Report from Dagstuhl Seminar 21362

Structure and Learning

Edited by

Tiansi Dong[1], Achim Rettinger[2], Jie Tang[3], Barbara Tversky[4], and Frank van Harmelen[5]

1 **Universität Bonn, DE,** tdong@uni-bonn.de
2 **Universität Trier, DE,** rettinger@uni-trier.de
3 **Tsinghua University – Beijing, CN,** jietang@tsinghua.edu.cn
4 **Columbia University – New York, US,** btversky@stanford.edu
5 **VU University Amsterdam, NL,** frank.van.harmelen@vu.nl

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 21362 "Structure and Learning", held from September 5 to 10, 2021. Structure and learning are among the most prominent topics in Artificial Intelligence (AI) today. Integrating symbolic and numeric inference was set as one of the next open AI problems at the Townhall meeting "A 20 Year Roadmap for AI" at AAAI 2019. In this Dagstuhl seminar, we discussed related problems from an interdiscplinary perspective, in particular, Cognitive Science, Cognitive Psychology, Physics, Computational Humor, Linguistic, Machine Learning, and AI. This report overviews presentations and working groups during the seminar, and lists two open problems.

Seminar September 5–10, 2021 – http://www.dagstuhl.de/21362

2012 ACM Subject Classification Computing methodologies → Artificial intelligence; Computing methodologies → Machine learning; Computing methodologies → Symbolic and algebraic manipulation

Keywords and phrases Knowledge graph, Machine learning, Neural-symbol unification

Digital Object Identifier 10.4230/DagRep.11.8.11

Edited in cooperation with Siba Mohsen

1 Executive Summary

Tiansi Dong (Universität Bonn, DE)
Achim Rettinger (Universität Trier, DE)
Jie Tang (Tsinghua University – Beijing, CN)
Barbara Tversky (Columbia University – New York, US)
Frank van Harmelen (VU University Amsterdam, NL)

Deep Learning systems are the hope of the fifth industrial revolution. However, recent studies have found that Deep Learning systems can be easily manipulated, i.e. in Natural Language Understanding, Object Recognition. How to introduce structures into Deep Learning systems to improve reliability and performance has become a hot topic in Natural Language Processing (NLP), Machine Learning (ML), Semantic Web (SW) communities around the world. The aim of the seminar is to bring together interdisciplinary researchers around the world for constructive discussions on this theme, in particular, it intends to establish international collaborations to promote computational Humor, with the hope to let AI bring more joy, more laugh into the world, and do more good for the society. The hybrid seminar is structured in the form of Talks, Working Groups, and Open questions. The seminar started with the talk "Hybrid AI for Humor". The dynamic semantics of humor is beyond the reach of the classic symbolic AI, the deep learning paradigm, and current neural-symbolic integration methods, but can be captured by the neural geometric embedding, in terms of rotating sphere embedding. This novel embedding is rooted in Qualitative Spatial Representation (QSR) in symbolic AI and Learning Representation (LR) in neural ML. The former tries to symbolically delineate the basic spatial knowledge that humans have and possible ways that this knowledge can be used as a reference for abstract knowledge in other domains. LR aims at learning latent feature knowledge from data. The motivation and a geometric approach to realizing the unification were introduced in the talk "Rotating Spheres – A New Wheel for Neuro-Symbolic Unification". The motion of rotating spheres in high-dimensional space is served as a computational model to simulate (1) the motion of the physical world, (2) the circular interaction among the mind, the body, and the world (called *spraction* – a contraction of space, action, and abstraction, in which actions in space create abstractions).

The motion of the physical world is vividly explained in the talk "Rotating Spheres in the Milky Way". This *spraction* process is explained in the talk "Thinking with the Body and the World", which can guide the design of novel cognitive robots, and promote novel cognitive architectures. Two topics were covered by the talk "Learning about Language and Action for Robots", and the talk "Neural-Symbolic Models, Dual-Process Theories, and Cognitive Architectures".

In primates, the same brain structures that support spatial thinking also support conceptual thinking, Single cells in hippocampus gather multi-media information from different memories in the brain to represent places in space, events in time, ideas in conceptual spaces. Update-to-date research of neural simulation is introduced by Volker Tresp with the talk "Knowledge Graph and Cognitive Learning: from Perception to Memory Embedding", which maps embedding models to various cognitive memory functions, in particular to semantic and concept memory, episodic memory, sensory memory, short-term memory, and working memory.

Spatial thinking is multi-modal and established and distorted by our actions and perceptions of the spaces we interact in. This raises two questions: What are good representations for video understanding? and how to compute symbolic rules that the models have learned from the training data? Juergen Gall introduced holistic video understanding and argued the potential of hybrid approaches that combine neural networks with symbolic AI for video understanding and reasoning. Cuenca Grau, Bernado gave the talk "Characterizing Graph Neural Networks Using Logical Rules". He formally defines what it means for a set of logical rules to characterize the behavior of a model and proposes a GNN-based architecture that admits a characterization in terms of Datalog rules.

Spatial thinking is evident in the ways we think and the ways we externalize thought, for example, through words. Our words act on thought the way we act on objects. The philosophy of spatial thinking challenges the computational approach to natural language processing and understanding. Roberto Navigli argued that Natural Language Understanding (NLU) is particularly challenging, as this requires the machine to go beyond processing strings to reach a semantic level. Recent developments and challenges were discussed through three key tasks in NLU, namely Word Sense Disambiguation, Semantic Role Labeling, and Semantic Parsing. Zhiyuan Liu argued that knowledge (including symbols, embeddings, or models) is the key to a deeper understanding of human languages and that big pretrained language models can be regarded as the most advanced approach to model knowledge and to capture knowledge (including commonsense) from plain text and that the key challenge is how to incorporate both open data and structural knowledge. Alexander Mehler reviewed problems of neural network-based language learning, suggested to introduce the concept of cognitive maps and spatial information processing, and sketched a synergistic model that relates the dynamics of distributed information processing to bias interaction. Jie Tang introduced Wu-Dao, China's first homegrown super-scale intelligent model system, with the goal of building an ultra-large-scale cognitive-oriented pretraining model to focus on essential problems in general artificial intelligence from a cognitive perspective. Wu-Dao substantially outperforms BERT on the SuperGLUE natural language understanding benchmark with the same amount of pre-training data. Alam Mehwish discusses the characteristics of the existing benchmark datasets for the task of KG Completion, and limitations of the existing benchmark datasets and targets those issues in the generation of LiterallyWikidata.

Another externalization of spatial thinking is through graphics. In the talk "Semi-Riemannian Graph Convolutional Networks", Steffen Staab introduced their new geodesic tools that allow for extending neural network operations into geodesically disconnected semi-Riemannian manifolds. Thomas Liebig introduced using p-adic coding and computation for structured domains or domains with inherent granularity.

