

Graph Embeddings: Theory meets Practice

Martin Grohe^{*1}, Stephan Günnemann^{*2}, Stefanie Jegelka^{*3}, and Christopher Morris^{*4}

¹ RWTH Aachen University, DE. grohe@informatik.rwth-aachen.de

² TU München, DE. guennemann@in.tum.de

³ MIT – Cambridge, US. stefje@csail.mit.edu

⁴ McGill University & MILA – Montreal. chris@christophermorris.info

Abstract

Vectorial representations of graphs and relational structures, so-called graph embeddings, make it possible to apply standard tools from data mining, machine learning, and statistics to the graph domain. In particular, graph embeddings aim to capture important information about, both, the graph structure and available side information as a vector, to enable downstream tasks such as classification, regression, or clustering. Starting from the 1960s in chemoinformatics, research in various communities has resulted in a plethora of approaches, often with recurring ideas. However, most of the field advancements are driven by intuition and empiricism, often tailored to a specific application domain. Until recently, the area has received little stimulus from theoretical computer science, graph theory, and learning theory. The Dagstuhl Seminar 22132 “Graph Embeddings: Theory meets Practice”, was aimed to gather leading applied and theoretical researchers in graph embeddings and adjacent areas, such as graph isomorphism, bio- and chemoinformatics, and graph theory, to stimulate an increased exchange of ideas between these communities.

Seminar March 27–30, 2022 – <http://www.dagstuhl.de/22132>

2012 ACM Subject Classification Mathematics of computing → Discrete mathematics; Computing methodologies → Machine learning

Keywords and phrases Machine Learning For Graphs, GNNs, Graph Embedding

Digital Object Identifier 10.4230/DagRep.12.3.141

1 Executive Summary

Martin Grohe (RWTH Aachen University, DE)

Stephan Günnemann (TU München, DE)

Stefanie Jegelka (MIT – Cambridge, US)

Christopher Morris (McGill University & MILA – Montreal)

License  Creative Commons BY 4.0 International license

© Martin Grohe, Stephan Günnemann, Stefanie Jegelka, and Christopher Morris

Graph-structured data is ubiquitous across application domains ranging from chemo- and bioinformatics to image and social network analysis. To develop successful machine learning algorithms or apply standard data analysis tools in these domains, we need techniques that map the rich information inherent in the graph structure to a vectorial representation in a meaningful way—so-called graph embeddings. Designing such embeddings comes with unique challenges. The embedding has to account for the complex structure of (real-world) networks and additional high-dimensional continuous vectors attached to nodes and edges in a (permutation) invariant way while being scalable to massive graphs or sets of graphs. Moreover, when used in supervised machine learning, the model trained with such embeddings

* Editor / Organizer



Except where otherwise noted, content of this report is licensed under a Creative Commons BY 4.0 International license

Graph Embeddings: Theory meets Practice, *Dagstuhl Reports*, Vol. 12, Issue 3, pp. 141–155

Editors: Martin Grohe, Stephan Günnemann, Stefanie Jegelka, and Christopher Morris



Dagstuhl Reports

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

must generalize well to new or previously unseen (graph) instances. Hence, more abstractly, designing graph embeddings results in a trade-off between expressivity, scalability, and generalization.

Starting from the 1960s in chemoinformatics, different research communities have worked in the area under various guises, often leading to recurring ideas. Moreover, triggered by the resurgence of (deep) neural networks, there is an ongoing trend in the machine learning community to design invariant/equivariant neural architectures that are capable of dealing with graph- and relational input, both (semi-)supervised and unsupervised, often denoted as graph neural networks. Although successful in practical settings, most of these developments are driven by intuition and empiricism and are geared towards specific application areas. There is no clear understanding of these approaches' limitations and their trade-offs in complexity, expressivity, and generalization. Researchers recently started to leverage connections to graph theory, group theory, logic, combinatorial algorithms, and (algorithmic) learning theory, leading to new theoretical insights and triggering new research in applications. Hence, in this seminar, we aimed to bring together leading applied and theoretical researchers in graph embeddings and adjacent areas, such as graph isomorphism, bio- and chemoinformatics, graph theory, to facilitate an increased exchange of ideas between these communities. Concretely, we aimed to understand what hinders recent theoretical developments being applied in application areas and worked towards a more practical theory. Further, we aimed at understanding the overarching challenges across applications and challenges inherent to specific areas to stimulate directions for further practical and theoretical research.

