

EMNLP 2024

**Proceedings of The Seventh Workshop on Computational Models of
Reference, Anaphora and Coreference (CRAC 2024)**

CRAC 2024, an EMNLP 2024 Workshop
November 15, 2024
Miami, FL 33131-2197
USA

©2024 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 979-8-89176-171-1

Message from the Program Chairs

Time flies. This is already the seventh edition of the Workshop on Computational Models of Reference, Anaphora and Coreference (CRAC). After two consecutive years in Asia, CRAC returned to North America this year.

Here is a bit of history for those of you who are participating in the workshop for the first time. CRAC was first held in New Orleans six years ago in conjunction with NAACL HLT 2018. The workshop series, however, dates back to its predecessor, Coreference Resolution Beyond OntoNotes (CORBON), which 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.

To facilitate the planning of the workshop, we decided in late 2019 that starting in 2020, CRAC would be held towards the end of each calendar year. Since then, we have received a healthy number of submissions every year, until this year when the number of submissions was lower than expected. Specifically, we received eight submissions, all of which were rigorously reviewed by three to four program committee members. Based on their recommendations, we accepted six 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. While we are still investigating the reasons for the decline in the number of submissions, we speculate that the proximity of the CRAC submission deadline to the COLING 2025 submission deadline might have played a role.

This year we continued to partner with our colleagues at Charles University, Prague and hosted the shared task on *Multilingual Coreference Resolution* for the third time at CRAC. The shared task 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. In a departure from last year, we decided to merge the shared task proceedings with the CRAC workshop proceedings this year. In other words, you can enjoy both the workshop papers and the shared task papers in this proceedings.

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 average reviewing load was three papers per reviewer. All of our program committee members did the incredible job of completing their reviews in a short reviewing period. Second, we thank Jackie Chi-Kit Cheung, an established researcher in Discourse and Coreference, for accepting our invitation to be this year's invited speaker. Those of us who care about the future of coreference would not want to miss our panel discussion on "Coreference resolution in the era of LLMs," which will be led by our colleagues from Prague. Finally, we would like to thank the workshop participants for joining us in this event.

We hope you will enjoy the workshop and sunny Miami as much as we do!

— Sameer Pradhan, Maciej Ogrodniczuk, Michal Novák, Massimo Poesio, and Vincent Ng

Organizers

Organizing Committee:

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

Program Committee:

Arie Cattan, Bar-Ilan University, Israel
Sobha Lalitha Devi, Anna University of Chennai, India
Elisa Ferracane, Abridge
Yulia Grishina, Amazon, USA
Christian Hardmeier, IT University of Copenhagen, Denmark
Lars Hellan, Norwegian University of Science and Technology, Norway
Veronique Hoste, Ghent University, Belgium
Ekaterina Lapshinova-Koltunski, University of Hildesheim, Germany
Sharid Loáiciga, University of Gothenburg, Sweden.
Costanza Navaretta, University of Copenhagen, Denmark
Michal Novák, Charles University in Prague, Czechia
Massimo Poesio, Queen Mary University of London, UK
Ian Porada, Mila–Quebec Artificial Intelligence Institute, Canada
Carolyn Rosé, Carnegie Mellon University, USA
Juntao Yu, University of Essex, UK
Amir Zeldes, Georgetown University, USA
Yilun Zhu, Georgetown University, USA

Invited Talk

Reference at the Heart of Natural Language Processing

Jackie Chi Kit Cheung, McGill University, Montreal, Canada

Abstract

Natural language is traditionally framed as a mapping from form to content, with reference being the connection between the two. Yet curiously, large language models have achieved impressive levels of performance and adoption through training on distributional signals, which concerns form alone. In this talk, I argue for the importance of reference and coreference in NLP, and discuss topics in NLP which are touched by these phenomena, including model "hallucinations" and factual errors, knowledge updating, common sense reasoning, and conversational agents. I discuss how existing evaluation practices based on large-scale benchmarking often masks the importance of reference-related phenomena, and present work from my lab that reflects on current evaluation practices and their validity. I call for more serious consideration of reference including targeted evaluation of reference-related phenomena as a necessary step towards achieving robust NLP systems.

Speaker Bio

Jackie Chi Kit Cheung is an associate professor at McGill University's School of Computer Science, where he co-directs the Reasoning and Learning Lab. He is a Canada CIFAR AI Chair and an Associate Scientific Co-Director at the Mila Quebec AI Institute. His research focuses on topics in natural language generation such as automatic summarization, and on integrating diverse knowledge sources into NLP systems for pragmatic and common-sense reasoning. He also works on applications of NLP to domains such as education, health, and language revitalization. He is motivated by how the structure of the world can be reflected in the structure of language processing systems. He is a consulting researcher at Microsoft Research Montreal.

