CSSA'20: Workshop on Combining Symbolic and Sub-Symbolic Methods and their Applications

Alam, M.; Groth, P.; Hitzler, P.; Paulheim, H.; Sack, H.; Tresp, V.

DOI
10.1145/3340531.3414072

Publication date
2020

Document Version
Final published version

Published in
CIKM '20

CSSA'20: Workshop on Combining Symbolic and Sub-Symbolic Methods and their Applications

Mehwish Alam
FIZ Karlsruhe - Leibniz Institute for Information Infrastructure
Karlsruhe, Germany
mehwish.alam@fiz-karlsruhe.de

Paul Groth
Universiteit van Amsterdam
Amsterdam, The Netherlands
p.groth@uva.nl

Pascal Hitzler
Kansas State University
Manhattan, KS, USA
hitzler@ksu.edu

Heiko Paulheim
University of Mannheim
Mannheim, Germany
heiko@informatik.uni-mannheim.de

Harald Sack
FIZ Karlsruhe - Leibniz Institute for Information Infrastructure
Karlsruhe, Germany
harald.sack@fiz-karlsruhe.de

Volker Tresp
Ludwig Maximillian University of Munich, Research Scientist Siemens
Munich, Germany
volker.tresp@siemens.com

ABSTRACT

There has been a rapid growth in the use of symbolic representations along with their applications in many important tasks. Symbolic representations, in the form of Knowledge Graphs (KGs), constitute large networks of real-world entities and their relationships. On the other hand, sub-symbolic artificial intelligence has also become a mainstream area of research. This workshop brought together researchers to discuss and foster collaborations on the intersection of these two areas.

CCS CONCEPTS

• **Computing methodologies** → **Knowledge representation and reasoning**; **Machine learning**.

KEYWORDS

Symbolic Artificial Intelligence, Sub-symbolic Artificial Intelligence, Neural Networks, Knowledge Graphs

ACM Reference Format:
Mehwish Alam, Paul Groth, Pascal Hitzler, Heiko Paulheim, Harald Sack, and Volker Tresp. 2020. CSSA'20: Workshop on Combining Symbolic and Sub-Symbolic Methods and their Applications. In *Proceedings of the 29th ACM International Conference on Information and Knowledge Management (CIKM '20), October 19–23, 2020, Virtual Event, Ireland*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3340531.3414072

1 INTRODUCTION

Since the beginning of 2000s, there has been an increasing number of studies and standards proposed for generating large scale symbolic representations of knowledge (known as Knowledge Graphs (KGs)) out of heterogeneous resources such as text, images, etc. Moreover, there have been many advances in the symbolic reasoning as well as their applications to various fields. Recently, sub-symbolic methods have gained momentum. These methods aim at generating distributed representations from several resources such as text or symbolic representations (Graph Neural Networks, KG embeddings, etc.). These sub-symbolic methods for symbolic representations mainly focus on the task of KG completion [1]. However, these have also recently been used for various tasks such as Natural Language Processing (NLP). The perspective future of these methods would be a combination of these approaches leading to neurosymbolic reasoning. Advances in the real world applications related to these methods will also serve as a stepping stone in the practicality of these methods.

CIKM '20, October 19–23, 2020, Virtual Event, Ireland

ACM ISBN 978-1-4503-6859-9/20/10.
https://doi.org/10.1145/3340531.3414072

2 OBJECTIVES AND TOPICS OF INTEREST

The focus of the workshop is to stimulate discussions and encourage new collaborations among the researchers working on the proposed topic. This workshop invited contributions on the methods combining symbolic and sub-symbolic approaches. These methods included generation of KG embeddings, dealing with temporality in KG embeddings, how the evolution of KGs can be captured in the KG embeddings methods, neurosymbolic reasoning, etc. The workshop also also invited other related topics such as using these methods for NLP or Biomedical domain, explainability and interpretatbility in KG embeddings.

3 WORKSHOP PROGRAM & ORGANIZATION

CSSA takes place as a half-day virtual workshop due to the COVID-19 pandemic. The program opens with a keynote talk and continues with the original contributions of research papers and the lightning talks which focus on the open problems. All paper submissions were reviewed by three members of the program committee with respect to scientific quality. Submissions were accepted if all three reviewers agreed to accept, or if their average score was above zero (scores ranged from strong accept (+3) to strong reject (-3)).

3.1 Keynote Speaker

The keynote speaker of this workshop was **Maximilian Nickel**. He is a research scientist at Facebook AI Research in New York,

USA. His research centers around geometric methods for learning and reasoning with relational knowledge representations and their applications in artificial intelligence and network science.

The keynote entitled "Geometric Representation Learning in Symbolic Domains" focuses on methods such as hyperbolic embeddings and Riemannian generative models [3] which show that non-Euclidean geometries can provide significant advantages for modeling relational data, e.g., with regard to interpretability, scalability, and latent semantics. This talk provides an overview of recent work on such geometric approaches to representation learning. It first discussed how structural properties of relational data (such as latent hierarchies) are connected to the geometry of the embedding space and how methods such hyperbolic embeddings allow us to learn parsimonious representations in these cases. Moreover, it also discussed how the embeddings can be used to discover latent hierarchies and be applied for diverse tasks in NLP [2] and bioinformatics. In addition, It also focused on discussing how a flexible probability distributions over such geometric representations through Riemmannian continous normalizing flows can be modeled.