The seminar brought together 33 researchers from (applied) mathematics, specifically harmonic analysis and (algebraic) topology, (theoretical) computer science, machine learning, bioinformatics, and network science. Eighteen researchers attended remotely owing to the global COVID-19 pandemic. In total, the participants presented 18 talks on their recent progress in a better understanding of graph embeddings, focusing on supervised machine learning, particularly graph neural networks. Many talks dealt with leveraging tools from graph isomorphism testing and related areas such as finite model theory and group theory. In particular, the Weisfeiler-Leman algorithm, a popular heuristic for the graph isomorphism problem, was used to measure the expressivity of the presented algorithms and neural architectures. The consensus was that the above algorithm leads to a too coarse-grained measure of expressivity, and new notions of expressivity are needed to develop a thorough understanding. Surprisingly, only a few talks dealt with developing a better understanding of generalization, indicating that the research community still lacks an understanding. Notably, Gitta Kutyniok showed how to leverage random graph models and graphons to analyze the generalization error of graph neural networks, while Bruno Ribeiro talked about the connection between causality and out-of-distribution generalization. Further, some talks used methods from (algebraic) topology and their connection to graph theory to devise provably expressive architectures and to better understand common problems with graph neural networks, e.g., the problem of “over-smoothing” of node representations faced when considering deep architectures. Moreover, two talks covered the challenges of applying graph neural networks to biomedical data and industrial applications at Google, respectively, indicating a gap between theoretical results and practical architectures.

Concluding Remarks

The seminar was well received, as witnessed by several positive comments from on-site participants. In general, there was an exciting atmosphere at the seminar, particularly among the large number of junior researchers attending the seminar on-site, also witnessed by many

lively discussions during on-site talks. However, this was not always the case during online talks, and the active participation of online participants was relatively low. Finally, the organizers wish to express their gratitude to the Scientific Directors of Schloss Dagstuhl – Leibniz Center for Informatics for their support of the seminar.

2 Table of Contents

Executive Summary

Martin Grohe, Stephan Günnemann, Stefanie Jegelka, and Christopher Morris . . . 141

Overview of Talks

Graph Neural Networks with Local Graph Parameters <i>Pablo Barcelo</i>	146
Probing Graph Representations <i>Aleksandar Bojchevski</i>	146
Graph Neural Networks and Graph Representation Learning Through the Lens of Curvature <i>Francesco Di Giovanni</i>	147
Graph Representation Learning on Simplicial and Cellular Complexes <i>Fabrizio Frasca</i>	147
Higher-order MPNNs: A Unifying Approach for Studying Expressiveness and Approximation Properties of GNNs <i>Floris Geerts</i>	148
Weisfeiler and Leman Go Walking: Random Walk Kernels Revisited <i>Nils Kriege</i>	148
Stability and Generalization Capabilities of Graph Neural Networks <i>Gitta Kutyniok</i>	149
Equivariant Subgraph Aggregation Networks <i>Haggai Maron</i>	149
Frame Averaging for Invariant and Equivariant Network Design <i>Yaron Lipman</i>	150
Challenges of Applying Graph Neural Networks <i>Bryan Perozzi</i>	151
Causal Graph Representation Learning <i>Bruno Ribeiro</i>	151
Topology-Based Graph Learning <i>Bastian Rieck</i>	152
Universal Graph Neural Networks via Random Data Augmentations Using Graph Isomorphism Tools <i>Pascal Schweitzer</i>	152
Combining Representation Learning and Logical Rule Reasoning for Knowledge Graph Inference <i>Yizhou Sun</i>	153
Graph Learning with 1D Convolutions on Random Walks <i>Jan Tönshoff</i>	153
Graph Neural Networks are Dynamic Programmers <i>Petar Velickovic</i>	154

Infusing Structure and Knowledge into Biomedical AI

Marinka Zitnik 154

Participants 155

Remote Participants 155

3 Overview of Talks

3.1 Graph Neural Networks with Local Graph Parameters

Pablo Barcelo (PUC – Santiago de Chile, CL)

License © Creative Commons BY 4.0 International license
© Pablo Barcelo

Main reference Pablo Barceló, Floris Geerts, Juan Reutter, Maksimilian Ryschkov: “Graph Neural Networks with Local Graph Parameters”, in Proc. of the Advances in Neural Information Processing Systems, Vol. 34, pp. 25280–25293, Curran Associates, Inc., 2021.