The ultimate form of spatial thinking is comics (a form of humor, the most creative form of storytelling), which typically show bodies acting in space. Humor is used as a testbed and lighthouse for the development of AI and machine learning. In the talk "Ethics of AI Humor" Kiki explained how humor has frustrated symbolic and statistic AI approaches; in the talk "Knowledge and Inferences Needed for Humor" Julia Rayz introduced recent advances in transformer-based approaches, and raised open questions.

Working groups are the main components of the seminar. The hybrid seminar provides an excellent chance to practice the situation that participants can continue to work together after this seminar, which is the main outcome of this seminar.

The seminar ended with the discussion "Boxology for Hybrid Learning and Reasoning Systems" chaired by Frank van Harmelen.

2 Table of Contents

Working groups

Open problems

3 Overview of Talks

3.1 LiterallyWikidata – A Benchmark for Knowledge Graph Completion using Literals

Mehwish Alam (FIZ Karlsruhe, DE)

Joint work of Genet Asefa Gesese, Harald Sack, Mehwish Alam

Recently many efforts have been made in automatically constructing Knowledge Graphs (KGs) from heterogeneous data sources such as text, image, etc. More specifically, cross-domain open KGs such as Freebase, Dbpedia, Wikidata, etc. are either extracted automatically from structured data, generated using heuristics, or are human-curated. Due to the Open World Assumption, KGs are never complete, i.e., there are always some facts missing. In order to solve this problem, recently different KG embedding models have been proposed for automated KG Completion.

This talk discusses the characteristics of the existing benchmark datasets for the task of KG Completion. It further discusses a set of benchmark datasets extracted from Wikidata and Wikipedia, named LiterallyWikidata. It also takes into account the limitations of the existing benchmark datasets and targets those issues in the generation of LiterallyWikidata.

LiterallyWikidata has been prepared with the main focus on providing benchmark datasets for multimodal KG Embedding (KGE) models, specifically for models using numeric and/or text literals. Hence, the benchmark is novel as compared to the existing datasets in terms of properly handling literals for those multimodal KGE models. LiterallyWikidata contains three datasets that vary both in size and structure. These datasets are analyzed based on their connectivity, density, and diameter. Moreover, the datasets also include textual information about the entities in multiple languages (in addition to English). Finally, the results of the benchmarking experiments on the task of link prediction were conducted on LiterallyWikidata.

Currently, LiterallyWikidata does not consider image literals. Moreover, the current results report the performance of existing models on the task of head, tail, and relation prediction. More experiments need to be conducted for the task of entity classification. As a future perspective, these points will be considered along with the bias analysis of this benchmark dataset.

3.2 Online Perceptual Learning and Natural Language Acquisition for Autonomous Robots

Anthony Cohn (University of Leeds, GB)

Joint work of This work is from the PhD thesis of Muhannad Alomari, jointly supervised with David Hogg.

To operate effectively, and to collaborate with humans, robots need to know much about the world, including the kinds of objects in the world, their properties, the spatial relationships between them and actions that can be performed on them, as well as how language is used to describe these things. In this work[1], the problem of bootstrapping knowledge in language and vision for autonomous robots is addressed through novel techniques in grammar induction and word grounding to the perceptual world.In particular, we demonstrate a system, called OLAV, which is able, for the first time, to (1) learn to form discrete concepts from sensory data; (2) ground language (n-grams) to these concepts; (3) induce a grammar for the language being used to describe the perceptual world; and moreover to do all this incrementally, without storing all previous data. The learning is achieved in a loosely supervised manner from raw linguistic and visual data. Moreover, the learnt model is transparent, rather than a black-box model and is thus open to human inspection. The visual data is collected using three different robotic platforms deployed in real-world and simulated environments and equipped with different sensing modalities, while the linguistic data is collected using online crowdsourcing tools and volunteers. The analysis performed on these robots demonstrates the effectiveness of the framework in learning visual concepts, language groundings and grammatical structure in these three online settings.

3.3 Characterising Graph Neural Networks Using Logical Rules

Bernado Cuenca Grau (University of Oxford, GB)

There are many practical situations where we would like to learn a function that completes a given dataset in graph format (over a fixed set of unary and binary predicates) with additional facts; these include, for instance, knowledge graph completion and recommendation systems.

Graph Neural Networks (GNNs) are a family of ML models that have proved well-suited for such situations; however, as it is the case with other ML models, it is difficult to explain its predictions, and there is a growing interest in computing "general patterns" (or rules) that the models have learnt from the training data as a form of symbolic explanation.

In this work, we formally define what it means for a set of logical rules to characterise the behaviour of a model, and propose a GNN-based architecture that admits a characterisation in terms of Datalog rules. Our architecture consists of three main elements: (1) an encoder, which transforms the input dataset into a graph annotated with numeric feature vectors; (2) a Monotonic Graph Neural Network (a GNN variant satisfying a property akin to that of monotonicity under homomorphisms of First Order Logic); and (3) a decoder, which transforms the result of GNN application into the output dataset.

Our architecture can be successfully trained in practice for tasks such as knowledge graph completion; furthermore, the corresponding set of rules can be extracted algorithmically from the trained model. Our experiments on well-known knowledge graph completion benchmarks show competitive performance with that of state-of-the-art rule learning methods such as AnyBURL and DRUM.

[1] The financial support provided by EU FP7 project 600623 (STRANDS) as well as the EU Horizon 2020 framework under grant agreement 825619 (AI4EU) is gratefully acknowledged, as is support from the Alan Turing Institute.

3.4 Rotating Spheres: A New Wheel for Neuro-Symbolic Unification

Tiansi Dong (Universität Bonn, DE)

The distinction between discrete symbolic representations and continuous vector embeddings (i.e. subsymbolic representation) separates AI researches into two seemingly incompatible paradigms. Cognitively, both of them are products of our minds. How can discrete symbolic representations and rigorous symbolic reasoning be carried out by our neural mind? Sun (1994) suggested a dual-process theory [1]: *…cognitive processes are carried out in two distinct "levels" with qualitatively different mechanisms. Each level encodes a … set of knowledge for its processing, and the coverage of the two sets … overlaps substantially. Two different "levels" can potentially work together synergistically, complementing and supplementing each other.*

Given a discrete tree structure and the vector embeddings of its nodes, we can promote these vectors into spheres, and let the containment relations among spheres capture the discrete tree structure. Following this intuition, I argue for a novel neuro-geometric approach for neuro-symbolic unification [2] as follows: (1) vector embeddings from classic neural-networks can be promoted into spheres in higher dimensional space; (2) symbolic structures shall be precisely encoded as topological relations among these spheres. To support this argument, I show the empirical experiments with tree structures and their vector embeddings [3, 4], and neural Euler diagram embeddings for syllogistic reasoning [5].

