

ACL 2024

**2nd Workshop on Natural Language Reasoning and
Structured Explanations (@ACL 2024)**

Proceedings of the Workshop

August 15, 2024

©2024 Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
317 Sidney Baker St. S
Suite 400 - 134
Kerrville, TX 78028
USA
Tel: +1-855-225-1962
acl@aclweb.org

ISBN 979-8-89176-142-1

Introduction

Welcome to NLRSE, the Second Workshop on Natural Language Reasoning and Structured Explanations, co-located with ACL 2024 in Bangkok, Thailand.

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 7 papers appear in the proceedings. 40 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already presented at other venues, such as EMNLP or the ACL main conference), and regular non-archival submissions (unpublished work).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/>

Six papers were featured as oral presentations. These papers represented a selection of strong work that the organizers felt would be of broad interest to workshop participants. In addition, we featured four invited talks: Sherry Wu, Karthik Narasimhan, Thomas Icard, and Lorraine Li.

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, Ben Lipkin, Danilo Ribeiro, Lio Wong, Xi Ye, and Wenting Zhao, NLRSE organizers

Organizing Committee

Organizers

Bhavana Dalvi Mishra, Allen Institute for Artificial Intelligence

Greg Durrett, UT Austin

Peter Jansen, University of Arizona

Ben Lipkin, Massachusetts Institute of Technology

Danilo Neves Ribeiro, Amazon AWS

Lio Wong, Massachusetts Institute of Technology

Xi Ye, UT Austin

Wenting Zhao, Cornell University

Program Committee

Program Chairs

Bhavana Dalvi Mishra, Allen Institute for Artificial Intelligence
Greg Durrett, University of Texas, Austin
Peter Jansen, University of Arizona
Ben Lipkin, Massachusetts Institute of Technology
Danilo Neves Ribeiro, Amazon AWS
Lionel Wong, Massachusetts Institute of Technology
Xi Ye, University of Texas, Austin
Wenting Zhao, Cornell University

Program Committee

Jun Araki, Bosch Research
Jinheon Baek, Korea Advanced Institute of Science & Technology
Kaj Bostrom, University of Texas, Austin
Faeze Brahman, Allen Institute for AI
Justin T Chiu, Cornell University
Yung-Sung Chuang, Massachusetts Institute of Technology
Shahriar Golchin, University of Arizona
Shashank Gupta, Allen Institute for Artificial Intelligence
Naoya Inoue, RIKEN and Japan Advanced Institute of Science and Technology
Uri Katz, Bar-Ilan University
Arun Balajiee Lekshmi Narayanan, University of Pittsburgh
Zhengzhong Liang, Google
Ziqian Luo, Oracle
Bodhisattwa Prasad Majumder, Allen Institute for Artificial Intelligence
Enrique Noriega-Atala, University of Arizona
Yasumasa Onoe, Google
Soumya Sanyal
Zayne Rea Sprague, University of Texas at Austin
Arvind Krishna Sridhar, Qualcomm Inc, QualComm
Oyvind Taffjord, Allen Institute for Artificial Intelligence
Harsh Trivedi, State University of New York, Stony Brook
Hoang Van
Ruoyao Wang, University of Arizona
Anne Wu, Cornell University
Li Zhang
Jiawei Zhou
Bowei Zou, A*STAR

Table of Contents

<i>GraphReason: Enhancing Reasoning Capabilities of Large Language Models through A Graph-Based Verification Approach</i>	
Lang Cao	1
<i>PROC2PDDL: Open-Domain Planning Representations from Texts</i>	
Tianyi Zhang, Li Zhang, Zhaoyi Joey Hou, Ziyu Wang, Yuling Gu, Peter Clark, Chris Callison-Burch and Niket Tandon	13
<i>Towards A Unified View of Answer Calibration for Multi-Step Reasoning</i>	
Shumin Deng, Ningyu Zhang, Nay Oo and Bryan Hooi	25
<i>Applying RLAIIF for Code Generation with API-usage in Lightweight LLMs</i>	
Sujan Dutta, Sayantan Mahinder, Raviteja Anantha and Bortik Bandyopadhyay	39
<i>SummEQuAL: Summarization Evaluation via Question Answering using Large Language Models</i>	
Junyuan Liu, Zhengyan Shi and Aldo Lipani	46
<i>LOGIC-LM++: Multi-Step Refinement for Symbolic Formulations</i>	
Shashank Kirtania, Priyanshu Gupta and Arjun Radhakrishna	56
<i>From Good to Great: Improving Math Reasoning with Tool-Augmented Interleaf Prompting</i>	
Nuo Chen, Hongguang Li, Baoyuan Wang and Jia Li	64

Program

Tuesday, August 15, 2023

- 08:00 - 09:00 *Virtual Poster Session*
- 08:55 - 09:00 *Opening Remarks*
- 09:00 - 09:45 *Invited Speaker - Sherry Wu*
- 09:45 - 10:30 *Invited Speaker - Karthik Narasimhan*
- 10:30 - 11:00 *Break*
- 11:00 - 11:45 *Invited Speaker - Thomas Icard*
- 11:45 - 12:30 *Oral Presentations 1*
- 12:30 - 14:00 *Lunch*
- 14:00 - 14:45 *Invited Speaker - Lorraine Li*
- 14:45 - 15:30 *Oral Presentations 2*
- 15:30 - 16:00 *Break*
- 16:00 - 17:30 *Poster Session*