DAGSTUHL
REPORTS

Volume 13, Issue 12, December 2023

ISSN 2192-5283

Published online and open access by
Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing, Saarbrücken/Wadern, Germany. Online available at https://www.dagstuhl.de/dagpub/2192-5283

Publication date
April, 2024

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at `https://dnb.d-nb.de`.

Digital Object Identifier: 10.4230/DagRep.13.12.i

Aims and Scope
The periodical *Dagstuhl Reports* documents the program and the results of Dagstuhl Seminars and Dagstuhl Perspectives Workshops.
In principal, for each Dagstuhl Seminar or Dagstuhl Perspectives Workshop a report is published that contains the following:

- an executive summary of the seminar program and the fundamental results,
- an overview of the talks given during the seminar (summarized as talk abstracts), and
- summaries from working groups (if applicable).

This basic framework can be extended by suitable contributions that are related to the program of the seminar, e. g. summaries from panel discussions or open problem sessions.

Contact
Schloss Dagstuhl – Leibniz-Zentrum für Informatik
Dagstuhl Reports, Editorial Office
Oktavie-Allee, 66687 Wadern, Germany
reports@dagstuhl.de
https://www.dagstuhl.de/dagrep

Report from Dagstuhl Seminar 23491

Scalable Graph Mining and Learning

Danai Koutra*[1], Henning Meyerhenke*[2], Ilya Safro*[3], and Fabian Brandt-Tumescheit†[4]

1 **University of Michigan – Ann Arbor, US & Amazon, US.** dkoutra@umich.edu
2 **Humboldt-Universität zu Berlin, DE.** meyerhenke@hu-berlin.de
3 **University of Delaware, Newark, US.** isafro@udel.edu
4 **Humboldt-Universität zu Berlin, DE.** fabian.brandt-tumescheit@hu-berlin.de

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 23491 "Scalable Graph Mining and Learning". The event brought together leading researchers and practitioners to discuss cutting-edge developments in graph machine learning, massive-scale graph analytics frameworks, and optimization techniques for graph processing. Besides the executive summary, the report contains abstracts of the 18 scientific talks presented, descriptions of three open problems, and preliminary results of three working groups formed during the seminar. In summary, the seminar successfully fostered discussions that bridged theoretical research and practical applications in scalable graph learning, mining, and analytics. Several potential outcomes include writing position and research papers as well as joint grant submissions.

Seminar December 03–08, 2023 – https://www.dagstuhl.de/23491
2012 ACM Subject Classification Theory of computation → Graph algorithms analysis; Computing methodologies → Machine learning algorithms; Computing methodologies → Parallel algorithms
Keywords and phrases Graph mining, Graph machine learning, (hyper)graph and network algorithms, high-performance computing for graphs, algorithm engineering for graphs
Digital Object Identifier 10.4230/DagRep.13.12.1

1 Executive Summary

Danai Koutra (University of Michigan – Ann Arbor, US & Amazon, US, dkoutra@umich.edu)
Henning Meyerhenke (Humboldt-Universität zu Berlin, US, meyerhenke@hu-berlin.de)
Ilya Safro (University of Delaware, Newark, US, isafro@udel.edu)

This Dagstuhl Seminar demonstrated a comprehensive exploration of the latest advancements and research directions in the field of scalable graph analytics and learning. The seminar featured a diverse array of talks that spanned foundational theory, innovative algorithms, and real-world applications. This event brought together leading researchers and practitioners to discuss cutting-edge developments in graph machine learning, massive-scale graph analytics frameworks, and optimization techniques for graph processing.

The seminar highlighted significant contributions to graph neural networks and representation learning, emphasizing such discussions as the problems related to the development of scalable algorithms, a proper benchmarking of the algorithms, several types of data reduction on graphs and various models including static, dynamic, and streaming graph data.

* Editor / Organizer
† Editorial Assistant / Collector

Scalable Graph Mining and Learning, *Dagstuhl Reports*, Vol. 13, Issue 12, pp. 1–23
Editors: Danai Koutra and Henning Meyerhenke and Ilya Safro
Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

On the algorithmic front, several participants presented frameworks and algorithms designed for efficient graph analytics at massive scales in regular and streaming contexts. The topics included such broad problems as graph sparsification, various versions of graph partitioning, and matchings in big graphs. They highlighted the ongoing efforts to address the scalability challenges inherent in processing large-scale graph data. Notably, some emphasis was on accelerating optimization on graphs by machine learning.

The seminar also served as a platform for discussing the practical implications and applications of graph mining and learning in various domains. Examples of talks and discussions included such topics as the role of knowledge graphs, causal discovery and inference from social networks, as well as industry level graph mining. All these topics illuminated the broad applicability of graph analytics in social science, industry, and beyond. Moreover, these case studies provided valuable insights into how theoretical advancements translate into applications.

One of the working groups explored using graph neural networks for smaller kernels with an example of the maximum weighted independent set problem, focusing on understanding GNNs' ability to learn existing reduction rules, discovering new rules through GNN insights, and potentially enhancing graph reduction beyond current capabilities. Another working group discussed the ways of formulating the FAIR (Findable, Accessible, Interoperable, and Reusable) principles in an algorithmic context and promoting the ways to generalize future algorithm developement to bridge the gap between specialized practical applications and fundamental algorithms. Another working group focused on causal representation learning, aiming to identify high-level causal variables from observational data within graphs. This area is important for improving machine learning models and understanding causal mechanisms in various networks like social and protein interaction networks. Despite the importance, there's limited research on applying causal representation learning to graph data, a gap that hinders our ability to discern causal relationships in complex systems. The group emphasized the need for advancements in graph representation learning to address this challenge effectively.

In summary, the "Scalable Graph Mining and Learning" Dagstuhl Seminar successfully fostered discussions that bridged theoretical research and practical applications in scalable graph learning, mining, and analytics. Several potential outcomes include writing position and research papers as well as joint grant submissions.

2 Table of Contents

3 Overview of Talks

3.1 Graph Machine Learning and Neural Network Research

George Karypis (University of Minnesota – Minneapolis, US, karypis@umn.edu)

This talk surveys recent accomplishments in graph machine learning and graph neural network research.

3.2 Arachne: An Open-Source Framework for Interactive Massive-Scale Graph Analytics

David A. Bader (NJIT – Newark, US, bader@njit.edu)

Joint work of David A. Bader, Oliver Alvarado Rodriguez, Zhihui Du, Joseph Patchett, Naren Khatwani, Fuhuan Li

A real-world challenge in data science is to develop interactive methods for quickly analyzing new and novel data sets that are potentially of massive scale. In this talk, Bader will discuss his development of graph algorithms in the context of Arkouda, an open-source NumPy-like replacement for interactive data science on tens of terabytes of data. Massive-scale analytics is an emerging field that integrates the power of high-performance computing and mathematical modeling to extract key insights and information from large-scale data sets. Productivity in massive-scale analytics entails quick interpretation of results through easy-to-use frameworks, while also adhering to design principles that combine high-performance computing and user-friendly simplicity. However, data scientists often encounter challenges, especially with graph analytics, which require the analysis of complex data from various domains, such as the cybersecurity, natural and social sciences. To address this issue, we introduce Arachne, an open-source framework that enhances accessibility and usability in massive-scale graph analytics. Arachne offers novel algorithms and implementations of graph kernels for efficient data analysis, such as connected components, breadth-first search, triangle counting, k-truss, among others. The high-performance algorithms are integrated into a back-end server written in HPE/Cray's Chapel language and can be accessed through a Python application programming interface (API). Arachne's back-end server is compatible with Linux supercomputers, is easy to set up, and can be utilized through either Python scripts or Jupyter notebooks, which makes it a desirable tool for data scientists who have access to high performance computers. In this talk, Bader presents an overview of the algorithms his research group has implemented into Arachne and, if applicable, the algorithmic innovations of each. Further, Bader will discuss improvements to our graph data structure to store extra information such as node labels, edge relationships, and node and edge properties. Arachne is built as an extension to the open-source Arkouda framework and allows for graphs to be generated from Arkouda dataframes.

The open-source code for Arachne can be found at `https://github.com/Bears-R-Us/arkouda-njit`. This is joint work with Oliver Alvarado Rodriguez, Zhihui Du, Joseph Patchett, Naren Khatwani, Fuhuan Li, Bader is supported in part by the National Science Foundation award CCF-2109988.

3.3 Data Reductions in Combinatorial Optimization

Ernestine Großmann (Universität Heidelberg, DE, e.grossmann@informatik.uni-heidelberg.de)

Combinatorial optimization problems play a pivotal role in various domains, challenging researchers to find optimal solutions within large solution spaces. One key strategy to tackle the complexity of these problems is data reduction, where instances are simplified to facilitate more efficient algorithms. In this talk, we delve into the world of combinatorial optimization, exploring classic data reduction techniques and their impact on problem-solving efficiency.

Additionally, we will explore different strategies of utilizing machine learning to reduce problem instances in the context of combinatorial optimization and ask how machine learning and exact data reductions can be combined more effectively. In particular, in future work we want to explore potential synergies between these approaches, considering the strengths of each and the challenges in their integration.

3.4 Targeted Branching for the Maximum Independent Set Problem Using Graph Neural Networks

Kenneth Langedal (University of Bergen, NO, kenneth.langedal@uib.no)

Identifying a maximum independent set is a fundamental NP-hard problem. This problem has several real-world applications and requires finding the largest possible set of vertices not adjacent to each other in an undirected graph. Over the past few years, branch-and-bound and branch-and-reduce algorithms have emerged as some of the most effective methods for solving the problem exactly. Specifically, the branch-and-reduce approach, which combines branch-and-bound principles with reduction rules, has proven particularly successful in tackling previously unmanageable real-world instances. This progress was largely made possible by the development of more effective reduction rules. Nevertheless, other key components that can impact the efficiency of these algorithms have not received the same level of interest. Among these is the branching strategy, which determines which vertex to branch on next. Until recently, the most widely used strategy was to choose the vertex of the highest degree. In this work, we present a graph neural network approach for selecting the next branching vertex. The intricate nature of current branch-and-bound solvers makes supervised and reinforcement learning difficult. Therefore, we use a population-based genetic algorithm to evolve the model's parameters instead. Our proposed approach results in a speedup on 73% of the benchmark instances with a median speedup of 24%.

3.5 What is the role of Knowledge Graphs in Graph Mining and Learning?

Davide Mottin (Aarhus University, DK, davide@cs.au.dk)

A knowledge graph (KG) represents a unique convergence of graph structures and natural language. On one side the graph structure is well defined by connection among entities. On the other hand, each entity and each relationship has a type and convey semantic meaning. Over time, research has emphasized different facets – while the machine learning community favors semantic and semi-structured information, the Data Mining and Database community views KGs as labeled graphs. Recently, the raise of Language Models and multimodal learning, seem to have found a new space for KGs to enhance the interpretability and improve the accuracy of models. This talk will survey some research on all sides of the coin and challenge the audience in finding a unification framework for KGs.

3.6 Causal Discovery from Social Networks

Elena Zheleva (University of Illinois – Chicago, US, ezheleva@uic.edu)

Social network data breaks a fundamental assumption of existing causal inference techniques, known as the Stable Unit Treatment Value Assumption (SUTVA), which states that the treatment of one unit cannot influence the outcome of other units. In real-world scenarios, it is common for related units to interact with each other which leads to interference (also known as spillover, or peer effects), in which the outcomes of units are interdependent. For example, the opinion of one person can influence the opinion of their friends, and the health status of one individual can impact the health status of others they interact with. In this talk, I will focus on recent work and problems related to discovering causal insights from social network data.

3.7 Topologies of Reasoning: Demystifying Chains, Trees, and Graphs of Thoughts

Maciej Besta (ETH Zürich, CH, maciej.besta@inf.ethz.ch)

The field of natural language processing has witnessed significant progress in recent years, with a notable focus on improving language models' performance through innovative prompting techniques. Among these, structure-enhanced prompting has emerged as a promising paradigm, with designs such as Chain-of-Thought (CoT) or Tree of Thoughts (ToT), in which the LLM reasoning is guided by a structure such as a tree. We refer to these structures as reasoning topologies, because their representation becomes to a degree spatial, as they are contained within the LLM context. In the first part of the talk, we overview this recent

field, focusing on fundamental classes of harnessed structures, the representations of these structures, algorithms executed with these structures, relationships to other parts of the generative AI pipeline such as knowledge bases or Graph Neural Networks, and others. Second, we introduce Graph of Thoughts (GoT): a framework that advances prompting capabilities in LLMs beyond those offered by CoT or ToT. The key idea and primary advantage of GoT is the ability to model the information generated by an LLM as an arbitrary graph, where units of information ("LLM thoughts") are vertices, and edges correspond to dependencies between these vertices. This approach enables combining arbitrary LLM thoughts into synergistic outcomes, distilling the essence of whole networks of thoughts, or enhancing thoughts using feedback loops. We illustrate that GoT offers advantages over state of the art on different tasks such as keyword counting while simultaneously reducing costs. We finalize with outlining research challenges in this fast-growing field.

3.8 Mining and Learning with Graphs at Google Scale

Bryan Perozzi (Google – New York, US, bperozzi@cs.stonybrook.edu)

For over a decade, the Graph Mining team in Google Research has been working on the mission to build the world's most scalable library for graph algorithms and analysis, and use it to improve Google products. Our goal is to make it easy for developers to use graph algorithms in their applications, and to provide them with the tools they need to build scalable and efficient graph processing systems. Our algorithms and systems are used in a wide array of Google products, such as Search, YouTube, AdWords, Play, Maps, and Social. The systems developed by our team have had a significant impact on the way that Google engineers build and deploy graph-based applications.

