LREC-COLING 2024

# The First Workshop on Bridging Neurons and Symbols for Natural Language Processing and Knowledge Graphs Reasoning @LREC-COLING-2024 (NeusymBridge 2024)

Workshop Proceedings

Editors
Tiansi Dong, Erhard Hinrichs, Zhen Han, Kang Liu, Yangqiu Song, Yixin Cao, Christian F. Hempelmann, Rafet Sifa

21 May, 2024
Torino, Italia

**Proceedings of NeusymBridge 2024: The First Workshop on Bridging Neurons and Symbols for Natural Language Processing and Knowledge Graphs Reasoning @ LREC-Coling 2024**



ISBN 978-2-493814-12-8
ISSN 2951-2093 (COLING); 2522-2686 (LREC)

Jointly organized by the ELRA Language Resources Association
and the International Committee on Computational Linguistics

# Message from the Workshop Organizers

Endowing machines with knowledge has long been regarded as one of the important goals of AI. Traditionally, symbols and their relations represent knowledge for natural language processing. Obviously, because of the limitations of the classical symbolic-based knowledge representation theory and knowledge acquisition technologies, symbolic knowledge bases have typical weaknesses, such as limited representation capacity, low acquisition efficiency, low coverage of multiple knowledge types, and applicable difficulties in reasoning scenarios. By contrast, large language models (LLMs) follow quite a different paradigm: the tradition of connectionism and neural networks. It employs distributional numerical vectors/matrices to represent the knowledge. This way, almost all knowledge types can be represented and embodied into a unified semantic space.

LLMs can be treated as knowledge base and provide an easier way to acquire and collect knowledge and inject knowledge into the downstream models or applications. However, compared with traditional symbolic knowledge bases, LLMs still have limitations including hallucination, and relying exclusively on fill-in-the-blank close tasks. A recent study showed that LLMs may miss more tail knowledge than head knowledge. LLMs still struggle to acquire negative knowledge. On the other hand, queries on knowledge graphs, in symbolic and/or neural ways, can vastly answer more complex logical queries, such as union, intersection, negation, counting, etc. Various strategies have been explored to improve the interpretability and reasoning performance of LLMs, for example, CoT, CoT-SC, Tree-of-Thoughts, or using external symbolic inference engines. unavailable. However, researchers still argue that LLMs are not good logical reasoners. One of the main reasons is that LLMs' reasoning is mostly non-rigorous — neither the reasoning process nor the result is guaranteed to be correct and complete. Despite these shortcomings, LLMs are becoming fundamental tools and have achieved great success in both academia and industry. They not only unify various NLP-related tasks in the form of text generation, but also have shown remarkable reasoning ability.

A cutting-edge research direction is to move from System I associative thinking to System II rational thinking – in the sense of D. Kahneman. Researchers are targeting novel machine learning systems for "slow, logical, sequential, conscious, linguistic, algorithmic, planning, and reasoning" problems. Knowledge graphs provide a natural way of connecting the dots across texts. Building an inherent linkage module for LLMs can provide a better global view of the world.

Moving from System I thinking to System II thinking demands traditional deep-learning to go beyond the statistical learning framework, and make qualitative extensions. A variety of new learning biases has been proposed to narrow the gap between higher-level cognition and traditional deep-learning. Language is embodied and schematizes space. The next generation of neural language system shall be a brain- and AI-inspired understanding system that explicitly represents situations, which roots in qualitative spatial representation, then extending to spatio-temporal and event representation, moving on to causality and emotion. Recent research proposes tensors as a unified representation for perception and memory, proposes spheres to explicitly unify symbolic structure with neural embedding for deterministic reasoning, neurosymbolic unification, and for humour understanding.

This workshop invited renowned scholars to give keynotes and active researchers to introduce their pioneering works in the fields, topics covering both academic researches and industrial applications. The state-of-the-art in deep learning for NLP and beyond shows that there are many open research questions to be addressed at the interface of symbolic and neural

approaches, and that bridging neurons and symbols may break the glass ceiling of deep learning for NLP.

The NeusymBridge 2024 Organizers

# Organizing Committee

- Tiansi Dong – Fraunhofer IAIS – Neurosymbolic Representation Learning Team
- Erhard Hinrichs – University of Tübingen – Department of Computational Linguistics
- Zhen Han – Amazon Inc.
- Kang Liu – Chinese Academy of Sciences – Research Group of Speech and Language Technology
- Yangqiu Song – The Hong Kong University of Science and Technology – Department of Computer Science and Engineering
- Yixin Cao – Singapore Management University – Department of Computer Science
- Christian F. Hempelmann – Texas A&M-Commerce – the Semantic Applied Linguistics and Creative Laboratory
- Rafet Sifa – University of Bonn – the Applied Machine Learning (AML) Lab

# Keynotes



# Table of Contents of Accepted Papers

# Bridging Neurons and Symbols for Natural Language Processing and Knowledge Graphs Reasoning @ LREC-Coling 2024

**Tuesday, May 21, 2024**

**09:00–09:30** **Welcome**

**09:30–10:30** **Keynote 1: Pascale Fung – Human Value Representation in Large Language Models - Bridging the Neural and the Symbolic**
Chair: Yangqiu Song

**10:30–11:00** **Morning coffee break**

**11:00–12:00** **Keynote 2: Juanzi Li – Neural-symbolic Programming for Explainable Knowledge-intensive Question Answering**
Chair: Yixin Cao

**12:00–13:00** **Paper Presentation**
Chair: Tiansi Dong

*Probing Large Language Models from a Human Behavioral Perspective*
Xintong Wang, Xiaoyu Li, Xingshan Li and Chris Biemann

*The Semantic Relations in LLMs: An Information-theoretic Compression Approach*
Yu-Hsiang Tseng, Pin-Er Chen, Da-Chen Lian and Shu-Kai Hsieh

**13:00–14:00** **Lunch Break**

**Tuesday, May 21, 2024 (continued)**

**14:00–15:00** **Keynote 3: Alessandro Lenci – The Semantic Gap in LLMs and How to Bridge It**
Chair: Erhard Hinrichs

**15:00–16:00** **Paper Presentation**
Chair: Yixin Cao

*Word Sense Disambiguation as a Game of Neurosymbolic Darts*
Tiansi Dong and Rafet Sifa

*Open Event Causality Extraction by the Assistance of LLM in Task Annotation, Dataset, and Method*
Kun Luo, Tong Zhou, Yubo Chen, Jun Zhao and Kang Liu

**16:00–16:30** **Afternoon coffee break**

**16:30–17:30** **Keynote 4: Volker Tresp – The Tensor Brain: A Unified Theory of Perception, Memory and Semantic Decoding**
Chair: Han Zhen

**17:30–18:00** **Paper Presentation**
Chair: Erhard Hinrichs

*The Need for Grounding in LLM-based Dialogue Systems*
Kristiina Jokinen

**18:00–18:30** **Conclusion**
Chair: Rafet Sifa

**19:30–21:30** **Workshop Dinner (Keynote speakers, paper presenters, workshop organisers)**