CoNLL 2016

The 20th SIGNLL Conference on Computational Natural Language Learning (CoNLL)

Proceedings of the Conference

August 11-12, 2016
Berlin, Germany

Best paper awards sponsored by Google Inc.

©2016 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-945626-19-7

Introduction

The 2016 Conference on Computational Natural Language Learning is the twentieth in the series of annual meetings organized by SIGNLL, the ACL special interest group on natural language learning. CoNLL 2016 will be held on August 11-12, 2016, and is co-located with the 54th annual meeting of the Association for Computational Linguistics (ACL) in Berlin, Germany.

In order to accommodate papers with extended proofs and experimental material, CoNLL 2016 accepted only long papers, allowing 9 pages of content plus unlimited pages of references and supplementary material. We received 186 submissions in total, out of which 13 had to be rejected for formal reasons, and 21 were withdrawn by the authors. Of the remaining 149 papers, 30 papers were chosen to appear in the conference program, resulting in an overall acceptance rate of 20%. All accepted papers appear here in the proceedings.

As in previous years, CoNLL 2016 features a shared task, this year on Shallow Discourse Parsing. Papers accepted for the shared task are collected in a companion volume of CoNLL 2016.

To fit the paper presentations in a 2-day program, 21 long papers were selected for oral presentation, and 9 papers were presented as posters. The papers selected for oral presentation are distributed in six main sessions. Poster presenters were given the chance to present their poster in a short oral spotlight presentation.

For the first time, CoNLL 2016 announced a special topic on statistical natural language learning beyond linear models and convex optimization. The special topic was embraced by several authors and is reflected by the invited talks given by Jürgen Schmidhuber and Fernanda Ferreira.

We would like to thank all the authors who submitted their work to CoNLL 2016, as well as the program committee for helping us select the best papers out of many high-quality submissions. We are also grateful to our invited speakers, who graciously agreed to give talks at CoNLL.

Special thanks are due to the SIGNLL board members, Xavier Carreras and Julia Hockenmaier, for their valuable advice and assistance in putting together this year's program, and to Ben Verhoeven, for maintaining the CoNLL 2016 web page. We are grateful to the ACL organization for helping us with the program, proceedings and logistics. Finally, our gratitude goes to our sponsor, Google Inc., for supporting the best paper award at CoNLL 2016.

We hope you enjoy the conference!

Yoav Goldberg and Stefan Riezler
CoNLL 2016 conference co-chairs

Organizers:

Yoav Goldberg, Bar Ilan University
Stefan Riezler, Heidelberg University

Invited Speakers:

Jürgen Schmidhuber, Institute for Artificial Intelligence (IDSIA), Lugano, Switzerland
Fernanda Ferreira, Department of Psychology, UC Davis, CA, USA

Program Committee:

Omri Abend, The Hebrew University of Jerusalem
Željko Agić, University of Copenhagen
Waleed Ammar, CMU
Yoav Artzi, Cornell University
Omid Bakhshandeh, University of Rochester
Timothy Baldwin, The University of Melbourne
Miguel Ballesteros, Pompeu Fabra University
Roy Bar-Haim, IBM Research
Yonatan Belinkov, MIT CSAIL
Jonathan Berant, Stanford University
Yonatan Bisk, ISI/USC
Michael Bloodgood, University of Maryland
Francesca Bonin, IBM Research Ireland
Chloé Braud, University of Copenhagen
Marie Candito, Univ. Paris Diderot - INRIA - Alpage
Gracinda Carvalho, Universidade Aberta
Francisco Casacuberta, Universitat Politècnica de València
Kai-Wei Chang, University of Illinois
Ming-Wei Chang, Microsoft Research
Danqi Chen, Stanford University
Colin Cherry, NRC
Christos Christodoulopoulos, University of Illinois at Urbana Champaign
Grzegorz Chrupała, Tilburg University
Volkan Cirik, Carnegie Mellon University
Alexander Clark, King's College London
Stephen Clark, University of Cambridge
Trevor Cohn, University of Melbourne
Paul Cook, University of New Brunswick
Benoit Crabbé, University Paris Diderot
Walter Daelemans, University of Antwerp, CLiPS
Dipanjan Das, Google
Georgiana Dinu, IBM Watson
Rotem Dror, Technion
Kevin Duh, Johns Hopkins University
Greg Durrett, UC Berkeley
Jacob Eisenstein, Georgia Institute of Technology
Meng Fang, The University of Melbourne

