

StarSEM 2024

**The 13th Joint Conference on Lexical and Computational
Semantics**

Proceedings of the Conference (*SEM 2024)

June 20-21, 2024

The StarSEM organizers gratefully acknowledge the support from the following sponsors.

Sponsored by



©2024 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 979-8-89176-106-3

Preface by the Conference Organizers

We are excited to welcome you to *SEM 2024, the 13th Joint Conference on Lexical and Computational Semantics! We are pleased to present this volume containing the accepted long and short papers. *SEM 2024 is being held from June 20 to 21, 2024, in Mexico City, Mexico, co-located with NAACL 2024.

Since its first edition in 2012, *SEM has become a major venue to present recent advances in all areas of lexical and computational semantics, including semantic representations, theoretical semantics, multilingual semantics, and others. *SEM is sponsored by SIGLEX, the ACL Special Interest Group on the Lexicon.

*SEM 2024 accepted both papers submitted directly to *SEM and those already reviewed through ARR (ACL Rolling Review). We received submissions in 11 areas:

- Lexical Semantics
- Semantic Composition and Sentence-level Semantics
- Discourse, Dialogue and Generation
- Commonsense Reasoning and NLU
- Resources and Evaluation
- Theoretical and Formal Semantics
- Multilinguality
- Semantics in NLP Applications
- Psycholinguistics, Cognitive Linguistics, and Semantic Processing
- Social Biases and Ethics
- Interpretability and Explainability

We had 65 submissions this year combining both direct submissions and ARR commits. We compiled an exciting and wide-ranging program, accepting a total of 35 papers (27 long papers and 8 short papers). The submitted papers were carefully evaluated by a program committee led by 14 area chairs, who coordinated a large team of reviewers. The reviews were almost all of very high-quality, and for that we are extremely grateful! Area chairs then added meta-reviews to explain their accept/reject decisions. The final selection was made by the program co-chairs after a careful check of the reviews, meta-reviews, and discussions with the area chairs. We are also very excited to have two excellent keynote speakers: Greg Durrett from the University of Texas at Austin and Heng Ji from the University of Illinois Urbana-Champaign.

We are honored to serve as the organizing committee for *SEM 2024, and we absolutely could not have made this happen without a huge amount of help. First, tremendous thanks to all area chairs and reviewers for their invaluable help in selecting the program, for their engagement in thoughtful discussions, and for providing valuable feedback to authors. Second, thanks to our Publicity Chair Yi Zhou (Cardiff University) for taking care of the website and social media updates. Next, thanks to our Publication Chair Tom McCoy (Yale University) for putting together the proceedings, and to the NAACL 2024 workshop organizers for help and support with all organizational aspects of the conference. Finally, thank you to the authors and presenters for making *SEM 2024 such an engaging and exciting event! We hope that you will find the content of these proceedings as engaging as we do, and we hope to see you at future iterations of *SEM!

Danushka Bollegala and Vered Shwartz, co-Program Chairs
Jose Camacho-Collados, General Chair

Organizing Committee

General Chair

Jose Camacho-Collados, Cardiff University, UK

Program Chairs

Danushka Bollegala, Amazon and University of Liverpool, UK

Vered Shwartz, University of British Columbia and Vector Institute, Canada

Publicity Chair

Yi Zhou, Cardiff University, UK

Publication Chair

Tom McCoy, Yale University, USA

Program Committee

Area Chairs

Jackie Chi Kit Cheung, McGill University and Mila
Annemarie Friedrich, University of Augsburg
Hila Gonen, University of Washington
Masahiro Kaneko, MBZUAI and Tokyo Institute of Technology
Mamoru Komachi, Hitotsubashi University
Debjit Paul, EPFL
Enrico Santus, Bloomberg LP
Ryohei Sasano, Nagoya University and RIKEN
Asad Sayeed, University of Gothenburg
Dominik Schlechtweg, University of Stuttgart
Procheta Sen, University of Liverpool
Gabriel Stanovsky, The Hebrew University of Jerusalem
Elior Sulem, Ben-Gurion University of the Negev
Chris Tanner, Kensho Technologies and Massachusetts Institute of Technology