URL <https://proceedings.neurips.cc/paper/2021/file/d4d8d1ac7e00e9105775a6b660dd3cbb-Paper.pdf>

Various recent proposals increase the distinguishing power of Graph Neural Networks (GNNs) by propagating features between k -tuples of vertices. The distinguishing power of these “higher-order” GNNs is known to be bounded by the k -dimensional Weisfeiler-Leman (WL) test, yet their nonlinear memory requirements limit their applicability. Other proposals infuse GNNs with local higher-order graph structural information from the start, thereby inheriting the desirable linear memory requirement from GNNs at the cost of a one-time, possibly non-linear, preprocessing step. We propose local graph parameter enabled GNNs as a framework for studying the latter kind of approaches. We precisely characterize their distinguishing power, in terms of a variant of the WL test, and in terms of the graph structural properties that they can take into account. Local graph parameters can be added to any GNN architecture, and are cheap to compute. In terms of expressive power, our proposal lies in the middle of GNNs and their higher-order counterparts. Further, we propose several techniques to aid in choosing the right local graph parameters. Our results connect GNNs with deep results in finite model theory and finite variable logics.

3.2 Probing Graph Representations

Aleksandar Bojchevski (CISPA – Saarbrücken, DE)

License © Creative Commons BY 4.0 International license
© Aleksandar Bojchevski

Today we have a good theoretical understanding of the representational power of Graph Neural Networks (GNNs). For example, their limitations have been characterized in relation to a hierarchy of Weisfeiler-Lehman (WL) graph isomorphism tests. Consequently, there is a large body of work proposing more powerful GNNs that mitigate these limitations. We argue that these findings are only part of the story since many other factors besides the model influence learning. To complete the picture we propose a probing framework to quantify the amount of (semantically meaningful) information captured in learned graph representations. Our preliminary findings on molecular representations highlight the potential of this framework for understanding the inductive biases in GNNs and the interplay between node features and graph structure

3.3 Graph Neural Networks and Graph Representation Learning Through the Lens of Curvature

Francesco Di Giovanni (Twitter – San Francisco, US)

License © Creative Commons BY 4.0 International license

© Francesco Di Giovanni

Joint work of Francesco Di Giovanni, Giulia Luise, Jake Topping, Benjamin Chamberlain, Xiaowen Dong, Michael Bronstein

Curvature is a fundamental object in the analysis of manifolds and intrinsically characterizes their geometry. It is not surprising then that synthetic notions of curvature have been introduced on graphs despite the lack of an underlying differentiable structure. In this talk, I will explore how these ideas have been recently investigated in the context of graph neural networks and graph representation learning. In the first case, curvature turns out to be the right tool to monitor the propagation of information inside message passing neural networks and allows us to properly analyse and formalize the problem of over-squashing. In the second one, we construct a family of graph embeddings into heterogeneous manifolds that are able to both match pairwise distances on the graph and the discrete graph curvature with the one on the ambient space leading to better preservation of higher order structures.

3.4 Graph Representation Learning on Simplicial and Cellular Complexes

Fabrizio Frasca (Twitter – London & Imperial College London)

License © Creative Commons BY 4.0 International license

© Fabrizio Frasca

Joint work of Cristian Bodnar, Fabrizio Frasca, Nina Otter, Yu Guang Wang, Pietro Liò, Guido Montúfar, Michael Bronstein

Main reference Cristian Bodnar, Fabrizio Frasca, Nina Otter, Yu Guang Wang, Pietro Liò, Guido Montúfar, Michael Bronstein: “Weisfeiler and Lehman Go Cellular: CW Networks”, arXiv, 2021.

URL <https://doi.org/10.48550/ARXIV.2106.12575>

Graphs represent flexible and convenient mathematical abstractions for the modelling of relational systems. However, pairwise interactions may fail to capture the multi-level system of relations of many complex systems, and computational schemes embodying such paradigm are of limited expressive power. We explore topological generalisation of graphs: Simplicial and Cellular Complexes. We show they constitute natural and valid frameworks to model higher-order interactions, and how their combinatorial structure lead to the design of novel hierarchical colouring procedures extending the Weisfeiler-Leman algorithm. Graphs can be lifted to Simplicial and Cellular Complexes with appropriate transformations, allowing the application of such colouring procedures for provably more expressive representations. Finally, these procedures inspire the design of neural counterparts implementing a form of higher-order message passing. These expressive architectures overcome several limitations of standard Graph Neural Networks; we show they excel on a variety of graph learning benchmarks and obtain state-of-the-art results on various molecular datasets.

