CoNLL-2013
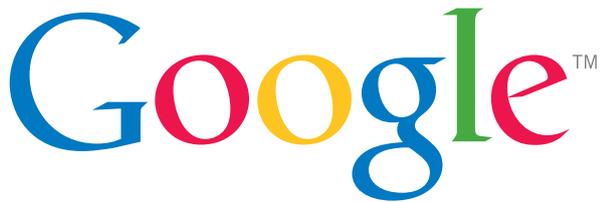
**Seventeenth Conference on
Computational Natural Language Learning**

Proceedings of the Conference

August 8-9, 2013
Sofia, Bulgaria

CoNLL 2013 Best Paper sponsor:

Order copies of this and other ACL proceedings from:

Preface

The 2013 Conference on Computational Natural Language Learning is the seventeenth in the series of annual meetings organized by SIGNLL, the ACL special interest group on natural language learning. CONLL-2013 will be held in Sofia, Bulgaria, Europe, August 8-9, 2013, in conjunction with ACL 2013.

For our special focus this year in the main session of CoNLL, we invited papers relating to compositional semantics. We received 107 submissions on this and other relevant topics, of which 7 were eventually withdrawn. Of the remaining 100 papers, 25 were selected to appear in the conference program as oral presentation. All accepted papers appear here in the proceedings. Each accepted paper was allowed eight content pages plus two pages containing only bibliographic references.

As in previous years, CoNLL-2013 has a shared task, Grammatical Error Correction. The Shared Task papers are collected in a companion volume of CoNLL-2013.

In contrast to previous conferences, we do not distinguish between long talks and posters. Instead, every CoNLL paper is allotted a 15 minute oral presentation slot as well as a poster. As a consequence, we have two poster sessions. Papers whose oral presentation is on Day 1 of the conference participate in the poster session on Day 1. The shared task posters and the CoNLL papers that are presented on Day 2 participate in the poster session on Day 2. This provides everybody with the opportunity to present their work in a plenary session, while also allowing more in-depth conversations during the two poster sessions.

We would like to thank all of the authors who submitted their work to CoNLL-2013, as well as the program committee for helping us select from among the many strong submissions. We are also grateful to our invited speakers, Ben Taskar and Roger Levy, who graciously agreed to give talks at CoNLL. Special thanks to the SIGNLL board members, Alexander Clark and Xavier Carreras, for their valuable advice and assistance in putting together this year's program, and to the SIGNLL information officer, Erik Tjong Kim Sang, for publicity and maintaining the CoNLL-2013 web page. We also appreciate the additional help we received from the ACL program chairs, workshop chairs, and publication chairs.

Finally, many thanks to Google for sponsoring the best paper award at CoNLL-2013.

We hope you enjoy the conference!

Julia Hockenmaier and Sebastian Riedel

CoNLL 2013 Conference Chairs

Conference Co-Chairs

Julia Hockenmaier (University of Illinois at Urbana-Champaign, USA)
Sebastian Riedel (University College London, United Kingdom)

Program Committee:

