

EMNLP 2020

**Spatial Language Understanding
(SpLU 2020)**

Proceedings of the Third Workshop

November 19, 2020

©2020 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 978-1-952148-72-9

Introduction

Welcome to the Third International Workshop on Spatial Language Understanding (SpLU 2020). The goal of this workshop is to bring together researchers working on computational models for spatial language understanding from different perspectives including cognitive linguistic studies and spatial information representations for machine learning and reasoning. Spatial semantics of the language is important in many real-world language understanding and language generation applications which makes it worth special attention. Examples of such applications are human-machine interaction and grounded dialogue systems, geographical information systems, systems that need combining vision and language such as visual question answering, navigation, and instruction-following agents. Since spatial semantics is related to the physical world, it can be visualized (except for the metaphoric usage). Thus, such semantics are important to be explored for text-to-scene conversion and vice-versa, that is, for text/caption generation for visual information such as images and videos. For this workshop, we called for research papers on spatial language meaning representations (continuous, symbolic), corpus collection and analysis, learning to extract spatial information, reasoning based on spatial language, spatial language grounding, and combining vision and language, and applications of spatial language in various downstream tasks.

This year, we received 13 submissions directly to the workshop (10 in the archival track and 3 in the non-archival track). After a rigorous review process, we accepted 7 submissions in the archival track and 3 in the nonarchival track. Moreover, based on the ACL new policies and the new Findings papers track, 11 papers from Findings were directed to the SpLU workshop from which we accepted 8 papers based on their relevance to the theme of the workshop. This volume is dedicated to the 7 papers from the archival track. The main contributions from the accepted papers in all the tracks (archival, non-archival, and Findings) range from all of the aforementioned topics. The workshop includes papers on cognitively motivated approaches as well as deep learning language models for spatial information extraction and location recognition, spatial question answering, navigation and grounding, instruction following and assembly tasks, and even applications of spatial information extraction in the medical domain and patient reports. It covers, data collection, introducing new corpora, annotation scheme and datasets, spatial information in Dialogue systems, and generating scenes based on spatial descriptions. The topics cover exciting research works that highlight the current efforts, challenges, and future endeavors for spatial language understanding research.

We thank our programming committee members, invited speakers, and authors, without whose help and contributions, this workshop would not have been possible. We look forward to seeing you virtually on November 19, 2020.

Organizers:

Parisa Kordjamshidi, Michigan State University
Archana Bhatia, Institute for Human and Machine Cognition
Malihe Alikhani, University of Pittsburgh
Jason Baldrige, Google
Mohit Bansal, UNC Chapel Hill
Marie-Francine Moens, KU Leuven

Program Committee:

Steven Bethard, The University of Arizona (USA)
Raffaella Bernardi, University of Trento (Italy)
Mehul Bhatt, Örebro University - CoDesign Lab (Sweden)
Yonatan Bisk, Carnegie Mellon University (USA)
Johan Bos, University of Groningen (The Netherlands)
Asli Celikyilmaz, Microsoft Research (USA)
Joyce Chai, University of Michigan (USA)
Angel Xuan Chang, Simon Fraser University (Canada)
Anthony Cohn, University of Leeds (England)
Guillem Collell, KU Leuven (Belgium)
Simon Dobnik, University of Gothenburg (Sweden)
Bonnie J. Dorr, Institute for Human and Machine Cognition (USA)
Ekaterina Egorova, University of Zurich (Switzerland)
Zoe Falomir, Universitat Bremen (Germany)
Francis Ferraro, University of Maryland Baltimore (USA)
Lucian Galescu, Institute for Human and Machine Cognition (USA)
Mehdi Ghanimifard, University of Gothenburg (Sweden)
Julia Hockenmaier, University of Illinois at Urbana-Champaign (USA)
Lei Li, Bytedance (China)
Bruno Martins, University of Lisbon (Portugal)
Srin Narayanan, Google Inc. (USA)
Mari Broman Olsen, Lionbridge AI (USA)
Martijn van Otterlo, Open University (The Netherlands)
Ian Perera, Institute for Human and Machine Cognition (USA)
Kirk Roberts, UT Health (USA)
Manolis Savva, Stanford University (USA)
Kristin Stock, Massey University (New Zealand)
Jesse Thomason, University of Washington (USA)
Clare Voss, ARL (USA)

Invited Speaker:

Yoav Artzi, Cornell University
Bonnie Dorr, Institute for Human and Machine Cognition
Julia Hockenmaier, University of Illinois at Urbana-Champaign
Douwe Kiela, Facebook
James Pustejovsky, Brandeis University

Table of Contents

<i>An Element-wise Visual-enhanced BiLSTM-CRF Model for Location Name Recognition</i> Takuya Komada and Takashi Inui	1
<i>BERT-based Spatial Information Extraction</i> Hyeong Jin Shin, Jeong Yeon Park, Dae Bum Yuk and Jae Sung Lee	10
<i>A Cognitively Motivated Approach to Spatial Information Extraction</i> Chao Xu, Emmanuelle-Anna Dietz Saldanha, Dagmar Gromann and Beihai Zhou	18
<i>They are not all alike: answering different spatial questions requires different grounding strategies</i> Alberto Testoni, Claudio Greco, Tobias Bianchi, Mauricio Mazuecos, Agata Marcante, Luciana Benotti and Raffaella Bernardi	29
<i>Categorisation, Typicality & Object-Specific Features in Spatial Referring Expressions</i> Adam Richard-Bollans, Anthony Cohn and Lucía Gómez Álvarez	39
<i>A Hybrid Deep Learning Approach for Spatial Trigger Extraction from Radiology Reports</i> Surabhi Datta and Kirk Roberts	50
<i>Retouchdown: Releasing Touchdown on StreetLearn as a Public Resource for Language Grounding Tasks in Street View</i> Harsh Mehta, Yoav Artzi, Jason Baldridge, Eugene Ie and Piotr Mirowski	56