3.2 Accepted Papers

Seven research papers were accepted, out of which four were long papers, three were short papers and two were lightning talks.

Long Papers. The long papers belonged to diverse topics within the field of symbolic and sub-symbolic artificial intelligence, i.e., simplifying architecture search for Graph Neural Networks, the negative impact of A-box materialization on RDF2vec KG Embeddings. One of the long papers used KGs along with language models for creative storytelling, i.e., using KGs for NLP task. Finally, another work focused on comparing the symbolic and sub-symbolic methods, i.e., whether they can help each other or using them together could be counter productive.

Short Papers. The three short papers discussed an adaptive semantic stream reasoning framework for Deep Neural Networks, leveraging an ontology in a neural learning process, and neuro-symbolic visual reasoning for multimedia event processing which gives an overview, some prospects and the current challenges.

Lightning Talks. Two extended abstracts were accepted to be presented as lightning talks (5-10 minutes slot) about WalkingTime: dynamic graph embedding using temporal-topological flows and learning with temporal KGs.

3.3 Program Committee

We thank the program committee for their on-time and thorough reviews and valuable support:

Federico Bianchi (Bocconi University, Milan, Italy), Danilo Dessi (FIZ Karlsruhe Leibniz Institute for Information Infrastructure, Germany), Stefan Dietze (L3S Hannover, Germany), Mauro Dragoni (Fondazione Bruno Kessler, Italy), Monireh Ebrahimi (DaSe Lab, Kansas State University, USA), Dagmar Gromann (University of Vienna, Austria), Jim Hendler (RPI, USA), Kim Hyeongsik (Robert Bosch, USA), Mayank Kejriwal (University of Southern California, USA), Luis Lamb (Instituto de Informatica, UFRGS, Brazil), Freddy Lecue (CortAIx Thales, Montreal, Canada and INRIA, France), Gerard de Melo (Rutgers University, USA), Pasquale Minervini (University College London, UK), Amedeo Napoli (CNRS, France), Finn Årup Nielsen (Technical University of Denmark, Denmark), Matteo Palmonari (University of Milan Bicocca, Italy), Ilaria Tiddi (Vrije Universiteit Amsterdam, The Netherlands), Lei Zhang (FIZ Karlsruhe Leibniz Institute for Information Infrastructure, Germany).

3.4 Organizing Committee

Mehwish Alam is a Postdoctoral Researcher at FIZ-Karlsruhe, Leibniz Institute for Information Infrastructure. Her main research interests include Knowledge Graphs, Machine/Deep Learning, Natural Language Processing and Knowledge Based Systems.

Paul Groth is a Professor of Data Science at the University of Amsterdam where he leads the Intelligent Data Engineering Lab (INDElab). His research focuses on intelligent systems for dealing with diverse contextualized knowledge with a particular focus on web and science applications. This includes research in data provenance, data integration and knowledge sharing.

Pascal Hitzler is a Professor and endowed Lloyd T. Smith Creativity in Engineering Chair and Director of the Center for Artificial Intelligence and Data Science (CAIDS) at the Department of Computer Science at Kansas State University. His research topic include diverse areas as semantic web, AI, neural-symbolic integration, knowledge representation and reasoning, machine learning, denotational semantics, and set-theoretic topology.

Heiko Paulheim is a Professor of Data Science at the University of Mannheim, Germany. His research focuses on the construction and usage of large-scale knowledge graphs. He explores methods for generating those KGs from various sources (such as Wikis and other structured Web sites), as well as techniques for automatically refining those graphs by inferring missing knowledge or finding errors by means of heuristic inference or machine learning.

Harald Sack is a Professor of Information Service Engineering at FIZ Karlsruhe, Leibniz Institute for Information Infrastructure and Karlsruhe Institute of Technology. His research is focused on semantic technologies, knowledge discovery, NLP, and machine learning, with applications in exploratory search and recommendation in unstructured and multimodal data.

Volker Tresp is a Research Scientist at Siemens and a Professor for Machine Learning at the Ludwig Maximilian University of Munich. His research focus has been "Machine Learning in Information Networks" for modelling KG, medical decision processes, perception, and cognitive memory functions.

REFERENCES

[1] Genet Asefa Gesese, Russa Biswas, Mehwish Alam, and Harald Sack. 2019. A Survey on Knowledge Graph Embeddings with Literals: Which model links better Literal-ly? *CoRR* abs/1910.12507 (2019). arXiv:1910.12507 http://arxiv.org/abs/1910.12507

[2] Matt Le, Stephen Roller, Laetitia Papaxanthos, Douwe Kiela, and Maximilian Nickel. 2019. Inferring Concept Hierarchies from Text Corpora via Hyperbolic Embeddings. In *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*, Anna Korhonen, David R. Traum, and Lluís Màrquez (Eds.). Association for Computational Linguistics, 3231–3241. https://doi.org/10.18653/v1/p19-1313

[3] Qi Liu, Maximilian Nickel, and Douwe Kiela. 2019. Hyperbolic Graph Neural Networks. In *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, 8-14 December 2019, Vancouver, BC, Canada.* 8228–8239.