In this talk I will give an overview of the team, what it works on, and what challenges we view as fundamental for large scale graph systems.

3.9 Bridging Nodes and Edges: The Interdisciplinary Landscape of Graph Mining, Complex Networks, Social Network Analysis and Network Science

Frank Takes (Leiden University, NL, f.w.takes@liacs.leidenuniv.nl)

Various disciplines, from computer science to physics to sociology to economics, have each approached "graphs" or "networks" from a different perspective or interest. Each of these fields has established its specific publishing venues, conferences and societies. Overall, the talk aims to present what (future) challenges and opportunities this division over disciplines brings in terms of the exchange of methods and findings. In addition to highlighting how insights from one field have positively influenced the other, the talk also describes several instances of "multiple discoveries" of network phenomena.

3.10 Matchings in Big Graphs: Approximation and Streaming Algorithms

Alex Pothen (Purdue University – West Lafayette, US, apothen@purdue.edu)

Matchings in graphs are classical problems in combinatorial optimization and computer science, significant due to their theoretical importance and relevance to applications. Polynomial time algorithms for several variant matching problems with linear objective functions have been known for fifty years. However, these algorithms fail to compute matchings in big graphs with billions of edges. They are also not concurrent and thus practical parallel algorithms are not known.

This has led to work in the last twenty years on designing approximation algorithms for variant matching problems with near-linear time complexity in the size of the graphs. Approximation has thus become a useful paradigm for designing parallel matching algorithms and also streaming algorithms. In this talk I will report on an approach to fast approximation algorithms and streaming algorithms for the maximization version of edge-weighted matching. It can be extended to edge-weighted b-matching, and the maximum k-disjoint weighted matching problems.

Matching and related problems could be applied to graph mining, where the matching objective is a submodular function.

3.11 Graph sparsification

Richard Peng (University of Waterloo, CA, peng@uwaterloo.ca)

Graph sparsification is the approximation of dense graphs by sparse ones. Work over the past two decades have exhibited a wide range of objectives that are sparsifiable, such as cut structure, dense subgraphs, and eigenvectors, but also objectives that are not sparsifiable, such as distances and matchings. This talk will overview recent results on sparsification, and also discuss algorithmic ways of blurring the distinctions between sparse and dense problems.

3.12 Towards Foundation Models for Knowledge Graph Reasoning

Mikhail Galkin (Intel AI Lab – San Diego, US, mikhail.galkin@intel.com)

Foundation models in graph learning are hard to design due to the lack of common invariances that transfer across different structures and domains. In this talk, I will give an overview of ULTRA, our new approach for creating foundation models for knowledge graph reasoning that captures relation interactions and does not require any input node or edge features. Experimentally, a single pre-trained ULTRA in the zero-shot inference mode outperforms supervised SOTA models on 50+ diverse graphs and can generalize to any multi-relational graph.

3.13 Improving Generalizability in Link Prediction with Application in Drug Discovery

Ayan Chatterjee (Northeastern University – Boston, US, chatterjee.ay@northeastern.edu)

State-of-the-art link prediction models leveraging neighborhood topology for node representation learning underperform in data-scarce regimes involving low-degree nodes and entities with insufficient metadata. To overcome this challenge, we have proposed a non-end-to-end training approach leveraging unsupervised pre-training on a corpus significantly different from and larger than the training graph to enhance the generalizability of link prediction models. Additionally, we have introduced AI-Bind, a model that combines network science-inspired negative sampling and unsupervised pre-training on molecular representation. AI-Bind not only addresses topological shortcuts and the generalization shortcomings of existing models in predicting drug-target interactions but also excels in identifying binding locations on proteins for novel molecular structures. Furthermore, its integration with AutoDock Tools expedites the molecular docking process.

3.14 Partitioning communication streams into graph snapshots

Cynthia Phillips (Sandia National Labs – Albuquerque, US, caphill@sandia.gov)

Joint work of Jeremy D. Wendt, Richard Field, Cynthia Philips, Arvind Prasadan, Tegan Wilson, Sucheta Soundarajan, Sanjukta Bhowmick

We give a technique for partitioning streaming communication data into static graph snapshots. We use combinatorial statistical models to adaptively find when a snapshot is stable, while watching for significant data shifts – indicating a new snapshot should begin. If snapshots are not found carefully, they poorly represent the underlying data – and downstream graph analytics can fail.

We demonstrate the method on several real-world datasets and show its accuracy against known-answer synthetic datasets. Not surprisingly, snapshot properties change from those created with other methods. For example, our snapshots do not generally "densify" over time, contradicting previous influential results that used simpler partitioning methods.

3.15 Mining Small Patterns in Dynamic Graphs

Kathrin Hanauer (Universität Wien, AT, kathrin.hanauer@univie.ac.at)

Detecting, counting, or even enumerating the copies of a specific subgraph pattern contained in a graph is an important tool in the analysis of various kinds of networks, from social sciences to biology and chemistry, with numerous further applications, e.g. in machine learning on graphs, building graph generators, or graph partitioning. In many use case scenarios, the input graphs are almost naturally subject to change, i.e., edges may newly arrive or disappear, vertices may be introduced or abandoned, or weights may change, which makes the input graph *dynamic*.

In this talk, we will focus on maintaining the number of occurrences of small patterns on three or four vertices in a dynamic graph that undergoes an a priori unknown sequence of edge insertions and deletions. In the first part, we will see how this problem can be solved algorithmically in theory. Afterwards, we will also take a look at the practical side and discuss how the algorithms behave on real-world instances and how they can be further improved experimentally.

The talk concludes with some challenges and open questions in algorithm engineering for dynamic graph problems.

3.16 Architectures, Algorithms, and Applications for Graph Mining and Learning

Johannes Langguth (Simula Research Laboratory – Oslo, NO, langguth@simula.no)

A major recent development in computer hardware was the rise of dedicated accelerator hardware for machine learning applications such as the Graphcore IPUs and Cerebras WSE. These processors have evolved from the experimental state into market-ready products, and they have the potential to constitute the next major architectural shift after GPUs saw widespread adoption a decade ago.

A salient feature of these devices is the use of SRAM for memory, which offers very low latency and high bandwidth, making them attractive for a wide range of graph algorithms. On the other hand, the wide parallelism employed in these devices makes it difficult to use them efficiently for irregular computations.

In this talk we will present the new hardware and discuss the programming techniques that are required to unlock their potential. We present implementations of basic graph algorithms and show early results on the attainable performance, as well as comparisons to other architectures. We follow up by discussing the wider implications of the architecture for algorithm design and programming, along with the wider implications of adopting such hardware, and we discuss some recent graph mining applications.

3.17 Approximating the Diagonal of a Directed Graph Laplacians's Pseudoinverse

Fabian Brandt-Tumescheit (Humboldt Universität zu Berlin, DE, brandtfa@hu-berlin.de)

The ubiquity of massive graph data sets in numerous applications requires fast algorithms for extracting knowledge from these data. We are motivated here by three electrical measures for the analysis of large small-world graphs $G = (V, E)$ – i.e., graphs with diameter in $O(\log |V|)$, which are abundant in complex network analysis. From a computational point of view, the three measures have in common that their crucial component is the diagonal of the graph Laplacian's Pseudoinverse, $L^\dagger$. Computing $diag(L^\dagger)$ exactly by pseudoinversion, however, is as expensive as dense matrix multiplication – and the standard tools in practice even require

cubic time. Moreover, the Pseudoinverse requires quadratic space – hardly feasible for large graphs. Resorting to approximation by, e.g., using the Johnson-Lindenstrauss transform, requires the solution of $O(\log |V|/\epsilon^2)$ Laplacian linear systems to guarantee a relative error, which is still costly for large inputs.

In the ongoing project, we work on extending a previous approximation algorithm that requires the solution of only one Laplacian linear system and sampling of uniform spanning trees, which then are related to $diag(L^\dagger)$ via effective resistances. This previous work so far only supports undirected graphs, and the extension is towards directed graphs. For that, we introduce a new sampling scheme, which relies on the sampling of escape random walks. In theory, the converted algorithm obtains an $\pm\epsilon$-approximation with high probability in a time that is nearly linear in $|E|$ and quadratic in $1/\epsilon$. In addition, the formulation makes the approximation suitable for GPU-based implementation, leading to massive parallelism.

3.18 Contextualizing protein representations learned on protein networks and single-cell data

Michelle Li (Harvard University – Boston, US, michelleli@g.harvard.edu)

Understanding protein function and developing molecular therapies require deciphering the cell types in which proteins act as well as the interactions between proteins. However, modeling protein interactions across diverse biological contexts, such as tissues and cell types, remains a significant challenge for existing algorithms. We introduce PINNACLE, a flexible geometric deep learning approach that is trained on contextualized protein interaction networks to generate context-aware protein representations. Leveraging a human multi-organ single-cell transcriptomic atlas, PINNACLE provides 394,760 protein representations split across 156 cell type contexts from 24 tissues and organs. PINNACLE's contextualized representations of proteins reflect cellular and tissue organization and PINNACLE's tissue representations enable zero-shot retrieval of the tissue hierarchy. Pretrained PINNACLE's protein representations can be adapted for downstream tasks: to enhance 3D structure-based protein representations for important protein interactions in immuno-oncology (PD-1/PD-L1 and B7-1/CTLA-4) and to study the effects of drugs across cell type contexts. PINNACLE outperforms state-of-the-art, yet context-free, models in nominating therapeutic targets for rheumatoid arthritis and inflammatory bowel diseases, and can pinpoint cell type contexts that predict therapeutic targets better than context-free models. PINNACLE is a graph-based contextual AI model that dynamically adjusts its outputs based on biological contexts in which it operates.

4 Open Problems

4.1 Open Problems in Benchmarking Graph Learning via Synthetic Graph Generation

Bryan Perozzi (Google – New York, US, bperozzi@cs.stonybrook.edu)

Graph learning algorithms have attained state-of-the-art performance on many graph analysis tasks such as node classification, link prediction, and clustering. It has, however, become hard to track the field's burgeoning progress. One reason is due to the very small number of datasets used in practice to benchmark the performance of graph learning algorithms. This shockingly small sample size used for standard evaluations allows for only limited scientific insight into how graph learning models perform.

My group has been working on addressing this deficiency through the use of synthetic graph generation for evaluating graph learning models [1]. This approach consists of generating synthetic graphs, and contrasting the behavior of different graph learning algorithms in these controlled scenarios. Our prior work in this area (e.g. GraphWorld [2]) introduced the "GraphWorld paradigm"

1. Define graph metrics to quantify the "universe" of graphs generated by the framework
2. Define a Task (graph generator, feature generator, and prediction) tuple that fully defines a graph learning problem
3. Vary the graphs generated, in order to cover the "universe" as well as possible
4. Evaluate a wide range of graph learning models on the Task to characterize the response surface for each model

Naturally, there is much more to do in the space! For example, our initial investigation focused on community-aware generators (like the Stochastic Block Model) for this evaluation. However, this design decision limited the scope of graphs which could be generated. Further analysis [3], has shown that better control of the degree distribution of the generated graph creates situations which further differentiate graph learning models. Aside from obvious improvements to the components of the framework above, this line of research enables new problem formulations. For example, we also have examined how privacy-aware graph generative models can generate anonymized graph datasets on which a model has similar performance to the real datasets [4]. This is a new and interesting regime for both graph generation and synthetic evaluation. Finally, in very recent work [5], we've illustrated how this style of synthetic analysis can also be used to quantify the reasoning capabilities of Large Language Models (LLMs), providing an interesting connection to the emerging field of Artificial Intelligence.

References

1 Anton Tsitsulin and Benedek Rozemberczki and John Palowitch and Bryan Perozzi. Synthetic Graph Generation to Benchmark Graph Learning. GLB'21, Graph Learning Benchmarks (WWW Workshops).
2 Palowitch, John and Tsitsulin, Anton and Mayer, Brandon and Perozzi, Bryan. GraphWorld: Fake Graphs Bring Real Insights for GNNs. KDD '22.
3 Yasir, Mustafa and Palowitch, John and Tsitsulin, Anton and Tran-Thanh, Long and Perozzi, Bryan. Examining the Effects of Degree Distribution and Homophily in Graph Learning Models. GLB'23, Graph Learning Benchmarks (KDD Workshops).

4 Yoon, Minji and Wu, Yue and Palowitch, John and Perozzi, Bryan and Salakhutdinov, Russ. Graph Generative Model for Benchmarking Graph Neural Networks. International Conference on Machine Learning, 2023.
5 Fatemi, Bahare and Halcrow, Jonathan and Perozzi, Bryan. Talk like a graph: Encoding graphs for large language models. arXiv preprint, 2023.

4.2 Causal Representation Learning

Elena Zheleva (University of Illinois – Chicago, US, ezheleva@uic.edu)

Causal representation learning refers to the discovery of high-level causal variables from low-level observational data and is an important problem in machine learning and causal inference. In machine learning, causal representation learning can help with robustness and learning reusable mechanisms [1]. In causal inference, it can help with discovering the causal mechanisms that gave rise to the observed data and identifying causal effects of interest. While initial work has been done in causal representation learning for IID data (e.g., [1, 2]), little research has been done on causal representation learning for graphs [1]. The proposed discussion is to identify a path towards causal graph representation learning, including challenges and open problems.

References

1 Schölkopf et al. Towards Causal Representation Learning. Proceedings of the IEEE 2021.
2 Lachapelle et al. Disentanglement via Mechanism Sparsity Regularization: A New Principle for Nonlinear ICA. CLeaR 2022.
3 Li et al. Out-Of-Distribution Generalization on Graphs: A Survey. Arxiv 2022.

