

ACL 2023

**1st Workshop on Natural Language Reasoning and  
Structured Explanations (@ACL 2023)**

**Proceedings of the Workshop**

June 13, 2023

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

## **Sponsor**



©2023 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](mailto:acl@aclweb.org)

ISBN 978-1-959429-94-4

## Introduction

Welcome to NLRSE, the First Workshop on Natural Language Reasoning and Structured Explanations, co-located with ACL 2023 in Toronto, Ontario, Canada.

With recent scaling of large pre-trained Transformer language models (LLMs), the scope of feasible NLP tasks has broadened. Significant recent work has focused on tasks that require some kind of natural language reasoning. A trajectory in question answering has led us from extraction-oriented datasets like SQuAD to “multi-hop” reasoning datasets like HotpotQA and StrategyQA. Although LLMs have shown remarkable performance on most NLP tasks, it is often unclear why their answers follow from what they know. To address this gap, a new class of explanation techniques has emerged which play an integral part in structuring the reasoning necessary to solve these datasets. For example, the chain-of-thought paradigm leverages explanations as vehicles for LLMs to mimic human reasoning processes. Entailment trees offer a way to ground multi-step reasoning in a collection of verifiable steps. Frameworks like SayCan bridge high-level planning in language and with low-level action trajectories. As a result, we see a confluence of methods blending explainable machine learning/NLP, classical AI (especially theorem proving), and cognitive science (how do humans structure explanations?). This workshop aims to bring together a diverse set of perspectives from these different traditions and attempt to establish common ground for how these various kinds of explanation structures can tackle a broad class of reasoning problems in natural language and beyond.

A total of 12 papers appear in the proceedings. Over 70 papers were presented at the workshop itself, with the rest being submitted under two archival options: cross-submissions (Findings papers or those already presented at other venues, such as ICLR or the ACL main conference), and regular non-archival submissions (unpublished work). The latter went through a normal peer review process. These papers can be found on the NLRSE website: <https://nl-reasoning-workshop.github.io/>

Four papers were featured as oral presentations. These were selected from the archival papers to represent a selection of strong work that the organizers felt would be of broad interest to workshop participants. In addition, we featured five invited talks: Peter Clark, Noah Goodman, Ellie Pavlick, Sarah Wiegrefe, and Denny Zhou.

We are thankful to all reviewers for their help in the selection of the program, for their readiness in engaging in thoughtful discussions about individual papers, and for providing valuable feedback to the authors. We would also like to thank the ACL workshop organizers for all the valuable help and support with organizational aspects of the conference. Finally, we would like to thank all our authors and presenters for making this such an exciting event!

Bhavana Dalvi, Greg Durrett, Peter Jansen, Danilo Ribeiro, Jason Wei, and Lio Wong, NLRSE organizers

# Organizing Committee

## Organizers

Bhavana Dalvi Mishra, Allen Institute for Artificial Intelligence

Greg Durrett, UT Austin

Peter Jansen, University of Arizona

Danilo Neves Ribeiro, Northwestern University

Jason Wei, OpenAI

Lio Wong, Massachusetts Institute of Technology

## Program Committee

### Chairs

Bhavana Dalvi Mishra, Allen Institute for Artificial Intelligence  
Greg Durrett, UT Austin  
Peter Jansen, University of Arizona  
Danilo Neves Ribeiro, Northwestern  
Jason Wei, OpenAI  
Lio Wong, Massachusetts Institute of Technology

### Program Committee

Jun Araki, Bosch Research  
Jinheon Baek, Korea Advanced Institute of Science and Technology  
Kaj Bostrom, University of Texas at Austin  
Faeze Brahman, Allen Institute for AI  
Kezhen Chen, Google, X  
Yung-sung Chuang, Massachusetts Institute of Technology  
Maxwell Crouse, IBM Research  
Yao Fu, The University of Edinburgh  
Shahriar Golchin, University of Arizona  
Shashank Gupta, Allen Institute for Artificial Intelligence  
Jie Huang, University of Illinois at Urbana-Champaign  
Naoya Inoue, Japan Advanced Institute of Science and Technology  
Uri Katz, Bar Ilan University  
Shiyang Li, UC Santa Barbara  
Shuyang Li, Meta AI  
Zhengzhong Liang, University of Arizona  
Pan Lu, University of California, Los Angeles  
Bodhisattwa Prasad Majumder, University of California San Diego  
Constantine Nakos, Northwestern University  
Enrique Noriega-atala, The University of Arizona  
Yasumasa Onoe, The University of Texas at Austin  
Shiva Kumar Pentyala, Salesforce AI  
Valentina Pyatkin, Bar-Ilan University  
Hossein Rajaby Faghihi, Michigan State University  
Juan Diego Rodriguez, The University of Texas at Austin  
Soumya Sanyal, University of Southern California  
Rebecca Sharp, University of Arizona  
Alexey Sorokin, Moscow State University  
Zayne Sprague, University of Texas at Austin  
Arvind Krishna Sridhar, Qualcomm Technologies R+D  
Oyvind Tafjord, AI2  
Harsh Trivedi, Stony Brook University  
Hoang Van, Harvard Medical School  
Ruoyao Wang, University of Arizona  
Xi Ye, The University of Texas at Austin  
Li Zhang, University of Pennsylvania

