch, Ido Dagan, and Iryna Gurevych

10:00–10:20 SemEval-2013 Task 13: Word Sense Induction for Graded and Non-Graded Senses
David Jurgens and Ioannis Klapaftis

10:20–10:30 AI-KU: Using Substitute Vectors and Co-Occurrence Modeling For Word Sense Induction and Disambiguation
Osman Baskaya, Enis Sert, Volkan Cirik, and Deniz Yuret

10:30–11:00 **Coffee Break**

(11:00–13:10) **Session SE6: Session 6**

11:00–11:10 unimelb: Topic Modelling-based Word Sense Induction
Jey Han Lau, Paul Cook, and Timothy Baldwin

11:10–11:30 SemEval-2013 Task 2: Sentiment Analysis in Twitter
Preslav Nakov, Sara Rosenthal, Zornitsa Kozareva, Veselin Stoyanov, Alan Ritter, and Theresa Wilson

11:30–11:50 NRC-Canada: Building the State-of-the-Art in Sentiment Analysis of Tweets
Saif Mohammad, Svetlana Kiritchenko, and Xiaodan Zhu

11:50–12:00 GU-MLT-LT: Sentiment Analysis of Short Messages using Linguistic Features and Stochastic Gradient Descent
Tobias Günther and Lenz Furrer

12:00–12:10 AVAYA: Sentiment Analysis on Twitter with Self-Training and Polarity Lexicon Expansion
Lee Becker, George Erhart, David Skiba, and Valentine Matula

12:10–12:30 SemEval-2013 Task 9 : Extraction of Drug-Drug Interactions from Biomedical Texts (DDIExtraction 2013)
Isabel Segura-Bedmar, Paloma Martínez, and María Herrero Zazo

12:30–12:50 FBK-irst : A Multi-Phase Kernel Based Approach for Drug-Drug Interaction Detection and Classification that Exploits Linguistic Information
Md. Faisal Mahbub Chowdhury and Alberto Lavelli

12:50–13:10 WBI-NER: The impact of domain-specific features on the performance of identifying and classifying mentions of drugs
Tim Rocktäschel, Torsten Huber, Michael Weidlich, and Ulf Leser

13:10–15:30 **Session SP3: Lunch Break + Poster Session 3 for Tasks 2, 3, 7, 9, 13**

- SemEval-2013 Task 2: Sentiment Analysis in Twitter
  Preslav Nakov, Sara Rosenthal, Zornitsa Kozareva, Veselin Stoyanov, Alan Ritter, and Theresa Wilson
- AMI&ERIC: How to Learn with Naive Bayes and Prior Knowledge: an Application to Sentiment Analysis
  Mohamed Dermouche, Leila Khouas, Julien Velcin, and Sabine Loudcher
- UNITOR: Combining Syntactic and Semantic Kernels for Twitter Sentiment Analysis
  Giuseppe Castellucci, Simone Filice, Danilo Croce, and Roberto Basili
- GU-MLT-LT: Sentiment Analysis of Short Messages using Linguistic Features and Stochastic Gradient Descent
  Tobias Günther and Lenz Furrer
- AVAYA: Sentiment Analysis on Twitter with Self-Training and Polarity Lexicon Expansion
  Lee Becker, George Erhart, David Skiba, and Valentine Matula
- TJP: Using Twitter to Analyze the Polarity of Contexts
  Tawunrat Chalothorn and Jeremy Ellman
- uOttawa: System description for SemEval 2013 Task 2 Sentiment Analysis in Twitter
  Hamid Poursepanj, Josh Weissbock, and Diana Inkpen
- UT-DB: An Experimental Study on Sentiment Analysis in Twitter
  Zhemin Zhu, Djoerd Hiemstra, Peter Apers, and Andreas Wombacher
- USNA: A Dual-Classifier Approach to Contextual Sentiment Analysis
  Ganesh Harihara, Eugene Yang, and Nate Chambers
- KLUE: Simple and robust methods for polarity classification
  Thomas Proisl, Paul Greiner, Stefan Evert, and Besim Kabashi
- SINAI: Machine Learning and Emotion of the Crowd for Sentiment Analysis in Microblogs
  Eugenio Martínez-Cámara, Arturo Montejo-Ráez, M. Teresa Martín-Valdivia, and L. Alfonso Ureña-López
- ECNUCS: A Surface Information Based System Description of Sentiment Analysis in Twitter in the SemEval-2013 (Task 2)
  Zhu Tiantian, Zhang Fangxi, and Man Lan
- Umigon: sentiment analysis for tweets based on terms lists and heuristics
  Clement Levallois
- [LVIC-LIMSI]: Using Syntactic Features and Multi-polarity Words for Sentiment Analysis in Twitter
  Morgane Marchand, Alexandru Ginsca, Romaric Besançon, and Olivier Mesnard
- SwatCS: Combining simple classifiers with estimated accuracy
  Sam Clark and Rich Wicentwoski
- NTNU: Domain Semi-Independent Short Message Sentiment Classification
  Øyvind Selmer, Mikael Brevik, Björn Gambäck, and Lars Bungum
- SAIL: A hybrid approach to sentiment analysis
  Nikolaos Malandrakis, Abe Kazemzadeh, Alexandros Potamianos, and Shrikanth Narayanan
- UMCC_DLSI-(SA): Using a ranking algorithm and informal features to solve Sentiment Analysis in Twitter
  Yoan Gutiérrez, Andy González, Roger Pérez, José I. Abreu, Antonio Fernández Orquín, Alejandro Mosquera, Andrés Montoyo, Rafael Muñoz, and Franc Camara
- ASVUniOfLeipzig: Sentiment Analysis in Twitter using Data-driven Machine Learning Techniques
  Robert Remus

- Experiments with DBpedia, WordNet and SentiWordNet as resources for sentiment analysis in micro-blogging
  Hussam Hamdan, Frederic Béchet, and Patrice Bellot
- OPTWIMA: Comparing Knowledge-rich and Knowledge-poor Approaches for Sentiment Analysis in Short Informal Texts
  Alexandra Balahur
- FBK: Sentiment Analysis in Twitter with Tweetsted
  Md. Faisal Mahbub Chowdhury, Marco Guerini, Sara Tonelli, and Alberto Lavelli
- SU-Sentilab : A Classification System for Sentiment Analysis in Twitter
  Gizem Gezici, Rahim Dehkharghani, Berrin Yanikoglu, Dilek Tapucu, and Yucel Saygin
- Columbia NLP: Sentiment Detection of Subjective Phrases in Social Media
  Sara Rosenthal and Kathy McKeown
- FBM: Combining lexicon-based ML and heuristics for Social Media Polarities
  Carlos Rodriguez-Penagos, Jordi Atserias Batalla, Joan Codina-Filbà, David García-Narbona, Jens Grivolla, Patrik Lambert, and Roser Saurí
- REACTION: A naive machine learning approach for sentiment classification
  Silvio Moreira, João Filgueiras, Bruno Martins, Francisco Couto, and Mário J. Silva
- IITB-Sentiment-Analysts: Participation in Sentiment Analysis in Twitter SemEval 2013 Task
  Karan Chawla, Ankit Ramteke, and Pushpak Bhattacharyya
- SSA-UO: Unsupervised Sentiment Analysis in Twitter
  Reynier Ortega Bueno, Adrian Fonseca Bruzón, Yoan Gutiérrez, and Andres Montoyo
- senti.ue-en: an approach for informally written short texts in SemEval-2013 Sentiment Analysis task
  José Saias and Hilário Fernandes
- teragram: Rule-based detection of sentiment phrases using SAS Sentiment Analysis
  Hilke Reckman, Cheyanne Baird, Jean Crawford, Richard Crowell, Linnea Micciulla, Saratendu Sethi, and Fruzsina Veress
- CodeX: Combining an SVM Classifier and Character N-gram Language Models for Sentiment Analysis on Twitter Text
  Qi Han, Junfei Guo, and Hinrich Schuetze
- sielers : Feature Analysis and Polarity Classification of Expressions from Twitter and SMS Data
  Harshit Jain, Aditya Mogadala, and Vasudeva Varma
- Kea: Expression-level Sentiment Analysis from Twitter Data
  Ameeta Agrawal and Aijun An
- NRC-Canada: Building the State-of-the-Art in Sentiment Analysis of Tweets
  Saif Mohammad, Svetlana Kiritchenko, and Xiaodan Zhu
- UoM: Using Explicit Semantic Analysis for Classifying Sentiments
  Sapna Negi and Michael Rosner
- bwbaugh : Hierarchical sentiment analysis with partial self-training
  Wesley Baugh
- Serendio: Simple and Practical lexicon based approach to Sentiment Analysis
  Prabu palanisamy, Vineet Yadav, and Harsha Elchuri

- SZTE-NLP: Sentiment Detection on Twitter Messages
  Viktor Hangya, Gabor Berend, and Richárd Farkas
- BOUNCE: Sentiment Classification in Twitter using Rich Feature Sets
  Nadin Kökciyan, Arda Çelebi, Arzucan Özgür, and Suzan Üsküdarlı
- nlp.cs.aueb.gr: Two Stage Sentiment Analysis
  Prodromos Malakasiotis, Rafael Michael Karampatsis, Konstantina Makrynioti, and John Pavlopoulos
- NILC_USP: A Hybrid System for Sentiment Analysis in Twitter Messages
  Pedro Balage Filho and Thiago Pardo
- SemEval-2013 Task 3: Spatial Role Labeling
  Oleksandr Kolomiyets, Parisa Kordjamshidi, Marie-Francine Moens, and Steven Bethard
- UNITOR-HMM-TK: Structured Kernel-based learning for Spatial Role Labeling
  Emanuele Bastianelli, Danilo Croce, Roberto Basili, and Daniele Nardi
- SemEval-2013 Task 7: The Joint Student Response Analysis and 8th Recognizing Textual Entailment Challenge
  Myroslava Dzikovska, Rodney Nielsen, Chris Brew, Claudia Leacock, Danilo Giampiccolo, Luisa Bentivogli, Peter Clark, Ido Dagan, and Hoa Trang Dang
- UKP-BIU: Similarity and Entailment Metrics for Student Response Analysis
  Omer Levy, Torsten Zesch, Ido Dagan, and Iryna Gurevych
- ETS: Domain Adaptation and Stacking for Short Answer Scoring
  Michael Heilman and Nitin Madnani
- EHU-ALM: Similarity-Feature Based Approach for Student Response Analysis
  Itziar Aldabe, Montse Maritxalar, and Oier Lopez de Lacalle
- CNGL: Grading Student Answers by Acts of Translation
  Ergun Bicici and Josef van Genabith
- Celi: EDITS and Generic Text Pair Classification
  Milen Kouylekov, Luca Dini, Alessio Bosca, and Marco Trevisan
- LIMSIILES: Basic English Substitution for Student Answer Assessment at SemEval 2013
  Martin Gleize and Brigitte Grau
- SOFTCARDINALITY: Hierarchical Text Overlap for Student Response Analysis
  Sergio Jimenez, Claudia Becerra, and Alexander Gelbukh
- CU : Computational Assessment of Short Free Text Answers - A Tool for Evaluating Students' Understanding
  IFEYINWA OKOYE, Steven Bethard, and Tamara Sumner
- CoMeT: Integrating different levels of linguistic modeling for meaning assessment
  Niels Ott, Ramon Ziai, Michael Hahn, and Detmar Meurers
- SemEval-2013 Task 9 : Extraction of Drug-Drug Interactions from Biomedical Texts (DDIExtraction 2013)
  Isabel Segura-Bedmar, Paloma Martínez, and María Herrero Zazo
- UC3M: A kernel-based approach to identify and classify DDIs in bio-medical texts.
  Daniel Sanchez-Cisneros
- UEM-UC3M: An Ontology-based named entity recognition system for biomedical texts.
  Daniel Sanchez-Cisneros and Fernando Aparicio Gali

- FBK-irst : A Multi-Phase Kernel Based Approach for Drug-Drug Interaction Detection and Classification that Exploits Linguistic Information
  Md. Faisal Mahbub Chowdhury and Alberto Lavelli
- WBI-DDI: Drug-Drug Interaction Extraction using Majority Voting
  Philippe Thomas, Mariana Neves, Tim Rocktäschel, and Ulf Leser
- WBI-NER: The impact of domain-specific features on the performance of identifying and classifying mentions of drugs
  Tim Rocktäschel, Torsten Huber, Michael Weidlich, and Ulf Leser
- UMCC_DLSI: Semantic and Lexical features for detection and classification Drugs in biomedical texts
  Armando Collazo, Alberto Ceballo, Dennys D. Puig, Yoan Gutiérrez, José I. Abreu, Roger Pérez, Antonio Fernández Orquín, Andrés Montoyo, Rafael Muñoz, and Franc Camara
- NIL_UCM: Extracting Drug-Drug interactions from text through combination of sequence and tree kernels
  Behrouz Bokharaeian and ALBERTO DIAZ
- UTurku: Drug Named Entity Recognition and Drug-Drug Interaction Extraction Using SVM Classification and Domain Knowledge
  Jari Björne, Suwisa Kaewphan, and Tapio Salakoski
- LASIGE: using Conditional Random Fields and ChEBI ontology
  Tiago Grego, Francisco Pinto, and Francisco M Couto
- UWM-TRIADS: Classifying Drug-Drug Interactions with Two-Stage SVM and Post-Processing
  Majid Rastegar-Mojarad, Richard D. Boyce, and Rashmi Prasad
- SCAI: Extracting drug-drug interactions using a rich feature vector
  Tamara Bobic, Juliane Fluck, and Martin Hofmann-Apitius
- UColorado_SOM: Extraction of Drug-Drug Interactions from Biomedical Text using Knowledge-rich and Knowledge-poor Features
  Negacy Hailu, Lawrence E. Hunter, and K. Bretonnel Cohen
- SemEval-2013 Task 13: Word Sense Induction for Graded and Non-Graded Senses
  David Jurgens and Ioannis Klapaftis
- UoS: A Graph-Based System for Graded Word Sense Induction
  David Hope and Bill Keller
- AI-KU: Using Substitute Vectors and Co-Occurrence Modeling For Word Sense Induction and Disambiguation
  Osman Baskaya, Enis Sert, Volkan Cirik, and Deniz Yuret
- unimelb: Topic Modelling-based Word Sense Induction
  Jey Han Lau, Paul Cook, and Timothy Baldwin

6

Workshops: Thursday–Friday, June 13–14

Thursday, June 13

Room	Name	Page
Vinings 2	Student Research Workshop	104
International G	The 8th Workshop on Innovative Use of NLP for Building Educational Applications (BEA8)	105
International C	Seventh Workshop on Syntax, Semantics and Structure in Statistical Translation (SSST-7)	108
International A	First workshop on Metaphor in NLP (Meta4NLP)	109
International B	9th Workshop on Multiword Expressions (MWE)	111
International H	Language Analysis in Social Media (LASM)	113

Friday, June 14

International C	Events: Definition, Detection, Coreference, and Representation	114
International H	Workshop on Vision and Natural Language Processing	115
International G	Computational Linguistics for Literature	117
International F	Natural Language Processing for Improving Textual Accessibility (NLP4ITA)	118
International B	9th Workshop on Multiword Expressions (continued)	111
International A	4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA 2013)	119

Student Research Workshop
Thursday, June 13

Venue: Vinings 2

9:00–9:15 **Opening remarks**

Session 1: Research paper presentations

9:15–9:30 Critical Reflections on Evaluation Practices in Coreference Resolution
Gordana Ilic Holen

9:30–9:45 Reducing Annotation Effort on Unbalanced Corpus based on Cost Matrix
Wencan Luo, Diane Litman, and Joel Chan

9:45–10:00 A Machine Learning Approach to Automatic Term Extraction using a Rich Feature Set
Merley Conrado, Thiago Pardo, and Solange Rezende

10:00–10:15 A Rule-based Approach for Karmina Generation
Franky Franky

10:15–10:30 From Language to Family and Back: Native Language and Language Family Identification from English Text
Ariel Stolerman, Aylin Caliskan, and Rachel Greenstadt

10:30–11:00 **Coffee break**

Session 2: Research paper presentations

11:00–11:15 Ontology Label Translation
Mihael Arcan and Paul Buitelaar

11:15–11:30 Reversing Morphological Tokenization in English-to-Arabic SMT
Mohammad Salameh, Colin Cherry, and Grzegorz Kondrak

Session 3: Thesis proposal presentations

11:30–12:00 Statistical Machine Translation in Low Resource Settings
Ann Irvine

12:00–12:30 Large-Scale Paraphrasing for Natural Language Understanding
Juri Ganitkevitch