4.3 Graphs and Matrix accelerators – opportunities and challenges?

Flavio Vella (University of Trento, IT, flavio.vella@unitn.it)

With the growing computational needs requested by deep learning and AI, new libraries and architectures have steadily improved the efficiency of dense matrix multiplication in recent years. Dense-matrix units, such as the Tensor Processing Unit (TPU) (TPU) [1], TensorCore (TC) [2], or, more recently, Advanced Matrix Extensions (AMX) with TMUL by Intel [3] among many others, are hardware accelerators designed to handle large volumes of multiply-accumulate operations efficiently.

The use of such accelerators for other computation is becoming a well-established practice, not only for designing newer building blocks (e.g, *scan and prefix-sum* [4]) and mixed-precision linear solvers for HPC scientific computing [5] but also for various other applications [6]. In the graph analytics realm, there are also preliminary studies on SRAM-centric architecture [7].

Focusing on matrix-accelerators is natural to think about graphs algorithms in terms of matrix operations over semirings [8, 9, 10]. For example, given a fixed semiring $\langle D, \bigoplus, \bigotimes, e\rangle$ (where D is the element set, $\bigoplus, \bigotimes$ the ring operations, and e identity element) and the adjacency matrix A of a graph G, we can compute connectivity, number of paths, shortest paths and graph matching as simple tensor operations.[1]

GraphBLAS potentially provides a suitable abstraction for expressing graph algorithms in terms of *semirings*, however, there are several issues and challenges to face in mapping the algorithm to the matrix accelerator, which range from the sparsity of the graphs to the algorithms we can express to enabling the usage of these newer architectures.

Challenge 1: Sparsity as usual.

As we delve into the field of graph analytics, the first challenge we encounter is how to efficiently handle sparsity. This issue remains prevalent in parallel architectures. Graphs are notoriously sparse or even hyper-sparse while matrix accelerators support dense block-based data structure or structured sparsity. The practice of blocking the adjacency matrix often results in suboptimal performance compared to traditional sparse data structures like CSR.

To mitigate this, two approaches are considered. The first aims to leverage the *inherent density* found in certain graph components (often hidden by their labelling). The second approach involves artificially increasing the density within a block through more sophisticated data structures [11, 12].

Regarding the "inherent density", reordering algorithms and graph partitioning (see Table 1) can be used [13]. Trotter et al. [13] found that reordering based on graph partitioning provides better SpMV performance than the alternatives for the majority of matrices. However, both approaches are not specific to our purpose: traditional reordering methods (such as Saad's rendering method etc. [14]) typically are used for finding dense blocks over the diagonal which leads to good properties for linear solvers, and partitioning methods find a cut over the edges by assuming a fixed number of partitions (you want to minimize the cost among the partitions, that in our context correspond to blocks). Nevertheless, matrix accelerators need small (e.g., 4×4) but really dense blocks (order of thousands for medium matrices) to obtain a significant improvement.

Preliminary works [15, 16] introduce a new family of reordering algorithms based on clustering where each cluster corresponds to a distinct block. This algorithm requires $\mathcal{O}(n^2)$ comparisons among rows in the worst case. The main issue of such an approach is still on the final density that depends on the graphs, therefore it seems that, for certain graphs, we need to use ad-hoc data structures that do not allow us to exploit existing vendor-based routines for tensor accelerators.

P1.1. Find a linear algorithm that provides theoretical guarantees on the final density.
P1.2. Find a reordering that fits a structured sparsity schema.

(To solve P1.1 hashing-based algorithms such as Local-Sensitive Hashing (LSH) may be considered [16].)

[1] connectivity $\rightarrow \langle\{0,1\}, \vee, \wedge, 0\rangle$, number of paths $\rightarrow \langle \mathbf{N}, +, \cdot, 0\rangle$, shortest paths $\rightarrow \langle \mathbf{R} \cup \{+\infty\}, \min, +, +\infty\rangle$, matching $\rightarrow \langle \mathbf{R} \cup \{-\infty\}, \max, +, -\infty\rangle$.

Table 1 State-of-the-art techniques.

Method	Parametric	Process Rectangular	Only symmetric	Objective function (target)	Archs
Reverse Cuthill–McKee [17]	×	×	×	Matrix Bandwidth	CPU
Approximate minimum degree [18]	×	×	×	Row Degree	CPU
Nested dissection[19]	×	×	×	Cholesky Factorization	CPU
Graph partitioning[19]	✓	×	✓	Edge Cut & Load Balance	CPU
Hypergraph partitioning[20]	✓	✓	×	Edge Cut & Load Balance	CPU
Gray code ordering[21]	×	✓	✓	Branch Mispredictions & Data Locality	CPU
Saad's algorithm[22]	×	✓	✓	Row Similarity	CPU

Challenge 2: Graph Algorithms that can benefit from matrix accelerators.

The second issue is related to what algorithm can exploit matrix accelerators. The matrix-centric formulation of typical traversal-based algorithms such as the Breadth-first search (BFS) requires Sparse Matrix-Vector multiplication among A and the source vector B. On matrix accelerators with large capacity for storing the operands, this is not efficient. Other algorithms such as Triangle Counting or unweighted Single(Multi)-source shortest paths fit better (B is this case is the matrix where each column is a search from different sources). However, the mapping is not as straightforward as it seems to be. Firstly, the second operand is typically sparse both on triangle-based (especially for the implementation based on triangular matrices [23]). The consequence is if we apply the reordering to A (adjacency Matrix of G) we need to apply the same permutation to B to make a coherent product. Alternatively, B should be represented in its dense blocked format with the consequence of obtaining poor performance.

Moreover, weighted graphs introduce further difficulties. The issue is the mapping of the tropical semirings (e.g. the $\langle \mathbf{R} \cup \{+\infty\}, \min, +, +\infty \rangle$ and $\langle \mathbf{R} \cup \{-\infty\}, \max, +, -\infty \rangle$ required by the weighted graph computation) on the matrix accelerator that natively only supports a simple algebra $+/\cdot$ (i.e. $\langle \mathbf{N}, +, \cdot, 0 \rangle$). In the general case, which means we assume G with arbitrary weights, this could lead to an approximation version of the algorithm. However, within some constraints (e.g., where the weights are positive integers), it is possible to have the exact algorithm through a graph subdivision. Formalize this and find the constraints that have not been solved yet.

So, to summarize, the open problems are:

P2.1. How to design efficient primitives to solve unweighted graph problems by using matrix accelerators.

P2.2. Weighted graphs do not fit the algebraic properties exposed by matrix accelerators, the problem is how to "approximate" a semiring over the LA.

For P2.2. we need to explore theoretical foundations on the equivalences between graphs and matrix [24].

Challenge 3: Dynamic graphs and semirings from a GraphDB perspective.

Modern graph databases employ efficient graph engines to perform queries, with recent interest growing in matrix-centric graph engines. For example, RedisDB [25] adopts a GraphBLAS-based engine.

Databases typically serve multiple queries, including nested ones, and often require or rely on similar information. Currently, query languages translate operations into linear algebra expressions and optimize them using fundamental properties, such as the associative and distributive nature of matrix multiplication. Similarly, can we apply additional optimizations, akin to those long used in relational algebra? The second problem arises when the graph's structure changes. Optimizing in the context of dynamic operations, such as edge insertion/deletion or weight updates, remains a topic of exploration, as highlighted in various works by Giuseppe F. Italiano et al. To the best of our knowledge, there have been no significant enhancements in the processing of dynamic graphs in a truly algebraic form. We have recently started working on dynamic transitive closure and APSP over semirings.

References

1. Norman P Jouppi, Cliff Young, Nishant Patil, David Patterson, Gaurav Agrawal, Raminder Bajwa, Sarah Bates, Suresh Bhatia, Nan Boden, Al Borchers, et al. In-datacenter performance analysis of a tensor processing unit. In Proceedings of the 44th Annual International Symposium on Computer Architecture, pages 1–12. ACM, 2017.
2. Nvidia. Nvidia Tesla V100 GPU architecture. `https://www.nvidia.com/content/PDF/nvidia-ampere-ga-102-gpu-architecture-whitepaper-v2.pdf`. Accessed: 2018-04-13.
3. Mummidi, Chandra Sekhar and Ferreira, Victor C and Srinivasan, Sudarshan and Kundu, Sandip. Highly Efficient Self-Checking Matrix Multiplication on Tiled AMX Accelerators. ACM Transactions on Architecture and Code Optimization, 2023.
4. Dakkak, Abdul and Li, Cheng and Xiong, Jinjun and Gelado, Isaac and Hwu, Wenmei. Accelerating reduction and scan using tensor core units. Proceedings of the ACM International Conference on Supercomputing, 2019.
5. Ltaief, Hatem and Hong, Yuxi and Dabah, Adel and Alomairy, Rabab and Abdulah, Sameh and Goreczny, Chris and Gepner, Pawel and Ravasi, Matteo and Gratadour, Damien and Keyes, David. Steering Customized AI Architectures for HPC Scientific Applications. International Conference on High Performance Computing, 2023.
6. Burchard, Luk and Zhao, Max Xiaohang and Langguth, Johannes and Buluç, Aydın and Guidi, Giulia. Space Efficient Sequence Alignment for SRAM-Based Computing: X-Drop on the Graphcore IPU. Association for Computing Machinery, 2023.
7. Burchard, Luk and Cai, Xing and Langguth, Johannes. iPUG for Multiple Graphcore IPUs: Optimizing Performance and Scalability of Parallel Breadth-First Search. 2021 IEEE 28th International Conference on High Performance Computing, Data, and Analytics (HiPC).
8. Kepner, Jeremy and Gilbert, John. Graph Algorithms in the Language of Linear Algebra. Society for Industrial and Applied Mathematics, 2011.
9. Solomonik, Edgar and Besta, Maciej and Vella, Flavio and Hoefler, Torsten. Scaling betweenness centrality using communication-efficient sparse matrix multiplication. Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, 2017.
10. Yiran Chen and Yuan Xie and Linghao Song and Fan Chen and Tianqi Tang. A Survey of Accelerator Architectures for Deep Neural Networks. Engineering, 2020.
11. Lu, Yuechen and Liu, Weifeng. DASP: Specific Dense Matrix Multiply-Accumulate Units Accelerated General Sparse Matrix-Vector Multiplication. Association for Computing Machinery, 2023.
12. Zachariadis, Orestis and Satpute, Nitin and Gómez-Luna, Juan and Olivares, Joaquín. Accelerating sparse matrix–matrix multiplication with GPU Tensor Cores. Computers & Electrical Engineering, 2020.

13 Trotter, James D. and Ekmekçibaşı, Sinan and Langguth, Johannes and Torun, Tugba and Düzakın, Emre and Ilic, Aleksandar and Unat, Didem. Bringing Order to Sparsity: A Sparse Matrix Reordering Study on Multicore CPUs. Association for Computing Machinery, 2023.

14 Chen, Jie and Saad, Yousef. Finding dense subgraphs for sparse undirected, directed, and bipartite graphs. IEEE Trans. Knowl. Data Eng., submitted

15 Labini, Paolo Sylos and Bernaschi, Massimo and Nutt, Werner and Silvestri, Francesco and Vella, Flavio. Blocking Sparse Matrices to Leverage Dense-Specific Multiplication. 2022 IEEE/ACM Workshop on Irregular Applications: Architectures and Algorithms (IA3).

16 Jiang, Peng. Hong, Changwan. and Agrawal, Gagan. 2020. A novel data transformation and execution strategy for accelerating sparse matrix multiplication on GPUs. In Proceedings of the 25th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming (PPoPP '20).

17 Liu, Wai-Hung and Sherman, Andrew H. Comparative Analysis of the Cuthill–McKee and the Reverse Cuthill–McKee Ordering Algorithms for Sparse Matrices. SIAM Journal on Numerical Analysis, 1976.

18 Amestoy, Patrick R. and Davis, Timothy A. and Duff, Iain S. Algorithm 837: AMD, an Approximate Minimum Degree Ordering Algorithm. Association for Computing Machinery, 2004.

19 Karypis, George and Kumar, Vipin. A Fast and High Quality Multilevel Scheme for Partitioning Irregular Graphs. Society for Industrial and Applied Mathematics, 1998.

20 Catalyurek, U.V. and Aykanat, C. Hypergraph-partitioning-based decomposition for parallel sparse-matrix vector multiplication. IEEE Transactions on Parallel and Distributed Systems.

21 Zhao, Haoran and Xia, Tian and Li, Chenyang and Zhao, Wenzhe and Zheng, Nanning and Ren, Pengju. Exploring Better Speculation and Data Locality in Sparse Matrix-Vector Multiplication on Intel Xeon. 2020 IEEE 38th International Conference on Computer Design (ICCD).

22 Saad, Yousef. Finding Exact and Approximate Block Structures for ILU Preconditioning. SIAM Journal on Scientific Computing.

23 Azad, Ariful and Buluç, Aydin and Gilbert, John. Parallel triangle counting and enumeration using matrix algebra. Parallel and Distributed Processing Symposium Workshop (IPDPSW), 2015 IEEE International.

24 Williams, Virginia Vassilevska and Williams, R Ryan. Subcubic equivalences between path, matrix, and triangle problems. Journal of the ACM (JACM).

25 Cailliau, Pieter and Davis, Tim and Gadepally, Vijay and Kepner, Jeremy and Lipman, Roi and Lovitz, Jeffrey and Ouaknine, Keren. Redisgraph graphblas enabled graph database. 2019 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW).