Table of Contents

<i>Major Entity Identification: A Generalizable Alternative to Coreference Resolution</i> Kawshik S. Manikantan, Shubham Toshniwal, Makarand Tapaswi and Vineet Gandhi	1
<i>Enriching Conceptual Knowledge in Language Models through Metaphorical Reference Explanation</i> Zixuan Zhang and Heng Ji	18
<i>Polish Coreference Corpus as an LLM Testbed: Evaluating Coreference Resolution within Instruction-Following Language Models by Instruction–Answer Alignment</i> Karol Saputa, Angelika Peljak-Łapińska and Maciej Ogrodniczuk	23
<i>MSCAW-coref: Multilingual, Singleton and Conjunction-Aware Word-Level Coreference Resolution</i> Houjun Liu, John Bauer, Karel D’Oosterlinck, Christopher Potts and Christopher D. Manning	33
<i>Unifying the Scope of Bridging Anaphora Types in English: Bridging Annotations in ARRAU and GUM</i> Lauren Levine and Amir Zeldes	41
<i>WinoPron: Revisiting English Winogender Schemas for Consistency, Coverage, and Grammatical Case</i> Vagrant Gautam, Julius Steuer, Eileen Bingert, Ray Johns, Anne Lauscher and Dietrich Klakow	52
<i>DeepHCoref: A Deep Neural Coreference Resolution for Hindi Text</i> Kusum Lata, Pardeep Singh, Kamlesh Dutta and Abhishek Kanwar	67
<i>Findings of the Third Shared Task on Multilingual Coreference Resolution</i> Michal Novák, Barbora Dohnalová, Miloslav Konopík, Anna Nedoluzhko, Martin Popel, Ondrej Prazak, Jakub Sido, Milan Straka, Zdeněk Žabokrtský and Daniel Zeman	78
<i>CorPipe at CRAC 2024: Predicting Zero Mentions from Raw Text</i> Milan Straka	97
<i>End-to-end Multilingual Coreference Resolution with Headword Mention Representation</i> Ondrej Prazak and Miloslav Konopík	107
<i>Multilingual coreference resolution as text generation</i> Natalia Skachkova	114

Workshop Program

Friday, November 15, 2024

Opening Remarks

09:00–09:15 *Opening and Welcome*
Vincent Ng, Maciej Ogrodniczuk

Invited Talk

09:15–10:30 *Reference at the Heart of Natural Language Processing*
Jackie Chi Kit Cheung

Short Break

10:30–11:00 *Coffee Break*

Findings Paper Session

11:00–11:20 *Challenges to Evaluating the Generalization of Coreference Resolution Models: A Measurement Modeling Perspective*
Ian Porada, Alexandra Olteanu, Kaheer Suleman, Adam Trischler and Jackie Chi Kit Cheung

11:20–11:40 *Any Other Thoughts, Hedgehog? Linking Deliberation Chains in Collaborative Dialogues*
Abhijnan Nath, Videep Venkatesha, Mariah Bradford, Avyakta Chelle, Austin Collin Youngren, Carlos Mabrey, Nathaniel Blanchard and Nikhil Krishnaswamy

11:40–11:50 *MMAR: Multilingual and Multimodal Anaphora Resolution in Instructional Videos*
Cennet Oguz, Pascal Denis, Simon Ostermann, Emmanuel Vincent, Natalia Skachkova and Josef van Genabith

Friday, November 15, 2024 (continued)

EMNLP Paper

11:50–12:10 *Major Entity Identification: A Generalizable Alternative to Coreference Resolution*
Kawshik S. Manikantan, Shubham Toshniwal, Makarand Tapaswi and Vineet Gandhi

Long Break

12:10–13:50 *Lunch Break*

Research Paper Session

13:50–14:00 *Enriching Conceptual Knowledge in Language Models through Metaphorical Reference Explanation*
Zixuan Zhang and Heng Ji

14:00–14:10 *Polish Coreference Corpus as an LLM Testbed: Evaluating Coreference Resolution within Instruction-Following Language Models by Instruction-Answer Alignment*
Karol Saputa, Angelika Peljak-Łapińska and Maciej Ogrodniczuk

14:10–14:30 *MSCAW-coref: Multilingual, Singleton and Conjunction-Aware Word-Level Coreference Resolution*
Houjun Liu, John Bauer, Karel D’Oosterlinck, Christopher Potts and Christopher D. Manning

14:30–14:50 *Unifying the Scope of Bridging Anaphora Types in English: Bridging Annotations in ARRAU and GUM*
Lauren Levine and Amir Zeldes

14:50–15:10 *WinoPron: Revisiting English Winogender Schemas for Consistency, Coverage, and Grammatical Case*
Vagrant Gautam, Julius Steuer, Eileen Bingert, Ray Johns, Anne Lauscher and Dietrich Klakow

15:10–15:30 *DeepHCoref: A Deep Neural Coreference Resolution for Hindi Text*
Kusum Lata, Pardeep Singh, Kamlesh Dutta and Abhishek Kanwar

Friday, November 15, 2024 (continued)

Short Break

15:30–16:00 *Coffee Break*

Shared Task Paper Session

16:00–16:30 *Findings of the Third Shared Task on Multilingual Coreference Resolution*
Michal Novák, Barbora Dohnalová, Miloslav Konopík, Anna Nedoluzhko, Martin Popel, Ondřej Prazak, Jakub Sido, Milan Straka, Zdeněk Žabokrtský and Daniel Zeman

16:30–16:50 *CorPipe at CRAC 2024: Predicting Zero Mentions from Raw Text*
Milan Straka

16:50–17:10 *End-to-end Multilingual Coreference Resolution with Headword Mention Representation*
Ondřej Prazak and Miloslav Konopík

17:10–17:20 *Multilingual coreference resolution as text generation*
Natalia Skachkova

Panel Discussion

17:20–17:50 *The future of coreference resolution in the era of LLMs*
Michal Novák, Ondřej Pražák and Martin Popel

Closing Remarks

17:50–18:00 *Closing of the workshop*
Maciej Ogrodniczuk