3.5 Higher-order MPNNs: A Unifying Approach for Studying Expressiveness and Approximation Properties of GNNs

Floris Geerts (University of Antwerp, BE)

License  Creative Commons BY 4.0 International license
© Floris Geerts

Joint work of Floris Geerts, Juan L Reutter

Main reference Floris Geerts, Juan L Reutter: “Expressiveness and Approximation Properties of Graph Neural Networks”, in Proc. of the International Conference on Learning Representations, 2022.

URL <https://openreview.net/forum?id=wIzUeM3TAU>

Characterizing the separation power of graph neural networks (GNNs) provides an understanding of their limitations for graph learning tasks. Results regarding separation power are, however, usually geared at specific GNN architectures, and tools for understanding arbitrary GNN architectures are generally lacking. We provide an elegant way to easily obtain bounds on the separation power of GNNs in terms of the Weisfeiler-Leman (WL) tests, which have become the yardstick to measure the separation power of GNNs. The crux is to view GNNs as expressions in a procedural tensor language describing the computations in the layers of the GNNs. Then, by a simple analysis of the obtained expressions, in terms of the number of indexes and the nesting depth of summations, bounds on the separation power in terms of the WL-tests readily follow. We use tensor language to define Higher-Order Message-Passing Neural Networks (or k-MPNNs), a natural extension of MPNNs. Furthermore, the tensor language point of view allows for the derivation of universality results for classes of GNNs in a natural way. Our approach provides a toolbox with which GNN architecture designers can analyze the separation power of their GNNs, without needing to know the intricacies of the WL-tests. We also provide insights in what is needed to boost the separation power of GNNs.

3.6 Weisfeiler and Leman Go Walking: Random Walk Kernels Revisited

Nils Kriege (Universität Wien, AT)

License  Creative Commons BY 4.0 International license
© Nils Kriege

Main reference Nils M. Kriege: “Weisfeiler and Leman Go Walking: Random Walk Kernels Revisited”, CoRR, Vol. abs/2205.10914, 2022.

URL <https://doi.org/10.48550/arXiv.2205.10914>

Random walk kernels have been introduced in seminal work on graph learning and were later largely superseded by kernels based on the Weisfeiler-Leman test for graph isomorphism. We give a unified view on both classes of graph kernels. We study walk-based node refinement methods and formally relate them to several widely-used techniques, including Morgan’s algorithm for molecule canonization and the Weisfeiler-Leman test. We define corresponding walk-based kernels on nodes that allow fine-grained parameterized neighborhood comparison, reach Weisfeiler-Leman expressiveness, and are computed using the kernel trick. From this we show that classical random walk kernels with only minor modifications regarding definition and computation are as expressive as the widely-used Weisfeiler-Leman subtree kernel but support non-strict neighborhood comparison. We verify experimentally that walk-based kernels reach or even surpass the accuracy of Weisfeiler-Leman kernels in real-world classification tasks.

3.7 Stability and Generalization Capabilities of Graph Neural Networks

Gitta Kutyniok (LMU München, DE)

License © Creative Commons BY 4.0 International license
© Gitta Kutyniok

Joint work of Gitta Kutyniok, Holger Boche, Michael M. Bronstein, Lorenzo Bucci, Adalbert Fono, Wei Huang, Yunseok Lee, Ron Levie, Sohir Maskey

The tremendous importance of graph structured data due to recommender systems or social networks led to the introduction of graph convolutional neural networks (GCN). We ask the question to which extent GCN are able to generalize to graphs, which describe a similar phenomenon as present in the training data set. We consider different notions of similarity, using random graph models as well as graphons, and analyze both spectral GCNs [1, 2] and message passing neural networks [3]. In these settings, we will then derive comprehensive non-asymptotic bounds on the related generalization error. We will finish with a word of caution when training graph neural networks on classical digital hardware, and present fundamental limitations [4, 5].

References

- 1 R. Levie, W. Huang, L. Bucci, M. M. Bronstein, and G. Kutyniok. Transferability of Spectral Graph Convolutional Neural Networks. *J. Mach. Learn. Res.*, to appear. (arXiv:1907.12972).
- 2 S. Maskey, R. Levie, and G. Kutyniok. Transferability of Graph Neural Networks: an Extended Graphon Approach. (arXiv:2109.10096)
- 3 S. Maskey, Y. Lee, R. Levie, and G. Kutyniok. Stability and Generalization Capabilities of Message Passing Graph Neural Networks (arXiv:2202.00645)
- 4 H. Boche, A. Fono and G. Kutyniok. Limitations of Deep Learning for Inverse Problems on Digital Hardware (arXiv:2202.13490)
- 5 H. Boche, A. Fono and G. Kutyniok. Inverse Problems Are Solvable on Real Number Signal Processing Hardware (arxiv:2204.02066)