Bowei Zou, Institute for Infocomm Research

## Table of Contents

<i>Knowledge Graph-augmented Language Models for Complex Question Answering</i> Priyanka Sen, Sandeep Mavadia and Amir Saffari .....	1
<i>Exploring the Curious Case of Code Prompts</i> Li Zhang, Liam Dugan, Hainiu Xu and Chris Callison-burch .....	9
<i>A smashed glass cannot be full: Generation of Commonsense Explanations through Prompt-based Few-shot Learning</i> Andrea Zaninello and Bernardo Magnini .....	18
<i>Saliency Map Verbalization: Comparing Feature Importance Representations from Model-free and Instruction-based Methods</i> Nils Feldhus, Leonhard Hennig, Maximilian Nasert, Christopher Ebert, Robert Schwarzenberg and Sebastian Mller .....	30
<i>Using Planning to Improve Semantic Parsing of Instructional Texts</i> Vanya Cohen and Raymond Mooney .....	47
<i>Reasoning Circuits: Few-shot Multi-hop Question Generation with Structured Rationales</i> Saurabh Kulshreshtha and Anna Rumshisky .....	59
<i>Knowledge-Augmented Language Model Prompting for Zero-Shot Knowledge Graph Question Answering</i> Jinheon Baek, Alham Fikri Aji and Amir Saffari .....	78
<i>Can In-context Learners Learn a Reasoning Concept from Demonstrations?</i> Michal Tefnik and Marek Kadlcik .....	107
<i>Effect Graph: Effect Relation Extraction for Explanation Generation</i> Jonathan Kobbe, Ioana Hulpu and Heiner Stuckenschmidt .....	116
<i>OPT-R: Exploring the Role of Explanations in Finetuning and Prompting for Reasoning Skills of Large Language Models</i> Badr Alkhamissi, Siddharth Verma, Ping Yu, Zhijing Jin, Asli Celikyilmaz and Mona Diab ..	128
<i>Deductive Additivity for Planning of Natural Language Proofs</i> Zayne Sprague, Kaj Bostrom, Swarat Chaudhuri and Greg Durrett .....	139
<i>Synthetic Dataset for Evaluating Complex Compositional Knowledge for Natural Language Inference</i> Sushma Anand Akoju, Robert Vacareanu, Eduardo Blanco, Haris Riaz and Mihai Surdeanu .	157



# Program

**Thursday, July 13, 2023**

08:00 - 09:00 *Virtual Poster Session*

09:00 - 09:10 *Opening Remarks*

09:10 - 09:50 *Invited Speaker - Ellie Pavlick*

09:50 - 10:30 *Invited Speaker - Noah Goodman*

10:30 - 11:00 *Break*

11:00 - 11:20 *Oral Presentations 1*

*Exploring the Curious Case of Code Prompts*

Li Zhang, Liam Dugan, Hainiu Xu and Chris Callison-burch

*Using Planning to Improve Semantic Parsing of Instructional Texts*

Vanya Cohen and Raymond Mooney

11:20 - 12:20 *Poster Session 1*

12:25 - 13:30 *Lunch*

13:30 - 14:30 *Poster Session 2*

14:30 - 14:50 *Oral Presentations 2*

*Knowledge-Augmented Language Model Prompting for Zero-Shot Knowledge Graph Question Answering*

Jinheon Baek, Alham Fikri Aji and Amir Saffari

*Can In-context Learners Learn a Reasoning Concept from Demonstrations?*

Michal Tefnik and Marek Kadlcik

**Thursday, July 13, 2023 (continued)**

14:50 - 15:30     *Invited Speaker - Peter Clark*

15:30 - 16:00     *Break*

16:00 - 16:40     *Invited Speaker - Denny Zhou*

16:40 - 17:20     *Invited Speaker - Sarah Wiegreffe*