Timothy Baldwin (University of Melbourne, Australia), Kedar Bellare (University of Massachusetts Amherst, USA), Yonatan Bisk (University of Illinois, USA), John Blitzer (Google, USA), Thorsten Brants (Google, USA), Chris Brew (Nuance, USA), Sabine Buchholz (SynapseWork, United Kingdom), Chris Callison-Burch (Johns Hopkins University, USA), Claire Cardie (Cornell University, USA), Xavier Carreras (Universitat Politècnica de Catalunya, Spain), Ming-Wei Chang (Microsoft Research, USA), Colin Cherry (National Research Council, Canada), Yejin Choi (Stony Brook University, USA), Stephen Clark (University of Cambridge, United Kingdom), Alexander Clark (Royal Holloway University of London, United Kingdom), James Clarke (Basis Technology, USA), Shay Cohen (Columbia University, USA), Trevor Cohn (University of Sheffield, United Kingdom), Aron Culotta (Northeastern Illinois University, USA), Walter Daelemans (University of Antwerp, Belgium), Laura Dietz (University of Massachusetts Amherst, USA), Mark Dras (Macquarie University, Australia), Markus Dreyer (SDL Language Weaver, USA), Gregory Druck (Yummly, USA), Kevin Duh (Nara Institute of Science and Technology, Japan), Chris Dyer (Carnegie Mellon University, USA), Jacob Eisenstein (Georgia Institute of Technology, USA), Micha Elsner (The Ohio State University, USA), Radu Florian (IBM Research, USA), Bob Frank (Yale University, USA), Michel Galley (Microsoft Research, USA), Kevin Gimpel (Toyota Technological Institute at Chicago, USA), Yoav Goldberg (Bar Ilan University, Israel), Nizar Habash (Columbia University, USA), John Hale (Cornell University, USA), Keith Hall (Google, USA), David Hall (University of California at Berkeley, USA), Mark Hasegawa-Johnson (University of Illinois at Urbana-Champaign, USA) Rebecca Hwa (University of Pittsburgh, USA), Mark Johnson (Macquarie University, Australia), Philipp Koehn (University of Edinburgh, United Kingdom), Mamoru Komachi (Tokyo Metropolitan University, Japan), Shalom Lappin (King's College London, United Kingdom), Gideon Mann (Google, USA), Lluís Màrquez (Universitat Politècnica de Catalunya, Spain), André Martins (Priberam Labs, Portugal), Yuji Matsumoto (Nara Institute of Science and Technology, Japan), David McClosky (IBM Research, USA), Ryan McDonald (Google, USA), Haitao Mi (IBM Research, USA), Yusuke Miyao (National Institute of Informatics, Japan), Alessandro Moschitti (Trento University, Italy), Mark-Jan Nederhof (University of St Andrews, United Kingdom), Joakim Nivre (Uppsala University, Sweden), Miles Osborne (University of Edinburgh, United Kingdom), Alexandre Passos (University of Massachusetts Amherst, USA), Ari Rappoport (The Hebrew University of Jerusalem, Israel), Sujith Ravi (Google, USA), Roi Reichart (The Hebrew University of Jerusalem, Israel), Alan Ritter (University of Washington, USA), Dan Roth (University of Illinois at Urbana-Champaign, USA), Rajhans Samdani (University of Illinois at Urbana-Champaign, USA), Mark Sammons (University of Illinois at Urbana-Champaign, USA), Anoop Sarkar (Simon Fraser University, Canada), Nathan Schneider (Carnegie Mellon University, USA), Sameer Singh (University of Massachusetts Amherst, USA),

Lucia Specia (University of Sheffield, United Kingdom), Valentin Spitkovsky (Stanford University, USA), Mark Steedman (University of Edinburgh, United Kingdom), Mihai Surdeanu (University of Arizona, USA), Hiroya Takamura (Tokyo Institute of Technology, Japan), Partha Talukdar (Carnegie Mellon University, USA), Ivan Titov (Saarland University, Germany), Antal van den Bosch (Radboud University Nijmegen, Netherlands), Andreas Vlachos (University of Cambridge, United Kingdom), Peng Xu (Google, USA), Charles Yang (University of Pennsylvania, USA), Limin Yao (University of Massachusetts Amherst, USA), Dani Yogatama (Carnegie Mellon University, USA), Chen Yu (Indiana University Bloomington, USA), Luke Zettlemoyer (University of Washington, USA)

Invited Speakers:

Ben Taskar (University of Washington, USA)
Roger Levy (University of California at San Diego, USA)

Table of Contents

Conference Program

Thursday August 8 2013

(8:30 AM - 10:30 AM) Session 1

8:30   Opening Remarks

8:45   *Online Active Learning for Cost Sensitive Domain Adaptation*
Min Xiao and Yuhong Guo

9:00   *Analysis of Stopping Active Learning based on Stabilizing Predictions*
Michael Bloodgood and John Grothendieck

9:15   *Improving Pointwise Mutual Information (PMI) by Incorporating Significant Co-occurrence*
Om Damani

9:30   *Supervised Morphological Segmentation in a Low-Resource Learning Setting using Conditional Random Fields*
Teemu Ruokolainen, Oskar Kohonen, Sami Virpioja and Mikko Kurimo

9:45   *Graph-Based Posterior Regularization for Semi-Supervised Structured Prediction*
Luheng He, Jennifer Gillenwater and Ben Taskar