Program Committee

Sweta Agrawal, Instituto de Telecomunicações
Shafiuddin Rehan Ahmed, University of Colorado, Boulder
Kenichiro Ando, RIKEN
Marianna Apidianaki, University of Pennsylvania
Mohamad Hardyman Barawi, Universiti Malaysia Sarawak
Valerio Basile, University of Turin
Gábor Berend, University of Szeged
Uri Berger, Hebrew University of Jerusalem
Alexander Berman, Göteborg University
Eduardo Blanco, University of Arizona
Michael Bloodgood, The College of New Jersey
Davide Buscaldi, Ecole polytechnique
Elena Cabrio, Université Côte d'Azur
Aoife Cahill, Dataminr
Pierluigi Cassotti, Göteborg University
Akshay Chaturvedi, IRIT, Toulouse, France
Zhousi Chen, Hitotsubashi University
Emmanuele Chersoni, The Hong Kong Polytechnic University
Patricia Chiril, University of Chicago
Bonaventura Coppola, SAP Security Research
Shaobo Cui, EPFL - EPF Lausanne
Noam Dahan, Hebrew University of Jerusalem
Jad Doughman, Mohamed bin Zayed University of Artificial Intelligence
Yo Ehara, Tokyo Gakugei University
Luis Espinosa-Anke, Cardiff University
Francis Ferraro, University of Maryland, Baltimore County
Simona Frenda, University of Turin
Richard Futrell, University of California, Irvine
Voula Giouli, Aristotle University of Thessaloniki

Colin S. Gordon, Drexel University
Venkata Subrahmanyam Govindarajan, University of Texas, Austin
Ivan Habernal, Ruhr-Universität Bochum
Udo Hahn, Friedrich-Schiller Universität Jena
Yi Han, Nagoya University
Yoshihiko Hayashi, Waseda University
Mete Ismayilzada, EPFL - EPF Lausanne
Abhik Jana, IIT Bhubaneswar
Abhishek Kaushik, DkIT
Najoung Kim, Boston University
Thomas H Kober, Zalando SE
Patrick Lee, Montclair State University
Els Lefever, Ghent University
Alisa Liu, University of Washington
Chunhua Liu, University of Melbourne
Renze Lou, Pennsylvania State University
Qing Lyu, University of Pennsylvania
Alda Mari, CNRS
Eugenio Martínez-Cámara, Universidad de Jaén
Sahisnu Mazumder, Intel Labs, USA
Yisong Miao, National University of Singapore
Timothee Mickus, University of Helsinki
Amita Misra, Amazon
Moran Mizrahi, Hebrew University of Jerusalem
Nona Naderi, Université Paris-Saclay
Christina Niklaus, Universität St. Gallen
Ayana Niwa, Megagon Labs
Alexis Palmer, University of Colorado, Boulder
Emerson Cabrera Paraiso, Pontificia Universidade Católica do Paraná
Alessandro Raganato, University of Milan - Bicocca
Sahithya Ravi, University of British Columbia
Michael Regan, University of Washington
Alla Rozovskaya, City University of New York
Irene Russo, Consiglio Nazionale delle Ricerche
Farig Sadeque, BRAC University
Steven Schockaert, Cardiff University
Sabine Schulte im Walde, University of Stuttgart
Sebastian Schuster, University College London, University of London
Esther Shizgal, Hebrew University of Jerusalem
Melanie Siegel, Darmstadt University of Applied Sciences
Egon Stemle, Masaryk University
Kevin Stowe, TU Darmstadt
Sara Stymne, Uppsala University
Aarne Talman, University of Helsinki
Andon Tchechmedjiev, IMT Mines Alès
Samia Touileb, University of Bergen
Enrica Troiano, Vrije Universiteit Amsterdam
Thinh Hung Truong, University of Melbourne
Sowmya Vajjala, National Research Council Canada
Eva Maria Vecchi, University of Stuttgart
Shira Wein, Georgetown University