5 Working Groups

5.1 Working Group on Causal Representation Learning

Elena Zheleva (University of Illinois – Chicago, US, ezheleva@uic.edu)
Michelle Li (Harvard University – Boston, US, michelleli@g.harvard.edu)

Causal representation learning refers to the discovery of high-level causal variables from low-level observational data and is an important problem in machine learning and causal inference. In machine learning, causal representation learning can help with robustness and learning reusable mechanisms [1]. In causal inference, it can help with discovering the causal mechanisms that gave rise to the observed data and with identifying causal effects of interest. While initial work has been done in causal representation learning for IID data (e.g., [1, 2]), little research has been done on causal representation learning for graphs [3, 4].

Many real-world systems can be represented as graphs in which nodes represent entities of interest, which can be of various types, and edges represent interactions between these entities. Some examples include social networks, protein interaction networks, molecular structure networks, and drug-protein-disease networks [4]. Graph representation learning has facilitated discoveries across real-world applications on structured data [4, 6]. However, despite the prevalence of causal relationships in real-world systems, existing graph representation learning algorithms are unable to extrapolate causal mechanisms.

References

1 Schölkopf et al. Towards Causal Representation Learning. Proceedings of the IEEE 2021.
2 Lachapelle et al. Disentanglement via Mechanism Sparsity Regularization: A New Principle for Nonlinear ICA. CLeaR 2022.
3 Li et al. Out-Of-Distribution Generalization on Graphs: A Survey. Arxiv 2022.
4 Li et al. Graph representation learning in biomedicine and healthcare. Nature Biomedical Engineering 2022. PDF.
5 Chalupka, Perona, Eberhardt. Multi-level cause-effect systems. AISTATS 2016.
6 Hamilton, William L., Rex Ying, and Jure Leskovec. Representation learning on graphs: Methods and applications. IEEE Data Engineering Bulletin 2017.

5.2 Working Group on Making Algorithms Research FAIR

David A. Bader (NJIT – Newark, US, bader@njit.edu)
Hannah Bast[2] *(University of Freiburg, DE, bast@informatik.uni-freiburg.de)*
Ümit V. Çatalyürek (Georgia Institute of Technology – Atlanta, US, umit@gatech.edu)
Joachim Giesen[3] *(Friedrich Schiller University Jena, DE, joachim.giesen@uni-jena.de)*
Kathrin Hanauer (Universität Wien, AT, kathrin.hanauer@univie.ac.at)
Ulrich Meyer (Goethe University – Frankfurt am Main, DE, umeyer@cs.uni-frankfurt.de)
Henning Meyerhenke (Humboldt-Universität zu Berlin, DE, meyerhenke@hu-berlin.de)
Cynthia A. Phillips (Sandia National Labs, Albuquerque, US, caphill@sandia.gov)
Ilya Safro (University of Delaware, Newark, US, isafro@udel.edu)

The need for efficient and tailor-made computer applications is dramatically increasing across all scientific areas. Even if research teams in application domains are acquiring their own computing expertise, they will still rely on powerful building blocks developed by experts. This is especially true for the fundamental algorithmic problems behind their applications – and is even more pronounced when these algorithms ought to exploit special hardware characteristics. In principle, algorithms research can provide these building blocks, but the relevant algorithms are often hard to find or understand or do not match the application needs well. Moreover, implementations are often missing or hard to employ. Users then usually resort to readily available or ad hoc solutions, often with markedly suboptimal performance.

5.2.1 Discussed Problems

Inspired by and based on ideas of an already ongoing initiative in Germany, this working group discussed the promotion of the goal to bridge this long-standing gap and make algorithms research more FAIR = Findable, Accessible, Interoperable, and Reusable; a set of principles originally defined for research data.

5.2.2 Possible Approaches

We argue that the natural way to achieve this is through problem abstractions with the following properties. Each abstraction should be lean, yet cover both a large body of algorithms research and a wide variety of applications. It should be comprehensible to a non-expert, remain stable over time, and come with an implementation.

5.2.3 Conclusions

It is intended to publish a position paper on the subject which presents success stories of the past and the main ideas how to achieve FAIRness in algorithmic research at large in the future.

[2] This co-author was not a seminar participant, but was invited to join the working group after the seminar.
[3] This co-author was not a seminar participant, but was invited to join the working group after the seminar.

5.3 GNNs for smaller kernels, finding the one rule to reduce them all

Kenneth Langedal (University of Bergen, NO, kenneth.langedal@uib.no)
Fredrik Manne (University of Bergen, NO, fredrik.manne@uib.no)
Ernestine Großmann (Universität Heidelberg, DE, e.grossmann@informatik.uni-heidelberg.de)
Christian Schulz (Universität Heidelberg, DE, christian.schulz@informatik.uni-heidelberg.de)
Matthias Schimek (KIT – Karlsruher Institut für Technologie, DE, schimek@kit.edu)
Fabian Brandt-Tumescheit (HU Berlin, DE, fabian.brandt-tumescheit@hu-berlin.de)

Finding the maximum weighted independent set (MWIS) in a graph is a fundamental NP-hard problem. Given an undirected graph $G = (V, E)$ with node weights $w : V \to \mathbb{R}^+$, MWIS is the problem of finding a set $\mathcal{I} \subseteq V$ of pair-wise nonadjacent vertices with the largest possible weight $\sum_{u \in \mathcal{I}} w(u)$. For both exact and heuristic solvers, reduction rules are crucial for tackling large instances. These reduction rules reduce the problem instance while ensuring that an optimal solution on the reduced instance can be lifted to an exact solution for the original graph.

We consider how graph neural networks (GNNs) can be used for solving the MWIS problem. In particular, we investigate how GNNs can be used to both design and apply reduction rules. Our discussions during the Dagstuhl Seminar 23491 mainly focused on three directions for further work.

What are the capabilities and limitations of current GNN architectures for learning existing reduction rules? Known reduction rules have a wide range of complexity [2]. The most straightforward rules compare weights between vertices or neighborhoods. More complicated rules search for patterns such as twins or dominated vertices. Eventually, some reduction rules even require solving the MWIS problem on a subgraph. As a first step, it makes sense to investigate what the existing GNN architectures are capable of in terms of learning known reduction rules. Furthermore, what features should be precomputed as initial input features?

Can we find new reduction rules with the help of a trained GNN model? Reduction rules should be provably exact, meaning we cannot rely on a machine-learning model to decide which vertices to reduce directly. However, a trained model could highlight interesting structures in otherwise irreducible graphs, where a human observer could then derive new reduction rules. We cannot train directly for this task since we do not know what structures could be interesting. Nevertheless, we could use a pre-trained model for predicting vertices that known reduction rules can reduce. To be clear, the outputs from the model should be negative for every vertex here since the graph is irreducible using known reductions. However, we can still rank the outputs and look into the vertices closest to a positive prediction. Alternatively, we could also train a model that predicts what vertices are part of some optimal solution. In this case, we would investigate the vertices that receive the most certain prediction, as in closest to positive or negative. Such models have been trained previously [3].

Can we train a GNN to reduce graphs further than what is currently possible? As mentioned earlier, some reduction rules must be applied cautiously due to their computational cost. Current implementations rely on simple heuristics to decide when to use these rules. Therefore, letting a GNN model decide instead could be a promising direction.

In addition to reduction rules that decrease the size of the problem, there are also transformations that increase the size of the graph [1]. The idea behind these rules is that alternating between phases where we reduce and expand the graph can eventually lead to smaller irreducible graphs. Combined with when to apply expensive rules, this becomes its own search problem. It also fits nicely into the reinforcement learning method with a clear set of actions, environment, and rewards. More specifically, an action would be a rule and area of the graph to apply it, and the goal would be to reduce the graph as much as possible. Alternatively, the goal could be to reduce it with the least amount of work.

References

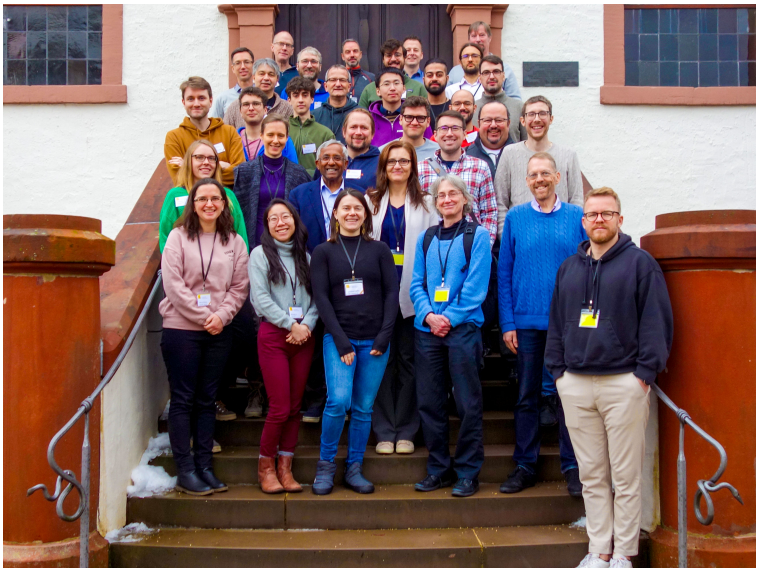
1 Alexander Gellner, Sebastian Lamm, Christian Schulz, Darren Strash, and Bogdán Zaválnij. "Boosting Data Reduction for the Maximum Weight Independent Set Problem Using Increasing Transformations". In: *2021 Proceedings of the Workshop on Algorithm Engineering and Experiments (ALENEX)*. SIAM. 2021, pp. 128–142.

2 Sebastian Lamm, Christian Schulz, Darren Strash, Robert Williger, and Huashuo Zhang. "Exactly Solving the Maximum Weight Independent Set Problem on Large Real-World Graphs". In: *2019 Proceedings of the Twenty-First Workshop on Algorithm Engineering and Experiments (ALENEX)*. SIAM. 2019, pp. 144–158.

3 Kenneth Langedal, Johannes Langguth, Fredrik Manne, and Daniel Thilo Schroeder. "Efficient Minimum Weight Vertex Cover Heuristics using Graph Neural Networks". In: *20th International Symposium on Experimental Algorithms (SEA 2022)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik. 2022.

Participants

- David A. Bader
NJIT – Newark, US
- Maciej Besta
ETH Zürich, CH
- Aleksandar Bojchevski
Universität Köln, DE
- Konstantinos Bougiatiotis
Demokritos – Athens, GR
- Fabian Brandt-Tumescheit
HU Berlin, DE
- Ümit V. Çatalyürek
Georgia Institute of Technology – Atlanta, US
- Ayan Chatterjee
Northeastern University – Boston, US
- Mikhail Galkin
Intel AI Lab – San Diego, US
- Ernestine Großmann
Universität Heidelberg, DE
- Kathrin Hanauer
Universität Wien, AT
- Mariya Ishteva
KU Leuven – Geel, BE
- George Karypis
University of Minnesota – Minneapolis, US
- Christine Klymko
LLNL – Livermore, US
- Danai Koutra
University of Michigan – Ann Arbor, US & Amazon, US
- Kenneth Langedal
University of Bergen, NO
- Johannes Langguth
Simula Research Laboratory – Oslo, NO
- Michelle Li
Harvard University – Boston, US
- Fredrik Manne
University of Bergen, NO
- Ulrich Carsten Meyer
Goethe University – Frankfurt am Main, DE
- Henning Meyerhenke
HU Berlin, DE
- Davide Mottin
Aarhus University, DK
- Richard Peng
University of Waterloo, CA
- Bryan Perozzi
Google – New York, US
- Cynthia A. Phillips
Sandia National Labs – Albuquerque, US
- Alex Pothen
Purdue University – West Lafayette, US
- Ilya Safro
University of Delaware – Newark, US
- Michael Schaub
RWTH Aachen, DE
- Matthias Schimek
KIT – Karlsruher Institut für Technologie, DE
- Christian Schulz
Universität Heidelberg, DE
- Christian Sohler
Universität Köln, DE
- Frank Takes
Leiden University, NL
- Elena Zheleva
University of Illinois – Chicago, US



Report from Dagstuhl Seminar 23492

Model Learning for Improved Trustworthiness in Autonomous Systems

Ellen Enkel*[1], Nils Jansen*[2], Mohammad Reza Mousavi*[3], and Kristin Yvonne Rozier*[4]

1 **Universität Duisburg-Essen, DE.** ellen.enkel@uni-due.de
2 **Ruhr-Universität Bochum, DE.** n.jansen@science.ru.nl
3 **King's College London, GB.** mohammad.mousavi@kcl.ac.uk
4 **Iowa State University – Ames, US.** kyrozier@iastate.edu

Abstract

The term of a model has different meanings in different communities, e.g., in psychology, computer science, and human-computer interaction, among others. Well-defined models and specifications are the bottleneck of rigorous analysis techniques in practice: they are often non-existent or outdated. The constructed models capture various aspects of system behaviours, which are inherently heterogeneous in nature in contemporary autonomous systems. Once these models are in place, they can be used to address further challenges concerning autonomous systems, such as validation and verification, transparency and trust, and explanation. The seminar brought together the best experts in a diverse range of disciplines such as artificial intelligence, formal methods, psychology, software and systems engineering, and human-computer interaction as well as others dealing with autonomous systems. The goal was to consolidate these understanding of models in order to address three grand challenges in trustworthiness and trust: (1) understanding and analysing the dynamic relationship of trustworthiness and trust, (2) the understanding of mental modes and trust, and (3) rigorous and model-based measures for trustworthiness and calibrated trust.

