

Technical Program Schedule

Monday, May 3

9:00 – 9:15 Conference Welcome

9:15 – 10:30 **Keynote Speaker I: Andrei Broder**
Ten Years of Web Search Technology

10:30-11:00 Break

Information Extraction I

11:00 – 11:25 *A Statistical Model for Multilingual Entity Detection and Tracking*
R Florian, H Hassan, A Ittycheriah, H Jing, N Kambhatla, X Luo, N Nicolov and S Roukos

11:25 – 11:50 *Cross-Document Coreference on a Large Scale Corpus*
Chung Heong Gooi and James Allan

11:50 – 12:15 *Robust Reading: Identification and Tracing of Ambiguous Names*
Xin Li, Paul Morie and Dan Roth

Dialogue and Discourse

11:00 – 11:25 *A Saliency-Based Approach to Gesture-Speech Alignment*
Jacob Eisenstein and C. Mario Christoudias

11:25 – 11:50 *Balancing data-driven and rule-based approaches in the context of a Multimodal Conversational System*
Srinivas Bangalore and Michael Johnston

11:50 – 12:15 *Optimizing Automated Call Routing by Integrating Spoken Dialog Models with Queuing Models*
Tim Paek and Eric Horvitz

12:15 – 1:30 Lunch

Question Answering I

1:30 – 1:55 *Answering Definition Questions with Multiple Knowledge Sources*
Wesley Hildebrandt, Boris Katz and Jimmy Lin

1:55 – 2:20 *Automatic Question Answering: Beyond the Factoid*
Radu Soricut and Eric Brill

2:20 – 2:45 *A Probabilistic Rasch Analysis of Question Answering Evaluations*
Rense Lange, Juan Moran, Warren R Greiff and Lisa Ferro

2:45 – 3:10 *Acquiring Hyponymy Relations from Web Documents*
Keiji Shinzato and Kentaro Torisawa

Syntax and Semantics I

- 1:30 – 1:55 *Sentence-Internal Prosody Does not Help Parsing the Way Punctuation Does*
Michelle Gregory, Mark Johnson and Eugene Charniak
- 1:55 – 2:20 *Ensemble-based Active Learning for Parse Selection*
Miles Osborne and Jason Baldridge
- 2:20 – 2:45 *Speed and Accuracy in Shallow and Deep Stochastic Parsing*
Ron Kaplan, Stefan Riezler, Tracy H King, John T Maxwell III, Alex Vasserman and Richard Crouch
- 2:45 – 3:10 *Training Tree Transducers*
Jonathan Graehl and Kevin Knight
- 3:15 – 4:00 Award and Demo
- 4:00 – 4:15 Break
- 4:15 – 6:00 Poster and Demo Session

Tuesday, May 4

Human Language Technology

- 9:00 – 9:25 *Catching the Drift: Probabilistic Content Models, with Applications to Generation and Summarization*
Regina Barzilay and Lillian Lee
- 9:25 – 9:50 *The Web as a Baseline: Evaluating the Performance of Unsupervised Web-based Models for a Range of NLP Tasks*
Mirella Lapata and Frank Keller
- 9:50 – 10:15 *Lattice-Based Search for Spoken Utterance Retrieval*
Murat Saraclar and Richard Sproat
- 10:15-10:45 Break

Speech, Summarisation and Discourse

- 10:45 – 11:10 *Detecting Structural Metadata with Decision Trees and Transformation-Based Learning*
Joungbum Kim, Sarah E Schwarm and Mari Ostendorf
- 11:10 – 11:35 *Evaluating Content Selection in Summarization: The Pyramid Method*
Ani Nenkova and Rebecca Passonneau
- 11:35 – 12:00 *Inferring Sentence-internal Temporal Relations*
Mirella Lapata and Alex Lascarides

