CORBON 2016

Coreference Resolution beyond OntoNotes

Proceedings of the Workshop

NAACL-HLT 2016 Workshop
June 16, 2016
San Diego, California, USA

©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 978-1-941643-90-7

Introduction

Many NLP researchers, especially those not working in the area of discourse processing, tend to equate coreference resolution with the sort of coreference that people did in MUC, ACE, and OntoNotes, having the impression that coreference is a well-worn task owing in part to the large number of papers reporting results on the MUC/ACE/OntoNotes corpora. This is an unfortunate misconception: the previous SemEval 2010 and CoNLL 2012 shared tasks on coreference resolution have largely focused on entity coreference, which constitutes only one of the many kinds of coreference relations that were discussed in theoretical and computational linguistics in the past few decades. In fact, by focusing on entity coreference resolution, NLP researchers have only scratched the surface of the wealth of interesting problems in coreference resolution.

The decision to focus on entity coreference resolution was initially made by information extraction (IE) researchers when coreference was selected as one of the tasks in the MUC-6 coreference in 1995. Many interesting kinds of coreference relations, such as bridging and reference to abstract entities, were left out not because they were not important, but because "it was felt that the menu was simply too ambitious". It turned out that this decision had an important consequence: the progress made in coreference research in the past two decades was largely driven by the availability of coreference-annotated corpora such as MUC, ACE, and OntoNotes, where entity coreference was the focus.

Given the plethora of work on entity coreference and aware of other fora gathering coreference-related papers (such as LAW, DiscoMT or EVENTS), we believed that time was ripe for a new workshop on the single topic of coreference resolution that would bring together researchers who were interested in under-investigated coreference phenomena, willing to contribute both theoretical and applied computational work on coreference resolution, especially for languages other than English, less-researched forms of coreference and new applications of coreference resolution.

Our call attracted 20 submissions out of which we have selected 4 long and 2 short papers for oral presentation and 7 papers for poster presentation based on reviewers' recommendations.

We would like to thank the Program Committee members who reviewed the submissions and the workshop participants for joining in!

— Maciej Ogrodniczuk and Vincent Ng

Organizing Committee and Proceedings Editors:

Maciej Ogrodniczuk, Institute of Computer Science, Polish Academy of Sciences
Vincent Ng, University of Texas at Dallas

Program Committee:

Anders Björkelund, University of Stuttgart
Antonio Branco, University of Lisbon
Dan Cristea, A. I. Cuza University of Iaşi
Sobha Lalitha Devi, AU-KBC Research Center, Anna University of Chennai
Lars Hellan, Norwegian University of Science and Technology
Veronique Hoste, Ghent University
Yufang Hou, Heidelberg University
Sandra Kübler, Indiana University
Sebastian Martschat, Heidelberg University
Ruslan Mitkov, University of Wolverhampton
Costanza Navaretta, University of Copenhagen
Anna Nedoluzhko, Charles University in Prague
Vincent Ng, University of Texas at Dallas
Michal Novak, Charles University in Prague
Maciej Ogrodniczuk, Institute of Computer Science, Polish Academy of Sciences
Constantin Orasan, University of Wolverhampton
Simone Paolo Ponzetto, University of Mannheim
Massimo Poesio, University of Essex
Sameer Pradhan, Boulder Language Technologies
Marta Recasens, Google Inc.
Agata Savary, François Rabelais University Tours
Olga Uryupina, University of Trento
Yannick Versley, Heidelberg University
Desislava Zhekova, Ludwig Maximilian University of Munich
Heike Zinsmeister, Universität Hamburg

Invited Speakers:

Andrew Kehler, University of California, San Diego
Michael Strube, Heidelberg Institute for Theoretical Studies GmbH

Table of Contents

Sense Anaphoric Pronouns: Am I One?
Marta Recasens, Zhichao Hu and Olivia Rhinehart 1

Experiments on bridging across languages and genres
Yulia Grishina 7

Bridging Relations in Polish: Adaptation of Existing Typologies
Maciej Ogrodniczuk and Magdalena Zawisławska 16