Genta Indra Winata, Bloomberg
Tak-Lam Wong, Douglas College
Hanzi Xu, Temple University
He Zhou, Hong Kong Polytechnic University

Table of Contents

<i>MASSIVE Multilingual Abstract Meaning Representation: A Dataset and Baselines for Hallucination Detection</i>	
Michael Regan, Shira Wein, George Baker and Emilio Monti	1
<i>How Does Stereotype Content Differ across Data Sources?</i>	
Kathleen C. Fraser, Svetlana Kiritchenko and Isar Nejadgholi	18
<i>Polysemy through the lens of psycholinguistic variables: a dataset and an evaluation of static and contextualized language models</i>	
Andrea Bruera, Farbod Zamani and Massimo Poesio	35
<i>Post-Hoc Answer Attribution for Grounded and Trustworthy Long Document Comprehension: Task, Insights, and Challenges</i>	
Abhilasha Sancheti, Koustava Goswami and Balaji Vasani Srinivasan	49
<i>A Benchmark Suite of Japanese Natural Questions</i>	
Takuya Uematsu, Hao Wang, Daisuke Kawahara and Tomohide Shibata	58
<i>ROUGE-K: Do Your Summaries Have Keywords?</i>	
Sotaro Takeshita, Simone Paolo Ponzetto and Kai Eckert	69
<i>Investigating Aspect Features in Contextualized Embeddings with Semantic Scales and Distributional Similarity</i>	
Yuxi Li, Emmanuele Chersoni and Yu-Yin Hsu	80
<i>WikiScenes with Descriptions: Aligning Paragraphs and Sentences with Images in Wikipedia Articles</i>	
Özge Alaçam, Ronja Utescher, Hannes Gröner, Judith Sieker and Sina Zarriß	93
<i>Relevance, Diversity, and Exclusivity: Designing Keyword-augmentation Strategy for Zero-shot Classifiers</i>	
Taro Yano, Kunihiro Takeoka and Masafumi Oyamada	106
<i>Lexical Substitution as Causal Language Modeling</i>	
Ning Shi, Bradley Hauer and Grzegorz Kondrak	120
<i>Paraphrase Identification via Textual Inference</i>	
Ning Shi, Bradley Hauer, Jai Riley and Grzegorz Kondrak	133
<i>Identifying Emotional and Polar Concepts via Synset Translation</i>	
Logan Woudstra, Moyo Dawodu, Frances Igwe, Senyu Li, Ning Shi, Bradley Hauer and Grzegorz Kondrak	142
<i>A Closer Look at Claim Decomposition</i>	
Miriam Wanner, Seth Ebner, Zhengping Jiang, Mark Dredze and Benjamin Van Durme	153
<i>Speedy Gonzales: A Collection of Fast Task-Specific Models for Spanish</i>	
José Cañete and Felipe Bravo-Marquez	176
<i>Exploring Factual Entailment with NLI: A News Media Study</i>	
Guy Mor-Lan and Effi Levi	190
<i>The Emergence of High-Level Semantics in a Signaling Game</i>	
Timotheé Bernard, Timothee Mickus and Hiroya Takamura	200