By representing features as rotating axes, I introduce the term *Rotating Spheres* as a neuro-symbolic building block, and illustrate how they can be used to computationally interpret symbolic humor theory [8] and to simulate "Spatial Humor" [6, 7] which helps to explore how spatial thinking can be computationally linked to non-spatial thinking [11]. This starts from a topological representation of spaces and events and moves on to represent expectations (as another space in mind) and emotions (as a rotating axis). The violation of expectation in humor [9, 10] is computationally simulated as the flip of a rotating axis of a sphere. The flip works like the turning of a button, which triggers the mind machine to laugh.

Continuous vector embeddings and discrete symbolic rules are images seen from the two traditional eyes of AI. Rotating spheres serve as basic building block to unify the two approaches in the mind of AI and to computationally embody the way of thinking. This shapes a new style of AI.

References

1 Ron Sun. *Integrating Rules and Connectionism for Robust Commonsense Reasoning*. John Wiley and Sons, New York, 1994
2 Tiansi Dong. *A Geometric Approach to the Unification of Symbolic Structures and Neural Networks*. Studies in Computational Intelligence 910, Springer, 2021.
3 Tiansi Dong, Christian Bauckhage, Hailong Jin, Juanzi Li, Olaf Cremers, Daniel Speicher, Armin B. Cremers, Joerg Zimmermann. *Imposing Category Trees Onto Word-Embeddings Using A Geometric Construction*. ICLR-19
4 Tiansi Dong, Zhigang Wang, Juanzi Li, Christian Bauckhage, Armin B. Cremers. *Triple Classification Using Regions and Fine-Grained Entity Typing*. AAAI-19
5 Tiansi Dong, Chengjiang Li, Christian Bauckhage, Juanzi Li, Stefan Wrobel, Armin B. Cremers. *Learning Syllogism with Euler Neural-Networks*. arXiv:2007.07320

6 Siba Mohsen, Tiansi Dong. *A Neuro-symbolic approach to Joke Understanding*. At Workshop "AI4Jokes: towards a Neuro-Symbolic Approach for the Analysis and the Generation of Jokes" at the 11th Humor Research Conference.
7 Bitasta Biswas, Tiansi Dong. *Using Blender for Joke Understanding*. At Workshop "AI4Jokes: towards a Neuro-Symbolic Approach for the Analysis and the Generation of Jokes" at the 11th Humor Research Conference.
8 Tiansi Dong. *Jokes, Collages, and AI*. Webinar Humor and Artificial Intelligence at the ISHS (The International Society for Humor Studies) Spring 2021 Webinar Series
9 Salvatore Attardo, Christian F Hempelmann, and Sara Di Maio. *Script Oppositions and Logical Mechanisms: Modeling Incongruities and their Resolutions*. Humor, 15(1):3 – 46, 2002.
10 Julia Taylor Rayz. *Scripts in the Ontological Semantic Theory of Humor*. In: Salvatore Attardo, editor, Script-Based Semantics. 2020.
11 Barbara Tversky. *Mind in Motion*. Basic Books, New York, USA, 2019.

3.5 What are Good Representations for Video Understanding?

Jürgen Gall (Universität Bonn, DE)

In this talk, I will give an overview of some representations that have been used for video understanding. The representation ranges from knowledge-based representations that have been used for robotics applications, fine-hierarchies for sports like gymnastics, joint representations for language and video, and verb-object representations. While all these representations have advantages for specific applications, it is still an open research problem how a universal representation of actions can be defined. In the second part of the talk, I introduce holistic video understanding. Instead of just representing a video by action labels or captions, holistic video understanding provides are more rich representation which contains labels for actions, objects, scenes, attributes, events, and concepts. In order to study this problem, we released the HVU dataset (`https://holistic-video-understanding.github.io/`). It consists of over 570k videos with over 9m annotations of 3142 different semantic labels. I will show a few examples like video retrieval, video captioning, and action recognition that demonstrate the benefits of having such rich semantic descriptions of the videos. Nevertheless, it is an open research questions how relations between objects, attributes, actions, and events can be best utilized for video understanding. In the last part of the talk, I will describe a hybrid approach that combines recurrent neural networks, hidden Markov models, and a context-free grammar for temporal action segmentation. I will show some results that demonstrate the advantage of combining grammars with neural networks and give a few examples for weakly supervised learning. The talk concludes that there is a large potential for hybrid models that combine neural networks with symbolic AI for video understanding and reasoning, which is a promising research direction.

References

1 A. Diba, M. Fayyaz, V. Sharma, M. Paluri, J. Gall, R. Stiefelhagen, L. van Gool. *Large Scale Holistic Video Understanding*. European Conference on Computer Vision (ECCV'20), Springer, LNCS 12350, 593-610, 2020.

2 S. Li, Y. Abu Farha, Y. Liu, M.-M. Cheng, J. Gall. *MS-TCN++: Multi-Stage Temporal Convolutional Network for Action Segmentation*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020.
3 H. Kuehne, A. Richard, J. Gall. *A Hybrid RNN-HMM Approach for Weakly Supervised Temporal Action Segmentation*. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 42, No. 4, 765-779, 2020.
4 A. Richard, H. Kuehne, J. Gall. *Action Sets: Weakly Supervised Action Segmentation without Ordering Constraints*. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'18), 5987-5996, 2018.
5 A. Richard, H. Kuehne, J. Gall. *Weakly Supervised Action Learning with RNN based Fine-to-Coarse Modeling*. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'17), 1273-1282, 2017.

3.6 Ethics of AI Humor

Christian Hempelmann (Texas A&M University – Commerce, US) and Max Petrenko (Amazon Web Services – Seattle, US)

To the layperson and self-delusional expert, AI has become frighteningly powerful and humanlike, not least in its application to natural language processing. This leads to ethical issues coming to the fore as to who is responsible for the output of these systems at various levels. General AI systems have to make life-and-death decisions when coupled with self-driving cars and weapons systems. Language-generating AI systems produce racist and sexist output reflecting the human-generated data the systems have learned from. These AI systems are also used to classify as well as generate humor, which raises the same general, but also specific ethical issues. The latter stem from the specific meaning constellations in humor leading not least to the deniability of its messages. We aim to outline the relevant key points to initiate a discussion that we think needs to happen now. I will present with an overview of past approaches to generating and analyzing humor computationally up to 2015. As in its parent discipline, computational linguistics, early approaches were symbolic, rule- and resource-based. Since the 1990s, the methodology came increasingly from computer sciences and was probabilistic, up to the unexplainable algorithms of machine learning.

