

COLING

**International Conference on
Computational Linguistics**

Proceedings of the Conference and Workshops

COLING

Volume 29 (2022), No. 7

**Proceedings of The Fifth Workshop on Computational
Models of Reference, Anaphora and Coreference
(CRAC 2022)**

**The 29th International Conference on
Computational Linguistics**

October 16–17, 2022
Gyeongju, Republic of Korea

Copyright of each paper stays with the respective authors (or their employers).

ISSN 2951-2093

Message from the Program Chairs

This is the fifth edition of the Workshop on Computational Models of Reference, Anaphora and Coreference (CRAC). CRAC was first held in New Orleans four years ago in conjunction with NAACL HLT 2018. But the workshop series dates back to its predecessor, the Coreference Resolution Beyond OntoNotes (CORBON) that started in 2016, and has 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 focused on under-investigated coreference phenomena, CRAC has a broader scope, covering all cases of computational modeling of reference, anaphora, and coreference.

CRAC 2022 continued to attract a large number of very high quality papers. Specifically, we received 14 submissions which were rigorously reviewed by three program committee members. Based on their recommendations, we accepted 10 papers and conditionally accepted one paper. The one conditionally accepted paper was 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. This time around we had a total of two papers that were withdrawn for various reasons.

This was the second year of the joint CODI-CRAC shared task on *Anaphora, Bridging, and Discourse Deixis in Dialogue*. In addition, this year debuted the CRAC shared task on *Multilingual Coreference Resolution*. Both these activities allowed researchers who did not participate in the workshop to disseminate their work to a smaller and more focused audience which should promote interesting discussions.

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. This year the reviewing load was on an average of two papers per reviewer. All of them did the incredible job of completing their reviews in a short reviewing period. Second, this is the first year where we have three invited talks. We thank Sharid Loáiciga, Juntao Yu, Michal Novák, Massimo Poesio and Lori Levin for accepting our invitation to be this year's invited speakers. This year we have a shorter duration for the continued panel on the Universal Anaphora (UA) effort, 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 idiosyncrasies among typologically different languages, support comparative evaluation of anaphora resolution systems and enable comparative linguistic studies. Finally, we would like to thank the workshop participants for joining us in this event.

We hope you will enjoy it as much as we do!

— Sameer Pradhan, Maciej Ogrodniczuk, Anna Nedoluzhko,
Massimo Poesio, and Vincent Ng

Organizers

Organizing Committee:

Maciej Ogrodniczuk, Institute of Computer Science, Polish Academy of Sciences, Poland
Sameer Pradhan, University of Pennsylvania and cemantix.org, USA
Anna Nedoluzhko, Charles University in Prague, Czechia
Massimo Poesio, Queen Mary University of London, UK
Vincent Ng, University of Texas at Dallas, USA

Program Committee:

Antonio Branco, University of Lisbon, Portugal
Arie Cattan, Bar-Ilan University, Israel
Haixia Chai, Heidelberg University, Germany
Stephanie Dipper, Ruhr-University, Germany
Elisa Ferracane, Abridge, USA
Yulia Grishina, Amazon, USA
Yansong Feng, Peking University, China
Christian Hardmeier, IT University of Copenhagen, Denmark
Lars Hellan, Norwegian University of Science and Technology, Norway
Veronique Hoste, Ghent University, Belgium
Yufang Hou, IBM, Dublin, Ireland
Ruihong Huang, Texas A&M University, USA
Sobha Lalitha Devi, Anna University of Chennai, India
Loic De Langhe, Ghent University, Belgium
Ekaterina Lapshinova-Koltunski, Saarland University, Germany
Sharid Loáiciga, University of Gothenburg, Sweden.
Costanza Navaretta, University of Copenhagen, Denmark
Michal Novák, Charles University in Prague, Czechia
Marta Recasens, Google, USA
Carolyn Rosé, Carnegie Mellon University, USA
Manfred Stede, University of Potsdam, Germany
Nobuhiro Ueda, Kyoto University, Japan
Yaqin Yang, Brandeis University, USA
Bonnie Webber, University of Edinburgh, UK
Juntao Yu, University of Essex, UK
Yilun Zhu, Georgetown University, USA
Heike Zinsmeister, University of Hamburg, Germany

Invited Talk (I)