12:30–14:00 **Lunch**

Session 4: Thesis proposal presentations

14:00–14:30 Domain-Independent Captioning of Domain-Specific Images
Rebecca Mason

14:30–15:00 Helpfulness-Guided Review Summarization
Wenting Xiong

15:00–15:30 Entrainment in Spoken Dialogue Systems: Adopting, Predicting and Influencing User Behavior
Rivka Levitan

15:30–16:00 **Coffee break**

Session 5: Thesis proposal presentation

16:00–16:30 User Goal Change Model for Spoken Dialog State Tracking
Yi Ma

16:30–17:30 **Panel**

Workshop 1:

The 8th Workshop on Innovative Use of NLP for Building Educational Applications (BEA8)

Thursday, June 13

Venue: International G

8:45–9:00 **Load Presentations**

9:00–9:15 **Opening Remarks**

Session 1

9:15–9:40 The Utility of Manual and Automatic Linguistic Error Codes for Identifying Neurodevelopmental Disorders
Eric Morley, Brian Roark, and Jan van Santen

9:40–10:05 Shallow Semantic Analysis of Interactive Learner Sentences
Levi King and Markus Dickinson

10:05–10:30 Building a Large Annotated Corpus of Learner English: The NUS Corpus of Learner English
Daniel Dahlmeier, Hwee Tou Ng, and Siew Mei Wu

10:30–11:00 **Break**

Session 2

11:00–11:25 Developing and testing a self-assessment and tutoring system
Øistein E. Andersen, Helen Yannakoudakis, Fiona Barker, and Tim Parish

11:25–11:45 Automated Essay Scoring for Swedish
Robert Östling, André Smolentzov, Björn Tyrefors Hinnerich, and Erik Höglin

11:45–12:10 A Report on the First Native Language Identification Shared Task
Joel Tetreault, Daniel Blanchard, and Aoife Cahill

12:10–1:50 **Lunch**

1:50–2:40 **BEA8 Poster Session A**

- Applying Unsupervised Learning To Support Vector Space Model Based Speaking Assessment
  Lei Chen
- Role of Morpho-Syntactic Features in Estonian Proficiency Classification
  Sowmya Vajjala and Kaidi Loo
- Automated Content Scoring of Spoken Responses in an Assessment for Teachers of English
  Klaus Zechner and Xinhao Wang

1:50–2:40 **NLI 2013 Poster Session A**

- Experimental Results on the Native Language Identification Shared Task
  Amjad Abu-Jbara, Rahul Jha, Eric Morley, and Dragomir Radev
- VTEX System Description for the NLI 2013 Shared Task
  Vidas Daudaravicius
- Feature Space Selection and Combination for Native Language Identification
  Cyril Goutte, Serge Léger, and Marine Carpuat
- Discriminating Non-Native English with 350 Words
  John Henderson, Guido Zarrella, Craig Pfeifer, and John D. Burger

- Maximizing Classification Accuracy in Native Language Identification
  Scott Jarvis, Yves Bestgen, and Steve Pepper
- Recognizing English Learners' Native Language from Their Writings
  Baoli LI
- NLI Shared Task 2013: MQ Submission
  Shervin Malmasi, Sze-Meng Jojo Wong, and Mark Dras
- NAIST at the NLI 2013 Shared Task
  Tomoya Mizumoto, Yuta Hayashibe, Keisuke Sakaguchi, Mamoru Komachi, and Yuji Matsumoto
- Cognate and Misspelling Features for Natural Language Identification
  Garrett Nicolai, Bradley Hauer, Mohammad Salameh, Lei Yao, and Grzegorz Kondrak
- Exploring Syntactic Representations for Native Language Identification
  Ben Swanson
- Simple Yet Powerful Native Language Identification on TOEFL11
  Ching-Yi Wu, Po-Hsiang Lai, Yang Liu, and Vincent Ng

2:40–3:30 **BEA8 Poster Session B**

- Prompt-based Content Scoring for Automated Spoken Language Assessment
  Keelan Evanini, Shasha Xie, and Klaus Zechner
- Automated Scoring of a Summary-Writing Task Designed to Measure Reading Comprehension
  Nitin Madnani, Jill Burstein, John Sabatini, and Tenaha O'Reilly
- Inter-annotator Agreement for Dependency Annotation of Learner Language
  Marwa Ragheb and Markus Dickinson

2:40–3:30 **NLI 2013 Poster Session B**

- Native Language Identification with PPM
  Victoria Bobicev
- Using Other Learner Corpora in the 2013 NLI Shared Task
  Julian Brooke and Graeme Hirst
- Combining Shallow and Linguistically Motivated Features in Native Language Identification
  Serhiy Bykh, Sowmya Vajjala, Julia Krivanek, and Detmar Meurers
- Linguistic Profiling based on General–purpose Features and Native Language Identification
  Andrea Cimino, Felice Dell'Orletta, Giulia Venturi, and Simonetta Montemagni
- Improving Native Language Identification with TF-IDF Weighting
  Binyam Gebrekidan Gebre, Marcos Zampieri, Peter Wittenburg, and Tom Heskes
- Native Language Identification: a Simple n-gram Based Approach
  Binod Gyawali, Gabriela Ramirez, and Thamar Solorio
- Feature Engineering in the NLI Shared Task 2013: Charles University Submission Report
  Barbora Hladka, Martin Holub, and Vincent Kriz
- Native Language Identification: A Key N-gram Category Approach
  Kristopher Kyle, Scott Crossley, Jianmin Dai, and Danielle McNamara
- Using N-gram and Word Network Features for Native Language Identification
  Shibamouli Lahiri and Rada Mihalcea
- LIMSI's participation to the 2013 shared task on Native Language Identification
  Thomas Lavergne, Gabriel Illouz, Aurélien Max, and Ryo Nagata

- Native Language Identification using large scale lexical features
  André Lynum
- The Story of the Characters, the DNA and the Native Language
  Marius Popescu and Radu Tudor Ionescu
- Identifying the L1 of non-native writers: the CMU-Haifa system
  Yulia Tsvetkov, Naama Twitto, Nathan Schneider, Noam Ordan, Manaal Faruqui, Victor Chahuneau, Shuly Wintner, and Chris Dyer

3:30–4:00 **Break**

Session 3

4:00–4:20 Evaluating Unsupervised Language Model Adaptation Methods for Speaking Assessment
Shasha Xie and Lei Chen

4:20–4:40 Improving interpretation robustness in a tutorial dialogue system
Myroslava Dzikovska, Elaine Farrow, and Johanna Moore

4:40–5:00 Detecting Missing Hyphens in Learner Text
Aoife Cahill, Martin Chodorow, Susanne Wolff, and Nitin Madnani

5:00–5:20 Applying Machine Translation Metrics to Student-Written Translations
Lisa Michaud and Patricia Ann McCoy

5:20–5:30 **Closing Remarks**

Workshop 2:

Seventh Workshop on Syntax, Semantics and Structure in Statistical Translation (SSST-7)

Thursday, June 13

Venue: International C

9:15–9:30 **Opening Remarks**

Session 1

9:30–10:00 A Semantic Evaluation of Machine Translation Lexical Choice
Marine Carpuat

10:00–10:30 Taste of Two Different Flavours: Which Manipuri Script works better for English-Manipuri Language pair SMT Systems?
Thoudam Doren Singh

10:30–11:00 **Break**

Session 2

11:00–11:30 Hierarchical Alignment Decomposition Labels for Hiero Grammar Rules
Gideon Maillette de Buy Wenniger and Khalil Sima'an

11:30–12:00 A Performance Study of Cube Pruning for Large-Scale Hierarchical Machine Translation
Matthias Huck, David Vilar, Markus Freitag, and Hermann Ney

12:00–12:30 Combining Word Reordering Methods on different Linguistic Abstraction Levels for Statistical Machine Translation
Teresa Herrmann, Jan Niehues, and Alex Waibel

12:30–2:00 **Lunch**

Session 3

2:00–3:00 **Panel discussion: Meaning Representations for Machine Translation, with Jan Hajic, Kevin Knight, Martha Palmer and Dekai Wu**

3:30–4:00 Combining Top-down and Bottom-up Search for Unsupervised Induction of Transduction Grammars
Markus Saers, Karteek Addanki, and Dekai Wu

3:30–4:00 **Break**

Session 4

4:00–4:30 A Formal Characterization of Parsing Word Alignments by Synchronous Grammars with Empirical Evidence to the ITG Hypothesis.
Gideon Maillette de Buy Wenniger and Khalil Sima'an

4:30–5:00 Synchronous Linear Context-Free Rewriting Systems for Machine Translation
Miriam Kaeshammer

Workshop 3:
First workshop on Metaphor in NLP
Thursday, June 13

Venue: International A

9:00–9:10 **Opening remarks**

9:10–10:05 **Invited talk: Srini Narayanan "From Metaphor to Action"**

Session 1

10:05–10:30 What metaphor identification systems can tell us about metaphor-in-language
Jonathan Dunn

10:30–11:00 **Coffee break**

Session 2

11:00–11:25 Argumentation-Relevant Metaphors in Test-Taker Essays
Beata Beigman Klebanov and Michael Flor

11:25–11:45 Relational words have high metaphoric potential
Anja Jamrozik, Eyal Sagi, Micah Goldwater, and Dedre Gentner

11:45–12:10 Semantic Signatures for Example-Based Linguistic Metaphor Detection
Michael Mohler, David Bracewell, Marc Tomlinson, and David Hinote

12:10–1:40 **Lunch**

1:40–2:20 **Invited talk: John Barnden "Computational Approaches to Metaphor Interpretation: Some Considerations arising from a Deep Reasoning System"**

Session 3

2:20–2:45 Automatic Metaphor Detection using Large-Scale Lexical Resources and Conventional Metaphor Extraction
Yorick Wilks, Adam Dalton, James Allen, and Lucian Galescu

2:45–3:10 Cross-Lingual Metaphor Detection Using Common Semantic Features
Yulia Tsvetkov, Elena Mukomel, and Anatole Gershman

3:10–3:30 Identifying Metaphorical Word Use with Tree Kernels
Dirk Hovy, Shashank Shrivastava, Sujay Kumar Jauhar, Mrinmaya Sachan, Kartik Goyal, Huying Li, Whitney Sanders, and Eduard Hovy

3:30–4:00 **Coffee break**

Session 4

4:00–4:25 Automatic Extraction of Linguistic Metaphors with LDA Topic Modeling
Ilana Heintz, Ryan Gabbard, Mahesh Srivastava, Dave Barner, Donald Black, Majorie Friedman, and Ralph Weischedel

4:25–4:50 Robust Extraction of Metaphor from Novel Data
Tomek Strzalkowski, George Aaron Broadwell, Sarah Taylor, Laurie Feldman, Samira Shaikh, Ting Liu, Boris Yamrom, Kit Cho, Umit Boz, Ignacio Cases, and Kyle Elliot

4:50–5:15 Annotating a Russian corpus of conceptual metaphor: a bottom-up approach
Yulia Badryzlova, Natalia Shekhtman, Yekaterina Isaeva, and Ruslan Kerimov

5:15–5:30 **Closing remarks**

Workshop 4:
9th workshop on Multiword Expressions (MWE)
Thursday, June 13–Friday, June 14

Venue: International B

09:00–09:15 **Opening Remarks**

Oral Session 1: Resources and Applications

09:15–09:40 Managing Multiword Expressions in a Lexicon-Based Sentiment Analysis System for Spanish
Antonio Moreno-Ortiz, Chantal Perez-Hernandez, and Maria Del-Olmo

09:40–10:05 Introducing PersPred, a Syntactic and Semantic Database for Persian Complex Predicates
Pollet Samvelian and Pegah Faghiri

10:05–10:30 Improving Word Translation Disambiguation by Capturing Multiword Expressions with Dictionaries
Lars Bungum, Björn Gambäck, André Lynum, and Erwin Marsi

10:30–11:00 **COFFEE BREAK**

Invited Talk 1

11:00–12:00 Complex Predicates are Multi-Word Expressions
Martha Palmer

Oral Session 2: Compositionality

12:00–12:25 The (Un)expected Effects of Applying Standard Cleansing Models to Human Ratings on Compositionality
Stephen Roller, Sabine Schulte im Walde, and Silke Scheible

12:30-14:00 **LUNCH BREAK**

Oral Session 2: Compositionality (continued)

14:05–14:30 Determining Compositionality of Word Expressions Using Word Space Models
Lubomír Krčmář, Karel Ježek, and Pavel Pecina

Invited Talk 2

14:30–15:30 Modelling the Internal Variability of MWEs
Malvina Nissim

15:30–16:00 **COFFEE BREAK**

Oral Session 3: Short Papers

16:00–16:15 Automatically Assessing Whether a Text Is Cliched, with Applications to Literary Analysis
Paul Cook and Graeme Hirst

16:15–16:30 An Analysis of Annotation of Verb-Noun Idiomatic Combinations in a Parallel Dependency Corpus
Zdenka Uresova, Jan Hajic, Eva Fucikova, and Jana Sindlerova

16:30–16:40 **Poster Boosters**

- Automatic Identification of Bengali Noun-Noun Compounds Using Random Forest
Vivekananda Gayen and Kamal Sarkar

- Automatic Detection of Stable Grammatical Features in N-Grams
  Mikhail Kopotev, Lidia Pivovarova, Natalia Kochetkova, and Roman Yangarber
- Exploring MWEs for Knowledge Acquisition from Corporate Technical Documents
  Bell Manrique-Losada, Carlos M. Zapata-Jaramillo, and Diego A. Burgos
- MWE in Portuguese: Proposal for a Typology for Annotation in Running Text
  Sandra Antunes and Amália Mendes
- Identifying Pronominal Verbs: Towards Automatic Disambiguation of the Clitic 'se' in Portuguese
  Magali Sanches Duran, Carolina Evaristo Scarton, Sandra Maria Aluísio, and Carlos Ramisch
- A Repository of Variation Patterns for Multiword Expressions
  Malvina Nissim and Andrea Zaninello

16:30–17:40 **Poster Session**

Friday, June 14, 2013

Oral Session 4: Identification and Classification

09:10–09:35 Syntactic Identification of Occurrences of Multiword Expressions in Text using a Lexicon with Dependency Structures
Eduard Bejček, Pavel Stranak, and Pavel Pecina

09:35–10:00 Combining Different Features of Idiomaticity for the Automatic Classification of Noun+Verb Expressions in Basque
Antton Gurrutxaga and Iñaki Alegria

Oral Session 5: Short Papers

10:00–10:15 Semantic Roles for Nominal Predicates: Building a Lexical Resource
Ashwini Vaidya, Martha Palmer, and Bhuvana Narasimhan

10:15–10:30 Constructional Intensifying Adjectives in Italian
Sara Berlanda

10:30–11:00 **COFFEE BREAK**

Invited Talk 3

11:00–12:00 The Far Reach of Multiword Expressions in Educational Technology
Jill Burstein

Oral Session 5: Short Papers (continued)

12:00–12:15 Construction of English MWE Dictionary and its Application to POS Tagging
Yutaro Shigeto, Ai Azuma, Sorami Hisamoto, Shuhei Kondo, Tomoya Kouse, Keisuke Sakaguchi, Akifumi Yoshimoto, Frances Yung, and Yuji Matsumoto

12:15–12:30 **Closing Remarks**

Workshop 5:
Language Analysis in Social Media (LASM)
Thursday, June 13

Venue: International H

9:00–9:15 **Introductions**

9:15–10:30 **Invited Key Note, Prof. Mor Naaman**

10:30–11:00 **Coffee Break**

Session 1

11:00–11:30 Does Size Matter? Text and Grammar Revision for Parsing Social Media Data
Mohammad Khan, Markus Dickinson, and Sandra Kuebler

11:30–12:00 Phonological Factors in Social Media Writing
Jacob Eisenstein

12:00–12:30 A Preliminary Study of Tweet Summarization using Information Extraction
Wei Xu, Ralph Grishman, Adam Meyers, and Alan Ritter

12:30–2:00 **Lunch**

Session 2

2:00–2:30 Really? Well. Apparently Bootstrapping Improves the Performance of Sarcasm and Nastiness Classifiers for Online Dialogue
Stephanie Lukin and Marilyn Walker

2:30–3:00 Topical Positioning: A New Method for Predicting Opinion Changes in Conversation
Ching-Sheng Lin, Samira Shaikh, Jennifer Stromer-Galley, Jennifer Crowley, Tomek Strzalkowski, and Veena Ravishankar

3:00–3:30 Sentiment Analysis of Political Tweets: Towards an Accurate Classifier
Akshat Bakliwal, Jennifer Foster, Jennifer van der Puil, Ron O'Brien, Lamia Tounsi, and Mark Hughes