3.8 Equivariant Subgraph Aggregation Networks

Haggai Maron (NVIDIA – Tel Aviv, IL)

License © Creative Commons BY 4.0 International license
© Haggai Maron

Joint work of Beatrice Bevilacqua, Fabrizio Frasca, Derek Lim, Balasubramaniam Srinivasan, Chen Cai, Gopinath Balamurugan, Michael M. Bronstein, Haggai Maron

Main reference Beatrice Bevilacqua, Fabrizio Frasca, Derek Lim, Balasubramaniam Srinivasan, Chen Cai, Gopinath Balamurugan, Michael M. Bronstein, Haggai Maron: “Equivariant Subgraph Aggregation Networks”, in *Proc. of the International Conference on Learning Representations*, 2022.

URL <https://openreview.net/forum?id=dFbKQaRk15w>

Message-passing neural networks (MPNNs) are the leading architecture for deep learning on graph-structured data, in large part due to their simplicity and scalability. Unfortunately, it was shown that these architectures are limited in their expressive power. This work proposes a novel framework called Equivariant Subgraph Aggregation Networks (ESAN) to address this issue. Our main observation is that while two graphs may not be distinguishable by an MPNN, they often contain distinguishable subgraphs. Thus, we propose to represent each graph as a set of subgraphs derived by some predefined policy, and to process it using a suitable equivariant architecture. We develop novel variants of the 1-dimensional Weisfeiler-Leman (1-WL) test for graph isomorphism, and prove lower bounds on the expressiveness of ESAN in terms of these new WL variants. We further prove that our approach increases the

expressive power of both MPNNs and more expressive architectures. Moreover, we provide theoretical results that describe how design choices such as the subgraph selection policy and equivariant neural architecture affect our architecture’s expressive power. To deal with the increased computational cost, we propose a subgraph sampling scheme, which can be viewed as a stochastic version of our framework. A comprehensive set of experiments on real and synthetic datasets demonstrates that our framework improves the expressive power and overall performance of popular GNN architectures.

3.9 Frame Averaging for Invariant and Equivariant Network Design

Yaron Lipman (*Weizmann Institute – Rehovot, IL*)

Joint work of Omri Puny, Matan Atzmon, Heli Ben-Hamu, Ishan Misra, Aditya Grover, Edward J. Smith, Yaron Lipman

License  Creative Commons BY 4.0 International license
© Yaron Lipman

Main reference Omri Puny, Matan Atzmon, Heli Ben-Hamu, Edward J. Smith, Ishan Misra, Aditya Grover, Yaron Lipman: “Frame Averaging for Invariant and Equivariant Network Design”, CoRR, Vol. abs/2110.03336, 2021.

URL <https://arxiv.org/abs/2110.03336>

Many machine learning tasks involve learning functions that are known to be invariant or equivariant to certain symmetries of the input data. However, it is often challenging to design neural network architectures that respect these symmetries while being expressive and computationally efficient. For example, Euclidean motion invariant/equivariant graph or point cloud neural networks.

In this work we introduce Frame Averaging (FA), a general purpose and systematic framework for adapting known (backbone) architectures to become invariant or equivariant to new symmetry types. Our framework builds on the well known group averaging operator that guarantees invariance or equivariance but is intractable. In contrast, we observe that for many important classes of symmetries, this operator can be replaced with an averaging operator over a small subset of the group elements, called a frame. We show that averaging over a frame guarantees exact invariance or equivariance while often being much simpler to compute than averaging over the entire group. Furthermore, we prove that FA-based models have maximal expressive power in a broad setting and in general preserve the expressive power of their backbone architectures. Using frame averaging, we propose a new class of universal Graph Neural Networks (GNNs), universal Euclidean motion invariant point cloud networks, and Euclidean motion invariant Message Passing (MP) GNNs. We demonstrate the practical effectiveness of FA on several applications including point cloud normal estimation, beyond 2-WL graph separation, and n -body dynamics prediction, achieving state-of-the-art results in all of these benchmarks.

3.10 Challenges of Applying Graph Neural Networks

Bryan Perozzi (Google – New York, US)

License © Creative Commons BY 4.0 International license
© Bryan Perozzi

Joint work of Bryan Perozzi, John Palowitch, Anton Tsitsulin, Brandon Mayer, Qi Zhu, Natalia Ponomareva, Jonathan Halcrow, Sam Ruth, Alexandru Mosoi, Jiawei Han

Main reference Jonathan Halcrow, Alexandru Mosoi, Sam Ruth, Bryan Perozzi: “Grale: Designing Networks for Graph Learning”, in Proc. of the KDD ’20: The 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Virtual Event, CA, USA, August 23-27, 2020, pp. 2523–2532, ACM, 2020.

