Helen Bolke-Hermanns, Jan-Christoph Kassing, Prof. Michael Schaub, Michael Scholkemper (editors)

Proceedings of the 2024 Joint Workshop of the German Research Training Groups in Computer Science

May 26 - May 29, 2024

DFG Deutsche Forschungsgemeinschaft

Imprint

Proceedings of the 2024 Joint Workshop of the German Research Training Groups in Computer Science
May 26 - May 29, 2024
RWTH Aachen University
Aachen, Germany

Editors
Helen Bolke-Hermanns
Jan-Christoph Kassing
Prof. Michael Schaub
Michael Scholkemper
RWTH Aachen University
RTG UnRAVeL
Ahornstrasse 55
52074 Aachen
Germany

RWTH Aachen University
Templergraben 55
52062 Aachen Germany

First published 2024

Preface

Since the early 2000s, the annual joint meeting of the German Research Training Groups (Graduiertenkollegs) funded by the German Research Foundation (DFG) have been held at Schloss Dagstuhl - Leibniz Centre for Informatics. Schloss Dagstuhl is one of the world’s leading venues for computer science seminars. The aim of the meeting is the interactive exchange of research results, ideas, and experiences in order to strengthen the German computer science community, also across different levels of seniority. This volume documents the abstracts of the research topics of the funded researchers in the participating RTGs and provides an insight into current research trends in Germany. This year’s meeting was jointly organised by RTG 2236: ’UnRAVeL Uncertainty and Randomness in Algorithms, Verification, and Logic’ and took place from 26. - 29. May 2024. (Dagstuhl event number 24223). The meeting consisted of a balanced mix of research workshop, a poster session, and a session where young researchers could ask senior researchers “Everything you always wanted to know about research and a research career". In addition, two keynotes were organised on important current topics in computer science: Christin Seifert, University of Marburg and Lisa Musculus, German Sport University Cologne. The organisers would like to thank all participants for their contribution to a successful and enlightening event in an open and welcoming atmosphere.

organized by

UnRAVeL

Research Training Group –
Uncertainty and Randomness
in Algorithms, Verification,
and Logic

RWTH AACHEN
UNIVERSITY

Contents

HPI Research Schools: Service-Oriented Systems Engineering & Data Science and Engineering **1**

GRK 2193: Adaption Intelligence of Factories in a Dynamic and Complex Environment **7**

HPI Research Schools: Service-Oriented Systems Engineering & Data Science and Engineering

Prof. Dr. Felix Naumann, Prof. Dr. Tilmann Rabl,
Prof. Dr. Robert Hirschfeld, Prof. Dr. Andreas Polze
Email: {firstname.lastname}@hpi.de
Hasso Plattner Institute at the University of Potsdam
Internet: hpi.de/en/research/cooperations-partners/research-schools

The two HPI Research Schools for "Service-Oriented Systems Engineering" and "Data Science and Engineering" are the HPI graduate schools. Our research schools devote themselves to current topics in the fields of IT systems engineering and data science with high potential in academic research as well as in industrial application. PhD students at our schools participate in joint activities such as lectures, seminars, winter schools, and workshops. With their interdisciplinary structure, our schools interconnect the HPI research groups and foster close and fruitful collaborations. Frequent workshops with the international branches of the graduate schools at the University of California in Irvine (USA) and the University of Cape Town (South Africa) encourage academic exchange.

Research School "Service-Oriented Systems Engineering"

Design and implementation of service-oriented architectures impose numerous research questions from the fields of software engineering, system analysis and modeling, adaptability, application integration and IT security. "Service-Oriented Systems Engineering" represents a symbiosis of best practices in object orientation, component-based development, distributed computing, and business process management. It provides integration of business and IT concerns.

Since 2005 the research school "Service-Oriented Systems Engineering" investigates this research topic in the field of IT-systems engineering with high potential in academic research as well as in industrial application.

Research School "Data Science and Engineering"

The increasing abundance of data in science and in industry creates many challenges and opportunities. Data science has grown to be a foundational discipline in information technology, allowing new insights from data and creating ever more intelligent applications. Simultaneously, it is becoming increasingly difficult to collect, clean and deliver the vast amounts of data and apply and maintain complex data science processes. Targeting these challenges, the discipline of data engineering has become equally foundational.

The 2019 newly established research school "Data Science and Engineering" unites top PhD students in all areas of data-driven research and technology, including scalable storage, stream processing, data cleaning, machine learning and deep learning, text processing, data visualization and more.

Enhancing Developer Experience through LLM-Assisted Mentorship

Lukas Böhme (lukas.boehme@hpi.de)
Supervisor: Prof. Dr. Robert Hirschfeld

Programmers encounter numerous challenges in software development when getting onboarded on a new project. Peter Naur's view of programming as theory building rather than mere code production highlights the importance of program comprehension and theory building for effective integration into software projects[1]. The resulting learning curve to fully comprehend the project hampers the swift adaption of new projects and risks introducing errors due to knowledge gaps. Current practices of onboarding newcomers to a project include studying documentation, exploring the codebase, and mentorship, each with shortcomings. Documentation is spread over multiple sources and is prone to becoming outdated. Codebases might grow too large, and thus, identifying relevant information for a concrete task becomes overwhelming. Mentoring requires a substantial amount of time from experienced programmers, and mentees must have the courage to ask questions (even simple ones).

Large language models (LLMs) hold the potential to overcome many of the mentioned shortcomings by assisting and guiding programmers, effectively acting as virtual mentor. Current LLM applications focus on generating code, but their application in parsing and explaining project-related information is essential for theory building. However, LLMs face limitations as virtual mentor due to their training on public data, often lacking specific information on new or private projects. To address this Retrieval Augmented Generation (RAG) allows to incorporate project-specific insights, improving answer quality[2]. Leveraging RAG requires identifying and implementing information retrieval algorithms to harvest pertinent details from documentation or the codebase. This research investigates information retrieval algorithms to extract information from codebase and documentation for a given query a mentee might ask. We explore practices such as (1) embedding-based and knowledge graph-based search as well as (2) incorporating static and dynamic information of the codebase. By introducing a novel set of strategies for information retrieval, we elevate LLMs to provide high-quality answers tailored to distinct projects. Thereby, novices can be supported in building a reliable understanding and a solid theory about their project, aligning with Peter Naur's view of programming as theory building.

[1] Peter Naur, Programming as Theory Building (1985), https://doi.org/10.1016/0165-6074(85)90032-8

[2] Patrick Lewis et al., Retrieval-augmented generation for knowledge-intensive NLP tasks. In Proceedings of the 34th International Conference on Neural Information Processing Systems (2020), https://doi.org/10.48550/arXiv.2005.11401

Programming In and For Virtual Reality

Leonard Geier (leonard.geier@hpi.de)
Supervisor: Prof. Dr. Robert Hirschfeld

As interest in virtual, augmented, and mixed reality (VR, AR, XR) increases, development tools for this domain become more and more important. However, the development of applications in this space can be cumbersome, especially if the application needs to be run frequently. Whenever a developer wants to switch between editing and testing an application hands-on, they need to switch between their desktop and their VR development environments, which necessarily entails a device switch involving the VR headset. This process creates developing friction.

Self-sustaining and live programming systems, which allow code to be easily modified while it is running, would be a promising way to achieve this goal—by both developing and testing an application in VR, the developer avoids costly context switches, which leads to shorter feedback loops, in turn reducing the cognitive load.

However, interaction in VR radically differs from the conventional desktop metaphor. It follows that the traditional mouse-and-keyboard interactions used in the most popular development environments cannot be transferred directly, or even at all, to VR. A set of new interactions that are 'VR-native' becomes necessary for development.

Text poses a major hurdle, since it is the main way that programs are represented. Text entry in VR carries with it harsh trade-offs; often, it is either slow (e.g. using a 6DOF controller) or tethers the user to a physical location (e.g. that of the physical keyboard)—both undesirable.

Alternatively, a VR programming system could be based on a rich set of presentations other than text, which creates new opportunities and exciting research questions.

Initial forays into this space proved promising. Live feedback in VR is desirable for programmers, and having a three-dimensional workspace appears to encourage code exploration. This also opens opportunities for novel workflows exclusively in VR and mixing desktop and VR, or asymmetric collaboration where one participant wears a VR device.

To explore this space, we are currently developing a self-sustaining programming system for virtual reality on the basis of WebXR and SqueakJS.

Delay-Robustness in Distributed Cyber-Physical System Models

Mustafa Ghani (Mustafa.Ghani@hpi.de)
Supervisor: Prof. Dr. Holger Giese

Distributed Cyber-Physical Systems (DCPSs) are omnipresent and their analysis against provided specifications is a central challenge. Distribution results in communication delays among agents that have to be adequately taken into account by software models to avoid race conditions. However, engineering systems at a higher level of detail by incorporating communication delays explicitly inflates model size and impedes analysis.