Machine Translation I

- 10:45 – 11:10 *A Smorgasbord of Features for Statistical Machine Translation*
Franz Josef Och, Daniel Gildea, Sanjeev Khudanpur, Anoop Sarkar, Kenji Yamada, Alex Fraser
Shankar Kumar, Libin Shen, David Smith, Katherine Eng, Viren Jain, Zhen Jin and Dragomir Radev
- 11:10 – 11:35 *Minimum Bayes-Risk Decoding for Statistical Machine Translation*
Shankar Kumar and William Byrne
- 11:35 – 12:00 *Discriminative Reranking for Machine Translation*
Libin Shen, Anoop Sarkar and Franz Josef Och
- 12:00 – 1:30 Lunch

Information Retrieval and Language Modelling

- 1:30 – 1:55 *Evaluating Multiple Aspects of Coherence in Student Essays*
Derrick Higgins, Jill Burstein, Daniel Marcu and Claudia Gentile
- 1:55 – 2:20 *A Language Modeling Approach to Predicting Reading Difficulty*
Kevyn Collins-Thompson and James P Callan
- 2:20 – 2:45 *Predicting Emotion in Spoken Dialogue from Multiple Knowledge Sources*
Kate Forbes-Riley and Diane Litman

Short Papers: Speech Recognition and Applications

- 1:30 – 1:45 *Feature-based Pronunciation Modeling for Speech Recognition*
Karen Livescu and James Glass
- 1:45 – 2:00 *Parsing Conversational Speech Using Enhanced Segmentation*
Jeremy G. Kahn, Mari Ostendorf and Ciprian Chelba
- 2:00 – 2:15 *A Comparison of Rule-Based and Statistical Methods for Semantic Language Modeling and Confidence Measurement*
Ruhi Srikaya, Yuqing Gao and Michael Picheny
- 2:15 – 2:30 *Correction Grammars for Error Handling in a Speech Dialog System*
Hirohiko Sagawa, Teruko Mitamura and Eric Nyberg
- 2:30 – 2:45 *Improving Automatic Sentence Boundary Detection with Confusion Networks*
D. Hillard, M. Ostendorf, A. Stolcke, Y. Liu and E. Shriberg

Short Papers: Language Modelling and Machine Translation

- 2:45 – 3:00 *Multi-Speaker Language Modeling*
Gang Ji and Jeffrey Bilmes
- 3:00 – 3:15 *Computational Linkuistics: Word Triggers across Hyperlinks*
Dragomir Radev, Hong Qi, Adam Winkel and Daniel Tam
- 3:15 – 3:30 *A Unigram Orientation Model for Statistical Machine Translation*
Christoph Tillman
- 3:30 – 3:45 *Morphological Analysis for Statistical Machine Translation*
Young-Suk Lee

Short Papers: Parsing and Ontologies

- 2:45 – 3:00 *Nearly-Automated Metadata Hierarchy Creation*
Emilia Soica and Marti A. Hearst
- 3:00 – 3:15 *A Lightweight Semantic Chunker Based on Tagging*
Kadri Hacioglu
- 3:15 – 3:30 *Automatic Construction of an English-Chinese Bilingual FrameNet*
Benfeng Chen and Pascale Fung
- 3:45-4:15 Break

Dialogue and Information Retrieval

- 4:15 – 4:40 *The Tao of CHI: Towards Effective Human-Computer Interaction*
Robert Porzel and Manja Baudis
- 4:40 – 5:05 *Non-Native Users in the Let's Go!! Spoken Dialogue System: Dealing with Linguistic Mismatch*
Antoine Raux and Maxine Eskenazi
- 5:05 – 5:30 *Comparison of two interactive search refinement techniques*
Olga Vechtomova and Murat Karamuftuoglu

Syntax and Semantics II

- 4:15 – 4:40 *Shallow Semantic Parsing using Support Vector Machines*
Sameer S Pradhan, Wayne H Ward, Kadri Hacioglu, James H Martin and Dan Jurafsky
- 4:40 – 5:05 *Paraphrasing Predicates from Written Language to Spoken Language Using the Web*
Nobuhiro Kaji, Masashi Okamoto and Sadao Kurohashi
- 5:05 – 5:30 *Shallow Semantc Parsing of Chinese*
Honglin Sun and Daniel Jurafsky
- 7:00 Banquet