3.7 Rotating spheres in the Milky Way

Michael Kramer (MPI für Radioastronomie, DE)

My research focuses on the study of "pulsars", i.e. rotating dense cores of exploded stars that emit a lighthouse beam of radio emission which makes them detectable as pulsating radio sources. Their rotation – and motion! – in the Milky Way can be tracked by their regular beacons that stand out from the usually irregular, random, and stochastic processes in the Milky Way and the Universe as a whole. This allows us to use them as reference points, charting our position and motion relative to them. They come with rotation frequencies from about once every second to almost thousand times a second. The physics is extreme, as the rotational speed at the equator of these about 25 km large objects reaches significant fractions of the speed of light. Finding and exploiting them allows us to use them as precise cosmic clocks, for instance, to study and test the predictions of general relativity. In the past, we had armies of Ph.D. students sifting through our telescope data. Today, by enlarging the parameter space and increasing our sensitivity, we get millions and millions of candidates, so that we need to deploy artificial intelligence and machine learning methods nowadays. One main area of concern is how to avoid throwing out or ignoring the discovery of a new type of signal, only because we didn't know its properties before. In 1967, when pulsars were discovered, it was a bright female Ph.D. student who noticed the unusual signature of the signal. Would artificial intelligence be as clever? If we make enough discoveries, we can expand our network (or "array") of pulsars to convert the Milky Way into a galaxy-sized gravitational wave detector. We can do this by comparing the arrival time of different pulsars and detecting tiny variations, correlated in direction on the sky, which are caused by a background of gravitational waves filling the Universe from past mergers and collisions of galaxies.

3.8 p-adic Coding & Computation for Structured Data

Thomas Liebig (TU Dortmund, DE)

For centuries, physical observations and sensor readings are foremost represented by real numbers. This particular choice of data representation and model selection poses assumptions on the geometry of the observed feature space.

Many currently applied models of physical processes suffer from the representation bias posed by the physical origin of most process models, which start modeling with ordinary or partial differential equations on the field of real numbers. Currently, lots of domain and model-specific literature exists, how to incorporate expert knowledge on processes, correlations, or physical behaviour in these models. But some of the data generating processes exhibit chaotic behaviour and we observe that
1) depending on granularity it is cumbersome to model bursts, and
2) scale matters.
Consider, as an example, a traffic flow prediction model. While it appears simple to predict the average daily traffic volumes on a street segment from surrounding observations, it gets hard at fine granularities, since traffic is controlled by external semaphores, and it is not the same predicting few cars more or less on a street when the total traffic flow is high or almost empty.

For such structured domains or domains with inherent granularity, p-adic coding and computations overcome the assumptions and provide a natural framework for structured data. We highlight approaches and challenges of p-adic modeling.

3.9 Knowledgeable Learning for Natural Language Processing

Zhiyuan Liu (Tsinghua University – Beijing, CN)

In this talk, we argue that knowledge is the key to deeper understanding of human languages. Knowledge can be represented in appropriate ways including symbols, embeddings or models. Natural language processing can be formalized as the acquisition, representation, and application of complicated knowledge for language understanding. Big pretrained language models can be regarded the most advanced approach to model knowledge and to capture knowledge from plain text including commonsense. The key challenge is how to incorporate existing knowledge and make PLMs learn from both open data and structural knowledge. In this talk, we summarize various promising approaches to knowledgeable learning for NLP, including knowledge augmentation over input, knowledge framework over neural architecture, and knowledge regularization over learning objectives. Prompt Tuning seems promising to stimulate model knowledge for diverse downstream tasks.

3.10 Learning Linguistic Representations: Some Challenges and Opportunities

Alexander Mehler (Goethe-Universität – Frankfurt am Main, DE)

In this short talk, I briefly review some of the problems of neural network-based language learning and suggestions for overcoming them. This includes an account of gaps arising from symbolic learning resources. This concerns, first, a so-called algorithmization bias, according to which the same corpus looks very different from the point of view of the output distributions of a set of NLP routines focusing on the same task (e.g., sentiment analysis), so that the application of these routines becomes predictable. Beyond that, polymorphic structuring of fragmented texts (using Twitter data as an example), aspects of distributed authorship and readership, and biased information processing are exemplified. To overcome problems related with these scenarios, the presentation builds on the concept of cognitive maps and spatial information processing. To this end, research on biases is combined with concepts of context-sensitive language learning (e.g., regarding the salience of landmarks, hierarchization effects, localization effects, and Zipfian tripartivity). This analysis is then used to distinguish between two roles of biases in distributional semantics: as OUT parameters for the purpose of hypothesis testing (asking which biases are reflected by which resource), and as IN parameters that constrain the generation of language representation models to reflect particular biases. The paper ends by sketching a synergetic model that relates the dynamics of distributed information processing to bias interaction.

3.11 The challenges of bringing together NLU, multilinguality and KG

Roberto Navigli (Sapienza University of Rome, IT)

Natural Language Processing (NLP) has seen an explosion of interest in recent years, with many industrial applications relying on key technological developments in the field. However, **Natural Language Understanding (NLU)** – which requires the machine to get beyond processing strings and involves a semantic level – is particularly challenging due to the **pervasive ambiguity of language**. In this talk I will present the **recent developments and challenges** of three key tasks in NLU, namely Word Sense Disambiguation, Semantic Role Labeling and Semantic Parsing.
Word Sense Disambiguation, the task of associating a word in context with its most appropriate sense from a predefined sense inventory, is one of the hardest tasks in NLP, however the advent of Deep Learning has led significant improvements, also thanks to the integration of explicit knowledge in the form of lexical knowledge graphs, leading to performances above the hard-to-beat threshold of 80% F1 on standard test sets. Still, several issues are open, including the availability of training data in languages other than English and the granularity of sense inventories. I will also mention the option of dropping the need of a sense inventory, an approach we called Generationary.

I will then move to sentence-level semantics, which is also hampered by the lack of large-scale annotated data. **Semantic Role Labeling**, aimed at extracting the predicate-argument structure of a sentence and identifying the semantic relationships between a predicate and its arguments, suffers from the existence of different, heterogeneous framesets for each language. Recently, we put forward a unifying neural architecture which, trained on data from different languages, outputs predicate senses and roles according to all the available inventories, and enables the use of previously unseen languages and the creation of a network of predicate-argument meanings. Finally, I will discuss the issues of **Semantic Parsing**, which moves from the predicate-argument structure to the overall structure of a sentence, typically as a semantic graph. I will also introduce two recent approaches to generative semantic parsing, based on graph linearization techniques and a pretrained encoder-decoder architecture, and cross-lingual parsing where we address the pervasive data paucity issue with the production of high-quality silver training data.