Manaal Faruqui, Carnegie Mellon University
Geli Fei, University of Illinois at Chicago
Katja Filippova, Google
George Foster, NRC
Annemarie Friedrich, Saarland University
Daniel Gildea, University of Rochester
Kevin Gimpel, Toyota Technological Institute at Chicago
Dan Goldwasser, Purdue University
Stephan Gouws, Stellenbosch University
Edward Grefenstette, Google DeepMind
Carolin Haas, Heidelberg University
Kenneth Heafield, University of Edinburgh
Felix Hieber, Amazon Research
Julian Hitschler, Heidelberg University
Julia Hockenmaier, University of Illinois Urbana-Champaign
Matthew Honnibal, Macquarie University
Andrea Horbach, Saarland University
Yufang Hou, Heidelberg University
Dirk Hovy, Center for Language Technology, University of Copenhagen
Laura Jehl, Heidelberg University
Charles Jochim, IBM Research – Dublin
Sariya Karimova, Heidelberg University
Fabio Kepler, L2F/INESC-ID and UNIPAMPA
Daniel Khashabi, University of Illinois, Urbana-Champaign
Eliyahu Kiperwasser, Bar-Ilan University
Philipp Koehn, Johns Hopkins University
Alexander Koller University of Potsdam
Mikhail Kozhevnikov, University of Saarland, Google
Julia Kreutzer, Heidelberg University
Germán Kruszewski, University of Trento
Marco Kuhlmann, Linköping University
Jonathan K. Kummerfeld, UC Berkeley
Angeliki Lazaridou, University of Trento
Richard Leibbrandt, Flinders University
Omer Levy, Bar-Ilan University
Mike Lewis, University of Washington
Wang Ling, Google DeepMind
Tal Linzen, ENS, PSL Research University
Daniel Marcu, ISI/USC
Héctor Martínez Alonso, University of Paris 7 - INRIA
Luis Marujo, Feedzai Research
Yuji Matsumoto, Nara Institute of Science and Technology
David McClosky, Google
Oren Melamud, Bar Ilan University
Marissa Milne, Flinders University of South Australia
Alessandro Moschitti, Qatar Computing Research Institute / University of Trento
Nasrin Mostafazadeh, University of Rochester
Preslav Nakov, Qatar Computing Research Institute, HBKU
Graham Neubig, Nara Institute of Science and Technology
Joakim Nivre, Uppsala University
Daniel Ortiz-Martínez, Technical University of Valencia

Miles Osborne, Bloomberg
Alexis Palmer, Heidelberg University
Denis Paperno, University of Trento
Xiaochang Peng, University of Rochester
Nghia The Pham, University of Trento
Yuval Pinter, Yahoo Labs
Barbara Plank, University of Groningen
David Powers, Flinders University
Chris Quirk, Microsoft Research
Roi Reichart, Technion - Israel Institute of Technology
Corentin Ribeyre, Univ. Paris Diderot - INRIA - ALPAGE
Brian Roark, Google
Michael Roth, University of Edinburgh
Alla Rozovskaya, Virginia Tech
Kenji Sagae, KITT.AI
Benôit Sagot, INRIA
Mark Sammons, University of Illinois at Urbana-Champaign
Shigehiko Schamoni, Heidelberg University
William Schuler, The Ohio State University
Roy Schwartz, The Hebrew University of Jerusalem
Djamé Seddah, Université Paris Sorbonne (Paris IV)
Wolfgang Seeker, spaCy GmbH
Vered Shwartz, Bar-Ilan University
Khalil Sima'an, ILLC, University of Amsterdam
Patrick Simianer, Heidelberg University
Artem Sokolov, Heidelberg University
Radu Soricut, Google
Vivek Srikumar, University of Utah
Simon Suster, University of Groningen
Anders Søgaard, University of Copenhagen
Partha Talukdar, Indian Institute of Science
Reut Tsarfaty, Open University of Israel
Oren Tsur, Harvard University
Ashish Vaswani, USC/ISI
Andreas Vlachos, University of Sheffield
Ivan Vulić, University of Cambridge
Ekaterina Vylomova, University of Melbourne
Taro Watanabe, Google
Sander Wubben, Tilburg University
Luke Zettlemoyer, University of Washington