In this thesis, we employ Timed Graph Transformation Systems (TGTSs) to model DCPSs and distinguish between local immediate and remote δ-delayed observations, requiring up to δ time units.

We then (a) demonstrate potential absence of δ-delay robustness for TGTSs models, (b) provide a procedure widening safe behavioral options of a verified 0-delay system model to derive a minimally restricted δ-delay robust TGTS model, and (c) analyze the resulting TGTSs model for new unsafe behavior.

Methods for characterization and landscape detection of post-translational modifications

Yannick Hartmaring (yannick.hartmaring@hpi.de)
Supervisor: Prof. Dr. Bernhard Renard, Dr. Christoph Schlaffner

A post-translational Modification (PTM) is an alteration on a protein that changes the amino acid sequence into a functional proteoform and therefore highly increase the complexity of the overall proteome. PTMs regulate not only the physical or chemical properties of proteins but also their structure, stability and cellular location. Overall, they affect almost all cellular processes. [1]

In addition to the most studied PTM, phosphorylation, more than 300 known PTMs exist. Since the individual modifications also often rely on each other, a much higher degree of combinatorial variation occurs. This type of interaction is called crosstalk. Discrepancies in these interaction is linked to the development of various diseases, e.g. Alzheimer's disease and diabetes. [2]

Through the shotgun approach of proteomic mass spectrometry, which got the method of choice to analyse PTMs, the protein sequences are divided into subsequences. Therefore, it is no longer clear which PTMs are located simultaneously on the same protein molecule.

To address this problem I am building a tool which is able to reconstruct the original proteoforms using statistical methods. Therefore, I simulate the process of the shutgun approach with a known ground truth and afterwards analyse which PTMs generally occur in the simulated data. I then utilize the PTM identifications and peptide quantitative information to initialise all possible proteoforms. An expectation maximization algorithm then calculates a probability score for each of the proteoforms. To exclude those proteoforms which are not supported by the data regularization is applied to only reconstruct proteoforms that are most likely represented in the sample.

The new knowledge gained from proteoform identification and PTM co-expression enables targeted functional experiments using synthesized versions. Also, the method can be applied post-hoc to all publicly available proteomics datasets and provide an overview of the most common proteoforms across tissues, conditions, and species. This will further provide new insights into cellular protein regulation and highlight modifications of interest in a variety of human tissues and conditions. Accordingly, it will be possible to develop new approaches to understanding and treating diseases.

[1] Aebersold et al., How many human proteoforms are there?. Nat Chem Biol 14, 206–214 (2018)

[2] Laarse et al., Crosstalk between phosphorylation and O-GlcNAcylation: friend or foe, FEBS J., vol. 17 (2018)

Multilingual Visual Metaphor Generation

Yindong Wang (yindong.wang@hpi.de)
Supervisor: Prof. Dr. Gerard de Melo

This research investigates the dual capabilities of state-of-the-art vision-language and diffusion models in both interpreting and generating visual metaphors across a diverse linguistic spectrum. It seeks to uncover the extent to which these advanced models can not only understand complex metaphorical constructs in different languages, but also creatively generate visual representations that embody these metaphors. By addressing key questions related to the models' ability to decode and visually articulate metaphors in multilingual contexts, the consistency of the images they generate across languages, and the methods for assessing similarities in visual output from identical metaphorical prompts, this study aims to illuminate the nuanced roles these models play in visual metaphor generation.

The challenge for vision-language and diffusion models lies not only in representing the literal elements of a metaphor, but also in capturing its abstract, often culturally nuanced meanings. This requires a deep understanding of textual input and an innovative approach to visual generation that goes beyond simple literal interpretations and embraces the complex interplay of cultural and contextual subtleties inherent in metaphors. The study emphasizes the importance of diversity and stylistic variation in the generated images, reflecting the intrinsic variability of metaphors.

Moreover, this research highlights the need for a systematic methodology to evaluate the coherence between visual representations generated from the same metaphorical cues, taking into account aspects such as thematic consistency, conceptual depth, and cultural accuracy. This study aims to contribute significantly to the understanding of how diffusion models can be used to enhance the interpretive and creative capabilities of AI systems in the domain of multilingual visual metaphor generation.

GRK 2193: Adaption Intelligence of Factories in a Dynamic and Complex Environment

Prof. Dr. Jakob Rehof
Email: jakob.rehof@cs.tu-dortmund.de
TU Dortmund
Internet: www.grk2193.tu-dortmund.de

Multidisciplinary approaches are indispensable in order to fully grasp the complexity of factory planning. By considering interdependencies of aspects related to completely separate disciplines, shorter reaction times and enhanced and robust adaption efficiency can be achieved. The Research Training Group 2193 - "Adaption Intelligence of Factories in a Dynamic and Complex Environment" enables an interdisciplinary education of doctoral researchers in adaption planning of factories and facilitates research that taps into the aforementioned potential. The Research Training Group is set up to promote cooperation, enabling the doctoral researchers to reach their interdisciplinary research aims by providing the interfaces to interdependent research fields.

The structure of the research program concentrates on a continuous and multidisciplinary reflection of the complete cause-activity-chain of adaption processes. The Research Training Group focuses on the goal-oriented compilation of integrative and consistent models and concepts, which continuously support the adaption processes and improve both collaboration and adaption intelligence in a dynamic workflow.

An Acceptance promoting Digital Twin for an Autonomous System under Consideration of Trust

Thomas Bömer (thomas.boemer@tu-dortmund.de)
Supervisor: Univ. -Prof. Dr. habil. M. Henke

This dissertation addresses the challenge of fostering acceptance for autonomous block-stacking warehouses (ABSW) through the development of a digital twin, focusing on the crucial role of calibrated trust in technology adoption. The rapid evolution of market demands necessitates efficient, adaptable logistics systems, where autonomy could offer significant advantages. However, the hesitancy towards adopting such systems is primarily due to doubts about their decision-making capabilities and technological maturity. By leveraging a digital twin—a virtual representation allowing for system testing and demonstration before implementation—this dissertation aims to mitigate these concerns by demonstrating feasibility and calibrating trust among potential users.

A literature review lays the groundwork, exploring autonomization in intralogistics, trust dynamics in autonomous systems, trust calibration mechanisms, and technology acceptance models, particularly the Unified Theory of Acceptance and Use of Technology (UTAUT). Identifying gaps in both the application of digital twins for ABSW and research on trust calibration for complex systems, this research proposes a novel approach: defining an ABSW system, developing trust calibration measures for transparency in system performance and processes, and integrating these into the digital twin for empirical evaluation.

The methodology involves a user study based on a trust-extended UTAUT model to assess how interaction with the digital twin impacts technology acceptance. The findings aim to contribute to the academic and practical understanding of autonomization in intralogistics by offering insights into digital twin design and operation, developing and testing trust calibration measures, and applying the UTAUT model to this context.

In essence, this dissertation explores how a digital twin can bridge the gap between technological capabilities and user acceptance in the realm of autonomous intralogistics, underscoring the importance of trust calibration and transparent communication. The research not only addresses theoretical and practical challenges in the adoption of autonomous systems but also sets a foundation for future studies in digital twins and industrial autonomous applications.

Synthesis of CAD Assemblies: Towards semi-automated Product Line Engineering

Constantin Chaumet (constantin.chaumet@tu-dortmund.de)
Supervisor: Prof. Dr. Jakob Rehof

Modern day products are rarely one-size-fits-all solutions. In midst of the twenty-first century paradigm shift towards smart and adaptable solutions, products have become configurable and customizable, on a per consumer basis. As such, engineering complete product lines instead of individual products is an emergent and increasingly important topic within that context. Despite this, CAD software packages usually offer no (or poor) automation of creating CAD assemblies for products, requiring engineers and designers to perform many manual, often redundant, tasks.

Modelling a product, or members of a product line, in CAD software is an integral part of the design process. It allows determining which parts need to be produced and which parts are off-the-shelf components, provides a complete bill of materials, and allows the design to be simulated, checked, and improved, repeatedly, without constructing a prototype for every iteration. While this is a large improvement over paper-and-pen drafting, the CAD process still entails a number of repetitive steps, e.g. inserting all screws needed for an assembly. This issue gets worse when engineering product lines, as the creation of each new member of the product line necessitates such repetition again. The ongoing shift towards highly customized products in wake of Industry 4.0 compounds this issue, necessitating even more design effort. While this can be mitigated for some problem domains where customization is sufficiently simple, this is not an option for product lines that differ in structure, i.e. robotic systems with varying degrees of freedom. Companies find themselves in a catch-22 dilemma: Customization is a competitive advantage, however it must also be achieved without significantly increasing costs.