3:30–3:45 **Coffee Break**

Session 3

3:45–4:15 A Case Study of Sockpuppet Detection in Wikipedia
Thamar Solorio, Ragib Hasan, and Mainul Mizan

4:15–4:45 Towards the Detection of Reliable Food-Health Relationships
Michael Wiegand and Dietrich Klakow

4:45–5:15 Translating Government Agencies' Tweet Feeds: Specificities, Problems and (a few) Solutions
Fabrizio Gotti, Philippe Langlais, and Atefeh Farzindar

5:15–5:30 **Closing Remarks**

Workshop 6:
Events: Definition, Detection, Coreference, and Representation (EVENTS)

Friday, June 14

Venue: International C

9:00–9:15 **Welcome**

9:15–9:30 **Working Session Instructions**

9:30–10:30 **Invited Talk: The Role of Event-based Representations and Reasoning in Language, James Pustejovsky**

10:30–11:00 **Break**

11:00–12:00 **Working Session I: What are events?**

12:00–1:00 **Poster Session**

- Coping With Implicit Arguments And Events Coreference
  Rodolfo Delmonte
- GAF: A Grounded Annotation Framework for Events
  Antske Fokkens, Marieke van Erp, Piek Vossen, Sara Tonelli, Willem Robert van Hage, Luciano Serafini, Rachele Sprugnoli, and Jesper Hoeksema
- Events are Not Simple: Identity, Non-Identity, and Quasi-Identity
  Eduard Hovy, Teruko Mitamura, Felisa Verdejo, Jun Araki, and Andrew Philpot
- Event representation across genre
  Lidia Pivovarova, Silja Huttunen, and Roman Yangarber
- A Semantic Tool for Historical Events
  Ryan Shaw
- Annotating Change of State for Clinical Events
  Lucy Vanderwende, Fei Xia, and Meliha Yetisgen-Yildiz

2:00–3:30 **Working Session II: When are two events the same? What relations are between events?**

3:30–4:00 **Break**

4:00–5:30 **Working Session III: How best to represent events? What aspects to annotate?**

5:30–6:00 **General Discussion**

Workshop 7:
Workshop on Vision and Natural Language Processing (WVL)
Friday, June 14

Venue: International H

Session 1

8:45–9:00 **Opening Remarks**

9:00–10:00 **Tutorial: Computational Visual Recognition for NLP (Alexander C Berg; Stony Brook University)**

10:00–10:30 **Invited talk: Modality Selection for Multimedia Summarization (Florian Metze; Carnegie Mellon University)**

10:30–11:00 **Coffee break**

Session 2

11:00–11:20 Annotation of Online Shopping Images without Labeled Training Examples
Rebecca Mason and Eugene Charniak

11:20–11:40 Generating Natural-Language Video Descriptions Using Text-Mined Knowledge
Niveda Krishnamoorthy, Girish Malkarnenkar, Raymond Mooney, Kate Saenko, and Sergio Guadarrama

11:40–12:00 Learning Hierarchical Linguistic Descriptions of Visual Datasets
Roni Mittelman, Min Sun, Benjamin Kuipers, and Silvio Savarese

12:00–12:30 **Invited talk: Joint Learning of Word Meanings and Image Tasks (Jason Weston, Google)**

12:30–2:00 **Lunch Break**

Session 3

2:00–2:15 **Invited student talk: Communicating with an Image Retrieval System via Relative Attributes (Adriana Kovashka and Kristen Grauman, University of Texas at Austin)**

2:15–2:30 **Invited student talk: Identifying Visual Attributes for Object Recognition (Caglar Tirkaz, Jacob Eisenstein, Berrin Yanikoglu and Metin Sezgin; Sabanci University, Georgia Institute of Technology, Koc University)**

2:30–2:45 **Invited student talk: Generating Visual Descriptions from Feature Norms of Actions, Attributes, Classes and Parts (Mark Yatskar and Luke Zettlemoyer; University of Washington)**

2:45–3:00 **Invited student talk: Bayesian modeling of scenes and captions (Luca del Pero and Kobus Barnard; University of Arizona)**

3:00–3:15 **Invited student talk: Data-Driven Generation of Image Descriptions (Vicente Ordonez and Tamara Berg; Stony Brook University)**

3:15–3:30 **Invited student talk: Framing image description as a retrieval problem (Micah Hodosh, Peter Young and Julia Hockenmaier; University of Illinois at Urbana-Champaign)**

Session 4

4:00–4:30 **Invited talk: Multimodal Semantics at CLIC (Elia Bruni; University of Trento)**

4:30–5:00 **Invited talk: Generating and Generalizing Image Captions (Yejin Choi; Stony Brook University)**

5:00–5:30 **Invited talk: Generating Descriptions of Visible Objects (Margaret Mitchell; Johns Hopkins University)**

5:30–6:00 **Panel discussion (Julia Hockenmaier, University of Illinois, and Tamara Berg, Stony Brook University)**

Workshop 8:
Computational Linguistics for Literature (CLFL)
Friday, June 14

Venue: International G

9:00–9:05 **Welcome**

9:05-10:00 **Invited talk by Livia Polanyi**

10:00-10:30 A Tale of Two Cultures: Bringing Literary Analysis and Computational Linguistics Together
Adam Hammond, Julian Brooke, and Graeme Hirst

10:30-11:00 **Coffee break**

11:00-11:30 Recognition of Classical Arabic Poems
Abdulrahman Almuhareb, Ibrahim Alkharashi, Lama AL Saud, and Haya Altuwaijri

11:30-12:00 Tradition and Modernity in 20th Century Chinese Poetry
Rob Voigt and Dan Jurafsky

12:00-12:30 Linguistic Resources and Topic Models for the Analysis of Persian Poems
Ehsaneddin Asgari and Jean-Cedric Chappelier

12:30-14:00 **Lunch break**

14:00-15:00 **Invited talk by Mark Riedl**

15:00-15:30 **Poster teasers**

- The desirability of a corpus of online book responses
Peter Boot
- Clustering Voices in The Waste Land
Julian Brooke, Graeme Hirst, and Adam Hammond
- An initial study of topical poetry segmentation
Chris Fournier
- Groundhog DAG: Representing Semantic Repetition in Literary Narratives
Greg Lessard and Michael Levison
- Exploring Cities in Crime: Significant Concordance and Co-occurrence in Quantitative Literary Analysis
Janneke Rauscher, Leonard Swiezinski, Martin Riedl, and Chris Biemann

15:30-16:00 **Coffee break**

16:00-16:30 **Poster session**

16:30-17:00 From high heels to weed attics: a syntactic investigation of chick lit and literature
Kim Jautze, Corina Koolen, Andreas van Cranenburgh, and Hayco de Jong

17:00-17:30 **An informal talk by He, Barbosa and Kondrak**

17:30–18:00 **Farewell**

Workshop 9:
Natural Language Processing for Improving Textual Accessibility (NLP4ITA)

Friday, June 14

Venue: International F

9:15–9:30 **Opening Remarks by Workshop Chairs**

9:30–10:00 A User Study: Technology to Increase Teachers' Linguistic Awareness to Improve Instructional Language Support for English Language Learners
Jill Burstein, John Sabatini, Jane Shore, Brad Moulder, and Jennifer Lentini

10:00–10:30 Open Book: a tool for helping ASD users' semantic comprehension
Eduard Barbu, Maria Teresa Martín-Valdivia, and Luis Alfonso Ureña-López

10:30–11:00 **Coffee Break**

11:00–11:30 Tools for non-native readers: the case for translation and simplification
Maxine Eskenazi, Yibin Lin, and Oscar Saz

11:30–12:30 **Invited Talk: *Information Accessibility: More than just text deep* (Kathleen F. McCoy, University of Delaware, USA)**

12:30–14:00 **Lunch Break**

14:00–14:30 Lexical Tightness and Text Complexity
Michael Flor, Beata Beigman Klebanov, and Kathleen M. Sheehan

14:30–15:00 A System for the Simplification of Numerical Expressions at Different Levels of Understandability
Susana Bautista, Raquel Hervás, Pablo Gervás, Richard Power, and Sandra Williams

15:00–15:30 A Two-Stage Approach for Generating Unbiased Estimates of Text Complexity
Kathleen M. Sheehan, Michael Flor, and Diane Napolitano

15:30–16:00 **Coffee Break**

16:00–17:00 **Final Discussion and Closing Remarks**

Workshop 10:

4th workshop on Computation Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA 2013)

Friday, June 14

Venue: International A

8:30–8:40 **Opening Remarks**

8:40–9:20 **Invited talk: Prof. Dr. Rosalind Picard**

- Recent adventures with emotion-reading technology
  Rosalind Picard

Session 1:Affect Recognition in Text (I)

9:20–9:45 Bootstrapped Learning of Emotion Hashtags #hashtags4you
Ashequl Qadir and Ellen Riloff

9:45–10:10 Fine-Grained Emotion Recognition in Olympic Tweets Based on Human Computation
Valentina Sintsova, Claudiu Musat, and Pearl Pu

10:10–10:30 Spanish DAL: A Spanish Dictionary of Affect in Language
Matías Dell' Amerlina Ríos and Agustin Gravano

10:30–11:00 **Break**

11:00–11:40 **Invited talk: Prof. Dr. Jonathan Gratch**

Session 2: Affect Recognition in Text (II)

11:40–12:05 The perfect solution for detecting sarcasm in tweets #not
Christine Liebrecht, Florian Kunneman, and Antal Van den Bosch

12:05–12:30 Using PU-Learning to Detect Deceptive Opinion Spam
Donato Hernández, Rafael Guzmán, Manuel Móntes y Gomez, and Paolo Rosso

12:30–12:55 Sexual predator detection in chats with chained classifiers
Hugo Jair Escalante, Esaú Villatoro-Tello, Antonio Juárez, Manuel Montes-y Gómez, and Luis Villaseñor

12:55–14:00 **Lunch Break**

14:00–14:40 **Invited talk: Dr. Theresa Wilson**

Session 3: Multilinguality in Social Media

14:40–15:05 Subjectivity and Sentiment Analysis of Modern Standard Arabic and Arabic Microblogs
Ahmed Mourad and Kareem Darwish

15:05–15:30 Sentiment Analysis in Czech Social Media Using Supervised Machine Learning
Ivan Habernal, Tomáš Ptáček, and Josef Steinberger

15:30–16:00 **Break**

Session 4: Subjectity, Sentiment and Social Media Analysis (I)

16:00–16:15 Tagging Opinion Phrases and their Targets in User Generated Textual Reviews
Narendra Gupta

16:15–16:30 From newspaper to microblogging: What does it take to find opinions?
Wladimir Sidorenko, Jonathan Sonntag, Nina Krüger, Stefan Stieglitz, and Manfred Stede

16:30–16:45 Bilingual Experiments on an Opinion Comparable Corpus
Eugenio Martínez-Cámara, M. Teresa Martín-Valdivia, M. Dolores Molina-González, and L. Alfonso Ureña-López

16:45–17:00 RA-SR: Using a ranking algorithm to automatically building resources for subjectivity analysis over annotated corpora
Yoan Gutiérrez, Andy González, Antonio Fernández, Andrés Montoyo, and Rafael Muñoz

17:00–17:15 **Break**

Session 5: Subjectivity, Sentiment and Social Media Analysis (II)

17:15–17:30 Sentiment analysis on Italian tweets
Valerio Basile and malvina nissim

17:30–17:45 Sentence-Level Subjectivity Detection Using Neuro-Fuzzy Models
Samir Rustamov, Elshan Mustafayev, and Mark Clements

17:45–18:00 Sentiment Classification using Rough Set based Hybrid Feature Selection
Basant Agarwal and Namita Mittal

18:00–18:15 Sentiment Analysis in Social Media Texts
Alexandra Balahur

18:15–18:30 **Closing remarks**

7

Joint NAACL/ICML Symposium on Natural Language Processing (NLP) and Machine Learning (ML)

Date: June 15, 2013
Website: https://sites.google.com/site/nlpml2013/

Description

Following the success of the joint NLP/ML symposium held in 2011, the second joint symposium on Natural Language Processing and Machine Learning will be held in Atlanta, Georgia. The goal of the symposium is to foster communication and collaboration between researchers in these synergistic areas, taking advantage of the nearby locations of NAACL-HLT 2013 (http://naacl2013.naacl.org/) and ICML 2013 (http://icml.cc/2013/). It will bring together members of the Association for Computational Linguistics (http://www.aclweb.org/), and the International Machine Learning Society (http://www.machinelearning.org/).

The symposium will be held on June 15, 2013 as part of several joint NAACL/ICML events running June 14–16. More information can be found on the NAACL (http://naacl2013.naacl.org/WorkshopsAccepted.aspx) and ICML (http://icml.cc/2013/?page_id=41) websites.

Topics

The symposium will feature a series of invited talks from top researchers covering several recent, exciting topics that are relevant to both Machine Learning and NLP. These areas include but are not limited to: structured prediction and inference, log-linear models, Bayesian modeling, deep learning, sparse representations, scalable probabilistic inference, semi-supervised and unsupervised training. Applications areas include learning language in context, language modeling, machine translation, learning semantic representations, topic modeling, among others.

Invited Speaker List

The list of confirmed speakers include:

- Raymond Mooney (University of Texas at Austin)
- Alexander Smola (Carnegie Mellon University / Google)
- Andrew McCallum (University of Massachusetts, Amherst)
- Daniel Hsu (Microsoft Research)
- Jeff Siskind (Purdue)
- Slav Petrov (Google)
- Andre Martins (Carnegie Mellon University / Priberam Labs)
- Scott Yih (Microsoft Research)
- Richard Socher (Stanford University)
- Vivek Srikumar (University of Illinois at Urbana-Champaign)
- Chris Dyer (Carnegie Mellon University)
- Kevin Gimpel (Toyota Technological Institute, Chicago)

For more updates or details regarding the program, please check the symposium website.

Organizing Committee

Dan Roth, University of Illinois at Urbana-Champaign (`http://l2r.cs.uiuc.edu/`)
Sujith Ravi, Google (`http://www.sravi.org`)

8

Venue information

Presenting Atlanta's
Favorite Restaurants

Downtown Atlanta ~ Peachtree Street Area

Alma Cocina, *191 Peachtree Street, N.E*

Cocina introduces Atlanta to a refreshing new approach to modern Mexican cuisine. Bright, fresh ingredients and traditional regional influences come together with other Latin American flavors in a variety of vibrant fresh dishes. We also feature the most unique tequila selections and a host of innovative Latin-influenced cocktails beyond a superior margarita. Lunch & Dinner daily. $$

Azio's, in Peachtree Center, *229 Peachtree Street, N.E*

This casual, trendy Italian restaurant specializes in creative pizzas. Lunch and dinner daily $

Barley's Sports Bar and Lounge , *338 Peachtree St.*

Barley's is a versatile entertainment venue. Patrons dine-in, play billiards, listen to live music, enjoy karaoke, and watch sporting events in a smoke-free environment.

Big Kahuna, *303 Peachtree Center Avenue.*

This is Atlanta's newest destination for lunch, dinner, and late night. Take a break from the mundane and relax at the Big Kahuna, the beach away from the beach. Flat screens, a great bar, and great food await you there.

Fire of Brazil*, 218 Peachtree Street*

A host of Gauchos parade through the restaurant with choice cuts of grilled Angus beef, chicken, lamb, pork and sausage. All selections are slow roasted over an open flame, seasoned to perfection and accompanied by an extravagant salad area. Fixed priced dining allows you eat until you've had your fill! Lunch and Dinner daily $$

Hard Rock Café, *215 Peachtree St NE*

Situated in downtown Atlanta's Cornerstone Building at the heart of Peachtree Street and Andrew Young International Boulevard, Hard Rock Cafe Atlanta could be considered a cornerstone of southern rock. With its southern charm and renowned hospitality, this location truly gives meaning to our "Love All Serve All" motto. Packed with memorabilia from Hard Rock's world-and serving hearty, fabulous food. Open daily, 11:00AM – 12:00AM $

Hsu's*, 192 Peachtree Center Avenue*

The faux-Chinatown exterior of this downtown restaurant gives way to a dimly lit, carpeted interior straight out of "Charlie Chan and the Curse of the Dragon Queen". The menu offers delicious, true flavors of China - Szechwan-style Chinese in a sophisticated setting. Lunch: Mon-Fri; dinner daily $$

Max Lager's*, 320 Peachtree Street*

The interior décor resembles an artist's loft with its brick walls, exposed light fixtures and colorful furniture. This brewery serves your basic American cuisine. Lunch: Mon-Fri. Dinner nightly $

Meehan's Public House, *180 Peachtree Street*

Meehan's Public House is an authentic Irish eatery that offers a full menu of classic Irish Dishes elevated to the gourmet level. Meehan's comfortable yet lively atmosphere is a draw for brunch, lunch, happy hour, dinner, and late night fun. Lunch & Dinner Mon-Friday 11:30am-2:00am; Saturday – 11:00am to 3:00am, Sunday 11:00am-12:00am $

This document is the intellectual property of Presenting Atlanta, LLC and is proprietary, confidential & privileged.