Table of Contents

Keynote Talk: Human Processing of Disfluent Speech: Basic Findings, Theoretical Approaches, and Implications for Natural Language Processing
Fernanda Ferreira xvi

Keynote Talk: RNNaissance
Jürgen Schmidhuber xvii

20 Years Retrospect: Computational Natural Language Learning: ±20years ±Data ±Features ±Multimodal ±Bioplausible
David Powers 1

Generating Sentences from a Continuous Space
Samuel R. Bowman, Luke Vilnis, Oriol Vinyals, Andrew Dai, Rafal Jozefowicz and Samy Bengio 10

Identifying Temporal Orientation of Word Senses
Mohammed Hasanuzzaman, Gaël Dias, Stéphane Ferrari, Yann Mathet and Andy Way 22

Semi-supervised Clustering for Short Text via Deep Representation Learning
Zhiguo Wang, Haitao Mi and Abraham Ittycheriah 31

Neighborhood Mixture Model for Knowledge Base Completion
Dat Quoc Nguyen, Kairit Sirts, Lizhen Qu and Mark Johnson 40

context2vec: Learning Generic Context Embedding with Bidirectional LSTM
Oren Melamud, Jacob Goldberger and Ido Dagan 51

Learning to Jointly Predict Ellipsis and Comparison Structures
Omid Bakhshandeh, Alexis Cornelia Wellwood and James Allen 62

Event Embeddings for Semantic Script Modeling
Ashutosh Modi 75

Beyond Centrality and Structural Features: Learning Information Importance for Text Summarization
Markus Zopf, Eneldo Loza Mencía and Johannes Fürnkranz 84

Incremental Prediction of Sentence-final Verbs: Humans versus Machines
Alvin Grissom II, Naho Orita and Jordan Boyd-Graber 95

A Data-driven Investigation of Corrective Feedback on Subject Omission Errors in First Language Acquisition
Sarah Hiller and Raquel Fernandez 105

Redefining part-of-speech classes with distributional semantic models
Andrey Kutuzov, Erik Velldal and Lilja Øvrelid 115

Analyzing Learner Understanding of Novel L2 Vocabulary
Rebecca Knowles, Adithya Renduchintala, Philipp Koehn and Jason Eisner 126

Modeling language evolution with codes that utilize context and phonetic features
Javad Nouri and Roman Yangarber 136

Harnessing Sequence Labeling for Sarcasm Detection in Dialogue from TV Series 'Friends'
Aditya Joshi, Vaibhav Tripathi, Pushpak Bhattacharyya and Mark J Carman 146

Leveraging Cognitive Features for Sentiment Analysis
Abhijit Mishra, Diptesh Kanojia, Seema Nagar, Kuntal Dey and Pushpak Bhattacharyya 156

Modelling Context with User Embeddings for Sarcasm Detection in Social Media
Silvio Amir, Byron C. Wallace, Hao Lyu, Paula Carvalho and Mario J. Silva 167

Learning when to trust distant supervision: An application to low-resource POS tagging using cross-lingual projection
Meng Fang and Trevor Cohn 178

