CRAC 2020

Third Workshop on Computational Models of Reference, Anaphora and Coreference

Proceedings of the Workshop

COLING 2020 Workshop
December 12, 2020
Barcelona, Spain (online)

ISBN 978-1-952148-35-4

Introduction

This is the third edition of the Workshop on Computational Models of Reference, Anaphora and Coreference (CRAC). CRAC was first held in New Orleans two years ago in conjunction with NAACL HLT 2018. CRAC and its predecessor, the Coreference Resolution Beyond OntoNotes (CORBON) workshop series that started in 2016, have arguably become the primary forum for coreference researchers to present their latest results since the demise of the Discourse Anaphora and Anaphor Resolution Colloquium series in 2011. While CORBON focuses on under-investigated coreference phenomena, CRAC has a broader scope, covering all cases of computational modeling of reference, anaphora, and coreference.

CRAC 2020 is by far the largest workshop in this workshop series in terms of the number of submissions and the number of accepted papers. Specifically, we received 23 submissions: 19 of them were from Europe, one was from the U.S. and the remaining two were from India. This geographical distribution of submissions provides suggestive evidence that coreference continues to be more actively researched in Europe than in other parts of the world. The submissions covered various aspects of reference-related topics, from resources for anaphora analysis, to different approaches to mention detection and anaphora resolution to applications. Each submission was rigorously reviewed by two to three program committee members. Based on their recommendations, we initially accepted 14 papers and conditionally accepted three papers. The three conditionally accepted papers were eventually accepted to the workshop after we made sure that the authors adequately addressed the reviewers' comments in the final camera-ready version. Overall, we were pleased with the large number of submissions as well as the quality of the accepted papers.

We started two new initiatives this year. First, with the goal of having a broad technical program, we introduced different paper categories. In addition to research papers, we welcomed survey papers, position papers, challenge papers, and demo papers. In each of these paper categories, authors were expected to report completed work. To encourage researchers to report work in progress and/or late-breaking results, we introduced another submission type, extended abstracts. Since we decided to introduce these new paper categories only in early October, many authors were not aware of these categories by the time they submitted their work. Of the 23 submissions we received, 20 were research papers, one was a position paper, one was a demo paper, and one was an extended abstract. While all of the accepted papers this year were research papers (13 long papers and 4 short papers), we are confident that these new paper categories will become increasingly popular in the future. Our second initiative involves recognizing outstanding research submitted to the workshop via a best paper award. The winner(s) will be announced at the closing session.

Due to the COVID-19 pandemic, CRAC will be organized for the first time as an online event. Authors of all accepted papers were asked to pre-record videos presenting their work, which will be made available for viewing by all attendees a week before the start of the workshop. At the workshop, each paper will be given a 5-minute oral presentation slot followed by a 5-minute discussion period.

We are grateful to the following people, without whom we could not have assembled an interesting program for the workshop. First, we are indebted to our program committee members. Owing to the unexpected increase in the number of submissions, each reviewer were assigned four papers to review. All of them did the incredible job of completing their reviews in a reviewing period that spanned less than two weeks. Second, we thank Juntao Yu for accepting our invitation to be this year's invited speaker. Juntao will give a talk on anaphora resolution beyond OntoNotes, which brings us back to the roots of our predecessor, the CORBON workshop. Third, we thank Massimo Poesio for agreeing to chair a plenary session that focuses on discussing the possibility of developing Universal Anaphora (UA), a unified, language-independent markup scheme that reflects common cross-linguistic understanding of reference-

related phenomena. Motivated by Universal Dependencies, UA aims to facilitate referential analysis of the similarities and idiosyncracies among typologically different languages, support comparative evaluation of anaphora resolution systems and enable comparative linguistic studies. If successful, UA will be used to produce annotated corpora for a joint shared task by CODI (The Workshop on Computational Approaches to Discourse) and CRAC next year. Finally, we would like to thank the workshop participants for joining in.

Despite these difficult times, we look forward to an exciting online workshop. We hope you will enjoy it as much as we do!