Morton's of Chicago, *303 Peachtree Center Avenue*

Cozy, modest, minimalist decor makes you feel comfortable upon entering Atlanta's favorite steak house. Dinner nightly $$$

Pacific Rim Bistro*, 303 Peachtree Center Avenue, NE*

Located on the ground floor of Peachtree Center, this attractive, modern bistro features floor-to-ceiling windows, lacquered wicker chairs and a large, horseshoe-shaped sushi bar. Superior sushi and a selection of fusion dishes are the highlights of the pan-Asian menu. Lunch: Mon-Fri. Dinner: Nightly $$

Ri Ra Irish Pub, *1080 Peachtree St.*

Beautifully restored old Irish pub, first class service, lunch & dinner, live entertainment, and banquet facilities is a MUST visit in midtown.

Rays in the City, *240 Peachtree Street*

This place has appetizing dishes to satisfy everyone. Along with their award winning wine list, you will have a very memorable dining experience at Ray's in the City. Lunch and Dinner Daily $$

Room*, 400 West Peachtree Street*

With a wide selection of small plates and skewered meats and fish with accompanying sauces, broiled and wood grilled meats and fresh fish, sushi and a la carte side dishes Executive Chef Nick Oltarsh highlights dishes with playful, international undercurrents. Room is anything but the typical steakhouse. Open for Lunch and Dinner $$$$

Sweet Georgia's Juke Joint, *200 Peachtree Street*

Sweet Georgia's Juke Joint is the culmination of a deep history of music, food, and drink, and socializing in the South. Located in the former Macy's building, in the heart of downtown Atlanta at the iconic 200 Peachtree address, Sweet Georgia's Juke Joint brings the city to a place to relax while enjoying live entertainment and exceptional cuisine. Lunch and Dinner Daily. $

White Oak Kitchen & Cocktails*, 270 Peachtree Street*

White Oak Kitchen & Cocktails features chef-inspired cuisine using ingredients fresh from the Southern region. The hospitality and ambiance are locally inspired and crafted, and the cocktails and wines are classic yet fresh. Let their hospitality lift you up! Lunch and Dinner Daily. $$

This document is the intellectual property of Presenting Atlanta, LLC and is proprietary, confidential & privileged.

Downtown, Luckie Street area

Der Biergarten, *300 Marietta Street Northwest*

An authentic experience is here, giving local residents and visitors a true taste of German culture. Located within the Luckie Marietta District, this happy dining space fills a citywide void for a beer garden and authentic German fare while providing world-class service in a festive, traditional atmosphere. Lunch and dinner daily $

Legal Seafood, *275 Baker Street, NW*

For three generations, Legal Sea Foods has served only the highest quality, freshest fish at a great value. We select the finest and most flavorful goods, demanding nutrition and value as well as great taste from every dish we prepare. And for service, you'll find no equal. Lunch and dinner: Monday – Saturday. Sunday Brunch $$

McCormick & Schmicks, *190 Marietta Street*

"Steak and seafood are the mainstay but the menu is broad... terrific happy-hour menu... a little more formal than your regular 'tot-friendly' restaurant, but the staff is great and goes out of their way to make sure you're comfortable... try to get one of the banquet rooms -good food for adults and more than enough for the little ones too." Lunch & Dinner Mon-Sat 11am-10pm; Lunch Sun 4pm-9pm $$

Peasant Bistro*, 250 Park Avenue West NW*

Peasant Bistro is a cosmopolitan, dramatic two-story restaurant overlooking Centennial Park and the Atlanta skyline. Featuring fresh, seasonal cuisine, the menu consists of delicious, traditional bistro favorites with French and Mediterranean influence. Lunch and Dinner Daily $

RISE Sushi*, 300 Marietta Street*

Rise Sushi lounge offers a full array of sushi and specialty rolls in a chic modern atmosphere. Conveniently located in the Lucky Marietta District of downtown Atlanta, Rise Sushi can meet all your dining needs Rise is situated just one block from the Georgia Aquarium and Centennial Park and offers complimentary parking for all guests. Lunch and Dinner $$

Stats, *300 Marietta Street*

Are you ready for some "food play?" Simply put, STATS is serious food and serious sports – Premium, food-focused, sports-centered, draught-direct eatery with a high-energy vibe, STATS. Adding to the excitement and anticipation, tap tasters will be mesmerized by the revolutionary beer system incorporated into the restaurant. With 70 high-definition televisions, gourmet American cuisine and beer that literally pours from the tables, the venue will be the definitive spot for sports lovers seeking good meals and libations. Lunch and Dinner $$

Thrive*, 101 Marietta Street*

Thrive is an upscale yet casual restaurant located in the heart of downtown. Thrive combines the elements of delicious food and extensive drink choices with a luxury setting. Lunch and Dinner Daily $$

This document is the intellectual property of Presenting Atlanta, LLC and is proprietary, confidential & privileged.

Downtown Atlanta ~ Old Fourth Ward, Inman Park

4th and Swift, *621 North Avenue*

Located in the Old Fourth Ward and occupying the space formerly the engine room, it retains many of the original elements. The farm-to-table menu features the freshest of ingredients from all natural and sustainable farmers. The atmosphere contains a sophisticated tavern, kitchen and dining space in which to enjoy the Southern hospitality. Dinner daily $$

Barcelona Atlanta Wine Bar & Restaurant, *240 N. Highland Avenue*

Located just two miles east of downtown Atlanta, Inman Park is not only the home of beautiful Victorian mansions and contemporary lofts, but also a culinary destination. Barcelona offers the neighborhood a delicious array of Spanish small and large plates that turn the mere act of dining out into a joyous communal experience amongst friends and family. $$

Il Localino*, 467 North Highland Avenue*

Il Localino, a charming family owned & operated Italian Restaurant. Delicious authentic food, fun atmosphere, & fabulous drinks!! The perfect spot for a casual dinner or a major celebration! Whether a romantic table for two, or a private party, Il Localino will exceed all of your dining expectations! Dinner Nightly. $$

Kevin Rathbun Steak, *154 Krog Street, Ste 200*

From the owners of the nationally acclaimed *Rathbun's* & *Krog Bar* in Atlanta, Kevin gets his motivation from where he came from .He pays homage to the big league steakhouses. He continues to bring his savvy Kansas City hospitality and his love of entertaining to the Atlanta area. Dinner daily, closed Sunday $$

Parish*, 240 North Highland Avenue*

Located on the main floor of the historic space, Parish Restaurant serves a variety of New Orleans inspired dishes with modern twists, alongside a classic fully stocked raw bar, seven nights a week. Dinner Daily $$

Rathbun's, *112 Krog Street*

This Modern-American menu "weaves the richest tapestry of new American cooking". Located in a 100-year-old building near Inman Park with a modern, warehouse feel, the open kitchen, vibrant colors and jazzy, urban tone excite one immediately after walking through the fuchsia door. Dinner: Mon-Sat $$

Serpas, *#501,659 Auburn Avenue, NE*

The first independent restaurant from revered Atlanta chef Scott Serpas, Serpas is located in reclaimed cotton storage facility at The Pavilion at Studioplex in the historic Old Fourth Ward. The food features cuisine that is true to its origin handcrafted by the New Orleans native chef and owner. Dinner: Tuesday – Sunday. Sunday Brunch $$

Wisteria*, 471 North Highland Avenue*

Located in the historic Atlanta neighborhood of Inman Park, our cuisine, contemporary American with a Southern twist, is sure to please every palate. Come in and enjoy an inviting and cozy atmosphere with Chef Jason Hill's seasonal menu.

This document is the intellectual property of Presenting Atlanta, LLC and is proprietary, confidential & privileged.

Downtown Atlanta ~ Castleberry Hill

No Mas Cantina, *180 Walker Street*

The best of Mexico! It features authentic favorites, outrageous specials, and luscious desserts. It also features handcrafted furnishings and accessories of over 300 artisans in Mexico.
Lunch 11-4: Dinner 4-10 (Fridays and Saturdays till 11): Brunch Saturday and Sunday 11-2.

Baltimore Crab Seafood, *253 Peters Street*

Located in trendy Castleberry Hills, this mostly takeout seafood spot is convenient for Clark Atlanta students. Order at the counter and get your seafood fried, broiled, steamed or sauteed. Home-style sides like mac and cheese, potato salad and corn-on-the-cob give your dinner a comfy finish....that is, unless you also have the very decadent banana pudding.

255 Tapas Lounge, *255 Peters Street*

This is one of Atlanta's newest hot spots. The music is a mix of old and new, classic and hip-hip. Emphasizing casual dining, social drinking, and a laissez-faire ambiance, 255 is Atlanta's premier lounge.

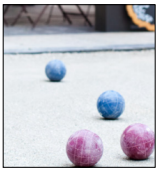
Presenting Atlanta
ATLANTA

This document is the intellectual property of Presenting Atlanta, LLC and is proprietary, confidential & privileged.

Midtown

Baraonda, *710 Peachtree Street, NE*

Shadowy hearth light, heavily lacquered plank tables, Italian liqueur posters and plenty of Gypsy Kings tunes make this cozy Midtown trattoria an instant romantic hit. Lunch: Mon-Fri. and Sun. Dinner nightly $



Davio's Northern Italian Steakhouse*, 3500 Peachtree Rd.*

Enjoy classic Italian cuisine at Atlanta's only Northern Italian Steakhouse. Davio's has a large display kitchen, and in-house bakery, vibrant bar and lounge and outdoor patio.

Empire State South, *999 Peachtree St NE # 140*

Empire State South takes a modern approach to creating authentic Southern dishes, relying on the beautiful foods of this region - a community restaurant that appeals to a broad range, providing a courtyard for bocce ball and an extensive coffee bar. Breakfast, Lunch and Dinner Daily $

Ecco, *40 7th Street, NE*

Located in the heart of Midtown, **Ecco** features seasonally inspired cuisine with Italian, Spanish and French influences. The atmosphere is welcoming, with a shot of sophisticated, casual elegance.With its soaring glass entrance and honed marble floors, exhibition kitchen and wood-burning oven, **Ecco** embraces classic roots while delivering an unmistakably modern and fresh dining experience. $$

Escorpion, *800 Peachtree Street*

Escorpion is a regionally inspired Mexican tequila bar and cantina capturing the authentic tastes and ingredients indigenous of Mexico.

Fig Jam Kitchen & Bar*, 3630 Peachtree Road NE*

Fig Jam Kitchen & Bar's vision is to be seen as a place where customers can create their own custom dining experiences. One of the ways they are doing that is by providing a warm environment where guests can enjoy a good conversation over a variety of dishes and hand crafted drinks with their friends, co-workers and family. Lunch and Dinner Daily. Closed Sunday. $$

Livingston*, 659 Peachtree Street*

The new dining destination takes it name frrom. Livingston Mims, who serverd as Atlanta's mayor during the early 20th.century. The restaurant is located across the street from The Fox Theatre and offers Atlantans an innotive reflection of flavor and history. Breakfast, Lunch and Dinner Daily $$

Lure*, 1106 Cresent Avenue*

Inspired by classic American and global coastal traditions, Lure presents Atlanta'a freshest fish in the heart of midtown on Cresent Avenue.

Marlow's Tavern (Plaza Midtown) *950 W. Peachtree St.*

This tavern features American tavern fare served in a modern atmosphere offering classic dished updated and elevated to gourmet level.

This document is the intellectual property of Presenting Atlanta, LLC and is proprietary, confidential & privileged.

***Nan**, 1350 Spring Street*

Nan creates a glamorous atmosphere for upscale, mild-mannered exquisite Thai food. The restaurant and the food are as sumptuous as any high-end restaurant in Bangkok – without the jet lag. Upscale but not stuffy, spectacular yet quietly restrained, the interior is mesmerizing with a menu to match. Lunch: Mon-Fri. Dinner daily $$

***Oceanaire Seafood Room**, 1100 Peachtree Street, NW*

With its red leather horseshoe-shaped booths, brass tacks and warm cherry hardwood floors, this lavish seafood restaurant trumpets a global selection of fresh fish and oysters. Lunch Mon-Fri. Dinner daily $$

***One Midtown Kitchen**, 559 Dutch Valley Road, NE*

An 80-foot curved counter bar, towering wine display and vast windows offering lovely views of the midtown skyline are a few reasons you should experience One Midtown. The seasonal menu focuses on small portions, encouraging the diners to share the modern-American cuisine. Dinner daily $$

***Olmsted**, 1180 Peachtree Street, NW*

The name Olmsted pays homage to Frederick Law Olmsted, who is known as the "father of landscape design." His designs and work greatly influenced the layout of Atlanta and its neighborhoods, but inspiration is not a one-way street. The restaurant Olmsted plays to the memory of Olmsted and his shaping of modern day Atlanta life with Incredible food, and an amazing atmosphere. Brunch and Lunch Daily. $$

***Shout,** 1197 Peachtree Road, NE*

This ultra hip, luxurious space houses a sushi lounge, intimate dining areas and several bars. Midtowners sip fruity cocktails while hipsters jockey for the best viewing tables, where they eye an eclectic menu of tapas, pizza and sushi. The rooftop lounge with a view of the city is one of the many draws to this great spot! Lunch: Mon-Fri. Dinner daily $$

***South City Kitchen,** 1144 Crescent Avenue, NE*

This energetic, popular spot offers modern gourmet creations with a southern flair. It's set in a renovated house accented with an interesting mix of new and old furnishings. Dinner nightly. Sunday brunch $

***The Spence**, 75 5th Street NW*

The Spence is poised to become one of Atlanta's, and the country's top restaurants. Helmed by Chef Richard Blais, recent "Top Chef All-Stars" winner, The Spence, serves classic American dishes as reinterpreted by Blais. Utilizing local ingredients on an ever-changing "pantry list", Blais displays his culinary strengths throughout. The Spence showcases the best in accessible atmosphere, food and beverage. Dinner and Lunch Daily. $$

***Tap**, 1180 Peachtree Street*

A simple and eclectic menu featuring bar snacks, items for sharing and entrees is prepared daily to the highest standards by Chef Todd Ginsberg. His impressive culinary background adds a touch of finesse to TAP's casual bill of fare. Brunch: Saturday and Sundays $. Lunch: Monday through Friday $. Dinner: Daily $

This document is the intellectual property of Presenting Atlanta, LLC and is proprietary, confidential & privileged.

Veni Vidi Vici, *41 14th Street, NW*

Situated in an unlikely location on the street level of a parking garage, this authentic home-style Italian restaurant brims with corporate executives, convention crowds, as well as midtown theater goers. Lunch: Mon-Fri. Dinner nightly. $$

This document is the intellectual property of Presenting Atlanta, LLC and is proprietary, confidential & privileged.

Westside

Abattoir, *1170 Howell Mill Road*

Abattoir is an American chophouse specializing in the use of locally-raised proteins to produce high-quality, affordable food. Offering freshly butchered whole fowl, fish, beef, pork and other game served in a variety of ways, the menu is based around "whole animal cuisine," a term referring to the use of every functional piece of the animal. Abattoir was named 2009 Restaurant of the Year by Atlanta Magazine. $$

Antico Pizza Napoletana, *1093 Hemphill Avenue*

Antico Pizza is an authentic Tuscan-style restaurant located in the heart of the Buckhead entertainment district of Atlanta. Marco Bettie invites you to taste the simple, flavorful dishes that have made Italian and Tuscan cooking the most appreciated cuisine in the world combining fresh ingredients with traditional recipes for a dining experience that until now, was only available at his restaurant in San CasCasiano,Italy. $$

Bacchanalia, *1198 Howell Mill Road*

Consistently voted number one for food, Bacchanalia, with its cutting edge décor, continues to be ahead of the pack. This chef-owned restaurant makes excellence look effortless. Set in a slick industrial space in an off-beat neighborhood, Bacchanalia has not skipped a culinary beat in achieving transition. Dominique Coyne's reconfiguring of this former meat packing plant built in the 1920s emphasizes clean spaces and retains the old engineering components. Gentle white draping creates a semi-private space for smaller parties. $$$$

Bocado, *887 Howell Mill Road, #2*

The design is a literal translation of the passion for sustainable foods. Keeping with our commitment to simplicity, Bocado's design reflects the love of all things natural. The restaurant utilizes organic, reused materials recreated in inventive and distinctive ways. $$

Five Seasons Brewery, 1000 Marietta Street, NW, # 204

Most of our produce comes from small, local, organic farms. Our chefs actually create 20-30 new specials every day with what comes in fresh from the farms (the "*5th Season*" is when things are ripe and in their peak condition!) We brew about 100 different beers every year - our beer list changes on a continual basis. Our kitchen makes almost everything from scratch. $

FLIP, *1587 Howell Mill Road, NW*

The menu of Flips features thoughtfully paired unique and homemade condiments and sides that complement the burgers. The full drink menu showcases a top selection of beer, wine, martini's and margaritas. There is a delicious selection of milkshakes at the unique milkshake bar. $$

JCT Kitchen, *1198 Howell Mill Road*

JCT. Kitchen opened in January 2007 in the Westside Urban Market of Atlanta's increasingly trendy West Midtown, JCT. Kitchen & Bar offers upscale, yet casual dining in an atmosphere inspired by the California wine "country" with an East Coast interpretation. This unique neighborhood restaurant features a menu reminiscent of traditional family favorites. Lunch and Dinner (closed on Sundays) $$

This document is the intellectual property of Presenting Atlanta, LLC and is proprietary, confidential & privileged.