Workshop Program

Thursday, November 19, 2020

8:00–9:00 **GatherTown I**

9:00–10:48 **Session 1**

9:00–9:10 *Opening Talk*

9:10–10:00 *Invited talk: James Pustejovsky*

10:00–10:56 *Paper Presentations (SpLU, Findings)*

10:00–10:14 *An Element-wise Visual-enhanced BiLSTM-CRF Model for Location Name Recognition*

Takuya Komada and Takashi Inui

10:14–10:28 *BERT-based Spatial Information Extraction*

Hyeong Jin Shin, Jeong Yeon Park, Dae Bum Yuk and Jae Sung Lee

10:28–10:42 *A Cognitively Motivated Approach to Spatial Information Extraction*

Chao Xu, Emmanuelle-Anna Dietz Saldanha, Dagmar Gromann and Beihai Zhou

10:42–10:56 *Language-Conditioned Feature Pyramids for Visual Selection Tasks (Findings)*

Taichi Iki and Akiko Aizawa

10:56–11:05 *Break*

11:05–12:51 **Session 2**

11:05–11:55 *Invited talk: Julia Hockenmeier*

11:55–12:51 *Paper Presentations (SpLU, Findings)*

11:55–12:09 *They are not all alike: answering different spatial questions requires different grounding strategies*

Alberto Testoni, Claudio Greco, Tobias Bianchi, Mauricio Mazuecos, Agata Marcante, Luciana Benotti and Raffaella Bernardi

12:09–12:23 *Categorisation, Typicality & Object-Specific Features in Spatial Referring Expressions*

Adam Richard-Bollans, Anthony Cohn and Lucía Gómez Álvarez

12:23–12:37 *A Linguistic Analysis of Visually Grounded Dialogues Based on Spatial Expressions (Findings)*

Takuma Udagawa, Takato Yamazaki and Akiko Aizawa

12:37–12:51 *Visually-Grounded Planning without Vision: Language Models Infer Detailed Plans from High-level Instructions (Findings)*

Peter A. Jansen

12:51–13:00 *Break*

13:00–14:46 **Session 3**

13:00–13:50 *Invited talk: Yoav Artzi*

13:50–14:46 *Paper Presentations (SpLU, Findings)*

13:50–14:04 *A Hybrid Deep Learning Approach for Spatial Trigger Extraction from Radiology Reports*

Surabhi Datta and Kirk Roberts

14:04–14:18 *Retouchdown: Releasing Touchdown on StreetLearn as a Public Resource for Language Grounding Tasks in Street View*

Harsh Mehta, Yoav Artzi, Jason Baldrige, Eugene Ie and Piotr Mirowski

Thursday, November 19, 2020 (continued)

14:18–14:32 *SpaRTQA: A Textual Question Answering Benchmark for Spatial Reasoning (Non-archival)*

Roshanak Mirzaee, Hossein Rajaby Faghihi and Parisa Kordjamshidi

14:32–14:46 *Decoding Language Spatial Relations to 2D Spatial Arrangements (Findings)*

Gorjan Radevski, Guillem Collell, Marie-Francine Moens and Tinne Tuytelaars

14:46–15:45 *Break*

15:45–17:31 Session 4

15:45–16:35 *Invited talk: Bonnie Dorr*

16:35–17:31 *Paper Presentations (SpLU, Findings)*

16:35–16:49 *Geocoding with multi-level loss for spatial language representation (Non-archival)*

Sayali Kulkarni, Shailee Jain, Mohammad Javad Hosseini, Jason Baldridge, Eugene Ie and Li Zhang

16:49–17:03 *Vision-and-Language Navigation by Reasoning over Spatial Configurations (Non-archival)*

Yue Zhang, Quan Guo and Parisa Kordjamshidi

17:03–17:17 *LiMiT: The Literal Motion in Text Dataset (Findings)*

Irene Manotas, Ngoc Phuoc An Vo and Vadim Sheinin

17:17–17:31 *ARRAMON: A Joint Navigation-Assembly Instruction Interpretation Task in Dynamic Environments (Findings)*

Hyounghun Kim, Abhay Zala, Graham Burri, Hao Tan and Mohit Bansal

17:31–17:45 *Break*

17:45–19:00 Session 5

17:45–18:35 *Invited talk: Douwe Kiela*

18:35–19:00 *Paper Presentations (Findings)*

18:35–18:49 *Robust and Interpretable Grounding of Spatial References with Relation Networks (Findings)*

Tsung-Yen Yang, Andrew S. Lan and Karthik Narasimhan

18:49–19:03 *RMM: A Recursive Mental Model for Dialogue Navigation (Findings)*

Homero Roman Roman, Yonatan Bisk, Jesse Thomason, Asli Celikyilmaz and Jianfeng Gao

19:03–20:00 Panel Discussion

20:00–21:00 GatherTown II