URL <https://doi.org/10.1145/3394486.3403302>

Graph Neural Networks are a tantalizing way of modeling data which doesn’t have a fixed structure. However, getting them to work as expected has had some twists and turns over the years. In this talk, I discuss three efforts from our group on important (and understudied) problems applying GNNs to real data including graph construction, model benchmarking, and model robustness:

Grale, is a scalable method we have developed to address the problem of graph design for graphs with billions of nodes. Grale operates by fusing together different measures of (potentially weak) similarity to create a graph which exhibits high task-specific homophily between its nodes. Grale is designed for running on large datasets. We have deployed Grale in more than 20 different industrial settings at Google, including datasets which have tens of billions of nodes, and hundreds of trillions of potential edges to score.

GraphWorld is a novel methodology and system for benchmarking GNN models on an arbitrarily-large population of synthetic graphs for any conceivable GNN task. GraphWorld allows a user to efficiently generate a world with millions of statistically diverse datasets. It is accessible, scalable, and easy to use. GraphWorld can be run on a single machine without specialized hardware, or it can be easily scaled up to run on arbitrary clusters or cloud frameworks. Using GraphWorld, a user has fine-grained control over graph generator parameters, and can benchmark arbitrary GNN models with built-in hyperparameter tuning

Shift-Robust GNN (SR-GNN) is designed to account for distributional differences between biased training data and a graph’s true inference distribution. SR-GNN adapts GNN models to the presence of distributional shift between the nodes labeled for training and the rest of the dataset. We illustrate the effectiveness of SR-GNN in a variety of experiments with biased training datasets on common GNN benchmark datasets for semi-supervised learning, where we see that SRGNN outperforms other GNN baselines in accuracy, addressing at least ~40% of the negative effects introduced by biased training data.

3.11 Causal Graph Representation Learning

Bruno Ribeiro (Purdue University – West Lafayette, US)

License © Creative Commons BY 4.0 International license
© Bruno Ribeiro

Joint work of Bruno Ribeiro, Beatrice Bevilacqua, Yangze Zhou, S Chandra Mouli

In this talk I discussed the challenges and opportunities in building graph representations for causal tasks (learning and prediction). We started with the question “Why is causality relevant for graph machine learning?”, expanding it into three threads: (a) Some graph tasks are causal, such as link prediction for recommender systems; (b) Extrapolation tasks in deep learning better defined through causality, since convex hull and other geometric definitions

in high dimensions tend to be meaningless for machine learning; (c) Out-of-distribution tasks are a mix of associational and counterfactual tasks (as the work of Bevilacqua et al. 2021 and Mouli et al. 2021 show). For out-of-distribution tasks we reviewed the concept of counterfactual invariant (graph) representations (Bevilacqua et al. 2021). Explaining why data augmentations for graphs are difficult to properly implement in practice (e.g., what it would look like if graph were larger without changing class label?). The talk ended stating that counterfactual-invariant representations are task-dependent and that, unlike associational graph tasks, there are provably no universal approximators for causal tasks.

3.12 Topology-Based Graph Learning

Bastian Rieck (Helmholtz Zentrum München, DE)

License  Creative Commons BY 4.0 International license
© Bastian Rieck

Joint work of Max Horn, Edward De Brouwer, Michael Moor, Yves Moreau, Bastian Rieck, Karsten Borgwardt
Main reference Max Horn, Edward De Brouwer, Michael Moor, Yves Moreau, Bastian Rieck, Karsten Borgwardt: “Topological Graph Neural Networks”, in Proc. of the International Conference on Learning Representations, 2022.
URL <https://openreview.net/forum?id=oxxUMeFwEHd>

Topological data analysis is starting to establish itself as a powerful and effective framework in machine learning, supporting the analysis of neural networks, but also driving the development of novel algorithms that incorporate topological characteristics. As a problem class, graph representation learning is of particular interest here, since graphs are inherently amenable to a topological description in terms of their connected components and cycles. This talk will provide an overview of how to address graph learning tasks using machine learning techniques, with a specific focus on how to make such techniques ‘topology-aware.’ We will discuss how to learn filtrations for graphs and how to incorporate topological information into modern graph neural networks, resulting in provably more expressive algorithms. This talk aims to be accessible to an audience of graph learning enthusiasts; prior knowledge of topological data analysis is helpful but not required.

