SIGHAN-10 2024

The 10th SIGHAN Workshop on Chinese Language Processing

Proceedings of the Workshop

August 16, 2024

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 979-8-89176-155-1

Introduction

We are excited to welcome you to SIGHAN-10, the 10th SIGHAN workshop on Chinese language processing. This year, the 10th SIGHAN workshop returned and co-located with the 62nd Annual Meeting of the Association for Computational Linguistics (ACL-2024) in Bangkok, Thailand on August 11–16, 2024. Furthermore, SIGHAN 10 provides a shared task, namely Chinese Dimensional Aspect-Based Sentiment Analysis, dimABSA.

In an increasingly interconnected world, the importance of Chinese language processing cannot be overstated. As one of the most widely spoken languages, Chinese presents unique challenges and opportunities in the current research of artificial intelligence. Effective processing of the Chinese language opens doors to vast markets and cultural exchanges, fostering global collaboration and understanding. It serves as a critical tool in bridging linguistic divides and unlocking the rich textual heritage and contemporary content in Chinese. The focus of this workshop delves into the challenges in processing of the Chinese language, especially within the technology explosion of large language model, to explore how the Chinese specific tasks can be optimised to effectively understand as well as generating Chinese text.

We received 29 submissions this year, comprising 21 papers from the main workshop, and 8 papers from the shared task (dimABSA). We had two Area Chair (AC) members for the main workshop and one AC for the shared task, guiding the discussion process and writing a meta-review. For the main workshop, we accepted 10 papers. The acceptance rate for main workshop papers is 47.6%.

This year, SIGHAN-10 held in a hybrid format. Kang Liu from Institute of Automation, Chinese Academy of Sciences presents a keynote on "Beyond Facts: Understanding and Inducing Rule-based Knowledge in LLMs". Further, there are also several oral sessions, including five oral papers from the main workshop and three oral papers from the shared task.

We thank our Program Committee members and all reviewers. We specially thank our three Area Chairs: Runcong Zhao (King's College London), Bin Liang (The Chinese University of Hong Kong), and Lung-Hao Lee (National Yang Ming Chiao Tung University). They did an excellent job in reviewing the submitted papers, and we thank them for their essential role in selecting the accepted papers and helping produce a high-quality program for the conference.

We extend special thanks to all authors who have submitted papers this year and those who have shown interest in SIGHAN-10. We also thank all attendees for their participation and support.

Kam-Fai Wong and Min Zhang
General Chairs
Ruifeng Xu and Lin Gui
Program Co-Chairs

Organizing Committee

General Chair

Kam-Fai Wong, The Chinese University of Hong Kong, China
Min Zhang, Harbin Institute of Technology, Shenzhen, China

Program Co-Chair

Ruifeng Xu, Harbin Institute of Technology, Shenzhen, China
Lin Gui, King's College London, UK

Program Committee

Program Chairs

Kam-Fai Wong, The Chinese University of Hong Kong
Min Zhang, Harbin Institute of Technology, Shenzhen
Ruifeng Xu, Harbin Institute of Technology
Lin Gui, King's College London, University of London

Area Chairs

Lung-Hao Lee, National Yang Ming Chiao Tung University
Bin Liang, The Chinese University of Hong Kong
Runcong Zhao, King's College London, University of London

Reviewers

Jianzhu Bao

Lei Chen, Shiwei Chen

Keyang Ding

Zhe Hu, Shijue Huang

Xianming LI, Yuhang Lai, Lung-Hao Lee, Taishan Li, Jian Liao, Shujun Liu, Zhuohan Long

Xintong Song, Yang Sun

Hanzhuo Tan, Geng Tu

Bingbing Wang, Qianlong Wang, Suge Wang

Yanzheng Xiang, Hainiu Xu, Luo Xuan

Erxin Yu, Liang-Chih Yu

Han Zhang, Wenjia Zhang, Yice Zhang, Libo Zhao

Keynote Talk

Beyond Facts: Understanding and Inducing Rule-based Knowledge in LLMs

Kang Liu
Institute of Automation, Chinese Academy of Sciences
2024-08-16 09:30:00 – Room: **TBD**

Abstract: Large language models (LLM) have been proven to be able to learn knowledge from massive data. Most research currently discusses the relationship between implicit knowledge in LLMs and symbolic factual knowledge in Knowledge Graphs. Besides facts, human knowledge contains more types, such as rules. How does a LLM understand a rule and promote reasoning ability? Whether a LLM induce new rules from the given data? This talk will introduce our latest research work on these questions.

Bio: Kang Liu is a full professor at Institute of Automation, Chinese Academy of Sciences. He is also a youth scientist of Beijing Academy of Artificial Intelligence and a professor of University of Chinese Academy of Sciences. His research interests include Knowledge Graphs, Information Extraction, Question Answering and Large Language Models. He has published over 80 research papers in AI conferences and journals, like ACL, EMNLP, NAACL, COLING, et al. His work has over 20,000 citations on Google Scholar. He received the Best Paper Award at COLING-2014, Best PosterDemo Award at ISWC-2023, and the Google Focused Research Award in 2015 and 2016.

Table of Contents