***Miller Union*,** 999 Brady Avenue

At Miller union, there is a partnership between chef and farmer. The menu is based around the week's harvest; we take care to preserve the true flavors of our ingredients with a simple rustic preparation. Miller Union's unique dining space sits on the site once occupied by the Miller Union Stockyards and contributes to the revitalization of Atlanta's dynamic Westide. $$

***The Optimist*,** 914 Howell Mill Road

The Optimist and Oyster Bar is simply "The Hamptons meets The South". They are passionate about sourcing sustainable seafood and preparing it in classic ways. Menu items range from the most authentic seafood gumbo with "pot luck" garlic bread to Maine lobster rolls, Gulf grouper with smoky spring Vidalias, and an ever changing selection of super fresh East and West coast oysters. Oh... and if you are not a seafood person, they LOVE meat as well! No worries, be optimistic! Lunch and Dinner Weekdays. Dinner Only on Weekends. $$$

Paschal's, *108-B Northside Drive*

Paschal's is the "soul" of Atlanta, serving favorites like Paschal's Fried Chicken, Fried Green Tomatoes, and our famous Peach Cobbler. Pashcal's Restaurant is where southern hospitality lives.

Taqueria del Sol, *1200-B Howell Mill Road*

Taqueria del Sol, is truly one of a kind, juxtaposing the quality of fine dining ingredients against fast-casual prices, and offering authentic Southwestern selections as well as a variety of Southern-inspired menu choices. Christiane Lauterbach of Atlanya Magazine writes, "to this day, here is no better bang for the buck in Atlanta." In 2007, BonApetit picked the Atlanta-based chain as a "Top American restaurant." $

This document is the intellectual property of Presenting Atlanta, LLC and is proprietary, confidential & privileged.

Virginia – Highlands

Dish, *870 North Highland Avenue, NE*
This urban, eclectic spot draws a trendy crowd for its flavors and combinations. Good looks and New American menu meet at this eye-catching cottage bistro. Dinner nightly $$

Fritti, *311 North Highland Avenue*
Both hip and comfortable, Fritti provides a most authentic Italian dining experience. Newly certified by the Verac Pizza Napoletana Association in Naples, Italy; Fritti features 26 varieties of Neapolitan pizzas prepared in hand-crafted wood-burning ovens. Fritti also offers a variety of Tuscan style wood roasted meats, great fried appetizers and refreshing salads $$

Harry & Sons, *820 North Highland Avenue, NE*
The simple flavors are nicely spicy with plenty of non-intimidating meats and heaps of veggies. Make this low-key Thai/sushi bar a must see. Lunch and dinner daily $$

La Tavola, *992 Virginia Avenue, NE*
This modern trattoria dishes up Italian favorites for chic Virginia-Highlanders. With its open kitchen, mirrored walls and delicious Italian cuisine, La Tavola fits right in to its stylish Virginia-Highlands environment. Lunch and dinner nightly. Brunch Sunday. $$

Murphy's, *997 Virginia Avenue, NE*
Located in the heart of Virginia Highlands, Murphy's offers contemporary American comfort food in an upbeat atmosphere. A wine shop and not-to-be missed bakery can also be found in the front of the restaurant. Lunch: Mon-Fri. Dinner nightly, Brunch Sat., Sun. and holidays $

Noche, *1000 Virginia Avenue, NE*
Brilliantly, breathlessly exudes 'southwestern'. Margaritas and fun will be found under the covered balcony at this friendly spot. Lunch: Sat. and Sun. Dinner nightly. $

Osteria, *832 North Highland Avenue, NE*
A smart and thoroughly unpretentious spot for Neapolitan-style pasta and thin crust pizzas that have become a big hit with Virginia Highlanders. Toppings range from gourmet-minded prosciutto and arugula to spicy salami and mushrooms. Lunch: Tues-Sun. Dinner daily. $

The Optimist, *914 Howell Mill Road*
Ford Fry's The Optimist brings a classic seafood experience to landlocked Atlantans. From Maine sea scallops and whole-roasted Georgia trout to garlicky clams and fresh oysters flown in from the East and West coasts, executive chef Adam Evans serves high-quality and sustainable seafood from the country's best suppliers. An extensive selection of beer, wine, and handcrafted cocktails pair with hickory-roasted fish and raw oysters alike. $$

The Original El Taco, *1186 North Highland Ave.*
Here you'll enjoy the simple, fresh Mexican food while sipping on our famous margaritas and frozen mojitos.

This document is the intellectual property of Presenting Atlanta, LLC and is proprietary, confidential & privileged.

Sotto Sotto, *313 North Highland Avenue, NE*

Subdued style and authentic Italian ingredients make this a great neighborhood spot and cross-town destination. Dinner: Mon-Sat. $$

Two Urban Licks (right)*, 820 Ralph McGill Boulevard*

From the creators of One Midtown Kitchen comes a Bar-B-Q joint Bubba never dreamt of. Scallops and duck along with traditional favorites on the menu taste as good as they sound. Fantastic, lively interior encased by a stunning view of the city. $$

Vine, *1190 North Highland Avenue*

Offering a world-class eclectic menu updated monthly, Vine is a perfect example of what makes Virginia Highlands so unique. Whether seated on the outside patio or next to the double sided fireplace, you'll enjoy comfort, character and an award-winning wine list. Dinner nightly. Lunch: Reservations only. Brunch Sunday $

This document is the intellectual property of Presenting Atlanta, LLC and is proprietary, confidential & privileged.

Atlantic Station

Lobby at Twelve, *361 17th Street NW*

Lobby is a high energy casual dining American Kitchen serving great food and drinks. With an open kitchen and great lobby bar with its wood fire pizza oven, Lobby becomes a great gathering place to graze, sip, wine or dine. Lunch and Dinner. $$

Rosa Mexicano, *245 18th Street NW*

The restaurant delivers the winning combination of authentic Mexican cuisine in an exciting, modern and friendly environment. A dramatic 18' high waterwall is the eye-catching centerpiece of the 200-seat restaurant, which also features a lounge area and a private dining area that will accommodate up to 60 guests. $$

Strip, *245 18th Street*,

Strip offers a menu of Steaks, Sushi and Teasers. Lunch and Dinner Daily $$

This document is the intellectual property of Presenting Atlanta, LLC and is proprietary, confidential & privileged.

Aja, 3500 Lenox Road, Suite 100

A modern Asian Kitchen, this two story environment consists of a private dining room which seats up to 30 guests, a 150 seat patio with views of the Buckhead skyline and sprawling lounge area where guests can enjoy dim sum, signature mai tai's and more! The restaurant is the home to a 10 ft brass Buddha that arrived straight from Thailand. Private dining available. Lunch: Mon-Sat. Dinner nightly. $$

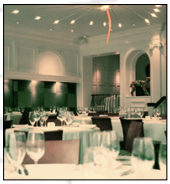
Anis, *2974 Grandview Avenue, NE*

This Buckhead French Bistro offers a wonderful meal at a very reasonable price. Its cozy setting in a charming bungalow attracts quite a following of devotees to this romantic spot. Private dining available. Lunch: Mon-Sat. Dinner nightly. $$

Antica Posta, *519 East Paces Ferry Road, NE*

This cozy bungalow with wooden beams warms the dining room of white-clothed tables. This little Italian main-stay bestows Atlantans with a taste of the Tuscan good life. Dinner nightly. $$

Aria, *490 East Paces Ferry Road, NE*



This new American savory menu includes fork-tender braised short ribs and Gorgonzola polenta, while seasonal offerings include enormous soft-shell crabs spiked with dill and delicate fish with direct-from-the-garden vegetables. The seductive surroundings include architecture from yesteryear and post modern art. Dinner: Mon-Sat. $$$

The Atlanta Fish Market, *265 Pharr Road, NE*

You are greeted by a most impressive outdoor bronze fish sculpture as you arrive at this upscale Buckhead fish house. An aromatic fish market display contains an extensive array of fresh seafood. Lunch: Mon-Sat. Dinner nightly. $$$

Bistro Niko, *3344 Peachtree Road*

Located in the spectacular Sovereign Building, Bistro Niko is an exciting French bistro, with cuisine that offers a lighter style of classic French cooking, as is the trend in Paris today. Lunch and dinner served seven days a week, a special brunch menu available on Sunday. $$

Blue Ridge Grill, *1261 West Paces Ferry Road*

The mountains meet the city in this rustic Buckhead home to winning American cuisine. Classic Southern and American dishes get the upscale treatment with largely winning results. Lunch: Sun-Fri. Dinner nightly. Sunday brunch. $$$

Bone's, *3130 Piedmont Road, N.E.*

Thickest and juiciest steaks in town. Atlantans have been gathering here since 1979 for the best prime rib, steaks, lamb chops, and live Maine lobsters (flown in daily). $$$

BrickTop's, *3280 Peachtree Street*

BrickTop's is a Nashville based company. It is the company's goal to balance high standards and quality with an unassuming atmosphere. BrickTop's is a place where it's pleasant to dine or have a drink after work. The restaurant is sophisticated, warm and friendly. From the greeters and service staff to the kitchen and management, everyone takes pride in their job. Lunch and Dinner daily $$

Buckhead Bottle Bar and Bistro, 268 East Paces Ferry Road

Buckhead Bottle Bar, a trendy and stylish restaurant, serves unique bites and well-crafted cocktails.$$

This document is the intellectual property of Presenting Atlanta, LLC and is proprietary, confidential & privileged.

The Buckhead Diner, *3073 Piedmont Road, NE*

This gleaming monument to American culinary chic of the late 1950's offers hearty American cuisine. "Comfort food gone gourmet" is the by-word at this well-known establishment. Lunch and dinner daily. Sunday brunch. $$

Capital Grille, *255 East Paces Ferry Road*

Boasting an atmosphere of energy, sophistication and power dining, The Capital Grille serves classic steak house offerings and raises the all around experience to a level of excellence. Best View in town! Lunch Mon – Fri. Dinner Nightly. $$$

Canoe, *4199 Paces Ferry Road, NW*

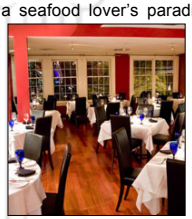
Situated on the banks of the Chattahoochee River, this is a favorite for "special occasion" destinations. It's a pure visual delight with its lovely patio and gorgeous gardens. Lunch: Mon-Fri. Dinner nightly. Sunday Brunch. $$

Cantina, *3280 Peachtree Street*

Here to Serve welcomes their first Mexican concept, Cantina, located at the Terminus building at the corner of Peachtree and Piedmont in Buckhead. The interior houses a modern design by Johnson Studios. Two outdoor patios: one for dining and one lounging and enjoying signature margaritas. Lunch and Dinner $$

Chops & Lobster Bar, *70 West Paces Ferry Road, NW*

Atlanta's "Tiffany of steakhouses" featuring seafood, as well as beef, it offers excellent service in a comfortable, clubby setting. The downstairs Lobster Bar (dinner only) is a seafood lover's paradise. Lunch: Mon-Fri. Dinner nightly. $$$



Coast, *111 West Paces Ferry Road (right)*

Coast Seafood and Raw Bar serves Atlanta's freshest seafood and island cocktails! Classics include crab and corn hush puppies, a signature seafood boil, and a variety of raw or steamed oysters, clams and mussels; along with signature fresh catch entrees. Large patio area for dining extended bar area, seating for over 100 and private dining. Dinner nightly - Sunday brunch $$

Eclipse di Luna, *764 Miami Circle, NE*

This festive, high-energy, interactive dining experience offers an artsy setting as a backdrop for award- winning specialty tapas and live music nightly. Belly up to the bar and enjoy authentic Spanish cuisine. Lunch: Tues-Sat. Dinner: Tues-Sun. $

F&B Atlanta, *3630 Peachtree Road NE*

A "local Neighborhood Bistro", located in the heart of Buckhead serving Americanfood with French twist. F&B is great for a casual drink, or a delicious meal. The atmosphere will captivate you. Lunch and Dinner Daily. $$

Fogo de Chao, *3101 Piedmont Road, NE*

This Brazilian churrascaria-style steakhouse is a meat lovers delight featuring fifteen types of perfectly roasted meat that is carved from skewers at your table by knife wielding gauchos. Lunch: Mon-Fri. Dinner nightly. $$$

This document is the intellectual property of Presenting Atlanta, LLC and is proprietary, confidential & privileged.

Hal's, 30 *Old Ivy Road*

A restaurant that triumphs in taste over trend, Hal's is not the typical highbrow restaurant so often equated with food of this quality. It is a place of white table cloths, flowers, warm brick, & candlelight. Dinner Nightly. Closed on Sunday. $$$

Horseradish Grill, *4320 Powers Ferry Road, NW*

A neighborhood "barn" and grill next to Chastain Park is a one of a kind eating establishment that makes you feel "at home". A creative menu that offers a most unique style of Southeastern cooking from North Carolina to the Georgia mountains Lunch: Mon-Fri. Dinner nightly. Brunch Sunday. $$

Kyma, *3085 Piedmont Road, NE*

Kyma offers magnificent authentic Greek cuisine in an uptown setting. This unique Buckhead restaurant is sure to be a "favorite spot"' to many! Dinner: Mon-Sat. $$$

La Fourchette, *3133 Piedmont Road Northeast*

A true Mediterranean delight - *la Fourchette* - picks up on warm notes from France, Italy, and Spain to create a cuisine unique to itself. French trained Chef Jeffrey Wall, a former member of the JOEL Brasserie team, brings a young and refreshing twist on traditional dishes allowing the food to speak for itself! $$

La Grotta, *2637 Peachtree Road, NE*

Ranked among Atlanta's finest restaurants old world charm and sophisticated cuisine come together to make for an elegant affair - Northern Italian dishes prepared with veal, seafood or pasta are uniformly excellent. Dinner: Mon-Sat. $$$

Local Three, *3290 Northside Parkway, Suite 120*

Using one the best kitchens in the country, Local Three presents a neighborhood spot with really good food, beverage and service! Join them for a come-as-you-are atmosphere and "foie gras in your flip flops" ambiance with the best quality cuisine at affordable prices alongside intriguing wine and beer lists and a full bar. Lunch and Dinner Weekdays. Brunch Sunday, Dinner on Weekends. $$

Market Buckhead, *W Hotel Buckhead, 3377 Peachtree Rd.*

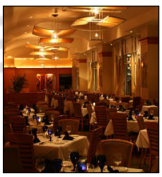
The concept of the hotel-as-a-home extends to these restaurants as its casual family kitchen by Jean-Georges.$

Ocean Prime, *3102 Piedmont Road Northeast*

Seafood or Steak? At Ocean Prime you can choose the best of both. Add in world class wines, award-winning handcrafted cocktails, timeless elegance and take-your-breath-away service. Their passion is to give you an extraordinary supper club experience. Dinner Nightly. $$$

Paul's, *10 King Circle*

You'll find luxurious American food with a continental flair and a full sushi bar in this Peachtree Hills neighborhood spot. A piano accompanies meals Wednesday through Saturday, and at Sunday brunch. Lunch and dinner Weds-Sat - Brunch Sunday $



This document is the intellectual property of Presenting Atlanta, LLC and is proprietary, confidential & privileged.