3.12 Hybrid Humor AI

Max Petrenko (Amazon Web Services – Seattle, US) and Christian Hempelmann (Texas A&M University – Commerce, US)

AI-oriented research on humor has been evolving in a pattern that can also be observed in the history of AI as a field. Case- and problem-specific projects, informed by specific methods, led to results that, while appropriate for the case in point, offered limited room for generalization beyond the discussed cases. These findings also struggled to advance the understanding of humor as a general purpose faculty of intelligence and what kind of AI systems would be required to model humor comprehension and generation. Similarly to the conceptual space, methodology-wise, much of the existing AI research on humor tends to adopt (often in an exclusionary fashion) a statistical/learning or a symbolic/knowledge-oriented view, which has yielded fragmentary dividends. We will offer insights on the properties of hybrid AI systems required to support general purpose humor research. We will first discuss the assumptions that AI research needs to treat humor in light of artificial general (as opposed to special purpose) intelligence theories, and that a hybrid, or neuro-symbolic, framework is an appropriate framework to pursue such research. We will then focus on the symbolic component of the AI research on humor and discuss the benefits (and challenges) of ontology modeling for the design and implementation of the symbolic component. We will work through an example of a joke modelled with the ontology of humor as an access point to its general architecture.

3.13 Knowledge and Inferences Needed for Humor

Julia Rayz (Purdue University – West Lafayette, US)

Humor is an interesting phenomenon in the context of AI because it relies on inferences and unstated knowledge. The state of the art in natural language processing can claim some (limited) successes in knowledge acquisition from text. It is also possible to access some information from text thought other means, such as asking questions and getting answers to them. This works for many applications but is not sufficient for humor, which – according to various theories – requires realization and resolution of something unexpected. How much, exactly, can we resolve or even approximate the needed unexpected situation with large corpora alone? This talk starts with a brief overview of linguistic theories of humor. We then look at one of the jokes extensively analyzed in the humor literature and discuss the type of knowledge (scripts) needed to understand and detect this joke and project this knowledge and inferences to the (relatively) recent advances in transformer-based approaches, which demonstrate activation of other scripts. The talk also covers some of the recent papers in (non-humorous) natural language processing that are relevant to humor processing. The question that remains unanswered is: would it be possible to retrieve low frequency but relevant information from enormous text corpora by rephrasing and extending the text of the joke or transferring it to a question/answering task? More importantly, if such low-frequency information could be retrieved, would it correspond to typical human-level script retrieval/analysis?

3.14 Semi-Riemannian Graph Convolutional Networks

Steffen Staab (Universität Stuttgart, DE)

Joint work of Bo Xiong, Shichao Zhu, Nico Potyka, Shirui Pan, Chuan Zhou, Steffen Staab
Main reference Bo Xiong, Shichao Zhu, Nico Potyka, Shirui Pan, Chuan Zhou, Steffen Staab: "Semi-Riemannian Graph Convolutional Networks", CoRR, Vol. abs/2106.03134, 2021.
URL https://arxiv.org/abs/2106.03134

Graph Convolutional Networks (GCNs) are typically studied through the lens of Euclidean geometry. Non-Euclidean Riemannian manifolds provide specific inductive biases for embedding hierarchical or spherical data, but cannot align well with data of mixed topologies. We consider a larger class of semi-Riemannian manifolds with indefinite metric that generalize hyperboloid and sphere as well as their submanifolds. We develop new geodesic tools that allow for extending neural network operations into geodesically disconnected semi-Riemannian manifolds. As a consequence, we derive a principled Semi-Riemannian GCN that first models data in semi-Riemannian manifolds of constant nonzero curvature in the context of graph neural networks. Our method provides a geometric inductive bias that is sufficiently flexible to model mixed heterogeneous topologies like hierarchical graphs with cycles. Empirical results demonstrate that our method outperforms Riemannian counterparts when embedding graphs of complex topologies.

3.15 Neural-Symbolic Models, Dual-Process Theories, and Cognitive Architectures

Ron Sun (Rensselaer Polytechnic – Troy, US)

In this talk, I address neural-symbolic models, dual-process theories, and cognitive architectures – their relationships and their relevance to each other. I provide some historical backgrounds and argue that dual-process theories have significant implications for developing neural-symbolic models. Computational cognitive architectures can help disentangle complex issues concerning dual-process theories and thus neural-symbolic models.

The notion of neural-symbolic models harkens back to the 1990s when such models first emerged (see, e.g., Sun & Bookman, 1994). There have been many different ways of structuring such models; the question remains: how should we best structure them? I argue that they should be structured in a cognitively motivated/justified way, based on human cognitive architecture. In particular, they should take into account dual-process theories concerning human cognitive architecture.

The distinction between "intuitive" and "reflective" thinking (i.e., system 1 and system 2) has been, arguably, one of the most important distinctions in cognitive science. There are currently many dual-process theories out there. One such theory was proposed early on in Sun (1994), where the two systems were characterized as follows: "... cognitive processes are carried out in two distinct "levels" with qualitatively different mechanisms. Each level encodes a ... set of knowledge for its processing, and the coverage of the two sets ... overlaps substantially." (Sun, 1994, p.44). That is, the two "levels" encode somewhat similar or overlapping content. But they encode their contents in different ways: Symbolic and subsymbolic representation are used, respectively. Different mechanisms are thus involved at these two "levels". It was hypothesized that these two different "levels" can potentially work together synergistically, complementing and supplementing each other.

However, some details of more recent dual-process theories are more questionable. Although the distinction is important, the terms involved have often been ambiguous. Not much fine-grained analysis has been done, especially not in a precise, mechanistic, process-based way. Therefore, we need a conceptual and computational framework in this regard. The Clarion cognitive architecture (Sun, 2002, 2016) may be used at a theoretical level as a conceptual tool for generating interpretations and explanations. Many empirical and simulation studies have been conducted within the Clarion framework that shed light on relevant issues. Based on the framework, we re-interpret some common folk psychological notions, to give them more clarity and precision.

In summary, dual-process theories have important implications for neural-symbolic models. If cognitive-psychological realism is what one wants to achieve in developing computational models, dual-process theories must be taken into consideration. However, some issues involved in dual-process theories are more complex than often assumed. These issues are crucial for developing cognitive architectures, and in turn cognitive architectures can help in disentangling these and other theoretically important issues. Together they can lead to better neural-symbolic models. Dual-process theories serve as the theoretical basis and justifications for many cognitively motivated neural-symbolic models.

References

1 Sun, R.. *Integrating Rules and Connectionism for Robust Commonsense Reasoning*. John Wiley and Sons, New York, 1994.
2 Sun, R.. *Duality of the Mind*. Lawrence Erlbaum Associates, Mahwah, NJ., 2002.
3 Sun, R.. *Anatomy of the Mind*. Oxford University Press, Oxford, UK., 2016.
4 Sun, R. & L. Bookman (eds.). *Computational Architectures Integrating Neural and Symbolic Processes*. Kluwer, Needham, MA., 1994.