<i>PDDLEGO: Iterative Planning in Textual Environments</i>	
Li Zhang, Peter Jansen, Tianyi Zhang, Peter Clark, Chris Callison-Burch and Niket Tandon .	212
<i>VOLIMET: A Parallel Corpus of Literal and Metaphorical Verb-Object Pairs for English–German and English–French</i>	
Prisca Piccirilli, Alexander Fraser and Sabine Schulte im Walde	222
<i>Improving Word Sense Induction through Adversarial Forgetting of Morphosyntactic Information</i>	
Deniz Ekin Yava, Timothée Bernard, Laura Kallmeyer and Benoît Crabbé	238
<i>What’s wrong with your model? A Quantitative Analysis of Relation Classification</i>	
Elisa Bassignana, Rob van der Goot and Barbara Plank	252
<i>Disambiguating Emotional Connotations of Words Using Contextualized Word Representations</i>	
Akram Sadat Hosseini and Steffen Staab	264
<i>Length-Aware Multi-Kernel Transformer for Long Document Classification</i>	
Guangzeng Han, Jack Tsao and Xiaolei Huang	278
<i>Investigating Wit, Creativity, and Detectability of Large Language Models in Domain-Specific Writing Style Adaptation of Reddit’s Showerthoughts</i>	
Tolga Buz, Benjamin Frost, Nikola Genchev, Moritz Schneider, Lucie-Aimée Kaffee and Gerard de Melo	291
<i>Multilingual and Code-Switched Sentence Ordering</i>	
Alexandre Salle and Shervin Malmasi	308
<i>HANS, are you clever? Clever Hans Effect Analysis of Neural Systems</i>	
Leonardo Ranaldi and Fabio Massimo Zanzotto	314
<i>Exploring Semantics in Pretrained Language Model Attention</i>	
Frédéric Charpentier, Jairo Cugliari and Adrien Guille	326
<i>Enhancing Self-Attention via Knowledge Fusion: Deriving Sentiment Lexical Attention from Semantic-Polarity Scores</i>	
Dongjun Jang, Jinwoong Kim and Hyopil Shin	334
<i>Handling Ontology Gaps in Semantic Parsing</i>	
Andrea Bacciu, Marco Damonte, Marco Basaldella and Emilio Monti	345
<i>PipeNet: Question Answering with Semantic Pruning over Knowledge Graphs</i>	
Ying Su, Jipeng Zhang, Yangqiu Song and Tong Zhang	360
<i>A Trip Towards Fairness: Bias and De-Biasing in Large Language Models</i>	
Leonardo Ranaldi, Elena Sofia Ruzzetti, Davide Venditti, Dario Onorati and Fabio Massimo Zanzotto	372
<i>Compositional Structured Explanation Generation with Dynamic Modularized Reasoning</i>	
Xiyan Fu and Anette Frank	385
<i>Inspecting Soundness of AMR Similarity Metrics in terms of Equivalence and Inequivalence</i>	
Kyung Seo Ki, Bugeun Kim and Gahgene Gweon	402
<i>Sõnajaht: Definition Embeddings and Semantic Search for Reverse Dictionary Creation</i>	
Aleksei Dorkin and Kairit Sirts	410

Do large language models and humans have similar behaviours in causal inference with script knowledge?

Xudong Hong, Margarita Ryzhova, Daniel Adrian Biondi and Vera Demberg 421

EDM3: Event Detection as Multi-task Text Generation

Ujjwala Anantheswaran, Himanshu Gupta, Mihir Parmar, Kuntal Kumar Pal and Chitta Baral 438

Program

Thursday, June 20, 2024

- 09:00 - 09:30 *Session 1: Opening Remarks*
- 09:30 - 10:30 *Invited Talk: Heng Ji*
- 10:30 - 11:00 *Break*
- 11:00 - 12:30 *Session 2: Applications*
- 12:30 - 14:00 *Lunch Break*
- 14:00 - 15:30 *Session 3: Discourse and Semantic Composition*
- 15:30 - 16:00 *Break*
- 16:00 - 17:30 *Session 4: In-Person Posters 1*

Friday, June 21, 2024

- 08:30 - 09:30 *Session 5: Virtual Poster Session*
- 09:30 - 10:30 *Invited Talk: Greg Durrett*
- 10:30 - 11:00 *Break*
- 11:00 - 12:30 *Session 6: Generation*
- 12:30 - 14:00 *Lunch Break*
- 14:00 - 15:30 *Session 7: In-Person Posters 2*
- 15:30 - 16:00 *Break*
- 16:00 - 17:00 *Session 8: Lexical Semantics and Closing Remarks*