Portofino, *3199 Paces Ferry Place*

Enjoy simpler dishes of rustic, regional Italian cooking at a restaurant whose prices are more than fair for the quality of food. The dining room and adjacent bar are awash with a warm glow that make you feel right at home. Lunch: Mon-Fri. Dinner nightly $$

Pricci, *500 Pharr Road*

Wood burning ovens grace the center of this establishment making it one of the most unique places in town. Here you can relish in authentic Italian cuisine while dining in this elegant, modern atmosphere. Lunch: Mon-Fri. Dinner nightly $$$

Prime, *3393 Peachtree Road, NE*

Steaks are supreme--and seafood and sushi fare well--at this attractive Buckhead hot spot. Blonde wood, marble floors, gauzy drape and flattering lighting invites a casually chic crowd. Prime's lively bar and posh dining room are located in the prestigious Lenox mall. Lunch: Mon-Sat. Dinner nightly $$$

Restaurant Eugene, *2277 Peachtree Road*

Award-winning Restaurant Eugene is a quiet, sophisticated fine dining restaurant located in South Buckhead on Peachtree Road. The cuisine is Seasonal New American, with a special emphasis placed on using fresh ingredients from local farmers and specialty purveyors from around the world. Recipient of the 2005 and 2006 Wine Specialist Award of Excellence! Lunch and Dinner $$

STG Trattoria, *102 West Paces Ferry*

STG presents local and seasonal culinary offerings hedged in Italian institution. Most prominently, the "Trattoria" will features an APN-approved, traditionally regulated Neapolitan pizza crafted "from scratch" and baked in 5,000 pound, wood-fired Acunto ovens shipped from Naples, Italy and a walk in wine cellar. The emphasis of STG is on traditional execution, ingredient quality, and a slightly disciplined creativity. Brunch and Dinner Daily. Closed on Mondays for Dinner. $$

Watershed, *1820 Peachtree Road Northwest*

Owner Emily Sailers (of the Grammy Award winning *Indigo Girls*), made the decision to join forces with Chef Scott Peacock and reopen Watershed in the heart of Buckhead. This Peachtree location continues the restaurant's long standing commitment to its founding principles with great attention to detail and community consciousness. The environment and farm to table approach will be sure to appease your pallet and create for a wonderful experience. Lunch and Dinner Daily. Open Sunday for brunch. Closed on Mondays. $$

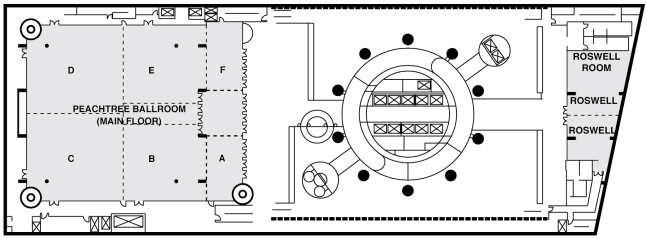
Woodfire Grill, *1782 Cheshire Bridge Road*

Executive Chef (and *Top Chef*) Kevin Gillespie incorporates the seasonal offerings of small local farms and purveyors as well as sustainable beef, lamb, pork, poultry and seafood for simply prepared cuisine. Foods are fire-roasted and grilled with a menu comprised of first, second and main courses in addition to chef's tasting menus. Dinner $$

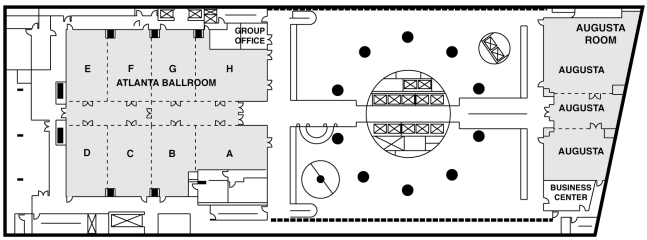
$25.00 and less	*$*
$50.00 and less	*$$*
$75.00 and less	*$$$*
More than $75.00	*$$$$*

This document is the intellectual property of Presenting Atlanta, LLC and is proprietary, confidential & privileged.

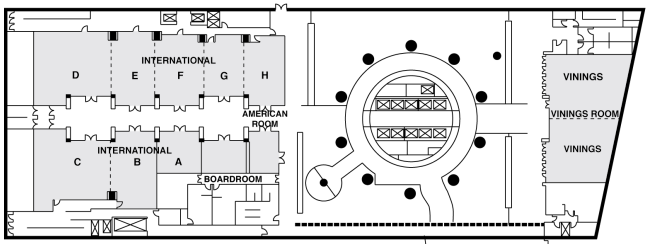
Hotel Maps



EIGHTH FLOOR

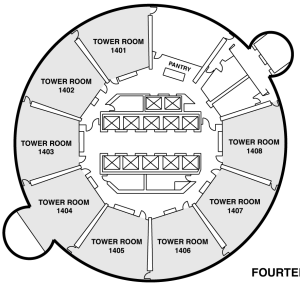


SEVENTH FLOOR

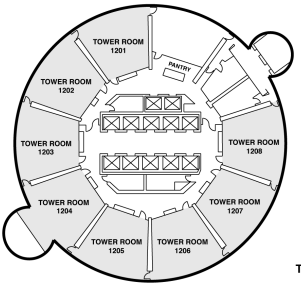


SIXTH FLOOR

THE
OVERLOOK

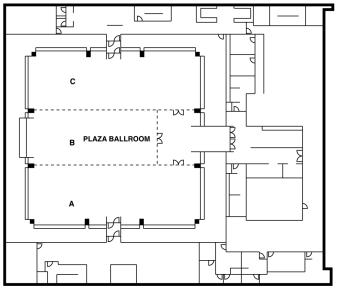


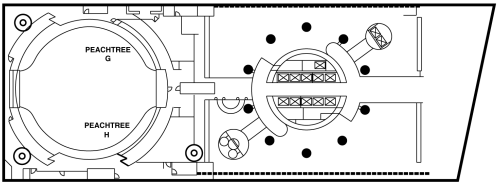
FOURTEENTH FLOOR



TWELFTH FLOOR

TENTH FLOOR





NINTH FLOOR

Index

Abney, Steven, 64
Abreu, José I., 78, 88, 94, 98, 101
Abreu, Jose I., 96
Abu-Jbara, Amjad, 34, 105
Addanki, Karteek, 108
Afantenos, Stergos, 94, 95
Agarwal, Basant, 120
Agichtein, Eugene, vii
Agirre, Eneko, 77, 78, 82, 89
Agrawal, Ameeta, 99
AL Saud, Lama, 117
Aldabe, Itziar, 100
Alegria, Iñaki, 112
Aleman, Yuridiana, 94
Aletras, Nikolaos, 22, 78, 89
Alkharashi, Ibrahim, 117
Alkuhlani, Sarah, 30
Allen, James, 93, 109
Almuhareb, Abdulrahman, 117
Altuwaijri, Haya, 117
Aluísio, Sandra Maria, 112
An, Aijun, 99
Andersen, Øistein E., 105
Andrade, Daniel, 43
Antunes, Sandra, 112
Aparicio Gali, Fernando, 100
Apers, Peter, 98
Apidianaki, Marianna, 95
Araki, Jun, 114
Arcan, Mihael, 35, 104
Artzi, Yoav, 51
Asgari, Ehsaneddin, 117
Atserias Batalla, Jordi, 99
Azab, Mahmoud, 30
Azuma, Ai, 112

Badryzlova, Yulia, 109
Baird, Cheyanne, 99
Bairi, Ramakrishna, 96
Bakliwal, Akshat, 113
Balage Filho, Pedro, 100
Balahur, Alexandra, 99, 120
Baldridge, Jason, 21
Baldwin, Tim, 75
Baldwin, Timothy, 72, 79, 92, 95–97, 101
Bandyopadhyay, Sivaji, 94
Banea, Carmen, 79, 92
Bangalore, Srinivas, vii, 23
Barbosa, Denilson, 52
Barbu, Eduard, 118
Barker, Fiona, 105
Barner, Dave, 109
Baroni, Marco, 72, 75
Barrón-Cedeño, Alberto, 78, 89
Barzilay, Regina, 51
Basile, Pierpaolo, 78, 90
Basile, Valerio, 120
Basili, Roberto, 77, 82, 98, 100
Baskaya, Osman, 97, 101
Bastianelli, Emanuele, 100
Batra, Arpita, 95
Baugh, Wesley, 99
Bautista, Susana, 118
Bean, David, vii
Becerra, Claudia, 79, 91, 93–95, 97, 100
Béchet, Frederic, 99
Becker, Lee, 97, 98
Behera, Satyabrata, 96

Beigman Klebanov, Beata, 109, 118
Bejček, Eduard, 112
Bellmore, Amy, 44
Bellot, Patrice, 99
Beltagy, Islam, 77, 81
Bentivogli, Luisa, 93, 94, 97, 100
Berend, Gabor, 100
Berg-Kirkpatrick, Taylor, 64
Bergler, Sabine, 79, 91
Bergsma, Shane, 61, 75
Berlanda, Sara, 112
Besançon, Romaric, 98
Bestgen, Yves, 106
Bethard, Steven, 78, 90, 93, 97, 100
Bhattacharyya, Pushpak, 46, 79, 92, 99
Bicici, Ergun, 79, 92, 100
Biemann, Chris, 60, 65, 75, 93, 94, 117
Bies, Ann, 32
Bigham, Jeffrey P., 23
Bikel, Dan, vi
Bilmes, Jeff, 45
Bisk, Yonatan, 65
Björne, Jari, 101
Black, Donald, 109
Blacoe, William, 51
Blanchard, Daniel, 105
Blanchon, Hervé, 96
Blanco, Eduardo, 75, 80, 86
Blunsom, Phil, vii, 58
Bobic, Tamara, 101
Bobicev, Victoria, 106
Bodoia, Max, 63
Bokharaeian, Behrouz, 101
Boleda, Gemma, 75, 77, 81
Bond, Francis, 75, 95
Boot, Peter, 117
Bosca, Alessio, 100
Bosch, Antal Van den, 119
Bosch, Antal van den, 95
Bouamor, Houda, 30
Boudin, Florian, 26
Boyce, Richard D., 101
Boyd-Graber, Jordan, vii, 39
Boyer, Kristy, vi
Boz, Umit, 109
Bracewell, David, 109
Brevik, Mikael, 98
Brew, Chris, vi, 97, 100
Briscoe, Ted, 28
Broadwell, George Aaron, 109
Brooke, Julian, 44, 106, 117
Brown, Gavin, 94
Buitelaar, Paul, 35, 104
Bunescu, Razvan, 77, 81
Bungum, Lars, 77, 83, 98, 111
Buntine, Wray, 22
Burchfiel, Benjamin, 44
Burger, John D., 105
Burgos, Diego A., 112
Burstein, Jill, 106, 112, 118
Buscaldi, Davide, 78, 90
Buy Wenniger, Gideon Maillette de, 108
Bykh, Serhiy, 106
Byrne, Lorna, 94

Çelebi, Arda, 45, 100
Cahill, Aoife, 31, 105, 107
Cai, Qingqing, 80, 87
Caliskan, Aylin, 34, 104
Callison-Burch, Chris, 31, 47, 52
Camara, Franc, 96, 98, 101
Can, Matthew, 53
Caputo, Annalina, 78, 90
Carbonell, Jaime, 25
Cardie, Claire, 31, 79, 92
Carenini, Giuseppe, 22
Carpuat, Marine, vii, 95, 105, 108
Carterette, Ben, 78, 90
Cases, Ignacio, 109
Castañeda, Yenier, 96
Castellanos, Malu, 61
Castelli, Vittorio, 52
Castellucci, Giuseppe, 98
Ceballo, Alberto, 101
Celikyilmaz, Asli, 29
Cer, Daniel, 77, 82
Chahuneau, Victor, 43, 67, 107
Chai, Joyce, vii
Chalothorn, Tawunrat, 98
Chambers, Nate, 94, 98
Chan, Joel, 37, 104
Chang, Angel, 94
Chang, Ming-Wei, 61
Chappelier, Jean-Cedric, 117
Charniak, Eugene, 20, 115
Chau, Cuong, 77, 81
Chávez, Alexander, 78, 88, 94
Chawla, Karan, 99
Che, Wanxiang, 19
Chen, Boxing, 58
Chen, John, 23
Chen, Lei, 105, 107
Chen, Wei-Te, 38
Chen, Zhiyuan, 61
Chengalvarayan, Rathinavelu, 23
Cherry, Colin, vi, 18, 36, 104
Cheung, Jackie Chi Kit, 51
Chikayama, Takashi, 94
Cho, Kit, 109
Chodorow, Martin, 107
Choi, Yoonjung, 79, 92

Chowdhury, Md. Faisal Mahbub, 48, 97, 99, 101
Christensen, Janara, 66
Cimino, Andrea, 106
Cirik, Volkan, 97, 101
Clark, Peter, 52, 97, 100
Clark, Sam, 98
Clements, Mark, 120
Clemmer, Alex, 2
Codina-Filbà, Joan, 99
Cohen, K. Bretonnel, 101
Cohen, Shay B., 21, 31
Cohen, Shay, 10
Cohen, William W., 44
Collazo, Armando, 78, 88, 94, 101
Collins, Michael, 10, 21, 31
Conrado, Merley, 35, 104
Cook, Paul, 75, 79, 80, 84, 85, 92, 96, 97, 101, 111
Costa-jussà, Marta R., 12
Couto, Francisco M, 101
Couto, Francisco, 99
Cranenburgh, Andreas van, 117
Crawford, Jean, 99
Croce, Danilo, 77, 82, 98, 100
Crossley, Scott, 106
Crowell, Richard, 99
Crowley, Jennifer, 113
Cruys, Tim Van de, 65, 94, 95

Dagan, Ido, 46, 97, 100
Dahlmeier, Daniel, 105
Dai, Jianmin, 106
Dalton, Adam, 109
Dan, Avishek, 79, 92
Dandapat, Sandipan, 38
Dang, Hoa Trang, 97, 100
Dara, Aswarth Abhilash, 38
Darwish, Kareem, 119
Daudaravicius, Vidas, 105
Daumé III, Hal, vi
Daumé III, Hal, 32
Dávila, Héctor, 78, 88, 94
Deemter, Kees van, 66
Dehkharghani, Rahim, 99
Del-Olmo, Maria, 111
Dell' Amerlina Ríos, Matías, 119
Dell'Orletta, Felice, 106
Delmonte, Rodolfo, 114
Demberg, Vera, vii
DeNero, John, 67
Deng, Lingjia, 79, 92
Derczynski, Leon, 93
Dermouche, Mohamed, 98
DeVault, David, 63
Diab, Mona, 46, 72, 75, 77, 82
DIAZ, ALBERTO, 101
Dickinson, Markus, 105, 106, 113
Dini, Luca, 100
Dolan, Bill, vii, 29
Dong, Daxiang, 33
Downey, Doug, vii, 33
Doyle, Gabriel, 20
Dras, Mark, 106
Dredze, Mark, vii, 22, 38, 44, 48, 61
Dreyer, Markus, vi
D'Souza, Jennifer, 53
Du, Lan, 22
Duh, Kevin, 58
Dunn, Jonathan, 109
Dunnion, John, 94
Durrani, Nadir, 18
Durrett, Greg, 67
Dyer, Chris, vi, 25, 28, 43, 67, 107, 124
Dzikovska, Myroslava, 97, 100, 107
Dziurzyński, Lukasz, 80, 86

Eichstaedt, Johannes, 80, 86
Eisenstein, Jacob, vi, 28, 49, 113
Ekbal, Asif, 94
Elchuri, Harsha, 99
Elliot, Kyle, 109
Ellman, Jeremy, 98
Elming, Jakob, 34
Erhart, George, 97, 98
Erk, Katrin, vi, 77, 81
Erp, Marieke van, 114
Escalante, Hugo Jair, 119
Eskander, Ramy, 29, 33
Eskenazi, Maxine, 118
Estrada, Rainel, 96
Etzioni, Oren, 66
Evanini, Keelan, 49, 106
Evert, Stefan, 78, 91, 98
Eyigöz, Elif, 18
Ezra, Jefferson, 34

Faghiri, Pegah, 111
Fangxi, Zhang, 98
Farkas, Richárd, 100
Farrow, Elaine, 107
Faruqui, Manaal, 107
Farzindar, Atefeh, 113
Feldman, Laurie, 109
Fenlon, Caroline, 94
Fernandes, Hilário, 99
Fernández Orquín, Antonio, 78, 88, 94, 96, 98, 101
Fernández, Antonio, 120
Filannino, Michele, 94
Filgueiras, João, 99
Filice, Simone, 98

Finin, Tim, 39, 77, 82
Flanigan, Jeffrey, 25
Flor, Michael, 109, 118
Florian, Radu, 52
Fluck, Juliane, 101
Fokkens, Antske, 114
Fonseca Bruzón, Adrian, 99
Forsyth, David, 75, 80, 84
Foster, Dean P., 10
Foster, Dean P., 21
Foster, George, 58
Foster, Jennifer, 113
Fournier, Chris, 117
Frank, Anette, 80, 87
Franky, Franky, 35, 104
Fraser, Alexander, 18
Freitag, Markus, 108
Friedman, Majorie, 109
Fucikova, Eva, 111
Fuentes, Maria, 78, 89
Furrer, Lenz, 97, 98