Seminar December 3–8, 2023 – https://www.dagstuhl.de/23492
2012 ACM Subject Classification General and reference → Reliability; General and reference → Validation; General and reference → Verification; Computing methodologies → Artificial intelligence; Applied computing → Psychology
Keywords and phrases artificial intelligence, automata learning, autonomous systems, cyber-physical systems, formal methods, machine learning, safety, safety-critical systems, self-adaptive systems, software evolution, technology acceptance, trust
Digital Object Identifier 10.4230/DagRep.13.12.24

1 Executive Summary

Ellen Enkel
Nils Jansen
Mohammad Reza Mousavi
Kristin Yvonne Rozier

This report documents the program and the outcomes of Dagstuhl Seminar 23492 "Model Learning for Improved Trustworthiness in Autonomous Systems". Autonomous systems increasingly enter our everyday life. Consequently, there is a strong need for safety, correctness,

* Editor / Organizer

Model Learning for Improved Trustworthiness in Autonomous Systems, *Dagstuhl Reports*, Vol. 13, Issue 12, pp. 24–47
Editors: Ellen Enkel, Nils Jansen, Mohammad Reza Mousavi, and Kristin Yvonne Rozier
Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

trust, and explainability. Well-defined models with clear semantics pose a convenient way to address these requirements. The area of model learning provides a structured way to obtain models from data. However, autonomous systems operate in the real world and pose challenges that go beyond the state-of-the-art in model learning. The technical challenges addressed in the seminar are system evolution and adaptations, learning heterogeneous models (addressing aspects such as discrete and continuous behaviours, stochastic, and epistemic uncertainty), and compositional learning. Our vision is that model learning is a key enabler solving the bottleneck of lack of specifications and models in various typical applications and hence, our seminar addressed fundamental challenges to enable impact in a number of application areas. In the seminar we brought together experts in (1) the domain of trust and technology acceptance, (2) the technical methods of model learning, and (3) the applications of model learning in robotics and autonomous systems. The first area includes domain experts in technology management, psychology, and trust; Technical methods include automata learning, synthesis of logical specifications, statistical model learning, machine learning, system identification, and process mining. Application experts include validation and verification, transparency and trust, and explainability, as well as experts in their application in robotics (planning, physical design and validation) and autonomous systems. With this seminar, we actively encouraged the interaction between experts and young researchers in the interdisciplinary areas of artificial intelligence, software engineering, autonomous systems, and human factors both from academia and industry. Content-wise, we emphasized the following directions: model learning techniques for AI-enabled autonomous systems: This involves recent techniques for learning models of evolving and variability-intensive systems; as well as application of model-learning to increase transparency and trust in robotics and autonomous system.

We discussed the following technical research questions during the seminar:

- How can we efficiently learn about system evolution and adaptation?
- How can we learn heterogeneous models, possibly by separating orthogonal concerns?
- How can we scale the mode learning?

Additionally, we discussed the following multi-disciplinary research questions:

- How can adaptive model learning be used to focus the validation and verification effort in evolving systems?
- How can learn model contribute to trust in autonomous systems?
- What types of models can be used to provide understandable explanations for AI-enabled and autonomous systems?

During the discussion four research questions and working groups emerged, that captured their discussion in scientific papers. The following is a short abstract of each paper that is currently in development.

Working group 1: Foundations of Learned Model Validation in the Age of AI

Models serve as the fundamental basis for the design, synthesis, verification, and implementation of software systems, yet before we can use the model for any of these, we must validate the model against the expectations on the system and/or against the real behavior of the system. In many development paradigms, an emerging trend is to move away from entirely human-designed models to models learned using automated techniques. We contribute a concrete roadmap for validating learned behavioral models comprised by a range of popular components. We pinpoint the current limits of model validation, provide insight into the reasons behind these limitations, and identify challenges that should serve as targets for

future research. By means of a running example of a cruise controller with different techniques for model learning, we show how guarantees derived from these techniques interplay with the validation challenges.

Working group 2: Mental Models, Human Models, Trust

Transposing the notion of interpersonal trust into the field of Computer Science, leads to the assumption that a high level of trust might also be pivotal in Human-Computer-Interaction (and even in interactions between two autonomous systems), since it enables the trustor (whether human or non-human) to make better predictions about the trustee. However, whereas humans possess an inherent "trust module," non-human agents lack such a component. Addressing this challenge, the present paper proposes a framework formalizing the trust relationship between a human and an autonomous system, aimed at creating more trustworthy Human-Computer-Interactions.

Working group 3: Research Agenda for Active Automata Learning

We conduct a survey of active automata learning methods, focusing on the different application scenarios (application domains, environment, and desirable guarantees) and the overarching goals that stem from them. We identify the challenges to achieve these goals. We organize a (short) bibliographic study highlighting the state-of-the-art and the technical challenges that are derived from the general goals and give some elements of answers related to these challenges.

Working group 4: Dynamic Interaction of Trust and Trustworthiness in AI-Enabled Systems

Trust is a user-centered notion, while trustworthiness pertains to the properties of the system. They dynamically influence each other, and interact with each other. We focus on AI-enabled systems, where establishing trustworthiness is challenging. In this paper we propose a framework for assessing trust and trustworthiness, and their alignment (calibration). We investigate factors that can influence them. We draw two case studies to illustrate our framework, and derive recommendations based on the insights we gain.

Besides interesting discussions, the four working groups focused on creating scientific articles, capturing their thoughts and insights. As a result, the four articles will be submitted to a special issue on Trust and Trustworthiness in Autonomous Systems International Journal of Software Tools for Technology Transfer (JSTTT) in 2024. Additionally, the organizers of this Dagstuhl Seminar will organize a track at the (A)ISoLa conference of October and November 2024 in Greece to deepen the discussion with the Dagstuhl attendees and with additional experts on this topic. Post-conference proceedings will document the insights gained.

2 Table of Contents

3 Overview of Talks

3.1 Learning-based Testing of Autonomous Systems

Bernhard K. Aichernig (TU Graz, AT)

Joint work of Martin Tappler, Edi Muškardin, Bernhard Aichernig, Bettina Könighofer
Main reference Martin Tappler, Edi Muškardin, Bernhard K. Aichernig, Bettina Könighofer: "Learning Environment Models with Continuous Stochastic Dynamics", CoRR, abs 2306.17204, 2023
URL https://doi.org/10.48550/arXiv.2306.17204

Learning-based testing combines model learning and model-based test-case generation in order to further automate the testing process. In our recent research we are aiming for testing autonomous systems, including autonomous cars (ADAS) and robots. In order to facilitate interaction with machine-learned systems, we have developed the AALpy library, an active automata learning library for Python. Next to standard automata learning algorithms, like L*, it supports the learning of stochastic models, like Markov decision processes or stochastic Mealy machines that are necessary to capture the policies of autonomous systems. Next to the algorithms implemented in AALpy, we also developed methods to learn timed models, like Timed Automata.

Our current research focuses on the scalability of automata learning to large state spaces in order to test autonomous systems. We aim to provide insights into the decisions faced by the agent by learning an automaton model of environmental behavior under the control of an agent. However, for most control problems, automata learning is not scalable enough to learn a useful model. In order to overcome this limitation, we compute an abstract state-space representation, by applying dimensionality reduction and clustering on the observed environmental state space. The stochastic transitions are learned via passive automata learning from observed interactions of the agent and the environment. In an iterative model-based reinforcement learning (RL) process, we sample additional trajectories to learn an accurate environment model in the form of a discrete-state Markov decision process (MDP). We applied our automata learning framework on popular RL benchmarking environments in the OpenAI Gym, including LunarLander, CartPole, Mountain Car, and Acrobot. Our results show that the learned models are so precise that they enable the computation of policies solving the respective control tasks.

3.2 Lifelong Learning of Reactive Systems

Bernhard Steffen (TU Dortmund, DE), Falk Howar (TU Dortmund, DE)

Joint work of Alexander Bainczyk, Bernhard Steffen, Falk Howar
Main reference Alexander Bainczyk, Bernhard Steffen, Falk Howar: "Lifelong Learning of Reactive Systems in Practice", in Proc. of the The Logic of Software. A Tasting Menu of Formal Methods – Essays Dedicated to Reiner Hähnle on the Occasion of His 60th Birthday, Lecture Notes in Computer Science, Vol. 13360, pp. 38–53, Springer, 2022.
URL https://doi.org//10.1007/978-3-031-08166-8_3

The talk presents our lifelong learning framework for continuous quality control that integrates automata learning, model checking, and monitoring into a six-phase continuous improvement cycle that is design to capture entire system life-cycles. Technical backbone of our framework is ALEX, our open source, web-based learning tool for defining adequate test blocks, as well

as for serving as test execution environment and as platfrom for learning Mealy machines. Key to the industrial success of our framework are a) the guarantee that the level of quality can only increase when using our framewok, b) the continuous improvement the originally customer-provided (regression) test suites, c) the maintenance of achieved quality levels even across system changes, and d) the visualization of system changes using automatically generated difference trees and difference automata. All this is illustrated using an adaptive cruise control system (ACC) that has been implemented in a one year students project.

3.3 Neuro-symbolic Model Learning for Controller Synthesis and Verification

Jyotirmoy Deshmukh (USC – Los Angeles, US)

Joint work of Navid Hashemi, Bardh Hoxha, Tomoya Yamaguchi, Danil V. Prokhorov, Georgios Fainekos, Jyotirmoy Deshmukh, Xin Qin, Lars Lindemann
Main reference Navid Hashemi, Bardh Hoxha, Tomoya Yamaguchi, Danil V. Prokhorov, Georgios Fainekos, Jyotirmoy Deshmukh: "A Neurosymbolic Approach to the Verification of Temporal Logic Properties of Learning-enabled Control Systems", in Proc. of the ACM/IEEE 14th International Conference on Cyber-Physical Systems, ICCPS 2023, (with CPS-IoT Week 2023), San Antonio, TX, USA, May 9-12, 2023, pp. 98–109, ACM, 2023.
URL https://doi.org//10.1145/3576841.3585928

Synthesizing control policies for high-dimensional, highly nonlinear/hybrid systems that guarantee satisfaction of safety and performance properties of the system is a significant challenge problem. In this talk, we will review some recent work on neuro-symbolic techniques, i.e., techniques that combine learned neural models of the system dynamics with symbolic techniques to synthesize control policies. We assume that safety/performance properties of the system are specified using bounded horizon Signal Temporal Logic (STL) formulas, and provide algorithms to automatically synthesize controllers that are guaranteed to satisfy these STL specifications. We also show how the learned neural surrogate models can be used for both deterministic and probabilistic verification of the closed-loop system. Our verification results on surrogate models can be mapped to probabilistic correctness guarantees on the original system, which allows a mechanism to address real-world issues such as the sim2real gap and distribution shifts between the design and deployment.

References

1 N. Hashemi, B. Hoxha, T. Yamaguchi, D. Prokhorov, G. E. Fainekos, J. V. Deshmukh. A Neurosymbolic Approach to the Verification of Temporal Logic Properties of Learning-enabled Control Systems. In Proceedings of the ACM/IEEE 14th International Conference on Cyber-Physical Systems, 2023.
2 N. Hashemi, X. Qin, J. V. Deshmukh, G. E. Fainekos, B. Hoxha, D. Prokhorov, and T. Yamaguchi, Risk-Awareness in Learning Neural Controllers for Temporal Logic Objectives. In Proc. of American Control Conference (ACC), 2023.
3 N. Hashemi, X. Qin, L. Lindemann, and J. V. Deshmukh, Data-Driven Reachability Analysis of Stochastic Dynamical Systems with Conformal Inference. To appear in the Proc. of IEEE Conf. on Decision and Control (CDC), 2023.

3.4 Learning Featured Transition Systems

Sophie Fortz (University of Namur, BE)

Main reference Sophie Fortz: "Variability-aware Behavioural Learning", SPLC (B) 2023: 11-15.
URL https://researchportal.unamur.be/en/studentTheses/learning-featured-transition-systems

I recently obtained my PhD from the University of Namur (Belgium). During these last four years, I worked on learning the behaviour of Variability-intensive Systems (VISs). In my thesis, entitled "Learning Featured Transition Systems", I combined classical learning algorithms (L^*) and deep learning techniques (recurrent neural networks) to learn behavioural models of VISs.

Variability-intensive Systems (VISs) are software-based systems whose characteristics and behaviour can be modified by the activation or deactivation of some options. Addressing variability proactively during software engineering (SE) activities means shifting from reasoning on individual systems to *reasoning on families of systems*. Adopting appropriate variability management techniques can yield important *economies of scale* and quality improvements. Conversely, variability can also be a curse, especially for Quality Assurance (QA), i.e. *verification and testing* of such systems, due to the combinatorial explosion of the number of software variants. Indeed, by combining only 33 Boolean options, we can define more variants of a system than the number of people on Earth. Verifying or testing each variant individually is thus impossible in most practical cases.

About a decade ago, *Featured Transition Systems (FTSs)* were introduced as a formalism to represent, and reason on, *the behaviour of VISs*. Instead of representing each variant by a (classical) transition system, an FTS bears annotations that relate transitions to options through *feature expressions*. FTSs thus make it possible to reason at the family level by modelling all the variants of a system in a single behavioural model. FTSs have been shown to significantly improve the possibilities and execution time of automated QA activities such as model-checking and model-based testing. They have also shown their usefulness to guide design exploration activities. Yet, as most model-based approaches, FTS modelling requires both *strong human expertise and significant effort* that would be unaffordable in many cases, in particular for large legacy systems with outdated specifications and/or systems that evolve continuously.

Therefore, in this thesis we aim *to automatically learn FTSs from existing artefacts*, to ease the burden of modelling FTS and support continuous QA activities. To answer this research challenge, we propose a two-phase approach. First, we rely on deep learning techniques to locate variability from execution traces. For this purpose, we implemented a tool called *VaryMinions*. Then, we use these annotated traces to learn an FTS. In this second part, we adapt the seminal L^* algorithm to *learn behavioural variability*. Both frameworks are open-source and we evaluated them separately on several datasets of different sizes and origins (e.g., software product lines and configurable business processes).