Aim of the proposed thesis is to identify shortcomings within the current state of product line engineering and provide solutions to increase automation, and thus improve efficiency and enhance creativity, throughout the process. This is achieved by developing a novel synthesis of CAD assemblies, as well as developing and integrating several tools which aid in setting up the synthesis. The synthesis and tools apply concepts from knowledge-based engineering and implement them by utilizing the CLS framework (a framework that employs combinatory logic to derive solutions that satisfy a specification by combining modular components), constructing results directly in CAD software.

Architecture of Digital Twins for Multi-AGVs application

Zhaoqing Gong (zhaoqing.gong@rt.rif-ev.de)
Supervisor: Prof. Dr. Jürgen Roßmann

In the realm of smart logistics, Automated Guided Vehicle Systems (AGVS) have been central to the enhancement of intralogistics systems. Their adoption has been propelled by factors including the rise in labor costs, labor shortages, the expansion of e-commerce, and the impact of the COVID-19 pandemic. The incorporation of AGVs into logistics operations is a key component of the broader Industrial 4.0 movement, focusing on automation and data exchange within manufacturing technologies.

Despite the benefits, AGVs face implementation challenges, notably high costs and substantial time investments, particularly during the initial commissioning phase. These challenges originate from complex system integration requirements, the necessity for accurate navigation technologies, and the extensive testing needed to ensure system reliability and efficiency. As industries progress towards Industry 4.0, factory owners encounter various technological challenges, which have somewhat hindered the full adoption of Industry 4.0 applications, especially among small and medium-sized enterprises (SMEs).

At the heart of Industry 4.0 lies the digitalization of manufacturing and logistics operations, with digital twin technology emerging as a particularly promising tool in this transition. Digital twins, dynamic virtual replicas of physical systems, provide a robust platform for simulation, analysis, and real-time monitoring. This technology is especially beneficial in the context of virtual commissioning for AGV systems, enabling system testing and refinement in a virtual or hybrid environment, thereby reducing the dependence on physical prototypes and streamlining the development process.

This dissertation delves into the critical role of digital twin technology in the development and operational phases of AGVS. During the development phase, digital twin-based virtual commissioning acts as an efficient and intuitive tool for identifying potential issues within the AGVS, controlled by the AGV master system, and subsequently verifying the system's functionality. Once the AGV system is operational, the digital twin runs in parallel to the actual AGVS, continually monitoring and synchronizing data. This facilitates the creation of an operational database, laying the groundwork for advanced applications such as data analysis and artificial intelligence implementations.

Strategic Decision-making for Energy Flexibility Investments in Industrial Processes: A Simulation-based Approach

Simon Kammerer (simon.kammerer@tu-dortmund.de)
Supervisor: Prof. Dr. Christian Rehtanz

The contemporary industrial landscape is confronted with unprecedented challenges and opportunities arising from escalating energy costs, sustainability mandates, and the transition towards renewable energy sources. These factors necessitate the exploration and utilization of energy flexibility within industrial operations to maintain competitiveness and sustainability. This dissertation presents an innovative approach to understanding, modeling, and exploiting energy flexibility in industrial processes through the development of a simulation tool designed for this purpose.

The core of this research lies in the creation of a versatile simulation environment that allows for the comprehensive modeling and optimization of site-specific energy and production systems. This tool employs an optimization-based modeling framework capable of representing complex energy supply systems, sector coupling technologies, and material flow-based production infrastructures. The environment facilitates the exploration of various strategies for infrastructure investments and the flexibilization of production processes to enhance energy efficiency, sustainability, and operational flexibility.

A significant contribution of this work is the development of the Factory Flexibility Model, an advanced simulation tool that integrates energy systems and production processes into a cohesive model. This tool leverages a node-edge approach for system layout representation, enabling the evaluation of demand-side management potentials, investment analysis in energy infrastructure, and the assessment of revenue opportunities in energy markets. The methodology empowers stakeholders to make informed decisions regarding investment and operational strategies by quantifying potential risks and opportunities.

Through a series of application scenarios and case studies, this dissertation demonstrates the tool's capacity to analyze the impact of various flexibility utilization strategies on industrial energy systems. These scenarios encompass a wide range of applications, including load management, optimization of self-consumption, evaluation of energy storage solutions, and exploration of market-based energy procurement and trading strategies.

Smart Contracting in digital twin in shared manufacturing

Larissa Krämer (larissa.kraemer@tu-dortmund.de)
Supervisor: Univ. -Prof. Dr. Dr. h. c. M. ten Hompel

The dissertation project focuses on the use of blockchain technology and smart contracting in shared manufacturing to reduce information asymmetry and to enhance trust between foreign parties. Multi-stakeholder processes often suffer from conflicting interests and a lack of transparency. A digital twin of a shared manufacturing system addresses these challenges by providing extensive visualization and simulation features. A digital twin relies on trustworthy data, which can be provided by blockchain technology due to its immutable and decentralized data storage.

In the dissertation project, a digital twin of a shared manufacturing system with smart contracting is developed. The research provides insights into how the efficiency of shared manufacturing can be increased and how such systems can be simulated. Additionally, the simulation provides insights into the feasibility of smart contracting in a digital twin. Experiments provide data on significant key figures and multiple possible configurations.

The dissertation project uses a design science research approach as the overarching methodology. A systematic literature review reveals crucial requirements for the development of the model. The model is realized as a configurable simulation model and linked with a physical production system in the laboratory to create a holistic and realistic digital twin. Quantitative experiments are executed in the simulation for several scenarios using multiple blockchain frameworks.

KI-gesteuerte Entscheidungsfindung Interaktionsbasierte Effekte und Implikationen im Führungskontext

Thorben Krokowski (thorben.krokowski@web.de)
Supervisor: Prof. Dr. Jürgen Howaldt

The aim of the dissertation is to find out to what extent the functional role played by executives and management in adaptation-intelligent factories undergoes a modification of their constitutive design, action and, above all, decision-making logics in the course of the influence of AI-supported decision-making systems. In this context, special attention is being paid to artificial intelligence - as a decision-making reference and benchmark. Taking into account the increasing spread of AI-supported or even AI-based decision-making processes in the corporate context and the associated assumption of decision-making competences, powers and responsibilities (AI-based decision-making process) through data-based, algorithmic processes and AI methods and technologies, "leadership" is confronted with a changed role and requirements profile.

Among other things, this can lead to a modification, reconfiguration or complete redefinition of the reciprocal interaction between management, employees and technology. As a result, the research project is also responsible for determining the interactionally driven understanding of leadership within the framework of the triadic relationship between leadership-management/ employees/ AI(technology) against the background of the concept of interaction work that guides the analysis.

Simulation-based analysis and development of an innovative clamping device for the post-machining of additively manufactured workpieces)

Jan Liß (jan.liss@tu-dortmund.de)
Supervisor: Prof. Dr.-Ing. Petra Wiederkehr

This dissertation project focuses on the simulation-assisted analysis and development of an innovative clamping device for post-machining of additively manufactured components. The project is motivated by the urgent need for manufacturing chains to swiftly and efficiently adapt their products and factory systems to the rapidly changing dynamics driven by increased cost pressures, intense international competition, and supply chain volatility. It aims to enhance the flexibility and adaptability of the post-process phase in additive manufacturing, focusing on the clamping strategies and automation possibilities for additively manufactured workpieces.

Additive Manufacturing (AM) has garnered significant industrial interest over the past 25 years for its unparalleled design freedom compared to traditional manufacturing processes. The project particularly examines the Laser Powder Bed Fusion (LPBF) technique, known for its ability to produce near-net-shape complex geometries and lightweight structures optimized for loading conditions. However, the process inherently produces parts with limited surface quality and dimensional accuracy, necessitating subsequent machining operations to meet precise tolerance requirements. The challenge is further compounded by the need for suitable fixture systems to securely hold often intricate and delicate AM parts during post-processing, without introducing deformations or affecting the part's integrity.

This research endeavors to develop an innovative clamping device that can be flexibly adjusted to the unique features of additively manufactured parts. The device aims to compensate for shape deviations and surface defects while preventing deformation from excessive clamping forces. A key aspect of this development is the integration of sensors within the clamping system to reproduce the clamping situation with minimal operator influence, thereby reducing setup times and associated indirect manufacturing times.

The methodology combines detailed empirical studies with simulation systems for machining to design and integrate the novel clamping device into existing process chains, aiming for a significant quality improvement in the machining results of AM parts. The research addresses a crucial gap in the integration of efficient post-processing into the additive manufacturing process chain, particularly in industrial AM series production. By focusing on the post-process costs, which can account for a significant portion of the total part costs depending on the geometry, the project aligns with the broader objective of the Research Training Group to support the adaptive adjustment of factory systems, thus enhancing productivity, reducing manufacturing costs, and increasing sustainability in the additive manufacturing field.