Gabbard, Ryan, 109
Gaikwad, Upasana, 96
Gal, Yarin, 58
Galescu, Lucian, 109
Gambäck, Björn, 77, 83, 98, 111
Ganitkevitch, Juri, 37, 47, 104
Gao, Jianfeng, 18, 30
Garcia Flores, Jorge J., 78, 90
García-Narbona, David, 99
Garrette, Dan, 21, 77, 81
Gaspari, Federico, 38
Gasser, Michael, 95, 96
Gayen, Vivekananda, 111
Gebre, Binyam Gebrekidan, 106
Gelbukh, Alexander, 79, 91, 93–95, 97, 100
Gella, Spandana, 79, 80, 84, 92
Genabith, Josef van, 38, 100
Gentner, Dedre, 109
Gershman, Anatole, 109
Gertz, Michael, 93, 94
Gervás, Pablo, 118
Gezici, Gizem, 99
Ghosh, Riddhiman, 61
Ghosh, Urmi, 95
Giampiccolo, Danilo, 93, 94, 97, 100
Gildea, Daniel, 18, 23
Giles, C. Lee, 25
Gimpel, Kevin, 28, 124
Ginsca, Alexandru, 98
Gleize, Martin, 100
Go, Junho, 52
Goebel, Randy, 52
Goldberg, Yoav, 80, 84
Goldwater, Micah, 109
Goldwater, Sharon, 62
Gómez, Helena, 94
Gomez, Manuel Móntes y, 119
Gómez, Manuel Montes-y, 79, 92, 119
Gompel, Maarten van, 95
Gondek, David, 48
González, Andy, 96, 98, 120
Gonzalez-Agirre, Aitor, 77, 78, 82, 89
Gormley, Matthew, 38
Gotti, Fabrizio, 113
Goulian, Jérôme, 96
Goutte, Cyril, 105
Goyal, Kartik, 109
Graham, Yvette, 95
Grau, Brigitte, 100
Gravano, Agustín, 31
Gravano, Agustin, 119
Greenstadt, Rachel, 34, 104
Grefenstette, Edward, 77, 81
Grego, Tiago, 101
Greiner, Paul, 78, 91, 98
Grieser, Karl, 79, 92
Grishman, Ralph, 48, 113
Grivolla, Jens, 99
Groves, Declan, 38
Guadarrama, Sergio, 115
Guerini, Marco, 99
Günther, Tobias, 97, 98
Guo, Junfei, 99
Guo, Stephen, 61
Guo, Weiwei, 46, 77, 82
Guo, Yufan, 53
Gupta, Narendra, 119
Gupta, Rajdeep, 94
Gurevych, Iryna, 65, 96, 97, 100
Gurrutxaga, Antton, 112
Gutiérrez, Yoan, 78, 88, 94, 96, 98, 99, 101, 120
Guzmán, Rafael, 119
Gyawali, Binod, 106

Habash, Nizar, 27, 29, 30, 33
Habernal, Ivan, 119
Haddow, Barry, 27
Hage, Willem Robert van, 114
Hagiwara, Masato, 39
Hahn, Michael, 100
Hailu, Negacy, 101
Hajic, Jan, 111
Hamdan, Hussam, 99
Hammond, Adam, 117
Han, Bo, 80, 84
Han, Lushan, 77, 82
Han, Qi, 99
Hancock, Jeffrey T., 31
Hangya, Viktor, 100

Hanneman, Greg, 26
Harabagiu, Sanda, vii
Harihara, Ganesh, 98
Hasan, Ragib, 113
Hashimoto, Chikara, 19
Hassan, Samer, 79, 92
Hauer, Bradley, 42, 106
Hayashibe, Yuta, 106
He, Hua, 27
He, Xiaodong, 30
Heafield, Kenneth, 58
Heilman, Michael, 78, 88, 97, 100
Heintz, Ilana, 109
Henderson, James, vii, 48
Henderson, John, 105
Hendrickx, Iris, 95
Herdagdelen, Amaç, 75
Hernández, Donato, 119
Herrero Zazo, María, 97, 100
Herrmann, Teresa, 108
Hervás, Raquel, 118
Heskes, Tom, 106
Hiemstra, Djoerd, 98
Higgins, Derrick, vi
Hinote, David, 109
Hirst, Graeme, 44, 106, 111, 117
Hisamoto, Sorami, 112
Hixon, Ben, 63
Hladka, Barbora, 106
Hockenmaier, Julia, vi, 65
Hoeksema, Jesper, 114
Hofmann-Apitius, Martin, 101
Höglin, Erik, 105
Holen, Gordana Ilic, 35, 104
Holub, Martin, 106
Hope, David, 101
Horbach, Andrea, 80, 86
Hoste, Véronique, 95, 96
Hou, Yufang, 53
Hovy, Dirk, 64, 109
Hovy, Eduard, 64, 109, 114
Hsu, Daniel, 124
Hsu, Meichun, 61
Hu, Yuening, 39
Huang, Liang, 28
Huang, Ruihong, 19
Huber, Torsten, 97, 101
Huck, Matthias, 108
Hughes, Mark, 113
Hunter, Lawrence E., 101
Huttunen, Silja, 114

Illouz, Gabriel, 106
Inkpen, Diana, 98
Ionescu, Radu Tudor, 107
Iosif, Elias, 78, 88
Irvine, Ann, 31, 36, 104
Isaeva, Yekaterina, 109
Ishikawa, Kai, 43

Jain, Harshit, 99
Jamrozik, Anja, 109
Jarvis, Scott, 106
Jauhar, Sujay Kumar, 109
Jautze, Kim, 117
Ježek, Karel, 111
Jha, Rahul, 105
Jiang, Jing, 29, 61
Jimenez, Sergio, 79, 91, 93–95, 97, 100
Johannsen, Anders, 34
Johansson, Richard, 21
Johnson, Mark, 22
Jong, Hayco de, 117
Joshi, Mahesh, 44
Joshi, Salil, 46
Joty, Shafiq, 22
Juárez, Antonio, 119
Judd, Joel, 30
Jung, Hyuckchul, 93, 94
Jurafsky, Dan, 117
Jurafsky, Daniel, 53, 63
Jurgens, David, 33, 96, 97, 101

Kabashi, Besim, 78, 91, 98
Kaeshammer, Miriam, 108
Kaewphan, Suwisa, 101
Kalita, Jugal, 30, 32
Kannan, Sampath, 21
Kanojia, Diptesh, 46
Karampatsis, Rafael Michael, 100
Karttunen, Lauri, 75
Kashefi, Elham, 51
Kawai, Takao, 19
Kazama, Jun'ichi, 19
Kazemzadeh, Abe, 98
Keller, Bill, 101
Kerimov, Ruslan, 109
Kern, Margaret L., 80, 86
Kern, Roman, 78, 89
Khan, Mohammad, 113
Khapra, Mitesh M., 26
Khouas, Leila, 98
Kiciman, Emre, 61
Kiela, Douwe, 78, 88
Kim, Mi-Young, 52
Kim, Young-Bum, 67
King, Ben, 64
King, Levi, 105
Kirchhoff, Katrin, 45
Kirchoff, Katrin, vi
Kiritchenko, Svetlana, 97, 99
Klakow, Dietrich, 32, 113

Klapaftis, Ioannis, 97, 101
Klerke, Sigrid, 34
Knight, Kevin, 76
Knowles, Rebecca, 38
Kochetkova, Natalia, 112
Koehn, Philipp, 58
Kökciyan, Nadin, 100
Kokhlikyan, Narine, 48
Kolomiyets, Oleksandr, 94, 97, 100
Kolya, Anup Kumar, 94
Komachi, Mamoru, 106
Kondo, Shuhei, 112
Kondrak, Grzegorz, 36, 42, 104, 106
Koolen, Corina, 117
Kopotev, Mikhail, 112
Kordjamshidi, Parisa, 97, 100
Korhonen, Anna, 53, 65, 75
Korkontzelos, Ioannis, 93, 94
Kosseim, Leila, 94
Kouhestani, Manouchehr, 26
Kouse, Tomoya, 112
Kouylekov, Milen, 100
Kozareva, Zornitsa, 95, 97, 98
Kozhevnikov, Mikhail, 80, 87
Krishnamoorthy, Niveda, 115
Krivanek, Julia, 106
Kriz, Vincent, 106
Kroch, Anthony, 32
Krüger, Nina, 119
Krčmář, Lubomír, 111
Kuebler, Sandra, 113
Kuhn, Roland, 58
Kuipers, Benjamin, 115
Kulick, Seth, 32
Kumar Naskar, Sudip, 38
Kundu, Amitava, 94
Kunneman, Florian, 119
Kuperberg, Gina R, 16
Kuperberg, Gina R., 16
Kushman, Nate, 51
Kyle, Kristopher, 106

L. Kashyap, Abhay, 77, 82
Lacalle, Oier Lopez de, 100
Lahiri, Shibamouli, 106
Lai, Po-Hsiang, 106
Lam, Khang Nhut, 32
Lamb, Alex, 48
Lambert, Patrik, 99
Lan, Man, 78, 89, 94, 98
Lang, Joel, 48
Langlais, Philippe, 113
Laokulrat, Natsuda, 94
Lapalme, Guy, vii
Lapata, Mirella, 51
Lapponi, Emanuele, 34
Lasecki, Walter, 23
Lau, Jey Han, 96, 97, 101
Lavelli, Alberto, 48, 97, 99, 101
Lavergne, Thomas, 106
Lavie, Alon, vii, 26, 58
Lawrie, Dawn, 39
Le Roux, Joseph, 78, 90
Leacock, Claudia, 97, 100
Lee, Heeyoung, 23
Lefever, Els, 95, 96
Léger, Serge, 105
Lehr, Maider, 23
Lentini, Jennifer, 118
León, Saul, 94
Leser, Ulf, 97, 101
Lessard, Greg, 117
Levallois, Clement, 98
Levison, Michael, 117
Levitan, Rivka, 36, 104
Levy, Omer, 97, 100
Levy, Roger, 20
LI, Baoli, 106
Li, Huying, 109
Li, Junhui, 32
Li, Ru, 77, 83
Li, Tao, 66
Liang, Percy, vii
Liberman, Mark, 45
Liebrecht, Christine, 119
Lin, Ching-Sheng, 113
Lin, Jimmy, vi, 27
Lin, Yibin, 118
Litman, Diane, 37, 49, 104
Liu, Bing, 61
Liu, Can, 95, 96
Liu, Fei, 66
Liu, Hongfang, 78, 90
Liu, Ting, 19, 109
Liu, Yang, 27, 49, 106
Liu, Ying, 39
Liu, Yuzong, 45
Ljolje, Andrej, 23
Llorens, Hector, 93
Loo, Kaidi, 105
Lopez, Adam, 27
Lotan, Amnon, 46
Loudcher, Sabine, 98
Louis, Annie, vi
Lu, Qin, 78, 88
Lui, Marco, 79, 92
Lukin, Stephanie, 113
Luo, Wencan, 37, 104
Luo, Xiaoqiang, 52
Lynum, André, 77, 83, 107, 111

Ma, Wei-Yun, 29

Ma, Yi, 37, 104
Maamouri, Mohamed, 32
Madnani, Nitin, 31, 78, 88, 97, 100, 106, 107
Magnini, Bernardo, 75
Makhoul, John, 34
Makrynioti, Konstantina, 100
Malakasiotis, Prodromos, 100
Malandrakis, Nikolaos, 78, 88, 98
Malkarnenkar, Girish, 115
Malmasi, Shervin, 106
Manandhar, Suresh, 74, 75
Manion, Steve L., 96
Manning, Christopher D., 2, 19, 76, 94
Manrique-Losada, Bell, 112
Mansour, Saab, 43
Marchand, Morgane, 98
Marchetti, Alessandro, 93, 94
Marcus, Mitchell, 21
Maritxalar, Montse, 100
Markert, Katja, 53
Markiewicz, Gretchen, 34
Marlin, Benjamin M., 19
Marneffe, Marie-Catherine de, 42
Màrquez, Lluís, 78, 89
Marsi, Erwin, 77, 83, 111
Martín-Valdivia, M. Teresa, 98, 120
Martín-Valdivia, Maria Teresa, 118
Martinez Alonso, Hector, 34
Martínez, Paloma, 97, 100
Martínez-Cámara, Eugenio, 98, 120
Martins, Andre, 124
Martins, Bruno, 99
Marton, Yuval, 75
Maskey, Sameer, 52
Mason, Rebecca, 36, 104, 115
Masuko, Soh, 39
Materna, Jiří, 31
Matsoukas, Spyros, 34
Matsumoto, Yuji, 106, 112
Matula, Valentine, 97, 98
Mausam, 66
Max, Aurélien, 106
Mayfield, James, 39, 77, 82
McCallum, Andrew, 19, 124
McCarthy, Diana, 75
McCoy, Patricia Ann, 107
McDonald, Ryan, 62
McInnes, Bridget, 39
McKenzie, Amber, 29
McKeown, Kathleen, 29, 56
McKeown, Kathy, 99
McNamara, Danielle, 106
McNamee, Paul, 39
Medero, Julie, 45
Meek, Christopher, 60
Mehdad, Yashar, 22, 93, 94
Mehta, Sameep, 6
Melton-Meaux, Genevieve, 39
Mendes, Amália, 112
Mesnard, Olivier, 98
Meurers, Detmar, 100, 106
Meyers, Adam, 113
Micciulla, Linnea, 99
Michaud, Lisa, 107
Mihalcea, Rada, 77, 79, 81, 92, 106
Mikolov, Tomas, 46, 60
Min, Bonan, 48
Minker, Wolfgang, 33
Mirroshandel, Seyed Abolghasem, 25
Mitamura, Teruko, 114
Mitchell, Margaret, 66
Mittal, Namita, 120
Mittelman, Roni, 115
Miwa, Makoto, 94
Mizan, Mainul, 113
Mizumoto, Tomoya, 106
Moen, Hans, 77, 83
Moens, Marie-Francine, 20, 94, 97, 100
Mogadala, Aditya, 99
Mohammad, Saif, 97, 99
Mohit, Behrang, 30, 43
Mohler, Michael, 79, 92, 109
Molina-González, M. Dolores, 120
Moloodi, Amirsaeid, 26
Montejo-Ráez, Arturo, 98
Montemagni, Simonetta, 106
Montoyo, Andrés, 78, 88, 94, 96, 98, 101, 120
Montoyo, Andres, 99
Mooney, Raymond, 77, 81, 115, 124
Moore, Johanna, vii, 107
Morante, Roser, 75
Moreira, Silvio, 99
Moreno-Ortiz, Antonio, 111
Morin, Emmanuel, 26
Morley, Eric, 105
Moschitti, Alessandro, 77, 82
Mosquera, Alejandro, 98
Mott, Justin, 32
Moulder, Brad, 118
Mourad, Ahmed, 119
Mukomel, Elena, 109
Muller, Philippe, 94, 95
Müller, Stefan, 80, 85
Muñoz, Rafael, 78, 88, 94, 96, 98, 101, 120
Muresan, Smara, 75
Musat, Claudiu, 119
Mustafayev, Elshan, 120

Nagata, Ryo, 106
Naim, Iftekhar, 23
Nakov, Preslav, 75, 95, 97, 98
Napolitano, Diane, 31, 118

Narasimhan, Bhuvana, 112
Narayanan, Shrikanth, 78, 88, 98
Nardi, Daniele, 100
Nasiruddin, Mohammad, 96
Nasr, Alexis, 25
Navigli, Roberto, 64, 75, 95, 96
Negi, Sapna, 99
Negri, Matteo, 93–95
Nenadic, Goran, 94
Nenkova, Ani, vii
Neves, Mariana, 101
Ney, Hermann, 43, 108
Ng, Hwee Tou, 75
Ng, Hwee Tou, 30, 105
Ng, Raymond T., 22
Ng, Vincent, 53, 106
Nguyen, Viet-An, 39
Nicolai, Garrett, 106
Nicosia, Massimo, 77, 82
Niehues, Jan, 108
Nielsen, Rodney, 97, 100
Nissim, Malvina, 75, 111, 112
nissim, malvina, 120
Niu, Zheng-Yu, 94
Nivre, Joakim, vii, 62
Noeman, Sara, 78, 91