10:00   *A Boosted Semi-Markov Perceptron*
Tomoya Iwakura

10:45   *Spectral Learning of Refinement HMMs*
Karl Stratos, Alexander Rush, Shay B. Cohen and Michael Collins

Thursday August 8 2013 (continued)

(10:30 AM - 11:00 AM) Coffee break

(11:00 AM - 12:30 PM) Session 2

11:00   *Sentence Compression with Joint Structural Inference*
         Kapil Thadani and Kathleen McKeown

11:15   *Learning Adaptable Patterns for Passage Reranking*
         Aliaksei Severyn, Massimo Nicosia and Alessandro Moschitti

11:30   *Documents and Dependencies: an Exploration of Vector Space Models for Semantic Composition*
         Alona Fyshe, Brian Murphy, Partha Talukdar and Tom Mitchell

11:45   *Hidden Markov tree models for semantic class induction*
         Edouard Grave, Guillaume Obozinski and Francis Bach

12:00   *Better Word Representations with Recursive Neural Networks for Morphology*
         Thang Luong, Richard Socher and Christopher Manning

12:15   *Separating Disambiguation from Composition in Distributional Semantics*
         Dimitri Kartsaklis, Mehrnoosh Sadrzadeh and Stephen Pulman

(12:30 PM - 2:00 PM) Lunch break

(2:00 PM - 3:30 PM) Session 3

2:00   *Frame Semantics for Stance Classification*
         Kazi Saidul Hasan and Vincent Ng

2:15   *Philosophers are Mortal: Inferring the Truth of Unseen Facts*
         Gabor Angeli and Christopher Manning

2:30   *Towards Robust Linguistic Analysis using OntoNotes*
         Sameer Pradhan, Alessandro Moschitti, Nianwen Xue, Hwee Tou Ng, Anders Björkelund, Olga Uryupina, Yuchen Zhang and Zhi Zhong

2:45   *Dynamic Knowledge-Base Alignment for Coreference Resolution*
         Jiaping Zheng, Luke Vilnis, Sameer Singh, Jinho D. Choi and Andrew McCallum

Thursday August 8 2013 (continued)

3:00   *A Non-Monotonic Arc-Eager Transition System for Dependency Parsing*
Matthew Honnibal, Yoav Goldberg and Mark Johnson

3:15   *Collapsed Variational Bayesian Inference for PCFGs*
Pengyu Wang and Phil Blunsom

        (3:30 PM - 5:00 PM) Poster session 1

        (5:00 PM - 6 PM) Keynote 1

5:00   Invited Talk by Ben Taskar

Friday August 9 2013

        (8:45 AM - 10:30 AM) Session 4

8:45   *Polyglot: Distributed Word Representations for Multilingual NLP*
Rami Al-Rfou, Bryan Perozzi and Steven Skiena

9:00   *Exploiting multiple hypotheses for Multilingual Spoken Language Understanding*
Marcos Calvo, Fernando García, Lluís-F. Hurtado, Santiago Jiménez and Emilio Sanchis

9:15   *Multilingual WSD-like Constraints for Paraphrase Extraction*
Wilker Aziz and Lucia Specia

9:30   *Topic Models + Word Alignment = A Flexible Framework for Extracting Bilingual Dictionary from Comparable Corpus*
Xiaodong Liu, Kevin Duh and Yuji Matsumoto

9:45   *Terminology Extraction Approaches for Product Aspect Detection in Customer Reviews*
Jürgen Broß and Heiko Ehrig

10:00   Shared Task Overview by Hwee Tou Ng

Friday August 9 2013 (continued)

       (10:30 AM - 11:00 AM) Coffee break

       (11:00 AM - 12:30 PM) Shared task orals

       (12:30 PM - 2:00 PM) Lunch break

       (2:00 PM - 3:00 PM) Keynote 2

2:00   Invited Talk by Roger Levy

       (3:00 PM - 3:30 PM) Best Paper Award

3:00   *Acquisition of Desires before Beliefs: A Computional Investigation*
       Libby Barak, Afsaneh Fazly and Suzanne Stevenson

       (3:30 PM - 5:00 PM) Poster session 2 (incl. Shared Task)