Visual Performance Management in Socio-Cyber-Physical Production Environments – With Special Consideration of Employee-Centered Incentives in Reporting Systems

Cornelia Regelmann (cornelia.regelmann@tu-dortmund.de)
Supervisor: Prof. Dr. Andreas Hoffjan

This dissertation project focuses on **Visual Performance Management (VPM)** in socio-cyber-physical production environments, with particular emphasis on employee-centered incentives in reporting systems. It delves into the digitalization, individualization, and automation of the manufacturing industry in Germany, within the context of Industry 4.0. This involves the integration of the Internet of Things and the implementation of Cyber-Physical Systems (CPS), aiming to merge the physical and digital worlds.

The project highlights the increased role of production employees as central elements of successful production in the era of socio-cyber-physical production systems. It suggests that modern production environments necessitate a shift from traditional performance measurement to performance management, emphasizing the management of performance over its mere measurement. The research posits that Big Data, characterized by its "5 V's" (Volume, Velocity, Variety, Veracity, and Value), plays a significant role in adapting performance management processes, especially in the phase of performance realization.

The research introduces three central hypotheses:

1. Big Data significantly influences the phases of the performance management process, particularly in performance realization, affecting the speed, accuracy of analysis, and representability of performance.

2. Integrating employees into the goal-setting process and providing access to real-time data on the shop floor enables self-responsible work, positively impacting employee engagement.

3. Appropriate preparation and visualization of Key Performance Indicators (KPIs) for shop floor employees are essential for the acceptance and motivational effect of these indicators.

VPM is defined as a system for organizational improvement focusing on what is important to enhance performance. It incorporates strong graphic visualization techniques to emphasize sustaining competitive advantage. The project extends this concept to operational KPIs, examining how visualizing performance indicators can motivate and engage employees on the shop floor.

Methodologically, the project adopts a qualitative, empirical research design with a multiperspective case study approach, conducting in-depth interviews with both management and shop floor employees. This case study aims to understand the different needs and perceptions of KPIs between these groups and how visualizations can be effectively used as a motivational tool.

Trustworthy and Efficient Sales Forecasting in Large Industrial Companies

Alina Timmermann (alina.timmermann@tu-dortmund.de)
Supervisor: Prof. Dr. Peter Buchholz

This dissertation explores the development of reliable and efficient sales forecasting for large industrial companies, aiming to improve both the accuracy and the trustworthiness of forecasts. The challenge of generating correct and trustworthy sales forecasts in industrial enterprises is formidable. This research addresses the algorithm aversion phenomenon, where people tend to distrust algorithms, by focusing on the calculation and presentation of forecast results to enhance end-user confidence. It aims to achieve optimal forecasting results through suitable prediction methods and to provide descriptive metrics that optimize user trust in forecasting software.

The study encompasses the current state of scientific knowledge on time series forecasting, the reliability of forecasts, and clustering of time series. Time series analysis, dealing with the stochastic process of observing random variables over time, forms the foundation of forecasting. The research also delves into the influence of algorithm aversion on the acceptance of decision-support software and the potential of clustering time series to identify trends, employing various clustering algorithms and distance measures to increase the similarity within groups while decreasing it between groups.

The dissertation sets out to make sales forecasts more accurate and trustworthy through three key aspects: enhancing the quality and efficiency of forecasts, identifying metrics to increase forecast trustworthiness, and estimating these metrics in real cases without test data. This involves grouping time series through clustering and selecting the best forecasting method for each cluster. Metrics for increasing trustworthiness will be determined through literature review and independent consideration.

Methodologically, the dissertation uses synthetic and real datasets for experiments, applying both supervised and unsupervised clustering, and a range of forecasting methods including ARIMA, Holt-Winters, and neural networks. The effectiveness of these methods is evaluated based on deviation from actual values and computational efficiency.

The Impact of Part Complexity on the Demand of Customer Data

Greta Tjaden (greta.tjaden@trumpf.com)
Supervisor: Prof. Dr.-Ing. Anne Meyer

In the recent past, the competitive pressure on machine tool manufacturers has risen, while shop floor data became more accessible. To meet this competitive pressure, machine tool manufacturers increasingly want to offer customized solutions, for which they need data-based information about their customers.

This information is required for customer-centric activities along the product life cycle. This dissertation aims to research part complexity as a contributor to this gap of data-based customer information along the product life cycle.
The literature review reveals different attempts at part complexity definitions as well as approaches to part complexity assessment. The findings of this literature review build the foundation for the derivation of a definition of part complexity, which is adapted to the exemplary use case of a machine tool manufacturer for sheet metal processing.

Moreover, a mixed-methods approach consisting of both empirical and qualitative research methods is developed for the assessment of part complexity. This approach covers both scientific and industrial feedback loops, the development of a supporting online labelling tool for the computer-administered self-interviews for the complexity labelling, the creation of a dataset of 80 geometries, and a rating scale.

Furthermore, a subset of geometries is repeated in each week of the labelling to evaluate the consistency of the experts' labelling. To put this method into practice, we implement it for an exemplary production unit, offering one production step and two part handling options. To add to the use cases of part complexity, experts were asked with a cross-departmental workshop.

Financial Digital Twin

Anjali Vaghani (anjali.vaghani@tu-dortmund.de)
Supervisor: Univ. -Prof. Dr. habil. M. Henke

Digital Twins have gained significant popularity in recent years due to their ability to optimize operational performance in various industries. However, it is crucial to consider a wider range of data when making significant decisions. One area that has been gaining attention is the use of financial data in conjunction with operational metrics to gain a more comprehensive and holistic view of the organization's performance. Financial data plays a vital role in this regard since it provides valuable insights into the company's financial health, including cash flow, profitability, and revenue generation. By combining financial information from different sources within the organization, such as accounting systems, financial statements, and budgeting tools, with operational metrics such as inventory levels, production schedules, and delivery times, a detailed and comprehensive view can be achieved.

The objective of the research is to explore the potential of using smart contracts to automatically activate financial parameters based on off-chain data from digital twins, such as inventory levels. This approach has several advantages, including the ability to create detailed scenarios that take into account both financial transactions and operational events. This enables decision-makers to gain a better understanding of how different financial parameters, such as payment terms, impact the overall efficiency and performance of the supply chain. Additionally, by automating financial management using smart contracts, organizations can reduce the risk of errors and improve the accuracy of financial data.

Algorithmic Evaluation and Decision Support for Agile Factory Planning in the Context of the EU Taxonomy Regulation Using BIM Method

Daniel Wentzek (daniel.wentzek@tu-dortmund.de)
Supervisor: Univ.-Prof. Dr.-Ing. M. Gralla

This dissertation project focuses on the algorithmic assessment and decision support for agile factory planning through the lens of the EU Taxonomy Regulation using the Building Information Modeling (BIM) method. As the globalization era and the decentralization of manufacturing demand increasingly agile factory infrastructures to accommodate short product cycles and dynamic market conditions, the integration of production and construction planning becomes crucial for optimizing efficiency in factory planning processes.

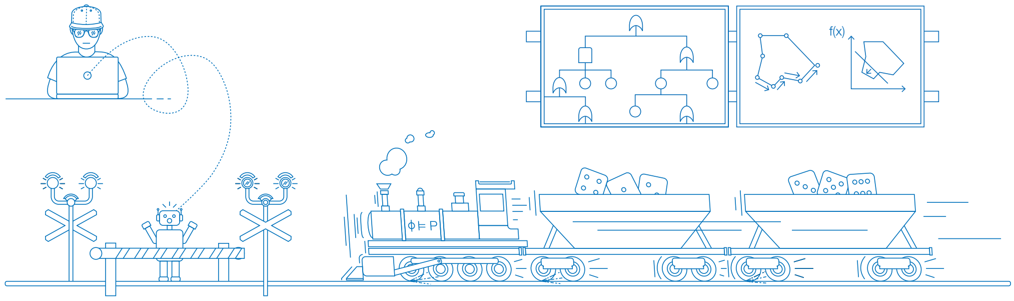
The project explores the role of the advancing digitalization in the industry, particularly the adoption of BIM, which is gaining prominence for its ability to create digital twins of properties, enhancing interoperability, transparency, and cooperation among project participants. This research investigates how factory layout planning can comply with existing data models and software used by stakeholders, enabling collaborative work within a dynamic BIM model. This approach is particularly relevant in light of the European Union's Green Deal and the accompanying EU Taxonomy Regulation, which aims to direct capital flows into sustainable economic activities, thereby significantly impacting the entire lifecycle of factory properties.