3.16 WuDao: Pretrain the World

Jie Tang (Tsinghua University – Beijing, CN)

Large-scale pre-trained model on web texts have substantially advanced the state of the art in various AI tasks, such as natural language understanding and text generation, and image processing, multimodal modeling. The downstream task performances have also constantly increased in the past few years. In this talk, I will first go through three families: autoregressive models (e.g., GPT), autoencoding models (e.g., BERT), and encoder-decoder models. Then, I will introduce China's first homegrown super-scale intelligent model system, with the goal of building an ultra-large-scale cognitive-oriented pretraining model to focus on essential problems in general artificial intelligence from a cognitive perspective. In particular, as an example, I will elaborate a novel pretraining framework GLM (General Language Model) to address this challenge. GLM has three major benefits: (1) it performs well on classification, unconditional generation, and conditional generation tasks with one single pretrained model; (2) it outperforms BERT-like models on classification due to improved pretrain-finetune consistency; (3) it naturally handles variable-length blank filling which is crucial for many downstream tasks. Empirically, GLM substantially outperforms BERT on the SuperGLUE natural language understanding benchmark with the same amount of pre-training data.

3.17 Knowledge Graph and Cognitive Learning: from Perception to Memories

Volker Tresp (Siemens – München, DE)

In this talk, I will present our work on learning representation about knowledge in different memories. A variety of cognitive memory functions are simulated, in particular those in semantic and concept memory, episodic memory, sensory memory, short-term memory, working memory, and the way that perception shapes semantic memory. I will also present and discuss our on-going researches, ranging from an integrated theoretical analysis, novel algorithms, to many new experimental results.

3.18 Thinking with the Body and the World

Barbara Tversky (Columbia University – New York, US)

All living things must move and act in space to survive, even plants, Without action in space, life ceases. Spatial thinking is the foundation of thought; not the entire edifice, but the foundation. In primates, the same brain structures that support spatial thinking also support conceptual thinking, Single cells in hippocampus gather multi-media information from all over the brain to represent places in space, events in time, ideas in conceptual spaces. The single cells are spatially arrayed in entorhinal cortex. Spatial thinking is acting in space with the things in space. Spatial thinking is multi-modal and established and distorted by our actions and perceptions of the spaces we interact in: the space of the body, the space immediately around the body; the larger space of navigation that must be pieced together from different multi-modal experiences.

Spatial thinking is evident in the ways we think and the ways we externalize thought, primarily through words, gestures, and graphics, Our minds go from thought to thought the ways our feet go from place to place, real or virtual paths from place to place or thought to thought. Our words and gestures act on thought the way we act on objects. We raise ideas, pull them together, tear them apart; those words are often accompanied by gestures of the physical actions, even though those actions aren't actually performed and don't need to be performed. Gestures and graphics bear more direct relations to meanings than symbolic words. They augment and alter our own thinking and that of others. When people are alone in a room studying complex texts for later testing, their hands often spontaneously create spatial-motor models of what they are trying to learn. When they do so, they remember better. Gestures we watch also augment comprehension and learning, of action, of time, of number, and many other concepts. Diagrams and sketches, both those provided to us and those we create, also augment comprehension and creativity. All of the above claims were substantiated by experiments. Gestures and graphics communicate more directly than purely symbolic words, but because they are in most cases neither discrete nor componential, are less tractable to AI than language.

Finally, we show that the world we live in is designed by our actions. The designs we create in the world reflect the ways we design our minds, into categories, orders, themes, 1-1 correspondences, symmetries, repetitions. The actions that create the designs are reflected in our words and our gestures; the designs are used to represent and communicate organizations of ideas, tables, networks, graphs. We call this cycle, spraction, a contraction of space, action, and abstraction. Actions in space create abstractions.

4 Working groups

4.1 Categorising Evaluation Instruments

Anthony Cohn (University of Leeds, GB)

I present work done jointly with Jose Hernandez-Orallo, for a project initiated by the OECD on the Future of Skills: Understanding the Educational Implications of AI and Robotics. I present and discuss an approach to categorising benchmarks, competitions, tests and evaluation standards as AI evaluation instruments (EI) via a set of 18 facets, which we believe will be valuable in distinguishing and evaluating different proposals for evaluating AI systems. These facets applied to two example AI evaluation instruments: the Arcade Learning Environment (ALE) and the Winograd Schema Challenge (WSC). We plan to conduct further evaluation on the validity and usefulness of these facets by applying them to many more EIs.

4.2 Spatial Humor

Tiansi Dong (Universität Bonn, DE) and Christian Hempelmann (Texas A&M University – Commerce, US)

Spatial Humor refers to the set of humor, such that spatial scene is the topic of two scripts and the spatial knowledge causes the logic opposition. The stimuli of *Spatial Humor* can either be cartoon without language description, or pure text description, or a mixture of both. We work on a novel neural-geometric computational model with rotating spheres as building blocks. This rotating sphere model serves as the semantic representation of *Spatial Humor* which synergistically unifies neural vector embedding and symbolic relations. Its representation power is demonstrated in the following aspects: (1) an object instance is represented by a rotating sphere in the embedding space; (2) object features are represented by rotating axes; (3) a snapshot view is represented by a configuration of spheres [3]; (4) an event [2] is schematized by a triple of a configuration of spheres, representing starting, middle, and end of the event; (5) the script opposition of a *spatial humor* is represented by two overlapped sphere configurations such that the overlapped spheres switch rotating axes. In this way, the motion of spheres embodies the dynamic process of humor understanding and the Script-based Semantic Theory of Humor [1].

References

1 Attardo, S., Hempelmann, C. F., and Di Maio, S. (2002). Script oppositionsand logical mechanisms: Modeling incongruities and their resolutions. Humor,15(1):3 – 46.

2 Zacks, J. and Tversky, B. (2001). Event structure in perception and conception. Psychological Bulletin, 127:3 – 21.

3 Dong, T. (2012). Recognizing Variable Environment – The Theory of Cognitive Prism. Studies in Computational Intelligence, Vol. 388, Springer.