Beyond Identity Coreference: Contrasting Indicators of Textual Coherence in English and German
Kerstin Kunz, Ekaterina Lapshinova-Koltunski and José Manuel Martínez 23

Exploring the steps of Verb Phrase Ellipsis
Zhengzhong Liu, Edgar Gonzàlez Pellicer and Daniel Gillick 32

Anaphoricity in Connectives: A Case Study on German
Manfred Stede and Yulia Grishina 41

Abstract Coreference in a Multilingual Perspective: a View on Czech and German
Anna Nedoluzhko and Ekaterina Lapshinova-Koltunski 47

Antecedent Prediction Without a Pipeline
Sam Wiseman, Alexander M. Rush and Stuart Shieber 53

Bridging Corpus for Russian in comparison with Czech
Anna Roitberg and Anna Nedoluzhko 59

Coreference Resolution for the Basque Language with BART
Ander Soraluze, Olatz Arregi, Xabier Arregi, Arantza Diaz de Ilarraza, Mijail Kabadjov and Massimo Poesio 67

Error analysis for anaphora resolution in Russian: new challenging issues for anaphora resolution task in a morphologically rich language
Svetlana Toldova, Ilya Azerkovich, Alina Ladygina, Anna Roitberg and Maria Vasilyeva 74

How to Handle Split Antecedents in Tamil?
Vijay Sundar Ram and Sobha Lalitha Devi 84

When Annotation Schemes Change Rules Help: A Configurable Approach to Coreference Resolution beyond OntoNotes
Amir Zeldes and Shuo Zhang 92

Workshop Program: June 16, 2016

Session 1

09:00–09:10 *Introduction*

09:10–10:10 *Invited talk*

The (Non)Utility of Semantics for Coreference Resolution (CORBON Remix)
Michael Strube

10:10–10:30 *Sense Anaphoric Pronouns: Am I One?*
Marta Recasens, Zhichao Hu and Olivia Rhinehart

Coffee break

Session 2

11:00–11:30 *Experiments on Bridging Across Languages and Genres*
Yulia Grishina

11:30–12:00 *Bridging Relations in Polish: Adaptation of Existing Typologies*
Maciej Ogrodniczuk and Magdalena Zawisławska

12:00–12:30 *Beyond Identity Coreference: Contrasting Indicators of Textual Coherence in English and German*
Kerstin Kunz, Ekaterina Lapshinova-Koltunski and José Manuel Martínez

Lunch break

Session 3

14:00–15:00 *Invited talk*

A Bayesian Model of Pronoun Production and Interpretation
Andrew Kehler

15:00–15:30 *Exploring the Steps of Verb Phrase Ellipsis*
Zhengzhong Liu, Edgar Gonzàlez Pellicer and Daniel Gillick

Coffee break

Session 4

16:00–16:20 *Anaphoricity in Connectives: A Case Study on German*
Manfred Stede and Yulia Grishina

16:20–16:30 *One minute madness for posters*

16:30–17:30 *Poster session*

Abstract Coreference in a Multilingual Perspective: a View on Czech and German
Anna Nedoluzhko and Ekaterina Lapshinova-Koltunski

Antecedent Prediction Without a Pipeline
Sam Wiseman, Alexander M. Rush and Stuart Shieber

Bridging Corpus for Russian in Comparison with Czech
Anna Roitberg and Anna Nedoluzhko

Coreference Resolution for the Basque Language with BART
Ander Soraluze, Olatz Arregi, Xabier Arregi, Arantza Diaz de Ilarraza, Mijail Kabadjov and Massimo Poesio

Error Analysis for Anaphora Resolution in Russian: New Challenging Issues for Anaphora Resolution Task in a Morphologically Rich Language
Svetlana Toldova, Ilya Azerkovich, Alina Ladygina, Anna Roitberg and Maria Vasilyeva

How to Handle Split Antecedents in Tamil?
Vijay Sundar Ram and Sobha Lalitha Devi

When Annotation Schemes Change Rules Help: A Configurable Approach to Coreference Resolution beyond OntoNotes
Amir Zeldes and Shuo Zhang