The dissertation seeks to develop practical recommendations for the categorization and evaluation of factory properties based on technical criteria from the EU Taxonomy, visualized and analyzed through the BIM method. The research includes an evaluation of digitalization implementation at factory sites, identification of potential research partners, and differentiation in data management across various scales of operation, ensuring referential integrity for algorithmic evaluation.

The research questions address the relevant technical criteria from the EU Taxonomy for project stakeholders, the necessary categorization and evaluation for visualizing the desired assessment, the lifecycle stages considered by the EU Taxonomy, its impact on collaborative factory planning, and the synchronization of digital tools and software between construction and process planning for holistic lifecycle evaluation.

GRK 2236: UNcertainty and Randomness in Algorithms, VErification, and Logic

Prof. Dr. Ir. Dr. h.c. Joost-Pieter Katoen (PDEng)
Email: katoen@cs.rwth-aachen.de
Rheinisch-Westfälische Technische Hochschule Aachen
Internet: www.unravel.rwth-aachen.de



Uncertainty is nowadays more and more pervasive in computer science. It is important both in big data and at the level of events and control. Applications have to treat large amounts of data, often from unreliable sources such as noisy sensors and untrusted web pages. Data may also be subject to continuous changes, may come in different formats, and is often incomplete. Robots, trains, and production machines have to deal with unpredictable environments. The growing use of machine-learning components — often providing weak guarantees — forms an additional factor of uncertainty. Probabilistic modelling and randomisation are key techniques for dealing with uncertainty.

Many trends witness this. Probabilistic programming exceeds the capabilities of probabilistic graphical models and automates statistical inference. Probabilistic databases deal with noisy data by associating probabilities to the possible worlds. Probabilistic model checking emerged as a key systems verification technique allowing to integrate correctness checking and performance analysis. Similar developments take place in automata, logic, and game theory.

The pervasiveness of uncertainty urges to make substantial enhancements in probabilistic modelling and reasoning so as to get deeper insight into, reason about, and master uncertainty. The aim of this RTG is and was to significantly advance various theoretical concepts (in algorithms, logic, verification) as well as their connection to deal with uncertainty and randomness, and to tailor and apply these techniques to problems in application areas such as railway engineering, network dynamics, and cyber-physical systems. This challenge is faced by a unique mixture of scientists from theoretical and applied computer science, management science, mechanical engineering, and railway engineering.

The qualification and supervision concept aims at offering the Ph.D. students an optimal environment to carry out their research. Every Ph.D. student has two supervisors; the rights and duties of the supervisors and students are laid down in a written supervision agreement. Progress and quality control is realised through regular individual meetings with the supervisors and regular talks at the RTG events. The curriculum consists of bi-weekly research seminars, soft-skill courses, reading groups, workshops (twice per year), a summer school in the first Ph.D. year, and (various new) advanced lectures.

Flows over Time and Their Applications

Emma Ahrens (ahrens@cs.rwth-aachen.de)
Supervisor: Prof. Dr. Joost-Pieter Katoen

This doctoral research investigates "weighted programming", a novel programming paradigm designed for describing and solving online optimization problems. The primary focus involves establishing mathematical foundations for automated, deductive procedures in the analysis of "weighted" programs. The study investigates the extendability of methods and techniques used in the analysis of probabilistic programs, particularly emphasizing the definition of proof rules for loops, demonstrating their correctness, and exploring the potential automation of loop invariant verification and synthesis.

Furthermore, the research explores the relationship of "weighted programming" to weighted Kleene algebras and weighted automata and investigates how suitable optimization problems may be modelled and solved via this new programming paradigm.

Robust Appointment Scheduling in Hospitals

Mariia Anapolska (anapolska@math2.rwth-aachen.de)
Supervisor: Prof. Dr. Christina Büsing

Introduction. As the demand for health care services increases each year, the need for efficient management of health care systems becomes more and more apparent. One of the most important health care providers are hospitals. Hospitals are under tremendous cost pressure and must achieve a balance between economic efficiency and a treatment that focuses on the patient. To improve clinical operations and patient safety, my research considers the appointment scheduling problem within a hospital.

Problem description. The problem aims to maximize the utilization of the hospital resources while minimizing the patients' inconveniences such as waiting time. Typically, an arriving patient needs to undergo several types of treatment. This means that several hospital resources will be needed either simultaneously or sequentially in a short time period. The treatments must be scheduled so that they satisfy the resource capacity restrictions. The hospital environment is very dynamic: The length of patients' treatments varies and arriving patients represent an uncertain demand for resources. The presence of emergency patients requires the schedule to be highly adaptable, i. e., robust and stable solutions are needed.

Envisioned work. Solutions of robust optimization problems depend on the uncertainty sets constituting the problem's input. In robust optimization, researchers assume these sets to be given by experts. However, experts often do not understand the dynamics within robust optimization, e.g., that integrating scenarios with high fluctuations leads to unpredictably high costs. Furthermore, especially in the hospital context, even for experts it is quite difficult to measure and obtain all data needed for presenting a scenario. To overcome this obstacle, we will use agent-based simulation to obtain all important parameters. To that end, the simulation framework "SiM-Care" [1] developed by Martin Comis needs to be extended and adapted. This agent-based simulation models interactions between the population and the physicians in a primary care system. It evaluates the input health care system by computing performance indicators that characterize the system's efficiency both from patients' and physicians' points of view. Moreover, the simulation allows us to assess the impact of changes in the system, such as changes in the patient-to-physician ratio or novel management strategies of physicians.

In order to obtain realistic input scenarios for the appointment scheduling problem, we plan to extend the model of `SiM-Care` further in order to integrate emergency and elective patients requiring hospital treatment. Since `SiM-Care` produces scenarios based on parameterized probability distributions, we will investigate the influence of the uncertainty sets for demands generated by `SiM-Care` on the resulting solutions for the robust appointment scheduling problem.

[1] Martin Comis, Catherine Cleophas, Christina Büsing, "Patients, Primary Care, and Policy: Simulation Modeling for Health Care Decision Support," arXiv.org (2019), no. 1910.11027, https://arxiv.org/abs/1910.11027

Automated Verification of Partially Observable Stochastic Models

Alexander Bork (alexander.bork@cs.rwth-aachen.de)
Supervisor: Prof. Dr. Joost-Pieter Katoen

Stochastic models like Markov decision processes (MDPs) and stochastic games are formalisms used in a wide array of domains to model systems where uncertainty and non-determinism are present. They assume perfect information about the state of the system at any time. Thus, these models often optimistically overestimate the amount of information available to a decision procedure to determine an optimal course of action for a given objective. In reality, a system's complete state is often hidden. *Partially observable* probabilistic systems extend the commonly used models with the notion that only *part* of the system's state is observable and decisions must be made based only on the observable information. These systems find application in fields like artificial intelligence, robotics and economics. Their analysis is significantly more involved compared to the fully observable case. Intuitively, this is due to the significantly increased dependence on the information history for making optimal decisions.

The research project focuses on the analysis and computer-aided verification of *partially observable MDPs (POMDPs)*. As perfect state information is not available, an optimal choice of action to satisfy a property is based on an estimate of the probabilities to be in states of the POMDP given the history of observations. These estimates are known as *beliefs*. These beliefs can be used to construct a fully observable, but typically infinite, *belief MDP* that captures the semantics of the POMDP.

Fully observable MDPs are well-studied in theory and many forms of analysis are tractable in practice, in particular for finite-state MDPs. However, even fundamental problems like the reachability problem (the question if a state of the system can be reached with a given probability at some point in time) are generally undecidable for finite-state POMDPs. As a consequence, related work typically restricts the considered properties to be bounded in time steps (*finite horizon*) or applies discounting in the computation to guarantee convergence. This, however, can severely distort results if a thorough analysis is desired.

The goal of the project is to develop novel, practically applicable verification methods for POMDPs in the *infinite horizon without discounting*. As such, finite abstractions of the belief MDP for a given POMDP are used to provide both upper[1] and lower bounds on the optimal value. These approximation algorithms leverage the existing knowledge in model checking finite MDPs to reason about POMDPs. An implementation of the approach as an extension of the probabilistic model checker Storm[2] shows its practical applicability.

Possible future directions include exploring compositional approaches to POMDP analysis, necessitating the definition of a compositional POMDP semantics and the development of model checking approaches for *partially observable stochastic games*.