4.3 Towards a Survey of Meaning Representation

Tiansi Dong (Universität Bonn, DE), Anthony Cohn (University of Leeds, GB), Christian Hempelmann (Texas A&M University – Commerce, US), Kanishka Misra (Purdue University – West Lafayette, , US), Jens Lehmann (Fraunhofer IAIS – Dresden, DE), Alexander Mehler (Goethe-Universität – Frankfurt am Main, DE), Tristan Miller (OFAI – Wien, AT), Siba Mohsen (Universität Bonn, DE), Roberto Navigli (Sapienza University of Rome, IT), Julia Rayz (Purdue University – West Lafayette, US) Stefan Wrobel (Universität Bonn, DE) Ron Sun (Rensselaer Polytechnic – Troy, US), and Volker Tresp (Siemens – München, DE)

After the working group on "What is missing in ML&AI to understanding Jokes?", we discussed the possibility to survey the expressiveness on existing models on meaning representation, contrasted by the forecast of existing theories in cognitive science about what is relevant cognitive activities and processes. Spatial stimuli activate the zoo of spatial cells in hippocampus, forming cognitive map or collage in the memory, producing spatial descriptions in languages. We need to survey existing models on Mental Spatial Representation (MSR) in the literature of cognitive psychology. On the other hand, we need to analyse vector embeddings of spatial entities and relations in the large-scaled pre-train world model, and find the gap between MSR and vector embedding via Machine Learning.

4.4 Rotating Sphere Model for NLP

Roberto Navigli (Sapienza University of Rome, IT), Tiansi Dong (Universität Bonn, DE), Thomas Liebig (TU Dortmund, DE), Yong Liu (Outreach Corporation – Seattle, US), Alexander Mehler (Goethe-Universität – Frankfurt am Main, DE), Tristan Miller (OFAI – Wien, AT), Siba Mohsen (Universität Bonn, DE), and Sven Naumann (Universität Trier, DE)

The existing n-ball embedding approach can precisely encode a large symbolic tree structure into tree node embeddings. In this working group, we discussed how to apply the idea of n-ball to solve NLP tasks, in particular, the Word Sense Disambiguation (WSD). WSD is a fundamental task in Natural Language Processing (NLP), which impacts a variety of downstream NLP applications. WSD determines the intended meaning of words in a context. To tackle the WSD task, researchers have been investigating knowledge-based approaches, supervised, semi-supervised, and unsupervised machine learning. However, those methods encounter a number of limitations, besides their costly computation. We let n-ball rotate, and result in the Rotating Spheres Model (RSM). Using RSM, embeddings of word senses work like gestures of a word. Given a context, the word chooses the best gesture. The WSD is to determine the best rotating axis in a given context. Each rating axis represents a sense that in the predefined sense inventory.

4.5 Joint deductive and inductive reasoning benchmarks

Achim Rettinger (Universität Trier, DE), Mehwish Alam (FIZ Karlsruhe, DE), Anthony Cohn (University of Leeds, GB), Bernado Cuenca Grau (University of Oxford, GB), Mateja Jamnik (University of Cambridge, GB), Thomas Liebig (TU Dortmund, DE), Roberto Navigli (Sapienza University of Rome, IT), and Steffen Staab (Universität Stuttgart, DE)

URL https://dblp.org/rec/journals/jwe/HarmelenT19

Data-driven empirical methods, specifically Deep Learning, has dominated the past decade of AI approaches, often to the extent that they allegedly achieve superhuman performance. In this working group we started from NLP benchmarks like GLUE and SuperGLUE and discussed how suited they are to measure language understanding, specifically in contrast to symbolic reasoning tasks. The goal of the working group was about benchmarks that require both statistical learning and deductive reasoning. The discussions in the group started from concrete benchmark tasks and capabilities of existing systems from all areas of AI, like Winograd Schema, Raven's Progressive Matrices, Digital Aristotle, Wolfram Alpha, Watson, humor detection and attempted to find criteria to describe characteristics and facets of benchmarks. Existing facets in related work on categorising evaluation instruments were discussed and an own list was extracted and compiled. Next, the discussions shifted from the benchmarks perspective to the systems perspective and existing approaches to categorizing hybrid systems were reviewed. While the questions of what makes a system a hybrid system, was not ultimately decided, it was agreed that a combination of the task perspective and the systems perspective would be a valuable contribution to the community. The next steps required to obtain more insights in this working groups topic would be a) an abstraction to classes of tasks, b) an abstraction to classes of concrete systems and c) and the extraction of relations between classes of tasks and classes of systems.

5 Open problems

5.1 What would be an aggregated neural model for syllogistic reasoning?

Tiansi Dong (Universität Bonn, DE), Pietro Lio (University of Cambridge, GB), and Ron Sun (Rensselaer Polytechnic – Troy, US)

Animals like monkeys can do syllogistic reasoning [2]. This suggests that syllogistic reasoning does not need language.

Given enough training data, vanilla neural networks can learn syllogism [3, 4]. Without training data, but equipped with a topological map, we can develop a novel neural network for rigorous syllogistic reasoning [5].

What would be an aggregated neural model of the two models satisfying all criteria in [1], such that it starts from learning patterns, making errors, then involving the capability in syllogistic reasoning. If one simulates System 1 for syllogism, and the other simulates System 2 for syllogistic reasoning. What could be the dual-process model [6, 7, 8, 9] for syllogistic reasoning?

References

1 S. Khemlani and P. N. Johnson-Laird (2012). *Theories of the Syllogism: A Meta-Analysis*. Psychological Bulletin 138(3) 427-457.
2 Ferrigno, Stephen and Huang, Yiyun and Cantlon, Jessica F (2021). *Reasoning Through the Disjunctive Syllogism in Monkeys*. Psychological Science 32(2) 1-9.
3 Wang, D., Jamnik, M., and Li'o, P. (2018). *Investigating diagrammatic reasoningwith deep neural networks*. In Diagrams 2018, pages 390 – 398.
4 Wang, D., Jamnik, M., and Li'o, P. (2020). *Abstract diagrammatic reasoningwith multiplex graph networks*. ICLR.
5 Dong, T., Li, C., Bauckhage, C., Li, J., Wrobel, S., and Cremers, A. B.(2020). *Learning syllogism with euler neural-networks*. arXiv preprintar Xiv:2007.07320.
6 Sun, R.. *Integrating Rules and Connectionism for Robust Commonsense Reasoning*. John Wiley and Sons, New York, 1994.
7 Sun, R.. *Duality of the Mind*. Lawrence Erlbaum Associates, Mahwah, NJ., 2002.
8 Sun, R.. *Anatomy of the Mind*. Oxford University Press, Oxford, UK., 2016.
9 Sun, R. & L. Bookman (eds.). *Computational Architectures Integrating Neural and Symbolic Processes*. Kluwer, Needham, MA., 1994.