Greedy, Joint Syntactic-Semantic Parsing with Stack LSTMs
Swabha Swayamdipta, Miguel Ballesteros, Chris Dyer and Noah A. Smith 187

Beyond Prefix-Based Interactive Translation Prediction
Jesús González-Rubio, Daniel Ortiz Martinez, Francisco Casacuberta
and Jose Miguel Benedi Ruiz 198

Exploring Prediction Uncertainty in Machine Translation Quality Estimation
Daniel Beck, Lucia Specia and Trevor Cohn 208

Cross-Lingual Named Entity Recognition via Wikification
Chen-Tse Tsai, Stephen Mayhew and Dan Roth 219

Coreference in Wikipedia: Main Concept Resolution
Abbas Ghaddar and Phillippe Langlais 229

Event Linking with Sentential Features from Convolutional Neural Networks
Sebastian Krause, Feiyu Xu, Hans Uszkoreit and Dirk Weissenborn 239

Joint Learning of the Embedding of Words and Entities for Named Entity Disambiguation
Ikuya Yamada, Hiroyuki Shindo, Hideaki Takeda and Yoshiyasu Takefuji 250

Entity Disambiguation by Knowledge and Text Jointly Embedding
Wei Fang, Jianwen Zhang, Dilin Wang, Zheng Chen and Ming Li 260

Substring-based unsupervised transliteration with phonetic and contextual knowledge
Anoop Kunchukuttan, Pushpak Bhattacharyya and Mitesh M. Khapra 270

Abstractive Text Summarization using Sequence-to-sequence RNNs and Beyond
Ramesh Nallapati, Bowen Zhou, Cicero dos Santos, Caglar Gulcehre and Bing Xiang 280

Compression of Neural Machine Translation Models via Pruning
Abigail See, Minh-Thang Luong and Christopher D. Manning 291

Modelling the Usage of Discourse Connectives as Rational Speech Acts
Frances Yung, Kevin Duh, Taku Komura and Yuji Matsumoto 302

Semi-supervised Convolutional Networks for Translation Adaptation with Tiny Amount of In-domain Data
Boxing Chen and Fei Huang 314

Conference Program

Thursday, August 11, 2016

9:00–9:10 Opening

9:10–10:30 Session 1: Word-Level Semantics

Generating Sentences from a Continuous Space
Samuel R. Bowman, Luke Vilnis, Oriol Vinyals, Andrew Dai, Rafal Jozefowicz and Samy Bengio

Identifying Temporal Orientation of Word Senses
Mohammed Hasanuzzaman, Gaël Dias, Stéphane Ferrari, Yann Mathet and Andy Way

Semi-supervised Clustering for Short Text via Deep Representation Learning
Zhiguo Wang, Haitao Mi and Abraham Ittycheriah

Neighborhood Mixture Model for Knowledge Base Completion
Dat Quoc Nguyen, Kairit Sirts, Lizhen Qu and Mark Johnson

10:30–11:00 Break

Thursday, August 11, 2016 (continued)

11:00–12:20 Session 2: Sentence-Level Semantics

context2vec: Learning Generic Context Embedding with Bidirectional LSTM
Oren Melamud, Jacob Goldberger and Ido Dagan

Learning to Jointly Predict Ellipsis and Comparison Structures
Omid Bakhshandeh, Alexis Cornelia Wellwood and James Allen

Event Embeddings for Semantic Script Modeling
Ashutosh Modi

Beyond Centrality and Structural Features: Learning Information Importance for Text Summarization
Markus Zopf, Eneldo Loza Mencía and Johannes Fürnkranz

12:20–12:50 SIGNLL Business Meeting

12:50–14:00 Lunch Break

14:00–15:40 Session 3: Human Language Processing

Incremental Prediction of Sentence-final Verbs: Humans versus Machines
Alvin Grissom II, Naho Orita and Jordan Boyd-Graber

A Data-driven Investigation of Corrective Feedback on Subject Omission Errors in First Language Acquisition
Sarah Hiller and Raquel Fernandez