[1] Bork, A., Junges, S., Katoen, J.-P., and Quatmann, T., "Verification of indefinite-horizon POMDPs", Automated Technology for Verification and Analysis. ATVA 2020. Lecture Notes in Computer Science, vol. 12302, pp. 288-304, 2020

[2] Hensel, C., Junges, S., Katoen, J.-P., Quatmann, T., and Volk, M., "The Probabilistic Model Checker Storm", STTT, 2021

Robust Hospital Management

Tabea Brandt (krabs@math2.rwth-aachen.de)
Supervisor: Prof. Dr. Christina Büsing

Hospitals are under tremendous cost pressure and must achieve a balance between economic efficiency and a treatment that focuses on the patient. To improve clinical operations and patient safety, methods from economics, mathematical optimization and IT-driven management systems are imported into the operational management of hospitals. The goal is to maintain the high quality in medical care while lowering the costs. A major challenge in this optimization process is the changing demand arising from emergencies or patients without appointments, which are difficult to forecast, and thus are, in general, not integrated into the planning process. In this part of the project we will focus on the integration of such uncertainties into three main areas of hospital management:

1. the operational planning and utilization of hospital beds,
2. the patient appointment scheduling, and
3. the transportation from patients to their appointments.

In the next subsection we will give a rough overview of existing scientific work in the mentioned subproblems. Finally, we will describe our approach to these problems in detail.

In 2012, Hulshof et al. [14] published a detailed bibliography and taxonomic classification on methods from operations research applied to problems in health care. Uncertainties are part of most decision problems in planning and controlling in health care. Mainly methods from queuing theory, Markov processes, and stochastic programming are used to include them into the optimization process, e.g., [1,2,3,9,13]. Besides dealing with uncertainties, [14] identifies the challenge for researchers to develop integral models of different hierarchical planning levels and services in health care.

The location of beds and the assignment of patients to these beds in a hospital is studied in operations research at the strategical, tactical and operational level. To support strategic planning queuing techniques, simulation and models from mathematical programming are already used. Traditionally, these planning decisions are based on target occupancy levels. However, Green [36] points out that, due to high fluctuations, different measurements such as patient waiting time [5] or patient refusal rate [18] need to be integrated into the optimization process. In [17], Ma and Demeulemeester combine the allocation of beds with the appointment of elective patients. In order to integrate emergencies, they reserve a fixed capacity. The Patient-to-Bed Assignment Problem on an operational level has been formalized in 2010 by Demeester et al. [8]. They use a combination of a patient-bed-suitability rating, the number of inpatient transfers and the number of mixed-gender-occupied rooms as the objective function and propose a hybrid tabu search algorithm for this problem. Later, the problem is reformulated to patient-to-room assignment, as it is generally assumed that all beds, located in the same room, are equal. Also more practical variants and other exact and heuristic approaches for patient-to-room assignment have been published, e.g., [6,7,16].

Vehicle routing problems are well-studied in discrete optimization [10]. In the context of patient routing within the hospital, Hanne et al. [12] designed a computer-based planning system. Johnson et al. [15] introduced a simulation tool, and Beaudry et al. [4] a two-phase heuristic to solve the dynamic problem. Schmid and Doerner [19] solved the combination of operating room scheduling and transportation with a hybrid metaheuristic.

So far, we concentrated on the operational patient-to-room assignment. Hospital beds are a special resource in a hospital. According to the number of beds the capacity of a hospital is measured and, thereby, the size of wards and clinics are given by this number and the corresponding budget on medical and nursing staff is determined by this number. Yet, the number of available beds fluctuates due to capacity changes in the nursing staff, patient demands and special needs of patients [11]. These fluctuations primarily affect the scheduling of elective patients and the daily allocation of emergency patients to different wards and rooms. In the case of a mismatch of available beds to admitted patients, a relocation of a bed or of a patient to a different clinic or ward, or the rejection of elective patients is possible. However, such means should only be used in extreme situations and not on a daily basis.

Contrary to all previously published work, we do not regard a weighted combination of the patient-bed-suitability rating, the number of inpatient transfers and the number mixed-gender-occupied rooms as the objective function. Choosing appropriate weights is very challenging and, also, no procedure has yet been proposed to check afterward if good weights have been chosen. Also, using a weighted combination prevents us from gaining better insights into how the different objectives influence each other. For this reason we keep the three different aspects separated and treat them as independent objective functions. We compare and develop exact and heuristic approaches to solve the multi-objective patient-to-room assignment problem with a focus on robust solutions.

References:

1. R. Akkerman and M. Knip. Reallocation of beds to reduce waiting time for cardiac surgery. Health Care Management Science, 7:119–126, 2004.
2. M. Asaduzzaman, T.J. Chaussalet, and N.J. Robertson. A loss network model with overflow for capacity planning of a neonatal unit. Annals of Operations Research, 178:67–76, 2010.
3. S. Batun, B.T. Denton, T.R. Huschka, and A.J. Schaefer. Operating room pooling and parallel surgery processing under uncertainty. INFORMS Journal on Computing, 23:220–237, 2011.
4. A. Beaudry, G. Laporte, T. Melo, and S. Nickel. Dynamic transportation of patients in hospitals. OR spectrum, 32:77–107, 2010.
5. P. Van Berkel and J. Blake. A comprehensive simulation for wait time reduction and capacity planning applied in general surgery. Health Care Management Science, 7:373–385, 2007.
6. S. Ceschia and A. Schaerf. Local search and lower bounds for the patient admission scheduling problem. Computers and Operations Research, 38(10):1452–1463, 2011
7. S. Ceschia and A. Schaerf. Modeling and solving the dynamic patient admission scheduling problem under uncertainty. Artificial Intelligence in Medicine, 56(3): 199–205, 2012.
8. P. Demeester, W. Souffriau, P. D. Causmaecker, and G. V. Berghe. A hybrid tabu search algorithm for automatically assigning patients to beds. Artificial Intelligence in Medicine, 48(1):61–70, 2010.
9. G. Dobson, H.H. Lee, and E. Pinker. A model of icu bumping. Operations Research, 58:1564–1576, 2010.
10. B.L. Golden, S. Raghavan, and E.A. Wasil, editors. The Vehicle Routing Problem: Latest Advances and New Challenges. Springer, 2008.
11. L.V. Green. Capacity planning and management in hospitals. In Operations Research and Health Care: A Handbook of Methods and Applications, pages 15–41. Kluwer Academic Publishers, Boston, 2004.
12. T. Hanne, T. Melo, and S. Nickel. Bringing robustness to patient flow management through optimized patient transports in hospitals. Interfaces, 39:241–255, 2009.

13. P.R. Harper, A.K. Shahani, J.E. Gallagher, and C. Bowie. Planning health services with explicit geographical considerations: a stochastic location-allocation approach. Omega, 33:141–152, 2005.

14. P. Hulshof, N. Kortbeek, R. Boucherie, E. Hans, and P. Bakker. Taxonomic classification of planning decisions in health care: a structured review of the state of the art in or/ms. Health Systems, 1:129–175, 2012.

15. K. Johnson, D. Kalowitz, J. Kellegrew, B. Kubic, J. Lim, J. Silberholz, A. Simpson, E. Sze, E. Taneja, and E. Tao. Emergency department efficiency in an academic hospital: A simulation study. Ph.D. Dissertation, Univ. of Maryland, 2010.

16. R. M. Lusby, M. Schwierz, T. M. Range, and J. Larsen. An adaptive large neighborhood search procedure applied to the dynamic patient admission scheduling problem. AI in Medicine, 74:21–31, 2016.

17. G. Ma and E. Demeulemeester. A multilevel integrative approach to hospital case mix and capacity planning. Computers and Operations Research, 40: 2198–2207, 2013.

18. A.K. Shahani P.R. Harper. Modelling for the planning and management of bed capacities in hospitals. Journal of the Operational Research Society, 53:11–18, 2002.

19. V. Schmid and K. Doerner. Examination and operating room scheduling including optimization of intrahospital routing. Transportation Science, 48: 59–77, 2013.

Adapting to Changing Environments in Control

Paul Brunzema (paul.brunzema@dsme.rwth-aachen.de)
Supervisor: Prof. Dr. Sebastian Trimpe

Model-based control uses a model of the system dynamics to find good control actions by predicting the evolution of the system over time. Modeling complex models from first principles can be difficult, so learning dynamics directly from data, utilizing e.g. neural networks, is desirable. This presents several new challenges, such as training data often being uncertain and sparse; in addition, training such a model can be costly. Moreover, real dynamical systems can undergo changes due to factors such as wear and tear, which can lead to poor control performance.

Such dynamical systems subject to uncertainty and change are at the core of my research. One aspect is the question of how to detect these changes in dynamics based on data, with the goal of (re)learning / adapting a dynamics model only when necessary. We do this by using event triggers and by exploiting uncertainty estimates, e.g. from Bayesian neural networks. Once these changes are detected, the question arises how to efficiently adapt the model to the changed dynamics to maximize the control performance over time. For this, architectures such as neural processes are promising. As not all control methods rely on a model for online decision making, I also aim to make classical control methods (such as PID control) adaptive to changing dynamics. Here, I combine the event-triggering ideas mentioned above with black-box optimization methods such as Bayesian optimization to optimize an unknown performance function in a changing environment.