— Maciej Ogrodniczuk, Sameer Pradhan, Yulia Grishina, and Vincent Ng

Organizing Committee and Proceedings Editors:

Maciej Ogrodniczuk, Institute of Computer Science, Polish Academy of Sciences
Sameer Pradhan, University of Pennsylvania and cemantix
Yulia Grishina, Amazon
Vincent Ng, University of Texas at Dallas

Programme Committee:

Antonio Branco, University of Lisbon
Dan Cristea, Alexandru Ioan Cuza University of Iași
Stephanie Dipper, University of Bochum
Yulia Grishina, Amazon
Veronique Hoste, Ghent University
Sandra Kübler, Indiana University
Sobha Lalitha Devi, AU-KBC Research Center, Anna University of Chennai
Emmanuel Lassalle, Machina Capital, Paris
Costanza Navaretta, University of Copenhagen
Anna Nedoluzhko, Charles University in Prague
Michal Novak, Charles University in Prague
Constantin Orasan, University of Surrey
Massimo Poesio, Queen Mary University of London
Marta Recasens, Google
Yannick Versley, Amazon
Heike Zinsmeister, University of Hamburg

Invited Speaker:

Juntao Yu, Queen Mary University of London

Table of Contents

Workshop Program: December 12, 2020

14:00–15:00 Welcome and Invited Talk

14:00–14:05 *Introduction*
Maciej Ogrodniczuk, Sameer Pradhan, Yulia Grishina and Vincent Ng

14:05–15:00 Invited talk: *Anaphora Resolution beyond OntoNotes*
Juntao Yu

15:00–15:10 Short Break

15:10–16:10 Paper Summary Session 1: Mention Detection and Deep Learning Approaches

15:10–15:20 *Anaphoric Zero Pronoun Identification: A Multilingual Approach*
Abdulrahman Aloraini and Massimo Poesio

15:20–15:30 *Partially-Supervised Mention Detection*
Lesly Miculicich and James Henderson

15:30–15:40 *E.T.: Entity-Transformers. Coreference Augmented Neural Language Model for Richer Mention Representations via Entity-Transformer Blocks*
Nikolaos Stylianou and Ioannis Vlahavas

15:40–15:50 *Neural Coreference Resolution for Arabic*
Abdulrahman Aloraini, Juntao Yu and Massimo Poesio

15:50–16:00 *Sequence-to-Sequence Networks Learn the Meaning of Reflexive Anaphora*
Robert Frank and Jackson Petty

16:00–16:10 *Sequence to Sequence Coreference Resolution*
Gorka Urbizu, Ander Soraluze and Olatz Arregi

16:10–16:20 Short Break

16:20–17:20 Paper Summary Session 2: Resources, Evaluation and Beyond the Identity of Reference

16:20–16:30 *A Benchmark of End-to-End and Deterministic Coreference Resolution of Dutch Novels and News*
Corbèn Poot and Andreas van Cranenburgh

16:30–16:40 *A Dataset for Anaphora Analysis in French E-mails*
Hani Guenoune, Cédric Lopez, Kevin Cousot, Melissa Mekaoui and Mathieu Lafourcade

16:40–16:50 *Integrating Knowledge Graph Embeddings to Improve Mention Representation for Bridging Anaphora Resolution*
Onkar Pandit, Pascal Denis and Liva Ralaivola

16:50–17:00 *Reference to Discourse Topics: Introducing "Global" Shell Nouns*
Fabian Simonjetz

Workshop Program (continued): December 12, 2020

17:00–17:10 *TwiConv: A Coreference-annotated Corpus of Twitter Conversations*
Berfin Aktaş and Annalena Kohnert

17:10–17:20 *Predicting Coreference in Abstract Meaning Representations*
Tatiana Anikina, Alexander Koller and Michael Roth

17:20–17:30 Short Break

17:30–18:20 Paper Summary Session 3: Applications

17:30–17:40 *It's Absolutely Divine! Can Fine-Grained Sentiment Analysis Benefit from Coreference Resolution?*
Orphee De Clercq and Veronique Hoste

17:40–17:50 *Enhanced Labelling in Active Learning for Coreference Resolution*
Vebjørn Espeland, Beatrice Alex and Benjamin Bach

17:50–18:00 *Coreference Strategies in English-German Translation*
Ekaterina Lapshinova-Koltunski, Marie-Pauline Krielke and Christian Hardmeier

18:00–18:10 *Reference in Team Communication for Robot-Assisted Disaster Response: An Initial Analysis*
Natalia Skachkova and Ivana Kruijff-Korbayova

18:10–18:20 *Resolving Pronouns in Twitter Streams: Context can Help!*
Anietie Andy, Chris Callison-Burch and Derry Tanti Wijaya

18:20–18:30 Short Break

18:30–19:15 Plenary Session on Universal Anaphora and Best Paper Award

18:30–19:10 *Universal Anaphora Discussion Panel*
Chair: Massimo Poesio

19:10–19:15 *Best Paper Award and Closing of the Workshop*
Maciej Ogrodniczuk, Sameer Pradhan, Yulia Grishina and Vincent Ng