Bringing together Anaphora Resolution and Linguistic Theory

Sharid Loáiciga, University of Gothenburg, Sweden

Abstract

Early work on anaphora resolution was intrinsically connected to linguistic theories of discourse interpretation. In later years, with the adoption of machine learning methods, great progress has been achieved in anaphora resolution as an independent task. This success has been even greater with current deep neural networks methods. However, the focus has been much more on solving the task than on acquiring new linguistic insights concerning anaphora resolution. In this talk, I present two ways in which we can gain and also utilize linguistic insights for anaphora resolution. First, I present experiments combining psycholinguistics with large-scale NLP tools. These show some of the complexities of hypothesis testing with corpus data. Second, I present the annotation of a multimodal corpus with anaphora information. The combination of images and text presents a unique opportunity to test our annotation schemes (i.e., our current linguistic knowledge) and to explore new ways to annotate what is unaccounted for in the same annotation schemes (i.e., new linguistic insights).

Speaker Bio

Sharid Loáiciga is a Researcher in the Department of Philosophy, Linguistics and Theory of Science at the University of Gothenburg, Sweden. She is also the Associate Director of CLASP (Centre for Linguistic Theory and Studies in Probability) in the same department. Her research is focused on discourse, and in particular on understanding human and machine interpretation of referring expressions. In recent work, she developed techniques for combining psycholinguistic methods with large-scale resources, and studied the discourse knowledge of pre-trained language models.

Invited Talk (II)

The Recent Developments in Universal Anaphora Scorer

Juntao Yu, University of Essex, UK

Michal Novák, Charles University, Prague, Czechia

Abstract

The Universal Anaphora initiative aims to push forward the state of the art in anaphora and anaphora resolution by expanding the aspects of anaphoric interpretation which are or can be reliably annotated in anaphoric corpora, producing unified standards to annotate and encode these annotations, deliver datasets encoded according to these standards, and developing methods for evaluating models carrying out this type of interpretation. Such expansion of the scope of anaphora resolution requires a comparable expansion of the scope of the scorers used to evaluate this work. Last year, we introduce an extended version of the Reference Coreference Scorer (Pradhan et al., 2014) that can be used to evaluate identity anaphora resolution (including singletons, split-antecedents), bridging reference resolution, non-referring expressions and discourse deixis. The scorer has been used in the two recent CODI-CRAC Shared Tasks on Anaphora Resolution in Dialogues. Recently, an extension of the UA scorer that supports also discontinuous markables has been used by Novák et al (2022) in the CRAC 2022 Shared Task on Multilingual Coreference Resolution. In this talk, we will introduce the details about the scorer on scoring the different aspects of anaphora resolutions and how has it been used in recently shared tasks. In addition, we will also discuss the work in progress for the scorers such as mention overlap ratio, anaphor-decomposable score and the adaptation for CRAFT shared task.

Speaker Bio

Juntao Yu is a Lecturer at the School of Computer Science and Electronic Engineering, University of Essex. Before joining Essex, He was a post-doctoral researcher at the Queen Mary University of London, working with Professor Massimo Poesio on his five-year DALI project (Disagreements and Language Interpretation, ERC-2015-AdG). He did his Ph.D. at the University of Birmingham, working on out-of-domain dependency parsing supervised by Dr Bernd Bohnet. His research interests include Deep Learning for NLP, Information Extraction, Coreference Resolution, Conversational AI, Dependency Parsing, Domain Adaptation, Semi-supervised Learning, and Multi-task Learning.

Michal Novák is a researcher at the Institute of Formal and Applied Linguistics at the Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic. He received his Ph.D. from the same university, exploring coreference and its resolution methods from cross-lingual perspective. Recently, he has co-authored the CorefUD dataset, which in its latest release harmonizes coreference of 17 corpora in 11 languages under the same annotation scheme. Besides coreference, his research also focuses on machine translation. He has participated on the Czech-Ukrainian translation system within the Charles Translator project, which aims to narrow the communication gap between Ukrainian refugees and other people in the Czech Republic.

Invited Talk (III)

The CODI/CRAC 2022 Shared Task Corpus of Anaphora Resolution in Dialogue

Massimo Poesio, Queen Mary University of London, UK
Lori Levin, Carnegie Mellon University, USA

Abstract

Most current research on anaphoric reference focuses on news text, in particular written, and on identity anaphora (coreference). This is largely due to the lack of annotated datasets of a sufficient size to train and evaluate models for other genres, and other types of anaphoric reference. Arguably the most important among the understudied genres is conversational language in dialogue. Anaphora resolution in dialogue requires systems to handle grammatically incorrect language suffering from disfluencies and mentions jointly created across utterances (Poesio & Rieser, 2010) or whose function is to establish common ground rather than refer (Clark & Brennan, 1990; Heeman & Hirst, 1995). Dialogue involves much more deictic reference, vaguer anaphoric and discourse deictic reference, speaker grounding of pronouns and long-distance conversation structure. These complexities are normally absent from news or Wikipedia articles, which constitute the bulk of current datasets for coreference resolution (Poesio et al, to appear).