3.13 Universal Graph Neural Networks via Random Data Augmentations Using Graph Isomorphism Tools

Pascal Schweitzer (TU Darmstadt, DE)

License  Creative Commons BY 4.0 International license
© Pascal Schweitzer

Joint work of Pascal Schweitzer, Markus Anders, Billy Joe Franks, Marius Kloft

Message-passing neural networks have provable limitations. Random data augmentations can be used to overcome these, resulting in provably universal graph neural networks. I will describe a solver from the realm of practical graph isomorphism testing that is based on so-called individualization-refinement techniques and uses random sampling. I will then describe how it can be employed to obtain efficient, scalable, universal graph neural networks.

3.14 Combining Representation Learning and Logical Rule Reasoning for Knowledge Graph Inference

Yizhou Sun (UCLA, US)

- License** © Creative Commons BY 4.0 International license
© Yizhou Sun
- Joint work of** Xuelu Chen, Ziniu Hu, Kewei Cheng, Ziqing Yang, Ming Zhang
- Main reference** Xuelu Chen, Ziniu Hu, Yizhou Sun: “Fuzzy Logic Based Logical Query Answering on Knowledge Graphs”, in Proc. of the Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelveth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 – March 1, 2022, pp. 3939–3948, AAAI Press, 2022.
URL <https://ojs.aaai.org/index.php/AAAI/article/view/20310>
- Main reference** Kewei Cheng, Ziqing Yang, Ming Zhang, Yizhou Sun: “UniKER: A Unified Framework for Combining Embedding and Definite Horn Rule Reasoning for Knowledge Graph Inference”, in Proc. of the 2021 Conference on Empirical Methods in Natural Language Processing, EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 7-11 November, 2021, pp. 9753–9771, Association for Computational Linguistics, 2021.
URL <https://doi.org/10.18653/v1/2021.emnlp-main.769>

Knowledge graph inference has been studied extensively due to its wide applications. It has been addressed by two lines of research, i.e., the more traditional logical rule reasoning and the more recent knowledge graph embedding (KGE). In this talk, we will introduce two recent developments in our group to combine these two worlds. First, we propose to leverage logical rules to bring in high-order dependency among entities and relations for KGE. By limiting the logical rules to be the definite Horn clauses, we are able to fully exploit the knowledge in logical rules and enable the mutual enhancement of logical rule-based reasoning and KGE in an extremely efficient way. Second, we propose to handle logical queries by representing fuzzy sets as specially designed vectors and retrieving answers via dense vector computation. In particular, we provide embedding-based logical operators that strictly follow the axioms required in fuzzy logic, which can be trained by self-supervised knowledge completion tasks. With additional query-answer pairs, the performance can be further enhanced. With these evidence, we believe combining logic with representation learning provides a promising direction for knowledge reasoning.

3.15 Graph Learning with 1D Convolutions on Random Walks

Jan Tönshoff (RWTH Aachen, DE)

- License** © Creative Commons BY 4.0 International license
© Jan Tönshoff
- Joint work of** Jan Tönshoff, Martin Ritzert, Hinrikus Wolf, Martin Grohe
- Main reference** Jan Tönshoff, Martin Ritzert, Hinrikus Wolf, Martin Grohe: “Graph Learning with 1D Convolutions on Random Walks”, CoRR, Vol. abs/2102.08786, 2021.
URL <https://arxiv.org/abs/2102.08786>

We propose CRaWl (CNNs for Random Walks), a novel neural network architecture for graph learning. It is based on processing sequences of small subgraphs induced by random walks with standard 1D CNNs. Thus, CRaWl is fundamentally different from typical message passing graph neural network architectures. It is inspired by techniques counting small subgraphs, such as the graphlet kernel and motif counting, and combines them with random walk based techniques in a highly efficient and scalable neural architecture. We demonstrate empirically that CRaWl matches or outperforms state-of-the-art GNN architectures across a multitude of benchmark datasets for graph learning.

3.16 Graph Neural Networks are Dynamic Programmers

Petar Velickovic (DeepMind – London, GB)

License © Creative Commons BY 4.0 International license
© Petar Velickovic

Joint work of Andrew Dudzik, Petar Veličković

Main reference Andrew Dudzik, Petar Veličković: “Graph Neural Networks are Dynamic Programmers”, CoRR, Vol. abs/2203.15544, 2022.