3.5 Trust in Conversational Agents

Effie Lai-Chong Law (Durham University, GB)

My long-term research focus centres on usability and elevating user experience (UX) methodologies, pivotal aspects within the realm of Human-Computer Interaction (HCI). My current research foci cover three areas: Conversational agents (CAs), Multisensory emotion recognition, and eXtended Reality (XR). Specifically, my investigations delve into the intricate dynamics of trust within customer service Conversational Agents (CAs). Through a series of empirical studies, I study the nuanced impact of various factors – ranging from CA conversational performance and humanlikeness to the task's nature – on the perceived trust in these AI-infused applications [1, 2]. Within an ongoing project, I am examining the role of inclusive design in augmenting the trust of older adults in online banking applications seamlessly integrated with ChatGPT. Simultaneously, I endeavour to examine the application of multisensory emotion recognition to provide support for the mental health and well-being of young adults. In the expansive domain of eXtended Reality (XR), which encompasses augmented, virtual, and mixed reality, I am exploring their educational potential across diverse sectors [3].

References

1 Law, E. L-C., Følstad, A., & Van As, N. (2022, October). Effects of Humanlikeness and Conversational Breakdown on Trust in Chatbots for Customer Service. In Nordic Human-Computer Interaction Conference (pp. 1-13).
2 Hobert, S., Følstad, A., & Law, E. L-C. (2023). Chatbots for active learning: A case of phishing email identification. International Journal of Human-Computer Studies, 179, 103108.
3 Thanyadit, S., Heintz, M., & Law, E. L-C. (2023, April). Tutor In-sight: Guiding and Visualizing Students' Attention with Mixed Reality Avatar Presentation Tools. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (pp. 1-20).

3.6 Reliable Learning for Safe Autonomy

Nicola Paoletti (King's College London, GB)

Main reference Francesca Cairoli, Nicola Paoletti, Luca Bortolussi: "Conformal Quantitative Predictive Monitoring of STL Requirements for Stochastic Processes", in Proc. of the 26th ACM International Conference on Hybrid Systems: Computation and Control, HSCC 2023, San Antonio, TX, USA, May 9-12, 2023, pp. 1:1–1:11, ACM, 2023.
URL https://doi.org//10.1145/3575870.3587113

With learning models increasingly being deployed in safety-critical systems, ensuring that predictions are reliable and providing guarantees for such models is paramount. In this talk, I will introduce conformal prediction, an assumption-free technique to obtain rigorous probabilistic guarantees on the predictions of any machine learning model. I will review recent work where my collaborators and I have applied this technique for data-driven verification of stochastic processes.

3.7 Robust and Reliable Reinforcement Learning

Marnix Suilen (Radboud University Nijmegen, NL)

My research focuses on planning and learning in probabilistic environments, typically modeled by Markov decision processes (MDPs) or partially observable MDPs (POMDPs). These MDPs are the standard models for decision-making problems where the outcome of an agent's choice is governed by a probability distribution. In the partially observable setting of POMDPs, the agent can only base its decisions on the given observations instead of the full state information. Recent results are in a model-based reinforcement learning (RL) setting, where the goal is to find an optimal decision-making policy by inferring a model from data and applying planning algorithms to find such a policy. Our results in this area can be split into two streams.

The first stream is the online RL setting, where we have sample access to the underlying model. By using robust dynamic programming as planning method, we can learn strategies that are probably approximately correct (PAC) optimal for the unknown true environment. Alternatively, linearly updating intervals (LUI) can find similarly performing strategies without any formal PAC guarantees. The advantage of LUI is that it is faster and requires less data than PAC learning, and it can easily adapt to new data that is inconsistent with previously encountered data, for example, due to a change in the probability distributions of the underlying true model. LUI was presented at NeurIPS 2022 [1].

The second stream is the offline RL setting, where only a fixed data set is given, and no further data can be collected. In the safe policy improvement (SPI) problem, we are given such a data set and the policy that collected it, known as the behavior policy. The goal is to, with a PAC guarantee, compute a new policy that outperforms this behavior policy with high confidence. We extended a standard method for SPI to the partially observable (POMDP) setting and devised alternative methods to compute the PAC guarantee such that it requires significantly less data. These results were published at AAAI 2023 [2] and IJCAI 2023 [3], respectively.

References

1 Marnix Suilen, Thiago Simão, David Parker, Nils Jansen. *Robust Anytime Learning of Markov Decision Processess*. NeurIPS 2022.
2 Thiago Simão, Marnix Suilen, Nils Jansen. *Safe Policy Improvement for POMDPs via Finite-State Controllers*. AAAI 2023.
3 Patrick Wienhöft, Marnix Suilen, Thiago Simão, Clemens Dubslaff, Christel Baier, Nils Jansen. *More for Less: Safe Policy Improvement with Stronger Performance Guarantees*.

3.8 Learning to avoid finding bugs

Thomas Arts (QuviQ AB – Gothenburg, SE)

Joint work of Thomas Arts, John Hughes, Ulf Norell, Nicholas Smallbone
Main reference John Hughes, Ulf Norell, Nicholas Smallbone, Thomas Arts: "Find more bugs with QuickCheck!", in Proc. of the 11th International Workshop on Automation of Software Test, AST@ICSE 2016, Austin, Texas, USA, May 14-15, 2016, pp. 71–77, ACM, 2016.
URL https://doi.org//10.1145/2896921.2896928

When testing industrial size software systems, one may be confronted with more than one fault. Clearly you want to report all faults found in a concise, understandable, to the software development team. The more faults they get at once, the better they can plan and fix errors in parallel.

With a QuickCheck state machine model, test cases are automatically generated and automatically shrunk to minimal test cases. The latter is important for size and understandability of the reported failure. However, there is an unexpected drawback when there are faults than can be demonstrated with a test sequence of only two or three commands. Such short sequences are very likely to appear as a subsequence of any larger generated test case. When shrinking the test case, it is very likely that just this sequence is reported, over and over again.

The reason to generate larger test cases is that in a realistic piece of software you may have 80 to 200 different API calls. To cover reasonably interesting sequences, one may need to cover quite a few to get to interesting states in the system. It turns out to be more effectiove to generate a long sequence and shrink when a failure is detected, then to explore subsets of the API in a more systematic way, by selecting a subset of the API and test with only that, after which one selects another subset... there are very many subsets.

Hughes et al [1] came up with some heuristics to avoid finding the same error twice by avoiding to generate it and avoiding to shrink to it. This resulted in being able to report more faults at once.

There is some clear open area for improvement in those heuristics as can be shown in a demo.

References

1 John Hughes, Ulf Norell, Nicholas Smallbone, and Thomas Arts. 2016. Find more bugs with QuickCheck! In Proceedings of the 11th International Workshop on Automation of Software Test (AST '16). Association for Computing Machinery, New York, NY, USA, 71–77.

3.9 Safe and Efficient Multi-agent Learning for Human–AI Collaboration

Mustafa Celikok

My work focuses on developing multi-agent learning methods that can learn to collaborate with both humans and AI agents with minimal prior coordination. The ideal outcome of this line of research is a learning agent who can be dropped in any team of agents which it has never seen before, and quickly learn how to collaborate with them towards a common goal. This capability is referred to as _ad hoc teamwork._ During my PhD, I have focused on

investigating model-based multi-agent reinforcement learning methods where the learning agent tries to infer an explicit model of its human partners. The model space in this work is derived from cognitive science research. My current work has three different paths.

(I) Bayesian Multi-agent Reinforcement Learning for Human–AI Collaboration
In this path, I address the following research questions.

RQ1. What assumptions can we make about the human partner in human–AI collaboration, and from where should they come?
RQ2. How can the AI take advantage of the assumptions and models of human behaviour?
RQ3. How can we guarantee safe and reasonable human–AI collaboration when our assumptions about the human partner are violated?

(II) Learning Dynamics of Continual Multi-agent Learning Agents for Long-Term Safety
There will come a point in time, where multi-agent learning systems are deployed in the real world to interact with humans and other learning systems, and learn continuously. Such systems might behave safely now, but since they are ever learning, their behaviour now does not guarantee future safety. They also create a complex system that is difficult to analyse. Dynamical systems theory allows us to study and _verify_ how the long-term learning dynamics of such agents will evolve. In our recent work, we stepped into this domain and studied what type of sets would learning agents in different ad hoc settings converge to.

(III) Sample Complexity and Non-asymptotic Results for Multi-agent Learning in Ad Hoc Collaboration
Even though there is a lot of work in terms of sample complexity when it comes to learning equilibria, most of these results either rely on all agents using similar learning algorithms, or that they are internally coupled. By relaxing these assumptions, I aim to complement the asymptotic results of Path II with finite-time results.

3.10 Chances and Challenges of AI-enhanced Automation for Automotive Health

Monique Dittrich (CARIAD – Berlin, DE)

Main reference Addam Mustapha, David Matusiewicz: "Automotive Health-A Systematic Overview of opportunities & boundaries" in Proceedings of the 12th edition of Numerical Analysis and Optimization Days.
URL https://hal.science/hal-02891566/document
Main reference Monique Dittrich: "Persuasive Technology to Mitigate Aggressive Driving : A Human-centered Design Approach", 2021.
URL https://doi.org//10.25972/OPUS-23022
Main reference Monique Dittrich: "The Car as a Personal Space to Improve your Health", in Proc. of the 18th International Conference on Persuasive Technology, Adjunct Proceedings co-located with PERSUASIVE 2023, Eindhoven University of Technology, Eindhoven, The Netherlands, April 19th – 21st, 2023, CEUR Workshop Proceedings, Vol. 3474, CEUR-WS.org, 2023.
URL https://ceur-ws.org/Vol-3474/paper23.pdf

Holding a PhD in Human-Computer-Interaction (Dittrich, 2020), I have been working in automotive research and development for around a decade, devoting myself to the human-centered design of Human-Machine-Interfaces of all kinds. By this, I dealt with overarching topics such as user experience and usability, emotion recognition, or behavior change.

Currently, my research concentrates on in-vehicle applications that are intended to improve the health and well-being of people the car. This area of research is scientifically manifested under the term "Automotive Health", whereby health (in the broader sense) is viewed through the lens of the automobile. Automotive Health deals with the questions of how health data can be sensed and interpreted in the vehicle environment, and how this information can be used to enable (or enrich) measures taken to maintain, restore, or strengthen occupants' well-being, (medical) health, and safety (Mustapha, & Matusiewicz, 2018). With a special focus on automatization through the use of Artificial Intelligence (AI) in this context, my research focusses on the following topics:

Data-driven decision making in the context of Automotive Health. Addressing the first pillar of Automotive Health, the sensing and interpretation of health data, inspiration can be found in the area of wearable technology. Wearables capture a wide range of vital signs (e.g., heart rate, respiratory rate, blood oxygen), and combine them into an meaningful information (e.g., stress, mood, sleeping quality), giving the user a quick and understandable indication of her state of health. For example, Garmin calculates the users performance condition based on her pace, heart rate, heart rate viability (HRV), and the maximum amount of oxygen that the body can take in (V02 max). However, compared to "lifestyle" use cases, such as sports or sleep, Automotive Health also addresses factors that are safety and medically relevant, such as fatigue, distraction, or sudden medical emergencies, and therefore require a higher level of accuracy and confidence. In this regard, the question of trustworthiness is primarily technical in nature. While the sensing and interpretation of health data is well established in the area of wearable technology and medicine, the topic is still in its infancy in the automotive domain.

Generative AI for mental health inventions. Generative AI, i.e., AI that generates texts, images, or other media content using generative models, existing knowledge, and prompts, is on everyone's lips. In the context of Automotive Health, I am primarily concerned with the generation of health-related content, more precise audio-guided mindfulness exercises (Dittrich, 2023). Mindfulness can be described as the process of bringing greater awareness to the present moment. Most practices to achieve this state are similar in their basic procedure, combining variations of breathing techniques, exercises for the awareness of body, mind, sensations and emotions, or mental images and thoughts. Mindfulness exercises can be also done while driving a car, since they are supposed to increase attention, lower stress, and reduce fatigue in order to promote driving safety. Generative AI would allow to adapt the exercise to the context in real time, i.e. the traffic situation or the person's current state of health, and to create all time new experiences for the user without the need to produce content through human speakers. However, in the context health one of the biggest problem is the generation of harmful content and the question of responsibility, when mental health deteriorates due to the content. Moreover, also trustworthiness plays an important role. When it comes to mental health practices, trust between the client and the counselors (even if embodied through audio content) is essential for the effectiveness of the intervention. So far, little is known about whether AI-generated content can overcome this hurdle.

3.11 Trust in and acceptance of technology

Ellen Enkel (Universität Duisburg-Essen, DE)

Joint work of Ellen Enkel, Magnus Liebherr, Monika Hengstler, Sabrina Dueli
Main reference Ellen Enkel: "To get consumers to trust AI, show them its benefits", Harvard Business Review Blog, 17 April 2017.
URL https://hbr.org/2017/04/to-get-consumers-to-trust-ai-show-them-its-benefits
Main reference Monika Hengstler, Ellen Enkel, Selina Duelli: "Applied artificial intelligence and trust–The case of autonomous vehicles and medical assistance devices", Technological Forecasting and Social Change, Vol. 105, pp. 105–120, 2016.
URL https://doi.org//10.1016/j.techfore.2015.12.014

Automation with inherent artificial intelligence (AI) is increasingly emerging in diverse applications, for instance, autonomous vehicles and medical assistance devices. However, despite their growing use, there is still noticeable skepticism in society regarding these applications. Drawing an analogy from human social interaction, the concept of trust provides a valid foundation for describing the relationship between humans and automation. Accordingly, we explore how firms systematically foster trust regarding applied AI. In the paper presented, the empirical analysis using nine case studies in the transportation and medical technology industries, illustrates the dichotomous constitution of trust in applied AI. Concretely, it emphasizes the symbiosis of trust in the technology as well as in the innovating firm and its communication about the technology. In doing so, it provides tangible approaches to increase trust in the technology and illustrate the necessity of a democratic development process for applied AI and provides a basis for our joined paper at this Dagstuhl Seminar.