5.2 How do we infuse human knowledge and machine learning to transform enterprise sales engagement and intelligence?

Yong Liu (Outreach Corporation – Seattle, US)

Joint work of Yong Liu, Joey Yip, Amit Sheth
Main reference Hong Yung Yip, Yong Liu, and Amit Sheth: "Using Contact, Content, and Context in Knowledge-Infused Learning: A Case Study of Non-Sequential Sales Processes in Sales Engagement Graphs", The Knowledge Graph Conference 2021 Workshop on Knowledge-Infused Learning (K-IL). May 2021
URL https://aiisc.ai/KiL2021/papers/K-iL2021_paper_2.pdf

The digital transformation is accelerating in the sales engagement domain with the applications of machine learning (ML) and artificial intelligence (AI). Given sales engagement is an inherently human-in-the-loop process, there is an increasing need from both business and technology capability sides to adopt a knowledge-infused learning approach that synthesizes human knowledge and statistical machine learning for explainable and trustworthy intelligence.

However, there are a lot of open questions about how we might realize such a vision. For example, how do we build a sales engagement knowledge graph incrementally by considering three types of data in the sales engagement (3Cs): Contacts (persons involved), Contents (multimedia materials including email messages, voice calls, demo videos, video conferences and transcripts, sales and purchase proposals, legal agreement, and contracts etc.) and Contexts (history of the engagement, customer pain points and success stories etc.)? How do we maintain and grow a temporal dynamic knowledge graph that allows time-traveling of the graph with flexibility to trace and incorporate the knowledge evolution? How do we provide turn-by-turn explainable recommendations along the process to guide the sales engagement? All these open questions are exciting areas to explore in the years to come.

5.3 What is missing in ML&AI to understand Jokes?

Alexander Mehler (Goethe-Universität – Frankfurt am Main, DE), Tiansi Dong (Universität Bonn, DE), Thomas Liebig (TU Dortmund, DE), Tristan Miller (OFAI – Wien, AT), Siba Mohsen (Universität Bonn, DE), and Sven Naumann (Universität Trier, DE)

Why current Machine Learning and AI (ML&AI) techniques cannot understand jokes as we humans do? What is missing? The knowledge that is needed to understand jokes is neither in the joke texts, nor in the neural networks. Acquisition and reasoning with commonsense knowledge is still an open problem for Machine Learning and AI. The meaning representation based on embeddings is insufficient. We need meaning representation formats that are beyond vector representations. Vectors are only shadows. Information processing and meaning understanding are embodied. The discussion guides us to develop novel embodied ML&AI techniques to understand *Spatial Jokes* first.

5.4 Can We Diagram the Understanding of Humour?

Tristan Miller (OFAI – Wien, AT), Anthony Cohn (University of Leeds, GB), Tiansi Dong (Universität Bonn, DE), Christian Hempelmann (Texas A&M University – Commerce, US), Siba Mohsen (Universität Bonn, DE), and Julia Rayz (Purdue University – West Lafayette, US)

Cartoons can be understood without language. That is, a suitably arranged scene of simple objects, with no accompanying text, is often enough to make us laugh – evidence that thinking (mental activity) happens before language [4]. This raises the question of non-linguistic diagrammatic representation of spatial humour, along with the mechanism of neural computation. In particular, we raise following questions: (1) How can we diagrammatically formalise spatial humour? (2) How can these diagrammatic formalisms be processed by neural networks? (3) How can this neural computation deliver high-level schema that are similar to the script-opposition semantic theory of humour [2, 1, 3]?

The spatial knowledge encoded in the scene can activate the necessary spatial and non-spatial knowledge. By what neural associative mechanism or process of reasoning do we put this all together to "get" the joke? During the seminar, we aimed to make some headway towards establishing (1) exactly what sort of scene-specific and common-sense knowledge is required to understand any given cartoon, (2) what part of this knowledge could in principle be acquired by existing machine learning (ML) techniques, and which could be acquired or encoded through symbolic structures, (3) what activation process acquires the rest of the knowledge required to interpret the humour, and (4) whether there is a unified representation that could represent this knowledge in a computer's working memory.

References

1 Salvatore Attardo. The General Theory of Verbal Humor. In Salvatore Attardo, editor, *The Routledge Handbook of Language and Humor*. Routledge, New York, NY, 2017.
2 Salvatore Attardo, Christian F. Hempelmann, and Sara Di Maio. Script oppositions and logical mechanisms: Modeling incongruities and their resolutions. *HUMOR: International Journal of Humor Research*, 15(1):3–46, 2002.
3 Julia Taylor Rayz. Scripts in the Ontological Semantic Theory of Humor. In Salvatore Attardo, editor, *Script-Based Semantics: Foundations and Applications*, pages 229–256. De Gruyter Mouton, Berlin, Boston, 2020.
4 Barbara Tversky. *Mind in Motion*. Basic Books, New York, NY, 2019.

Participants

- Tiansi Dong
Universität Bonn, DE
- Thomas Liebig
TU Dortmund, DE
- Alexander Mehler
Goethe-Universität – Frankfurt am Main, DE
- Sven Naumann
Universität Trier, DE
- Achim Rettinger
Universität Trier, DE
- Steffen Staab
Universität Stuttgart, DE
- Volker Tresp
Siemens – München, DE

Remote Participants

- Mehwish Alam
FIZ Karlsruhe, DE
- Christian Bauckhage
Fraunhofer IAIS – Sankt Augustin, DE
- Sven Behnke
Universität Bonn, DE
- Jian-Jia Chen
TU Dortmund, DE
- Anthony Cohn
University of Leeds, GB
- Armin B. Cremers
Universität Bonn, DE
- Bernado Cuenca Grau
University of Oxford, GB
- Asja Fischer
Ruhr-Universität Bochum, DE
- Xiaoming Fu
Universität Göttingen, DE
- Jürgen Gall
Universität Bonn, DE
- Christian Hempelmann
Texas A&M University – Commerce, US
- Mateja Jamnik
University of Cambridge, GB
- Markus Knauff
Universität Gießen, DE
- Michael Kramer
MPI für Radioastronomie, DE
- Jens Lehmann
Fraunhofer IAIS – Dresden, DE
- Juanzi Li
Tsinghua University – Beijing, CN
- Pietro Lio
University of Cambridge, GB
- Yong Liu
Outreach Corporation – Seattle, US
- Zhiyuan Liu
Tsinghua University – Beijing, CN
- Xiaoxing Ma
Nanjing University, CN
- Tristan Miller
OFAI – Wien, AT
- Siba Mohsen
Universität Bonn, DE
- Adithya Murali
University of Illinois – Urbana-Champaign, US
- Roberto Navigli
Sapienza University of Rome, IT
- Max Petrenko
Amazon Web Services – Seattle, US
- Julia Rayz
Purdue University – West Lafayette, US
- Ron Sun
Rensselaer Polytechnic – Troy, US
- Jian Tang
Mila – Montreal, CA
- Jie Tang
Tsinghua University – Beijing, CN
- Barbara Tversky
Columbia University – New York, US
- Frank van Harmelen
VU University Amsterdam, NL
- Ran Yu
Universität Bonn, DE