URL <https://doi.org/10.48550/arXiv.2203.15544>

Recent advances in neural algorithmic reasoning with graph neural networks (GNNs) are propped up by the notion of algorithmic alignment. Broadly, a neural network will be better at learning to execute a reasoning task (in terms of sample complexity) if its individual components align well with the target algorithm. Specifically, GNNs are claimed to align with dynamic programming (DP), a general problem-solving strategy which expresses many polynomial-time algorithms. However, has this alignment truly been demonstrated and theoretically quantified? Here we show, using methods from category theory and abstract algebra, that there exists an intricate connection between GNNs and DP, going well beyond the initial observations over individual algorithms such as Bellman-Ford. Exposing this connection, we easily verify several prior findings in the literature, and hope it will serve as a foundation for building stronger algorithmically aligned GNNs.

3.17 Infusing Structure and Knowledge into Biomedical AI

Marinka Zitnik (Harvard University – Boston, US)

License © Creative Commons BY 4.0 International license
© Marinka Zitnik

Joint work of Xiang Zhang, Kexin Huang, Michelle Li, Qianwen Wang, Joshua Pan, Theodoros Tsiligkaridis, Emily Alsentzer, Sam Finlayson, Matthew McDermott, Joe Loscalzo, Jure Leskovec, Laszlo Barabasi, Marinka Zitnik

Artificial intelligence has enabled scientific breakthroughs in diverse areas of biology and medicine. However, biomedical data present unique challenges, including limited annotations for supervised learning, the need to generalize to new scenarios not seen during training, and the need for trustworthy representations that lend themselves to actionable hypotheses in the laboratory. This talk describes our efforts to address these challenges by infusing structure and knowledge into biomedical AI. First, I outline subgraph neural networks that can disentangle distinct aspects of subgraph structure. I will then present a general-purpose approach for few-shot learning on graphs. At the core is the notion of local subgraphs that transfer knowledge from one task to another, even when only a handful of labeled examples are available. This principle is theoretically justified as we show that the evidence for predictions can be found in subgraphs surrounding the targets. Finally, to illustrate the benefits of modeling structure in non-graph datasets, I will introduce Raindrop, a graph neural network that embeds complex time series while also learning the dynamics of sensors purely from observational data. This research creates new avenues for accelerating drug discovery, fusing biomedical knowledge and patient data, and giving the right patient the right treatment at the right time to have effects that are consistent from person to person and with results in the laboratory.

Participants

- | | | |
|--|--|--|
| ■ Francesco Di Giovanni
Twitter – San Francisco, US | ■ Stephan Günnemann
TU München, DE | ■ Gaurav Rattan
RWTH Aachen, DE |
| ■ Federico Errica
NEC Laboratories Europe –
Heidelberg, DE | ■ Nils Kriege
Universität Wien, AT | ■ Bruno Ribeiro
Purdue University – West
Lafayette, US |
| ■ Fabrizio Frasca
Twitter – London & Imperial
College London, UK | ■ Gitta Kutyniok
LMU München, DE | ■ Bastian Rieck
Helmholtz Zentrum
München, DE |
| ■ Floris Geerts
University of Antwerp, BE | ■ Haggai Maron
NVIDIA – Tel Aviv, IL | ■ Pascal Schweitzer
TU Darmstadt, DE |
| ■ Martin Grohe
RWTH Aachen University, DE | ■ Christopher Morris
McGill University & MILA –
Montreal | ■ Jan Tönshoff
RWTH Aachen, DE |



Remote Participants

- | | | |
|---|---|---|
| ■ Pablo Barcelo
PUC – Santiago de Chile, CL | ■ Stefanie Jegelka
MIT – Cambridge, US | ■ Siamak Ravanbakhsh
McGill University –
Montréal, CA |
| ■ Aleksandar Bojchevski
CISPA – Saarbrücken, DE | ■ Elias Khalil
University of Toronto, CA | ■ Joshua Robinson
MIT – Cambridge, US |
| ■ Joan Bruna Estrach
New York University, US | ■ Benny Kimelfeld
Technion – Haifa, IL | ■ Yizhou Sun
UCLA, US |
| ■ Tina Eliassi-Rad
Northeastern University –
Boston, US | ■ Yaron Lipman
Weizmann Institute –
Rehovot, IL | ■ Petar Velickovic
DeepMind – London, GB |
| ■ Matthias Fey
TU Dortmund, DE | ■ Andreas Loukas
Prescient Design – Schlieren, CH | ■ Ulrike von Luxburg
Universität Tübingen, DE |
| ■ Barbara Hammer
Universität Bielefeld, DE | ■ Bryan Perozzi
Google – New York, US | ■ Marinka Zitnik
Harvard University – Boston, US |