Design and Analysis of Algorithms for Combinatorial Optimization Problems under Uncertainties

Katharina Eickhoff (katharina.eickhoff@oms.rwth-aachen.de)
Supervisor: Prof. Dr. Britta Peis

Matching Markets

Matchings appear in many combinatorial optimization models of applications where assignments between two parties (sellers and buyers, students and courses, ...) have to be found. In these examples each player has preferences to which he would like to be matched. Often, prices might be used to regulate imbalances between supplies and demands.

A possible aim is to find assignments and prices such that everyone is happy, i.e., with these prices no one prefers to trade with someone else instead of the assigned person. These prices are called equilibrium prices. If furthermore as much as possible is sold we call the prices market-clearing. Prices which are competitive and market-clearing describing the set of Walrasian prices. One possibility to find Walrasian prices is by a price raising auction[1] [2] [3]. To find the set of objects whose prices should be raised is quite complicated in the general case. We could show that these sets could be found by a max-flow computation in case of linear valuations. For the general case of gross substitute valuation functions, we simplified it to polymatroid sum computation and reachability in an exchange graph. Furthermore, we could give sensitivity results for the prices if the demand or supply in the matching-market changes.

There are many ways to expand this approach which we might consider in the future. One example are two-stage variants. In the first stage, agents decide on a strategy based on probability distributions of the agents' valuations. The agents are allowed to switch their strategies in a given neighborhood in the second stage when the true valuations are common knowledge. For example, the agents decide on a strategy based on guesses of the valuations and they can adapt their strategies in a given scope in the real scenario. The objective of an agent is to maximize the expected profit.

Furthermore, we like to consider the setting with risk-averse agents. They prefer a robust solution within all possible situations which means that the profit they receive in the worst-case scenario should be maximized. The strategies of the agents are in equilibrium if each strategy is the best robust response given the other strategies. We study the existence of equilibria in such markets. If equilibria exist, we like to analyze the complexity and design algorithms to compute or approximate them.

Stackelberg Network Pricing Games

Consider a game with to players. The leader can choose prices for some items of an underling network in the first stage. Afterwards, in the second stage, the follower chooses the items which yields a min cost solution of his optimization problem (e.g. matching, vertex cover, closure). Most of these problems are NP-hard in general, but if the underling

[1] L.M. Ausubel, "An efficient dynamic auction for heterogeneous commodities." American Economic Review, vol. 96(3), p. 602–629, 2006

[2] G. Demange, D. Gale and M. Sotomayor, "Multi-item auctions." Journal of political economy, vol. 94.4, p. 863-872, 1986

[3] K. Murota, A. Shioura, Z. Yang, "omputing a walrasian equilibrium in iterative auctions with multiple differentiated items." International Symposium on Algorithms and Computation, p. 468–478, 2013

network or the set of priceable items is restricted there might be a polynomial algorithm to solve it.
We consider the Stackelberg Bipartite Vertex Cover Problem, which is NP-hard[4]. It is known that the problem is polynomial solvable if the pricable vertices are on one side of the bipartition[5]. We like to show similar results for pricable vertices on both sides but if the underling graph has a special structure, e.g. if it is a path or a tree.

[4] K. Jungnitsch, B. Peis, M.Schröder: "Stackelberg Max Weight Closure with Multiple Followers." Mathematics of Operations Research 47.4, INFORMS, 2022

[5] P. Briest, M. Hoefer, P. Krysta: "Stackelberg network pricing games." Algorithmica 62.3, p. 733-753, 2012.

Calculating the capacity of railway systems considering microscopic infrastructure constraints

Tamme Emunds (emunds@via.rwth-aachen.de)
Supervisor: Prof. Dr. Nils Nießen

With rising demand on public transportation due to political and environmental influences, railway transportation is subject to rising requirements on quality, quantity and efficiency of the used infrastructure. To fulfill those enhanced needs, infrastructure managers are required to identify bottlenecks in existing infrastructure and precisely estimate the capacity of newly constructed or extended infrastructure. In many cases the stations turn out to be the bottleneck in the railway network.

Developing methods for the analyzation of railway infrastructure capacity has therefore been one of the major suspects of interest to railway researchers. While sufficient methods for the estimation of railway lines have already been developed and heavily used for planning purposes, the capacity analysis of entire railway stations remains a challenging research question.

In this project, new methods for quantifying the capacity of railway stations will be developed and analyzed. The primary focus will be laid on the development of efficient algorithms to estimate the theoretical capacity of railway infrastructure in stations. Further, methodologies to quantify the influence of disturbances from multiple sources – in example technical failures, maintenance work or peaks in the transported traffic volume – to the infrastructure capacity will be developed and analyzed.

Accelerating Robust Optimisation

Felix Engelhardt (engelhardt@combi.rwth-aachen.de)
Supervisor: Prof. Dr. Christina Büsing

Robust optimisation with budget uncertainty is both interpretable and lends itself to elegant mathematics. However, solving these problems computationally still provides major challenges for practitioners. As part of my PhD, I work on generalising theoretical results on branching and cutting planes for robust optimisation with budget uncertainty in order to speed up computations for large and complex real-world optimisation problems. Specifically, this is applied to energy system modelling, where short-term modelling is already being done using robust/stochastic optimisation, but long term planning is done deterministically, as current algorithms and computers can not solve realistically sized instances within reasonable timeframes.

Representing and Reasoning about Uncertain and Incomplete Beliefs in Multi-agent Domains

Qihui Feng (feng@kbsg.rwth-aachen.de))
Supervisor: Prof. Dr. Gerhard Lakemeyer

Many scenarios require expressive languages that support the ability to represent an agent's incomplete knowledge and beliefs as well as probabilistic uncertainty about beliefs. Devising knowledge bases (KBs) of uncertain and incomplete beliefs in multi-agent domains allows for the explicit representation of an agent's complex beliefs about the world and other agents' mental states, which paves the way to further develop approaches for automated reasoning, epistemic planning and games.

This research project is concerned with two goals:

The first is to design suitable formalism for probabilistic KBs in multi-agent domains and to dynamically update a KB as the result of actions. In the static case, we already proposed a logical framework that precisely captures the beliefs and non-beliefs of agents in a hierarchical manner [1]. For future work, we consider extending this to dynamic cases and exploring mechanisms for projection reasoning such as progression and regression.

Furthermore, it is widely recognized that problems such as the validity of multi-agent modal sentences are intractable, even for the non-probabilistic propositional fragment. It is of great interest to explore tractable fragments of the language or to develop approaches for reasoning with a practical utility. Currently, we are pursuing two directions: The first involves developing an epistemic variant of binary decision diagrams to represent and reason about propositional modal sentences in a compact manner. The second approach is to translate first-order modal formulae into classical FOL, thereby reducing the SAT and validity problem of modal formulae to non-modal cases, which allows the use of existing first-order theorem provers to prove modal theorems.

[1] Probabilistic Multi-agent Only-believing. Feng, Q. and Lakemeyer, G. In Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems, 2024

Optimization under Uncertainty

Dennis Fischer (fischer@algo.rwth-aachen.de)
Supervisor: Prof. Dr. Christina Büsing

Many optimization algorithms make the assumption that the input to problem is completely known in advance. This is not always true in practice. In practice we often have to make decisions before all the data about the problem are known. A further problem is that we may not have complete information since the data we have to work with are not completely accurate. This is due to the way data is acquired which introduces uncertainty, for example, perhaps the sensor used only gives us an approximation of the actual value.

Nonetheless, we want to be able to make decisions in these cases. It is clear that we cannot hope to always find the optimal solution that fits the actual data but we want to find a solution that gives guarantees about objective value in comparison to the achievable value if the input is known.

In my work I study these kinds of robust optimization problems.

One way of approaching robust optimization problems is to consider a 2 player game. The first player (the algorithm) is presented with a (possibly infinite) set of possible inputs. The algorithm has to fix an output. Now, in a second stage, the second player (the adversary) picks one of those inputs from the set that causes the worst possible performance for the algorithm.

One of these 2-player problems is the Continuous Knapsack Problem (CKP). In the CKP, player 1, the leader, packs some items (or fractional parts of items) into their knapsack. In the second stage, player 2, the follower, chooses items (or fractions of items) from the set of items already chosen by player 1 to pack into their knapsack thereby trying to optimize their gain. The leader's objective is to minimize the follower's objective. In a recent paper it has been shown to be solvable in time $O(n^2)$ [1]. We were able to improve this running time in [2] to $O(n \log n)$.