Redefining part-of-speech classes with distributional semantic models
Andrey Kutuzov, Erik Velldal and Lilja Øvrelid

Analyzing Learner Understanding of Novel L2 Vocabulary
Rebecca Knowles, Adithya Renduchintala, Philipp Koehn and Jason Eisner

Modeling language evolution with codes that utilize context and phonetic features
Javad Nouri and Roman Yangarber

Thursday, August 11, 2016 (continued)

15:40–16:00 Break

16:00–17:15 **Keynote 1: Human Processing of Disfluent Speech: Basic Findings, Theoretical Approaches, and Implications for Natural Language Processing**
Fernanda Ferreira

17:15–18:15 Session 4: Sarcasm / Sentiment

Harnessing Sequence Labeling for Sarcasm Detection in Dialogue from TV Series 'Friends'
Aditya Joshi, Vaibhav Tripathi, Pushpak Bhattacharyya and Mark J Carman

Leveraging Cognitive Features for Sentiment Analysis
Abhijit Mishra, Diptesh Kanojia, Seema Nagar, Kuntal Dey and Pushpak Bhattacharyya

Modelling Context with User Embeddings for Sarcasm Detection in Social Media
Silvio Amir, Byron C. Wallace, Hao Lyu, Paula Carvalho and Mario J. Silva

Friday, August 12, 2016

9:00–9:15 **SIGNLL Steering Committee Address**
David Powers: 20 Years Retrospect

9:15–10:30 **Keynote 2: RNNaissance**
Jürgen Schmidhuber

10:30–11:00 **Break**

11:00–12:40 **Session 5: Syntax, Named Entities, and Machine Translation**

Learning when to trust distant supervision: An application to low-resource POS tagging using cross-lingual projection
Meng Fang and Trevor Cohn

Greedy, Joint Syntactic-Semantic Parsing with Stack LSTMs
Swabha Swayamdipta, Miguel Ballesteros, Chris Dyer and Noah A. Smith

Beyond Prefix-Based Interactive Translation Prediction
Jesús González-Rubio, Daniel Ortiz Martinez, Francisco Casacuberta and Jose Miguel Benedi Ruiz

Exploring Prediction Uncertainty in Machine Translation Quality Estimation
Daniel Beck, Lucia Specia and Trevor Cohn

Cross-Lingual Named Entity Recognition via Wikification
Chen-Tse Tsai, Stephen Mayhew and Dan Roth

12:40–14:00 **Lunch Break**

14:00–15:30 **Session 6: Shared Task on Shallow Discourse Parsing**

15:30–16:00 **Break**

Friday, August 12, 2016 (continued)

16:00–16:20 Session 7: Spotlight Presentations

16:20–17:20 Session 8: Posters (+ Shared Task Posters)

Coreference in Wikipedia: Main Concept Resolution
Abbas Ghaddar and Phillippe Langlais

Event Linking with Sentential Features from Convolutional Neural Networks
Sebastian Krause, Feiyu Xu, Hans Uszkoreit and Dirk Weissenborn

Joint Learning of the Embedding of Words and Entities for Named Entity Disambiguation
Ikuya Yamada, Hiroyuki Shindo, Hideaki Takeda and Yoshiyasu Takefuji

Entity Disambiguation by Knowledge and Text Jointly Embedding
Wei Fang, Jianwen Zhang, Dilin Wang, Zheng Chen and Ming Li

Substring-based unsupervised transliteration with phonetic and contextual knowledge
Anoop Kunchukuttan, Pushpak Bhattacharyya and Mitesh M. Khapra

Abstractive Text Summarization using Sequence-to-sequence RNNs and Beyond
Ramesh Nallapati, Bowen Zhou, Cicero dos Santos, Caglar Gulcehre and Bing Xiang

Compression of Neural Machine Translation Models via Pruning
Abigail See, Minh-Thang Luong and Christopher D. Manning