Wednesday, May 5

9:00 – 10:15 **Keynote Speaker II: Jill Burstein**
Automated Essay Evaluation: From NLP Research through Deployment as a Business

10:15 – 10:45 Break

Machine Translation II

10:45 – 11:10 *Improvements in Phrase-Based Statistical Machine Translation*
Richard Zens and Hermann Ney

11:10 – 11:35 *Improved Machine Translation Performance via Parallel Sentence Extraction from Comparable Corpora*
Dragos Stefan Munteanu, Alexander Fraser and Daniel Marcu

11:35 – 12:00 *What's in a translation rule?*
Michel Galley, Mark Hopkins, Kevin Knight and Daniel Marcu

12:00 – 12:25 *Improving Named Entity Translation Combining Phonetic and Semantic Similarities*
Fei Huang, Stephan Vogel and Alex Waibel

Information Extraction II

10:45 – 11:10 *The (Non)Utility of Predicate-Argument Frequencies for Pronoun Interpretation*
Andrew Kehler, Douglas Appelt, Lara Taylor and Aleksandr Simma

11:10 – 11:35 *Unsupervised Learning of Contextual Role Knowledge for Coreference Resolution*
David Bean and Ellen Riloff

11:35 – 12:00 *Exponential Priors for Maximum Entropy Models*
Joshua Goodman

12:00 – 12:25 *Multiple Similarity Measures and Source-Pair Information in Story Link Detection*
Francine Chen, Ayman Farahat and Thorsten Brants

12:25 – 1:45 Lunch

1:45 – 2:45 NAACL Business Meeting

Short Papers: Statistical Techniques, Summarisation and IR

2:45 – 3:00 *Competitive Self-Trained Pronoun Interpretation*
Andrew Kehler, Douglas Appelt, Lara Taylor and Aleksandr Simma

3:00 – 3:15 *Augmenting the kappa statistic to determine interannotator reliability for multiply labeled data points*
Andrew Rosenberg and Ed Binkowski

3:15 – 3:30 *Using N-Grams To Understand the Nature of Summaries*
Michele Banko and Lucy Vanderwende

3:30 – 3:45 *Web Search Intent Induction via Automatic Query Reformulation*
Hal Daume III and Eric Brill

Short Papers: Multimodal Representation and Processing

- 2:45 – 3:00 *Performance Evaluation and Error Analysis for Multimodal Reference Resolution in a Conversation System*
Joyce Y. Chai, Zahar Prasov and Pengyu Hong
- 3:00 – 3:15 *Advances in Children's Speech Recognition within an Interactive Literacy Tutor*
Andreas Hagen, Bryan Pellom, Sarel van Vuuren and Ronald Cole
- 3:15 – 3:30 *A Speech-in List-out Approach to Spoken User Interfaces*
Vijay Divi, C. Forlines, Jan Van Gemert, Bhiksha Raj, B. Schmidt-Nielsen, Kent Wittenburg
Joseph Woelfel and Fang-Fang Zhang
- 3:30 – 3:45 *Converting a Text into Agent Animations: Assigning Gestures to a Text*
Yukiko I. Nakano, Masashi Okamoto, Daisuke Kawahara, Qing Li and Toyoaki Nishida
- 3:45 – 4:15 Break

Question Answering II

- 4:15 – 4:40 *Automatically Labeling Semantic Classes*
Patrick Pantel and Deepak Ravichandran
- 4:40 – 5:05 *Accurate Information Extraction from Research Papers using Conditional Random Fields*
Fuchun Peng and Andrew McCallum
- 5:05 – 5:30 *Name Tagging with Word Clusters and Discriminative Training*
Scott Miller, Jethran Guinness and Alex Zamanian
- 5:30 Conference End