The series of CODI/CRAC Shared Tasks in Anaphora Resolution in Dialogue (Khosla et al, 2021; Yu et al, 2022) was organized to address this issue by creating datasets that our community could use to study anaphoric reference in different types of conversational setups, and to tackle less studied forms of anaphoric reference such as bridging reference or discourse deixis. The annotated corpus created for the CODI/CRAC series consists of conversations from four well-known conversational datasets: the AMI corpus (Carletta, 2006), the LIGHT corpus (Urbanek et al, 2019), the PERSUASION corpus (Wang et al, 2019) and SWITCHBOARD (Godfrey et al, 1992). These documents were annotated according to the annotation scheme for the ARRAU 3 corpus, which includes guidelines for identifying discontinuous markables and annotating split antecedent plurals, bridging reference, and discourse deixis. For this second edition, we created new test sets, but also systematically checked the data annotated for the first edition. As this annotation effort also involved annotators that had not been previously involved in the ARRAU 3 annotation, this work also involved extensive discussions about the scheme; new reliability tests of the annotation scheme were carried out, and the annotation guidelines were substantially revised.

Speaker Bio

Massimo Poesio is a full professor in Computational Linguistics at the School of Electronic Engineering and Computer Science, Queen Mary University of London, and a member of the University's Cognitive Science and Games and AI research groups. He is also a Fellow of the Turing Institute, a supervisor in the IGGI Doctoral training centre in Intelligent Games and Game Intelligence and the Wellcome Trust's Ph.D. programme in Health Data in Practice. He is co-founder and have been Associate Editor of Dialogue and Discourse since its foundation and he recently became co-editor of the Computational and Mathematical section of Language and Linguistics Compass.

Lori Levin has a Ph.D. in linguistics and has been working in the fields of computational linguistics and natural language processing since the 1980's, where she uses her expertise in linguistics in the annotation of corpora and the design of meaning representations. She specializes in NLP for low-resource and endangered languages. She is the co-founder and co-chair of the North American Computational Linguistics Open competition.

Table of Contents

<i>Quantifying Discourse Support for Omitted Pronouns</i> Shulin Zhang, Jixing Li and John Hale	1
<i>Online Neural Coreference Resolution with Rollback</i> Patrick Xia and Benjamin Van Durme	13
<i>Analyzing Coreference and Bridging in Product Reviews</i> Hideo Kobayashi and Christopher Malon	22
<i>Anaphoric Phenomena in Situated dialog: A First Round of Annotations</i> Sharid Loáiciga, Simon Dobnik and David Schlangen	31
<i>Building a Manually Annotated Hungarian Coreference Corpus: Workflow and Tools</i> Noémi Vadász	38
<i>NARC – Norwegian Anaphora Resolution Corpus</i> Petter Mæhlum, Dag Haug, Tollef Jørgensen, Andre Kåsen, Anders Nøklestad, Egil Rønningstad, Per Erik Solberg, Erik Velldal and Lilja Øvreid	48
<i>Evaluating Coreference Resolvers on Community-based Question Answering: From Rule-based to State of the Art</i> Haixia Chai, Nafise Sadat Moosavi, Iryna Gurevych and Michael Strube	61
<i>Improving Bridging Reference Resolution using Continuous Essentiality from Crowdsourcing</i> Nobuhiro Ueda and Sadao Kurohashi	74
<i>Investigating Cross-Document Event Coreference for Dutch</i> Loic De Langhe, Orphee De Clercq and Veronique Hoste	88
<i>The Role of Common Ground for Referential Expressions in Social Dialogues</i> Jaap Kruijt and Piek Vossen	99

Workshop Program

Sunday, October 16, 2022

Welcome

13:30–13:40 *Opening and Welcome*

CRAC—Invited Talk

13:40–14:40 *Bringing together Anaphora Resolution and Linguistic Theory*
Sharid Loáiciga

Paper Session I

14:40–15:00 *Quantifying Discourse Support for Omitted Pronouns*
Shulin Zhang, Jixing Li and John Hale