3.12 Extracting Behavioral Models of Users/Applications by Combining Automata and Machine Learning

Fatemeh Ghassemi (University of Tehran, IR)

Joint work of Fatemeh Ghassemi, Zeynab Sabahi-Kaviani
Main reference Zeynab Sabahi-Kaviani, Fatemeh Ghassemi: "An Enhanced Encrypted Traffic Classifier via Combination of Deep Learning and Automata Learning", 2023.
URL https://doi.org//10.21203/rs.3.rs-3290610/v1
Main reference Fatemeh Marzani, Fatemeh Ghassemi, Zeynab Sabahi-Kaviani, Thijs van Ede, Maarten van Steen: "Mobile App Fingerprinting through Automata Learning and Machine Learning", in Proc. of the IFIP Networking Conference, IFIP Networking 2023, Barcelona, Spain, June 12-15, 2023, pp. 1–9, IEEE, 2023.
URL https://doi.org//10.23919/IFIPNETWORKING57963.2023.10186420
Main reference Zeynab Sabahi-Kaviani, Fatemeh Ghassemi, Zahra Alimadadi: "Combining Machine and Automata Learning for Network Traffic Classification", in Proc. of the Topics in Theoretical Computer Science – Third IFIP WG 1.8 International Conference, TTCS 2020, Tehran, Iran, July 1-2, 2020, Proceedings, Lecture Notes in Computer Science, Vol. 12281, pp. 17–31, Springer, 2020.
URL https://doi.org//10.1007/978-3-030-57852-7_2

My research focuses on the verification of real-world systems using model-based techniques such as model checking. I am concerned about approaches to make this technique feasible through reduction techniques or runtime monitoring. In recent years, I became interested in using the automata-based models as the behavioral fingerprint of applications for classification by combining automata learning and machine deep learning.

Machine/deep learning approaches are the main used technique to extract the fingerprint of applications from a set of sample data. The false positive rates of these methods are not negligible if they do not consider the temporal relations among events although they

are fast to generate and predict. On the other hand, learned models through automata learning are very precise with high true positives but too sensitive to the order of the events which leads to overfitting and not tolerable to noise and events loss in the distributed setting. So, the false negative rates of these methods are also high while they are not fast to either generate or use. By combining these approaches, we can take advantage of both; we use the FSM-learned models of applications to train our machine learning classifier.

We are using our approach to extract the behavioral models of users in the social network domain. We can interpret the events generated by a group of users with similar age, interest, and gender, ... identified as a specific user cluster in a social network, and generate a predictive behavioral model for those users. The following are possible directions for my research:

- Extending such combination for learned register automata
- Deriving the behavioral model of applications run within the cloud for predictive resource management
- Extending and using such behavioral models as digital twins

3.13 Automation challenges to human perception, emotion, and well-being

Heiko Hecht (Universität Mainz, DE)

Joint work of Heiko Hecht, Carina Röckel, Elisabeth Wögerbauer, Robin Welsch, Marlene Wessels, John Stins
Main reference Carina Röckel, Heiko Hecht: "Regular looks out the window do not maintain situation awareness in highly automated driving", Transportation Research Part F: Traffic Psychology and Behaviour, Vol. 98, pp. 368–381, 2023.
URL https://doi.org//10.1016/j.trf.2023.09.015

My field is human-machine interaction in the broadest sense. I am a cognitive experimental psychologist by training, with current specialization in the field of human factors and engineering psychology. In this context, the following current research might be of interest.

(1) Cyborg perception: On the road to fully autonomous vehicles, driver assistance systems enter a new phase in which human perception can be augmented. We explore how camera-monitor systems can replace traditional mirrors and how they can be used to compensate perceptual deficiencies by enhancing the digital mirror image with traffic-relevant information. Here the reduction of cognitive load is essential.
(2) Situation awareness maintenance: As more and more daily action become partially automated, the user is likely to become engrossed in secondary tasks, which makes it difficult if not impossible to switch back to manual control in critical cases where the automation fails. We explore how situation awareness can be best maintained or regained.
(3) Social robots: We have investigated how people regulate interpersonal distance among each other. The needs for trust, personal space, feelings of crowding, etc. likewise apply to robots. We explore how they may change as a function of robot appearance, situational factors, and habituation.
(4) Cybersickness: Car sickness is an unresolved problem in autonomous driving. We investigate the causes and potential remedies of car sickness and the related phenomena of simulator sickness, sea sickness, etc. Among others we found that even with see-through displays, moving virtual scenes can induces strong cybersickness.

References

1 Wögerbauer, E. M., Hecht, H., Wessels, M. (2023). Camera–monitor systems as an opportunity to compensate for perceptual errors in time-to-contact estimations. Vision, 7, 65. doi.org/10.3390/ vision7040065
2 Röckel, C., & Hecht, H. (2023). Regular looks out the window do not maintain situation awareness in highly automated driving. Transportation Research Part F: Traffic Psychology and Behaviour, 98, 368–381. doi.org/10.1016/j.trf.2023.09.015
3 Welsch, R., Hecht, H., & Stins, J. (2023). Task-relevant social cues affect whole-body approach-avoidance behavior. Scientific Reports 13:8568. doi.org/10.1038/s41598-023-35033-7
4 Kaufeld, M., Mundt, M., Forst, S., & Hecht, H. (2022). Optical see-through augmented reality can induce severe motion sickness. Displays, 74, doi.org/10.1016/j.displa.2022.102283

3.14 Learning and Testing of Real Systems

Leo Henry (University College London, GB)

After a PhD in the SUMO team in Rennes (France) I have been working as a Research Fellow at UCL.

I am mainly interested by how to use complex models to advance the verification (and general understanding) of real systems, which includes tackling questions relating to the availability of models, the complexity to interact with a real system or to control it.

My research stands on two legs: (1) verification of complex systems, with an emphasis on timed systems and timed automata stemming from my PhD. My main work in this area concerns the use of advanced game-strategies to create test case for systems modelled as input-out timed automata; (2) active learning, which I have mostly focused on at UCL. I have worked on practical questions, such as the handling of noise and changes to the target system during learning, or learning of a network of synchronizing models.

Ultimately, I would love to see off-the-shelf tools for learning, testing and verifying systems that developers and engineers could use without a great investment in theoretical matters.

3.15 Enhancing Human Interaction with Automated Technologies: The Role of Adaptability, Trust, and Cognition

Magnus Liebherr (Universität Duisburg-Essen, DE)

I hold a Ph.D. in psychology and currently lead the 'Human Factors' working group at the Chair of Mechatronics at the University of Duisburg-Essen. This position reflects my deep interest in exploring the impact of human factors within the realms of digital technologies, highly automated systems, and their interaction with our ever-evolving environment. Within the scope of my work, I concentrate on three fundamental areas: adaptability, trust, and cognition. Additionally, I delve into associated constructs such as mental workload. In the pursuit of my interdisciplinary research initiatives, I employ these themes across a diverse spectrum of technologies. This encompasses automation within the automotive and shipping sectors, collaborative robot systems, electromobility, m-health, and digital media.

Adapting to new technologies empowers individuals to effectively navigate changing environments or conditions, resulting in heightened performance, reduced mental workload, as well as increased comfort. Drawing from evolutionary biology, we distinguish between three facets of adaptability: Trait Adaptability, which represents a stable characteristic; State Adaptability, signifying the extent of immediate adjustment; and Adaptation, encapsulating the ongoing process of adaptation. Our previous research indicates that individuals' trait adaptability significantly influences the level of trust placed in a technology. Furthermore, we identified significant differences in state adaptability between subjectively perceived and objectively measured. Intriguingly, neither age nor prior experience with the technology accounted for this discrepancy. However, individuals exhibiting limited adaptation to a simulated environment spent considerably less time in the simulator, primarily due to the onset of simulator sickness.

Trust in technology encompasses individuals' confidence, belief, and reliance on the integrity, reliability, and security of a technological system. Findings on the interaction between humans and automation highlight the dependency of trust on perceived control over automated operations. Our contention is that heightened perceived control fosters trust in technology by diminishing users' uncertainties about transaction outcomes. Additionally, our findings reveal that familiarity and experience with automated technologies play a pivotal role in increasing trust.

The term cognition basically means the use of the brain but includes various complex activities. Within my work I focus on working memory, cognitive flexibility, and attentional processes. In the use of new technologies, working memory is crucial for grasping and remembering system features, navigating complex interfaces, and executing multitasking operations. Its role in problem-solving, decision-making, and managing cognitive load is essential for users to effectively engage with and adapt to evolving technological interfaces and functionalities. Cognitive flexibility enables rapid adaptation to new technologies and therefore improved performances, decreased error rate, and less mental workload. In terms of attentional processes, our findings highlight the greater impact of environmental complexity on attention during cognitive task performance compared to the difficulty of simultaneously executed motor tasks.

References

1 Vasile, L., Seitz, B., Staab, V., Liebherr, M., Däsch, C., & Schramm, D. (2023). Influences of Personal Driving Styles and Experienced System Characteristics on Driving Style Preferences in Automated Driving. Applied Sciences, 13(15), 8855.

2 Liebherr, M., Kohler, M., Zerr, M., Brand, M., & Antons, S. (2022). Effects of digital media use on attention subdomains in children aged 6 to 10 years. Children, 9(9), 1393.

3 Sauce, B., Liebherr, M., Judd, N., & Klingberg, T. (2022). The impact of digital media on children's intelligence while controlling for genetic differences in cognition and socioeconomic background. Scientific Reports, 12(1), 1-14.

4 Liebherr, M.*, Corcoran, A.W.*, Alday, P.M., Coussens, S., Bellan, V., Howlett, C.A., Immink, M.A., Kohler, M., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2021). EEG and behavioral correlates of attentional processing while walking and navigating naturalistic environments. Scientific Reports, 11, 22325. *Both authors contributed equally to this work.

5 Liebherr, M., Müller, M. S., Schweig, S., Maas, N., Schramm, D., & Brand, M. (2021). Stress and simulated environments – Insights from physiological marker. Frontiers in Virtual Reality, 2(18), 1-10.

6 Liebherr, M., Schweig, S., Brandtner, A., Averbeck, H., Maas, N., Schramm, D., & Brand, M. (2020). When virtuality becomes real: Relevance of mental abilities and age in system adaptation and the occurrence of simulator sickness. Ergonomics, 63(10), 1271-1280.

7 Brandtner, A., Liebherr, M., Schweig, S., Maas, N., Schramm, D., & Brand, M. (2019). Subjectively estimated vs. objectively measured adaptation to driving simulators – Effects of age, driving experience, and previous simulator adaptation. Transportation Research Part F: Psychology and Behaviour, 64, 440-446.

3.16 Compositional Learning of Synchronous Systems

Thomas Neele (TU Eindhoven, NL)

Joint work of Thomas Neele, Matteo Sammartino
Main reference Thomas Neele, Matteo Sammartino: "Compositional Automata Learning of Synchronous Systems", in Proc. of the Fundamental Approaches to Software Engineering – 26th International Conference, FASE 2023, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2023, Paris, France, April 22-27, 2023, Proceedings, Lecture Notes in Computer Science, Vol. 13991, pp. 47–66, Springer, 2023.
URL https://doi.org//10.1007/978-3-031-30826-0_3

My main research in the field of automata learning is efficient learning of concurrent systems, in the shape of synchronously communicating automata [1]. The large amount of interleaving behaviour of these automata can cause significant scalability issues when applying traditional automata learning techniques in a monolithic fashion. Our approach is to instantiate one learner per automaton inside the composite SUL. These learners are able to independently learn significant parts of their respective automata, and only synchronise when proposing a hypothesis (which is the parallel composition of their individual hyptheses).

Furthermore, I have an extensive background in the areas of model checking, parity games, bisimulation and partial-order reduction (see e.g. [2, 3, 4]).

References

1 Thomas Neele and Matteo Sammartino. *Compositional Automata Learning of Synchronous Systems*. In *FASE 2023*, volume 13991, pages 47–66. 2023. doi:10.1007/978-3-031-30826-0_3.

2 Thomas Neele, Antti Valmari, Wieger Wesselink, and Tim A. C. Willemse. *Partial-Order Reduction for Parity Games and Parameterised Boolean Equation Systems. Software Tools for Technology Transfer*, 24:735–756, 2022. doi:10.1007/s10009-022-00672-0.

3 Thomas Neele, Antti Valmari, and Tim A. C. Willemse. *A Detailed Account of The Inconsistent Labelling Problem of Stutter-Preserving Partial-Order Reduction. Logical Methods in Computer Science*, 17:8:1–8:27, 2021. doi:10.46298/lmcs-17(3:8)2021.

4 Thomas Neele, Tim A. C. Willemse, and Jan Friso Groote. *Finding Compact Proofs for Infinite-Data Parameterised Boolean Equation Systems. Science of Computer Programming, FACS 2018 special issue*, 188:102389, 2020. doi:10.1016/j.scico.2019.102389.