Ó Séaghdha, Diarmuid, 95
Oard, Douglas, 39
Oates, Tim, 39
O'Brien, Ron, 113
O'Connor, Brendan, 28
Oflazer, Kemal, 18, 30, 43
OKOYE, IFEYINWA, 100
Omuya, Adinoyi, 49
Onishi, Takashi, 43
Ordan, Noam, 107
O'Reilly, Tenaha, 106
Ortega Bueno, Reynier, 99
Orwant, Jon, 80, 84
Ostendorf, Mari, 45
Östling, Robert, 105
Otrusina, Lubomir, 78, 89
Ott, Myle, 31
Ott, Niels, 100
Owoputi, Olutobi, 28
Özgür, Arzucan, 100

Pakhomov, Serguei, 39
palanisamy, Prabu, 99
Palmer, Alexis, 80, 86
Palmer, Martha, 8, 75, 76, 111, 112
Pang, Bo, vii
Pantel, Patrick, vi
Pardo, Thiago, 35, 100, 104
Parish, Tim, 105
Park, Se-Young, 52
Park, Seong-Bae, 52
Passonneau, Becky, 75
Passonneau, Rebecca J., 63
Pate, John K, 62
Paul, Michael J., 22, 48
Paul, Soma, 95
Pavlopoulos, John, 100
Pecina, Pavel, 111, 112
Pedersen, Ted, 39, 96
Pepper, Steve, 106
Pérez, Roger, 96, 98, 101
Perez-Hernandez, Chantal, 111
Petrov, Slav, 124
Pfeifer, Craig, 105
Philpot, Andrew, 114
Picard, Rosalind, 119
Pilehvar, Mohammad Taher, 64
Pinkal, Manfred, 80, 86
Pinto, David, 94
Pinto, Francisco, 101
Pitler, Emily, 21
Piug, Dennys D., 96
Pivovarova, Lidia, 112, 114
Poibeau, Thierry, 65
Polajnar, Tamara, 78, 88
Poon, Hoifung, 51
Popescu, Adrian, 78, 90
Popescu, Marius, 107
Post, Matt, vi, 2
Potamianos, Alexandros, 78, 88, 98
Potts, Christopher, 42, 63
Poursepanj, Hamid, 98
Power, Richard, 118
Prabhakaran, Vinodkumar, 49
Prasad, Rashmi, 101
Preiss, Judita, 38, 44, 77, 83
Proisl, Thomas, 78, 91, 98
Prokopi, Vassiliki, 78, 88
Prud'hommeaux, Emily, 23, 45
Ptáček, Tomáš, 119
Pu, Pearl, 119
Puig, Dennys D., 101
Puil, Jennifer van der, 113
Pustejovsky, James, 93

Qadir, Ashequl, 119
Qian, Xian, 27, 49
Qiu, Minghui, 29, 61
Quinn, Kevin, 52
Quirk, Chris, 12, 18

Radev, Dragomir, 34, 105
Raghavan, Hema, 52
Ragheb, Marwa, 106
Ramakrishnan, Ganesh, 96

Ramanathan, Ananthakrishnan, 26
Rambow, Owen, 29, 33, 49, 76
Ramirez, Gabriela, 106
Ramisch, Carlos, 112
Ramones, Stephanie, 80, 86
Ramteke, Ankit, 99
Rangarajan Sridhar, Vivek Kumar, 23
Rasooli, Mohammad Sadegh, 26
Rassmussen, Priscilla, vi
Rastegar-Mojarad, Majid, 101
Rauscher, Janneke, 117
Ravi, Sujith, vi, 124
Ravishankar, Veena, 113
Recasens, Marta, 42, 53
Reckman, Hilke, 99
Rei, Marek, 28
Reichart, Roi, 53
Reiter, Ehud, 66
Remus, Robert, 98
Remus, Steffen, 60
Resnik, Philip, 32, 39
Rezende, Solange, 35, 104
Riedel, Sebastian, 19
Riedl, Martin, 117
Rigau, German, 78, 89
Riloff, Ellen, 19, 119
Rimell, Laura, 75, 78, 88
Ringger, Eric, vi
Ritter, Alan, 97, 98, 113
Roark, Brian, 23, 45, 105
Rocktäschel, Tim, 97, 101
Rodríguez Zivic, Pablo, 31
Rodriguez, Horacio, 78, 89
Rodriguez-Penagos, Carlos, 99
Roller, Stefan, 80, 85
Roller, Stephen, 111
Rosé, Carolyn P., 44
Rosenthal, Sara, 97–99
Rosner, Michael, 99
Rosso, Paolo, 79, 92, 119
Roth, Dan, 124
Roth, Dan, 60, 76
Roth, Michael, 80, 87
Roth, Ryan, 29, 30
Rouhizadeh, Masoud, 45
Rudnick, Alex, 95, 96
Ruppenhofer, Josef, 32
Rustamov, Samir, 120

Sabatini, John, 106, 118
Sachan, Mrinmaya, 109
Saeger, Stijn De, 19
Saenko, Kate, 115
Saers, Markus, 108
Saggion, Horacio, 25
Sagi, Eyal, 109
Sagot, Benoît, 25
Saias, José, 99
Sainudiin, Raazesh, 96
Sakaguchi, Keisuke, 106, 112
Salakoski, Tapio, 101
Salameh, Mohammad, 36, 104, 106
Salehi, Bahar, 79, 80, 85, 92, 95
Salloum, Wael, 27
Samvelian, Pollet, 111
Sanches Duran, Magali, 112
Sanchez-Cisneros, Daniel, 100
Sánchez-Vega, Fernando, 79, 92
Sanders, Whitney, 109
Sankaran, Baskaran, 58
Santen, Jan van, 45, 105
Santorini, Beatrice, 32
Saraçlar, Murat, 45
Sarkar, Anoop, 58
Sarkar, Kamal, 111
Satta, Giorgio, 31
Saurí, Roser, 99
Savarese, Silvio, 115
Sayeed, Asad, 44
Saygin, Yucel, 99
Saz, Oscar, 118
Scarton, Carolina Evaristo, 112
Scheible, Silke, 111
Schijndel, Marten van, 20
Schmid, Helmut, 18
Schmitt, Alexander, 33
Schneider, Nathan, 28, 43, 107
Schuetze, Hinrich, 99
Schuler, William, 20
Schwab, Didier, 96
Schwartz, Hansen Andrew, 80, 86
Schwartz, Richard, 34
Segura-Bedmar, Isabel, 97, 100
Seligman, Martin, 80, 86
Selmer, Øyvind, 98
Semeraro, Giovanni, 78, 90
Serafini, Luciano, 114
Sérasset, Gilles, 96
Sert, Enis, 97, 101
Sethi, Saratendu, 99
Setiawan, Hendra, 27
Severyn, Aliaksei, 77, 82
Shafran, Izhak, 23
Shafran, Zak, vii
Shaikh, Samira, 109, 113
Shareghi, Ehsan, 79, 91
Shaw, Ryan, 114
Sheehan, Kathleen M., 118
Shekhtman, Natalia, 109
Shen, Chao, 66
Shen, Hui, 77, 81
Shigeto, Yutaro, 112

Shore, Jane, 118
Shriberg, Elizabeth, 23
Shrivastava, Shashank, 109
Shutova, Ekaterina, 60, 80, 85
Siblini, Reda, 94
Sidorenko, Wladimir, 119
Silcox, Lucian, 32
Silva, Mário J., 99
Sima'an, Khalil, 108
Sindlerova, Jana, 111
Singh, Thoudam Doren, 108
Sintsova, Valentina, 119
Siskind, Jeff, 124
Sizov, Gleb, 77, 83
Skiba, David, 97, 98
Smith, Noah A., 28, 43, 67
Smola, Alexander, 124
Smolentzov, André, 105
Smrz, Pavel, 78, 89
Snyder, Benjamin, 67
Snyder, Justin, 38
Socher, Richard, vi, 2, 124
Soderland, Stephen, 66
Søgaard, Anders, 34, 43
Solorio, Thamar, 106, 113
Song, Hyun-Je, 52
Sonntag, Jonathan, 119
Sporleder, Caroline, 75
Sprugnoli, Rachele, 114
Srikumar, Vivek, 60, 124
Srivastava, Mahesh, 109
Stede, Manfred, 119
Stede, M., 4
Steinberger, Josef, 119
Stent, Amanda, 93, 94
Stern, Asher, 46
Stevenson, Mark, 22, 38, 44, 77, 78, 83, 89
Stieglitz, Stefan, 119
Stolcke, Andreas, 23, 45
Stolerman, Ariel, 34, 104
Storch, Valerio, 77, 82
Stoyanov, Veselin, 97, 98
Stranak, Pavel, 112
Stratos, Karl, 10, 21
Stromer-Galley, Jennifer, 113
Strötgen, Jannik, 93, 94
Strube, Michael, 53
Strzalkowski, Tomek, 109, 113
Styler, Will, 38
Subramaniam, L. Venkata., 6
Sultan, Md., 78, 90
Sumner, Tamara, 78, 90, 100
Sun, Lin, 60
Sun, Min, 115
Surtani, Nitesh, 95
Swanson, Ben, 20, 106
Swiezinski, Leonard, 117
Szarvas, György, 65
Szpakowicz, Stan, 95

Täckström, Oscar, 62
Tan, Liling, 95
Tanev, Hristo, 94
Tapucu, Dilek, 99
Taylor, Sarah, 109
Tchechmedjiev, Andon, 96
Tetreault, Joel, 31, 105
Thomas, Philippe, 101
Thomason, Jesse, 49
Tian, Hao, 33
Tiantian, Zhu, 78, 89, 98
Titov, Ivan, 8, 12, 80, 87
Tomai, Emmett, 32
Tomeh, Nadi, 29, 33
Tomlinson, Marc, 109
Tonelli, Sara, 99, 114
Toral, Antonio, 38
Torisawa, Kentaro, 19
Tounsi, Lamia, 113
Toutanova, Kristina, 18
Traum, David, vii, 63
Trevisan, Marco, 100
Trivedi, Rakshit, 49
Tsuchida, Masaki, 43
Tsuruoka, Yoshimasa, 94
Tsvetkov, Yulia, 107, 109
Turchi, Marco, 95
Turmo, Jordi, 78, 89
Turney, Peter, vii
Twitto, Naama, 107
Tyrefors Hinnerich, Björn, 105

Ultes, Stefan, 33
Ungar, Lyle, 10, 21, 80, 86
Ureña-López, L. Alfonso, 98, 120
Ureña-López, Luis Alfonso, 118
Uresova, Zdenka, 111
Üsküdarlı, Suzan, 100
UzZaman, Naushad, 93

Vaidya, Ashwini, 112
Vajjala, Sowmya, 105, 106
Van Durme, Benjamin, 47, 52, 61
Van Genabith, Josef, 79, 92
Vanderwende, Lucy, vi, 51, 114
Vannella, Daniele, 95, 96
Varma, Vasudeva, 99
Vaswani, Ashish, 64
Veale, Tony, 95
Velcin, Julien, 98
Venturi, Giulia, 106
Verdejo, Felisa, 114

Veress, Fruzsina, 99
Verhagen, Marc, 93
Versley, Yannick, 95
Vilar, David, 108
Vilariño, Darnes, 94
Villaseñor, Luis, 119
Villaseñor-Pineda, Luis, 79, 92
Villatoro-Tello, Esaú, 119
Violante, Luisina, 31
Visweswariah, Karthik, 26
Vogel, Adam, 63
Voigt, Rob, 117
Volkova, Svitlana, 29
Vossen, Piek, 114
Vreeke, Joris, 38
Vulić, Ivan, 20

Waibel, Alex, 48, 108
Walde, Sabine Schulte im, 80, 85, 111
Walker, Marilyn, 113
Wan, Li, 48
Wang, Chang, 48
Wang, Mengqiu, 19
Wang, Pidong, 30
Wang, Ruibo, 77, 83
Wang, Sai, 77, 83
Wang, Wen, 45
Wang, Xinhao, 49, 105
Wang, Zhiqiang, 77, 83
Wartena, Christian, 93, 94
Weese, Jonathan, 77, 82
Wei, Kai, 45
Weidlich, Michael, 97, 101
Weischedel, Ralph, 109
Weissbock, Josh, 98
Weng, Fuliang, 66
Wicentwoski, Rich, 98
Wiebe, Jan, 79, 92
Wiegand, Michael, 32, 113
Wilks, Yorick, 109
Williams, Sandra, 118
Wilson, Theresa, vii, 61, 97, 98
Wintner, Shuly, 107
Wittenburg, Peter, 106
Wolfe, Travis, 38
Wolff, Susanne, 107
Wombacher, Andreas, 98
Wong, Sze-Meng Jojo, 106
Wu, Ching-Yi, 106
Wu, Dekai, 108
Wu, Shumin, 8
Wu, Siew Mei, 105
Wu, Stephen, 78, 90
Wu, Zhaohui, 25

Xia, Fei, 114
Xie, Shasha, 106, 107
Xiong, Wenting, 36, 104
Xu, Jian, 78, 88
Xu, Jun-Ming, 44
Xu, Tan, 39
Xu, Wei, 113
Xu, Ying, 52
Xue, Nianwen, 62

Yadav, Vineet, 99
Yamrom, Boris, 109
Yan, Yulan, 19
Yang, Bishan, 79, 92
Yang, Eugene, 98
Yang, Liu, 29
Yang, Yaqin, 62
Yang, Yi, 33
Yangarber, Roman, 112, 114
Yanikoglu, Berrin, 99
Yannakoudakis, Helen, 105
Yao, Lei, 106
Yao, Limin, 19
Yao, Xuchen, 52
Yarowsky, David, 61
Yates, Alexander, 33, 80, 87
Yatskar, Mark, 29
Yeh, Eric, 78, 90
Yetisgen-Yildiz, Meliha, 114
Yih, Scott, 124
Yih, Wen-tau, 46, 60
Yoshimoto, Akifumi, 112
Yu, Dianhai, 33
Yu, Mo, 33
Yuan, Jiahong, 45
Yung, Frances, 112
Yuret, Deniz, 74, 75, 97, 101

Zampieri, Marcos, 106
Zaninello, Andrea, 112
Zanzotto, Fabio Massimo, 75
Zanzotto, Fabio Massimo, 93, 94
Zapata-Jaramillo, Carlos M., 112
Zarrella, Guido, 105
Zavarella, Vanni, 94
Zbib, Rabih, 34
Zechner, Klaus, 49, 105, 106
Zell, Julian, 93, 94
Zesch, Torsten, 93, 94, 97, 100
Zettlemoyer, Luke, vi, 29, 51
Zhang, Hui, 18
Zhang, Joy Ying, 48
Zhang, Xia, 77, 83
Zhang, Yuqi, 48
Zhao, Jiang, 94
Zhao, Kai, 28
Zhao, Tiejun, 33

Zhila, Alisa, 60
Zhou, Bowen, 27
Zhu, Dongqing, 78, 90
Zhu, Xiaodan, 97, 99
Zhu, Xiaojin, 44
Zhu, Zhemin, 98
Ziai, Ramon, 100
Ziak, Hermann, 78, 89
Zorn, Hans-Peter, 96
Zulu, Peleira Nicholas, 26
Zweig, Geoffrey, 46, 60

ETS

Listening. Learning. Leading.®

At ETS, education is more than our business. It's our mission.

As part of our mission, we administer over 50 million tests, such as the *TOEFL*® and *GRE*® assessments, to help advance quality and equity in education worldwide. The Natural Language Processing & Speech Group at ETS is integral to fulfilling that mission. It develops systems to provide reliable and valid analysis of text-based and spoken responses, as well as student and teacher support tools.

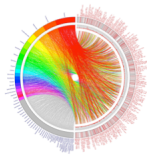
Learn more at: **www.ets.org/research/capabilities/automated_scoring**

Copyright © 2013 by Educational Testing Service. All rights reserved. ETS, the ETS logo, LISTENING. LEARNING. LEADING., GRE and TOEFL are registered trademarks of Educational Testing Service (ETS). 22841

AI @ ISI

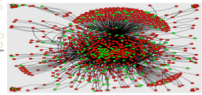
Natural Language Processing

Discovery Informatics



Information Integration

Social Networks



Join us!

Sabbatical Researcher
Visitor • PhD Student
Research Fellow
Research Scientist
Programmer • Summer
Intern • Research Faculty

And Much More!

USC

Appen Butler Hill provides high-quality, cost-effective solutions for expanding language technologies into global markets. We support the development of NLP systems, providing linguistic data and expertise, data collections, transcriptions and search engine relevance evaluations.
Please visit us at: http://www.appenbutlerhill.com

Presenting

ATLANTA

NAACL HLT 2013 gratefully acknowledges the following sponsors for their support.

Gold Level

Bronze Level

Student Best Paper Award and Student Lunch Sponsor

Student Volunteer Sponsor

Conference Bags