15:00–15:10 *Online Neural Coreference Resolution with Rollback*
Patrick Xia and Benjamin Van Durme

15:10–15:20 *Analyzing Coreference and Bridging in Product Reviews*
Hideo Kobayashi and Christopher Malon

15:20–15:30 *Anaphoric Phenomena in Situated dialog: A First Round of Annotations*
Sharid Loáiciga, Simon Dobnik and David Schlangen

Short Break

15:30–16:00 *Coffee Break*

Sunday, October 16, 2022 (continued)

Paper Session II

- 16:00–16:20 *Building a Manually Annotated Hungarian Coreference Corpus: Workflow and Tools*
Noémi Vadász
- 16:20–16:40 *NARC – Norwegian Anaphora Resolution Corpus*
Petter Mæhlum, Dag Haug, Tollef Jørgensen, Andre Kåsen, Anders Nøklestad, Egil Rønningstad, Per Erik Solberg, Erik Velldal and Lilja Øvrelid
- 16:40–17:00 *Evaluating Coreference Resolvers on Community-based Question Answering: From Rule-based to State of the Art*
Haixia Chai, Nafise Sadat Moosavi, Iryna Gurevych and Michael Strube
- 17:00–17:20 *Improving Bridging Reference Resolution using Continuous Essentiality from Crowdsourcing*
Nobuhiro Ueda and Sadao Kurohashi
- 17:20–17:40 *Investigating Cross-Document Event Coreference for Dutch*
Loic De Langhe, Orphee De Clercq and Veronique Hoste
- 17:40–18:00 *The Role of Common Ground for Referential Expressions in Social Dialogues*
Jaap Kruijt and Piek Vossen

Closing

- 13:30–13:40 *Closing Remarks*

Monday, October 17, 2022

CRAC Shared Task Session

- 9:00–9:30 *Findings of the Shared Task on Multilingual Coreference Resolution*
Michal Novák, Zdeněk Žabokrtský, Miloslav Konopík, Anna Nedoluzhko, Michal Novák, Maciej Ogrodniczuk, Martin Popel, Ondřej Pražák, Jakub Sido, Daniel Zeman and Yilun Zhu
- 9:30–9:45 *Coreference Resolution for Polish: Improvements within the CRAC 2022 Shared Task*
Karol Saputa
- 9:45–10:00 *End-to-end Multilingual Coreference Resolution with Mention Head Prediction*
Ondřej Pražák and Miloslav Konopík
- 10:00–10:15 *ÚFAL CorPipe at CRAC 2022: Effectivity of Multilingual Models for Coreference Resolution*
Milan Straka and Jana Straková

Discussion

- 10:15–10:30 *Open Discussion*

Short Break

- 10:30–11:00 *Coffee Break*

CRAC Shared Task—Invited Talk

- 11:00–12:00 *The Recent Developments in Universal Anaphora Scorer*
Juntao Yu and Michal Novák

Monday, October 17, 2022 (continued)

Panel

12:00–12:30 *Panel Discussion—Universal Anaphora*
Sameer Pradhan

Long Break

12:30–14:00 *Lunch Break*

CODI-CRAC Joint Shared Task Session

14:00–14:05 *Welcome*

14:05–14:30 *The CODI-CRAC 2022 Shared Task on Anaphora, Bridging, and Discourse Deixis in Dialogue*

Juntao Yu, Sopan Khosla, Ramesh Manuvinakurike, Lori Levin, Vincent Ng, Massimo Poesio, Michael Strube and Carolyn Rosé

14:35–14:50 *Anaphora Resolution in Dialogue: System Description—CODI-CRAC 2022 Shared Task*

Tatiana Anikina, Natalia Skachkova, Joseph Renner and Priyansh Trivedi

14:50–15:05 *Pipeline Coreference Resolution for Anaphoric Identity in Dialogues*

Damrin Kim, Seongsik Park, Mirae Han and Harksoo Kim

15:05–15:30 *Neural Anaphora Resolution in Dialogue Revisited*

Shengjie Li, Hideo Kobayashi and Vincent Ng

Monday, October 17, 2022 (continued)

CODI-CRAC Joint Shared Task—Invited Talk

16:00–16:45 *The CODI/CRAC 2022 Shared Task Corpus of Anaphora Resolution in Dialogue*
Massimo Poesio and Lori Levin

CODI-CRAC Joint Shared Task Discussion and Closing Remarks

16:45–17:45 *Open Discussion*

17:45–18:00 *Closing Remarks*