3.17 Approximation, abstraction and apartness in automata learning

Jurriaan Rot (Radboud University Nijmegen, NL)

URL http://jurriaan.creativecode.org/approximation-abstraction-and-apartness-in-automata-learning-apple/

My research focuses on the analysis and specification of models of computation, with the use of algebraic and coalgebraic techniques as a central theme. In recent years I have become interested in automata learning, and have worked on categorical generalisations of learning, extensions to various types of models, and algorithmic aspects. I currently lead the NWO VIDI project "Approximation, Abstraction and Apartness in Automata Learning", where the overall aim is to enhance the scope and scalability of automata learning through the development of learning algorithms that feature approximation and abstraction.

3.18 Automata learning algorithms for concurrent, infinite-state, and hardware systems

Matteo Sammartino (Royal Holloway, University of London, GB)

Joint work of Thomas Neele, Joshua Moerman, Alexandra Silva, Bartek Klin, Michal Szynwelski, Loris D'Antoni, Thiago Ferreira, Amir Naseredini, Martin Berger, Shale Xiong
Main reference Thomas Neele, Matteo Sammartino: "Compositional Automata Learning of Synchronous Systems", in Proc. of the Fundamental Approaches to Software Engineering – 26th International Conference, FASE 2023, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2023, Paris, France, April 22-27, 2023, Proceedings, Lecture Notes in Computer Science, Vol. 13991, pp. 47–66, Springer, 2023.
URL https://doi.org//10.1007/978-3-031-30826-0_3

My research is at the intersection of formal methods and AI. I am interested in both theoretical aspects, such as algebraic/coalgebraic approaches, formals semantics, and automata models with specific expressivity and decidability properties; and practical aspects, such as the automated verification of those systems.

Recent work focuses on learning algorithms for systems with "real-world" features, namely: concurrent/distributed systems, infinite-state systems, and physical hardware systems:

- *Compositional automata learning:* traditional model learning algorithms (eg, L*) are monolithic, and do not scale well when learning composite systems. In [1] we provide a compositional version of L*, where each component of a given synchronous system is learned independently, and independent learners cooperate in the presence of synchronisations. Experiments show that our approach may require up to six orders of magnitude fewer membership queries and up to ten times fewer equivalence queries than L* (applied to the monolithic system). Ongoing work focuses on relaxing assumption on the target system, for instance knowledge of the number of components.
- *Learning infinite-state models:* model learning algorithms exist for a wide range of automata models. In [2] we have introduced an extension for *nominal* automata, which are automata over infinite alphabets admitting a rich theory. Unlike the finite-alphabet case, non-deterministic nominal automata are strictly more expressive than deterministic ones (language equivalence is undecidable!), hence learning algorithm cannot deal with the full class. In [3] we have characterised and investigated learning of *residual* nominal

automata, which retain the required decidability properties for model learning to work. Future work will focus on richer classes of infinite-state automata, for instance [4], nominal ω-regular acceptors, and automata equipped with global freshness.

- *Learning from physical hardware:* due to their complexity, hardware components such as CPUs and DRAMs are hard to model and reason about, and often models are "idealised versions" that do not reflect the actual behaviour. Model learning can help in deriving models that are closer to the implementation. In [5] we have developed *ALARM*, a tool that uses learning to reverse-engineer undisclosed security features of DRAMs (= Dynamic Random-Access Memory). This was a proof-of-concept implementation targeting a software-simulated DRAM. Ongoing work is about enabling learning from physical hardware via a an FPGA-based testing harness.

References

1 Neele, T. & Sammartino, M. Compositional Automata Learning of Synchronous Systems. *Fundamental Approaches To Software Engineering – 26th International Conference, FASE*. **13991** pp. 47-66 (2023)

2 Moerman, J., Sammartino, M., Silva, A., Klin, B. & Szynwelski, M. Learning nominal automata. *Proceedings Of The 44th ACM SIGPLAN Symposium On Principles Of Programming Languages, POPL*. pp. 613-625 (2017)

3 Moerman, J. & Sammartino, M. Residuality and Learning for Nondeterministic Nominal Automata. *Log. Methods Comput. Sci.*. **18** (2022)

4 D'Antoni, L., Ferreira, T., Sammartino, M. & Silva, A. Symbolic Register Automata. *Computer Aided Verification – 31st International Conference, CAV*. **11561** pp. 3-21 (2019)

5 Naseredini, A., Berger, M., Sammartino, M. & Xiong, S. ALARM: Active LeArning of Rowhammer Mitigations. *Proceedings Of The 11th International Workshop On Hardware And Architectural Support For Security And Privacy, HASP*. pp. 1-9 (2022)

3.19 Safe, reliable and trustworthy AI

Philipp Sieberg (Schotte Automotive GmbH & Co KG – Hattingen, DE)

My research focuses on creating reliable, safe and trustworthy artificial intelligence.

During my doctoral research, I focused on the use of AI in safety-critical areas, where I believe reliability is the necessary basis for acceptance and trust. Consequently, I developed innovative hybrid approaches that integrate knowledge-based techniques with artificial intelligence methods. Through the implementation of hybrid methods, the benefits of incorporating artificial intelligence – including improved performance and accuracy – can be realized while maintaining a reliable process that instills a proven level of trust. In my doctoral research, hybrid methods were applied to the domain of automated driving, allowing for reliable implementation of virtual sensors and centralized predictive control.

As a principal investigator at the Chair of Mechatronics of the University of Duisburg-Essen, I am conducting further research on the subject of trustworthy and reliable artificial intelligence. For instance, an interdisciplinary project concentrates on utilizing AI to build a toolbox for automatically identifying wear mechanisms. The research initiative is organized into two distinct phases. During the first phase, the necessary technical methods will be developed to achieve reliable and comprehensible results. In the second stage of the research

project, the outcomes will be integrated into a toolbox in order to make the specialized materials engineering knowledge available to a wide range of users in industry and society. Interaction with users is critical in this context. The topics of acceptance and trust in this toolbox and in artificial intelligence are very important. Besides this project, I am also actively involved in the technical development of reliable and trustworthy AI methods in the field of automated driving. The focal point is the integration of physical knowledge to promote transparency and ensure the safety of AI.

As the General Manager of Schotte Automotive GmbH & Co. KG, one of my main goals is to drive the digitalization of the company. I concentrate on supporting and improving operational processes through the implementation of AI systems. In cooperation with Prof. Enkel and Dr. Liebherr from the University of Duisburg-Essen, a research project has been initiated that focuses on improving disposition with the help of artificial intelligence. In addition to predicting order quantities and times, achieving acceptance and trust in an AI-based tool by employees in the disposition department poses significant challenges.

References

1 Sieberg P.M. and Hanke S. (2023) Challenges and Potentials in the Classification of Wear Mechanisms by Artificial Intelligence. Wear, Vol. 522, 204725. https://doi.org/10.1016/j.wear.2023.204725

2 Stockem Novo, A., Hürten, C., Baumann, R., & Sieberg, P. (2023). Self-evaluation of automated vehicles based on physics, state-of-the-art motion prediction and user experience. Scientific Reports, 13(1), 12692.

3 Sieberg, P.M. and Schramm, D. (2022) Ensuring the Reliability of Virtual Sensors Based on Artificial Intelligence within Vehicle Dynamics Control Systems. Sensors, Vol. 22, 3513. https://doi.org/10.3390/s22093513

4 Hürten C., Sieberg P.M. and Schramm D. (2022) Generating a Multi-fidelity Simulation Model Estimating the Models' Applicability with Machine Learning Algorithms. In Proceedings of the 12th International Conference on Simulation and Modeling Methodologies, Technologies and Applications 2022: SIMULTECH – Vol. 1, pp. 131-141. ISBN 978-989-758-578-4, https://doi.org/10.5220/0011318100003274

5 Sieberg P.M., Blume S., Reicherts S., Maas N. and Schramm D. (2021) Hybrid State Estimation – A Contribution Towards Reliability Enhancement of Artificial Neural Network Estimators. In IEEE Transactions on Intelligent Transportation Systems, Vol. 23, pp. 6337-6346. https://doi.org/10.1109/TITS.2021.3055800

6 Sieberg P.M., Hürten C. and Schramm D. (2020) Representation of an Integrated Non-Linear Model-Based Predictive Vehicle Dynamics Control System by a Co-Active Neuro-Fuzzy Inference System. In Proceedings of the IEEE Intelligent Vehicles Symposium (IV) 2020, pp. 572-577, Las Vegas, USA – ISBN 978-1-7281-6673-5

4 Open problems

4.1 Sound Control Synthesis with Logics and Data

Alessandro Abate (University of Oxford, GB)

We are witnessing an inter-disciplinary convergence between scientific areas underpinned by model-based reasoning and by data-driven learning. Original technical work across these areas is justified by numerous applications, where access to information-rich data has to be traded off with a demand for safety criticality: cyber-physical systems are exemplar applications.

Within OXCAV, the Oxford Control and Verification group, I focus on control synthesis for complex objectives, and study how techniques from formal verification (logics and SAT, automata theory, abstractions) and from learning (sample-driven approaches and neural architectures) can be together leveraged to attain both sound and effective synthesis outcomes.

More broadly, I argue that, on the one hand, control theory and formal methods can provide certificates to learning algorithms and, on the other hand, that learning can bolster formal verification and strategy synthesis objectives.

4.2 Automated generation of efficient and systematic test suites

Robert M. Hierons (University of Sheffield, GB)

My research interests largely concern the automated generation of efficient, systematic test suites on the basis of models or specifications, although I have done some work on generating test cases based on the code. Within this, I have had a particular interest in testing based on state-based models, typically expressed either as a form of state-machine or using a process algebra such as CSP. Within testing, I have looked at how the nature of the interaction between the tester(s) and the system under test affect any notion of correctness (conformance) used, including work on distributed testing and asynchronous testing. In recent years, I have also become interested in the use of concepts regarding causality to support testing. Finally, I have longstanding interest in the relationship between testing and learning, with this including recent work on the use of reinforcement learning.

4.3 Testing causal properties, and reasoning about uncertainty

Neil Walkinshaw (University of Sheffield, GB)

My research focusses on model inference and statistical reasoning for the purpose of testing. My current research interests focus on two primary areas: (1) Testing for causal relationships between input and output variables, and (2) reasoning about uncertainty in reverse-engineered models.

Testing causality. I am the PI on the CITCOM project, which is concerned with the development of techniques to test for the presence or absence of causal relationships in complex systems. For this we have adapted and applied a family of statistical reasoning techniques called Causal Inference. Here, the developer provides a model (in the form of a DAG) setting out the causal relationships between the inputs and outputs they are interested in, and our testing technique can test for them in data.

We have shown how this form of testing can be framed as Metamorphic Testing. However, it also has the added benefit that we are able to test for causal relationships purely from observed data, without needing to control the inputs.

Reasoning about uncertainty. One strand of this research has focussed on the inference of models that explicitly incorporate second-order uncertainty. Second order uncertainty pertains our confidence (or lack thereof) in a particular model element or value. This is particularly important when dealing with inferred models, where some aspects of the inferred model will invariably be better supported by data than others.

Much of my research has used Subjective Logic – a way of reasoning about probabilities whilst explicitly incorporating second-order uncertainty. Most recently, we have developed a generalisation of probabilistic finite state machines, called Subjective Opinion State Machines. For this, we have also developed an inference approach that can be "layered over" any conventional state machine inference algorithm to produce these state machines with explicit second-order uncertainty.

References

1 `https://github.com/CITCOM-project/CausalTestingFramework`
2 Andrew G. Clark, Michael Foster, Benedikt Prifling, Neil Walkinshaw, Robert M. Hierons, Volker Schmidt, and Robert D. Turner. 2023. Testing Causality in Scientific Modelling Software. ACM Trans. Softw. Eng. Methodol. 33, 1, Article 10 (January 2024).Reference:
3 Walkinshaw, Neil, and Robert M. Hierons. "Modelling Second-Order Uncertainty in State Machines." IEEE Transactions on Software Engineering (2023).

Participants

- Alessandro Abate
University of Oxford, GB
- Wolfgang Ahrendt
Chalmers University of Technology – Göteborg, SE
- Bernhard K. Aichernig
TU Graz, AT
- Thomas Arts
QuviQ AB – Gothenburg, SE
- Thorsten Berger
Ruhr-Universität Bochum, DE
- Jyotirmoy Deshmukh
USC – Los Angeles, US
- Monique Dittrich
CARIAD – Berlin, DE
- Ellen Enkel
Universität Duisburg-Essen, DE
- Sophie Fortz
University of Namur, BE
- Fatemeh Ghassemi Esfahani
University of Tehran, IR
- Heiko Hecht
Universität Mainz, DE
- Leo Henry
University College London, GB
- Robert M. Hierons
University of Sheffield, GB
- Falk Howar
TU Dortmund, DE
- Nils Jansen
Ruhr-Universität Bochum, DE
- Effie Lai-Chong Law
Durham University, GB
- Magnus Liebherr
Universität Duisburg-Essen, DE
- Mohammad Reza Mousavi
King's College London, GB
- Thomas Neele
TU Eindhoven, NL
- Nicola Paoletti
King's College London, GB
- Jurriaan Rot
Radboud University Nijmegen, NL
- Kristin Yvonne Rozier
Iowa State University – Ames, US
- Matteo Sammartino
Royal Holloway, University of London, GB
- Philipp Sieberg
Schotte Automotive GmbH & Co KG – Hattingen, DE
- Bernhard Steffen
TU Dortmund, DE
- Marnix Suilen
Radboud University Nijmegen, NL
- Neil Walkinshaw
University of Sheffield, GB