Modelling the Usage of Discourse Connectives as Rational Speech Acts
Frances Yung, Kevin Duh, Taku Komura and Yuji Matsumoto

Semi-supervised Convolutional Networks for Translation Adaptation with Tiny Amount of In-domain Data
Boxing Chen and Fei Huang

17:20–18:00 Session 9: Best Paper Awards + Closing

Keynote Talk

Human Processing of Disfluent Speech: Basic Findings, Theoretical Approaches, and Implications for Natural Language Processing

Fernanda Ferreira

Abstract Disfluencies occur in human speech at the rate of about one per minute; therefore, any adequate theory of human language comprehension must explain how listeners process utterances containing them. Our theoretical approach is based on a 15-year program of research that has uncovered a number of fundamental mechanisms enabling humans to process disfluencies efficiently, including mechanisms that are backward looking (reanalysis of the input) and ones that are anticipatory or forward looking (prediction). This presentation will review the theory, the evidence that supports it, and the outstanding questions that are currently being investigated. I will also consider implications for refining NLP systems, which must be robust to speaker error and which should be capable of adapting to characteristics of particular speakers and language communities.

Biography of Speaker Fernanda Ferreira is Professor of Psychology and Member of the Graduate Group in Linguistics at the University of California, Davis. She obtained her Ph.D. in Cognitive Psychology in 1988 from the University of Massachusetts, Amherst, and prior to moving to UC Davis in 2015, she held faculty positions at Michigan State University and the University of Edinburgh. She has published over 100 papers and her research has been funded by the NSF and the NIH in the US, and the ESRC in the UK. She served as Editor in Chief of the Journal of Experimental Psychology: General, and she is currently an Associate Editor of Cognitive Psychology and of Collabra, an Open Access journal recently launched by University of California Press. She is a Fellow of the American Psychological Society and the Royal Society of Edinburgh, and she is currently an elected member of the Psychonomic Society's Governing Board.

Keynote Talk

RNNaissance

Jürgen Schmidhuber

Abstract Our deep learning artificial neural networks have won numerous contests in pattern recognition and machine learning. They are now widely used by the world's most valuable public companies. In particular, Long Short-Term Memory (LSTM) Recurrent Neural Networks (RNNs) are very useful not only for speech recognition but also for Computational Language Learning. I will discuss state-of-the-art results in numerous applications.

Biography of Speaker Since age 15 or so, the main goal of professor Jürgen Schmidhuber has been to build a self-improving Artificial Intelligence (AI) smarter than himself, then retire. He has pioneered self-improving general problem solvers since 1987, and Deep Learning Neural Networks (NNs) since 1991. The recurrent NNs developed by his research groups at the Swiss AI Lab IDSIA (USI & SUPSI) & TU Munich were the first to win official international contests. They have revolutionized handwriting recognition, speech recognition, machine translation, image captioning, and are now available to over a billion users through Google, Microsoft, IBM, Baidu, and many other companies. DeepMind is heavily influenced by his lab's former students (including 2 of DeepMind's first 4 members and their first PhDs in AI, one of them co-founder, one of them first employee). His team's Deep Learners were the first to win object detection and image segmentation contests, and achieved the world's first superhuman visual classification results, winning nine international competitions in machine learning & pattern recognition (more than any other team). They also were the first to learn control policies directly from high-dimensional sensory input using reinforcement learning. His research group also established the field of mathematically rigorous universal AI and optimal universal problem solvers. His formal theory of creativity & curiosity & fun explains art, science, music, and humor. He also generalized algorithmic information theory and the many-worlds theory of physics, and introduced the concept of Low-Complexity Art, the information age's extreme form of minimal art. Since 2009 he has been member of the European Academy of Sciences and Arts. He has published 333 peer-reviewed papers, earned seven best paper/best video awards, the 2013 Helmholtz Award of the International Neural Networks Society, and the 2016 IEEE Neural Networks Pioneer Award. He is also president of NNAISENSE, which aims at building the first practical general purpose AI.