One other robust optimization problem is the Recoverable Robust Assignment problem in which on a balanced bipartite graph with $2n$ vertices for given linear cost functions c_1 and c_2 the task is to find matchings M_1 and M_2 that have at least k edges in common while minimizing $c_1(M_1)+c_2(M_2)$. In joint work with Hartmann, Lendl, and Woeginger we were able to show W[1] hardness for parameter k and parameter $n-k$ even in very restricted special cases. We also showed that it is polynomial time solvable if the cost functions are restricted to being Monge and Anti-Monge. In the case where one of the matchings is fixed we showed that the Recoverable Robust Assignment problem is contained in RNC2 while being at least as hard as the well-known Exact Matching in Red-Blue Bipartite Graphs whose complexity is a long-standing open problem. These results are not published yet.

Another problem is the bilevel bottleneck assignment problem. In this problem a bipartite graph is given. The edges are split into a leader and follower set. The leader and follower have (different) cost functions for the edges. First the leader selects edges that form a matching from their leader set. Then the follower selects edges from the follower

[1] Margarida Carvalho, Andrea Lodi, PatriceMarcotte, "A polynomial algorithm for a continuous bilevel knapsack problem Oper," Res. Lett, vol. 46(2), p. 185–188, 2018

[2] Dennis Fischer, Gerhard J. Woeginger, "A faster algorithm for the continuous bilevel knapsack problem," Oper. Res. Lett., vol. 48(6), p. 784–786, 2020

set to complete the leader matching to a perfect matching. The goal of the leader is to minimize the largest used edge according to the leader cost function. The goal for the follower is to minimize the largest used edge according to the follower's objective function. In joint work with Muluk and Woeginger we showed that this problem is NP complete.

Another project is joint work with the UnRAVeL members Tauer, Fuchs, Koch, and Ziegler in which we looked at complexity results in a train routing problem [3]. This train routing problem is a generalization of packet routing without buffers. We distinguished the case where the train depots are part of the network or not and showed various complexity results on different networks.

[3] Bjoern Tauer, Dennis Fischer, Janosch Fuchs, Laura Vargas Koch, Stephan Zieger, "Waiting for Trains: Complexity Results," CALDAM 2020, p. 282–303, 2020

Probabilistic Hyperproperties

Carolina Gerlach (gerlach@cs.rwth-aachen.de)
Supervisor: Prof. Dr. Erika Ábrahám

Model checking is traditionally concerned with analyzing whether a certain model satisfies a specific trace propertyd. A trace is a sequence of observable facts about the states visited along a system execution. A trace property, formally specified using temporal logics like LTL or CTL, can be viewed as a set of model traces, encoding requirements on the system executions.

However, many interesting requirements, such as information-flow security policies, cannot be expressed as trace properties. For example, the property of non-interference for a deterministic system requires that the values of secret input variables do not influence the observable values of public output variables. In order to check whether this property holds, we need to compare different program executions that start in observationally equivalent initial configurations, i.e., with the same initial public variable values but possibly different secret inputs. Non-interference is satisfied if we can make the same observations along all executions with observationally equivalent initial configurations, i.e., generating the same public variable values. Hence, non-interference cannot be expressed as a set of traces and is therefore not a trace property. Instead, security policies like noninterference are hyperproperties, which are sets of sets of traces.

Since established temporal logics can only capture sets of traces, but not sets of sets of traces, several temporal logics have been extended to hyperlogics. Clarkson et al. generalized LTL and CTL* to HyperLTL and HyperCTL*, respectively, by adding explicit quantification over multiple traces. Even though CTL* already permits quantification over traces, it does not allow quantifying over several traces at the same time, which HyperCTL* now does.

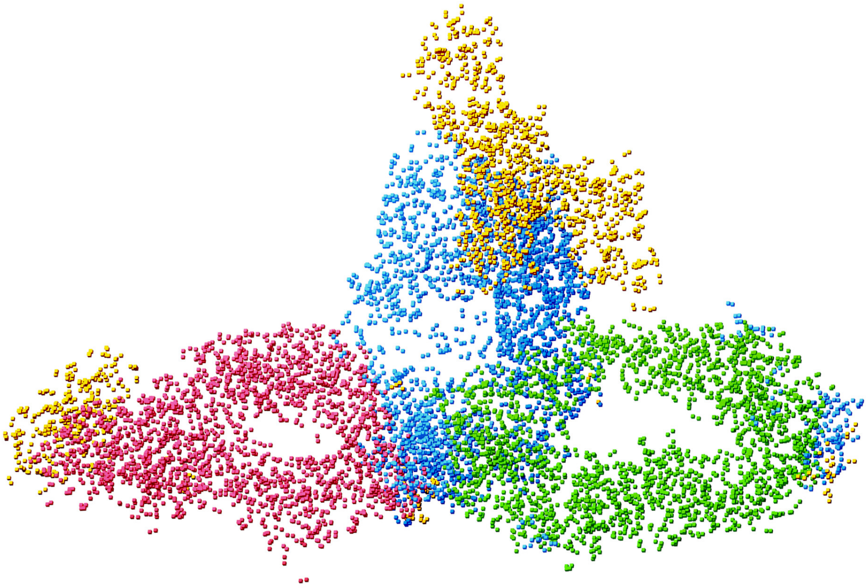
Probabilistic Hyperlogics. HyperPCTL was the first temporal logic for probabilistic hyperproperties. HyperPCTL extends PCTL by quantification over states to express probabilistic relations between several computation trees. Like PCTL, HyperPCTL formulas are evaluated over Markov chains. They can express information-flow security policies like probabilistic noninterference, which stipulates that, for all possible values of the public variables, the probability of observing these values should be the same for all program executions with observationally equivalent initial states.

HyperPCTL was lifted to Markov decision processes (MDPs) by additionally adding quantification over schedulers that resolve nondeterministic choices probabilistically. Model checking HyperPCTL for discrete-time Markov chains is decidable. For MDPs, the model checking problem becomes in general undecidable, but restricting scheduler quantification to deterministic memoryless schedulers makes it decidable again.

*Research.*We investigate the relationship of HyperPCTL to other logics, examine the complexity of HyperPCTL model-checking, and study possible extension of HyperPCTL to, e.g., asynchronous semantics or stochastic games.

Randomness and Uncertainty in Signal Processing on Topological Spaces

Vincent Grande (grande@cs.rwth-aachen.de)
Supervisor: Prof. Dr. Michael Schaub



Proteins are molecules that consist of long strings of amino acid residues. They play an integral role in almost every cellular process from metabolism, DNA replication, to intra-cell logistics. Their diverse functions are hugely influenced by their complex 3d geometry, which arises by folding the chains of amino acid residues. In the figure: Clustering of NALCN Channelosome, a channel membrane protein, data by Kschonsak et al., 2022. The topological structure influencing the function of the protein consists of three loops, which our methods are able to identify correctly.

Higher-order Information Encoded in Networks and Point Clouds

My research deals with the analysis of higher-order information in networks and point clouds. A central motive is to use a blend of techniques from algebra, topology, and homotopy theory to explore the relationship between small scale connectivity data and other localised information, and large-scale behaviour and global properties of data sets.

Over the last years, availability of large, complex data sets has increased dramatically with recent advances in collection, storage and processing techniques. Extracting human-interpretable information from these is a challenging task across application scenarios. Standard methods of data analysis include projecting the data set onto subspaces with the highest variance (PCA), clustering the data points using closeness- or density-based clustering algorithms (k-means, spectral clustering, DBSCAN, etc.), or fitting polynomials or other functions to the data points.
However in many cases, the information encoded in the point cloud cannot be retrieved by these methods: The detection of loops, voids, and holes in the point cloud requires

tools that consider more than just proximity of points or local density. For example, these features are relevant when trying to analyse complex proteins consisting of multiple loops, or trying to understand the structure of a manifold in high-dimensional space from where the point cloud is sampled from.

From a different perspective, higher-order information are global features of the data-set. We cannot extract these from the **local** neighbourhood of a point of interest. However by combining all local information, higher-order **global** features arise. Because we are interested in information on the individual points, we finally need to find a way to **localise** these global features of the point cloud.

Combining Tools from Algebraic Topology and Network Science

Algebraic Topology is an area of Mathematics that was established trying to provide tools for capturing the "essence" of topological spaces. Topological features of a space are global and by design robust to local perturbations and noise, and are somewhat emergent properties of all the local connectivity data of the individual points. This is ideal for the setting of modern data analysis, where the goal is to extract information out of data sets where local perturbations are likely to occur because of noisy data collection. Current tools of topological data analysis provide a good way of generating these global invariants (Betti numbers, persistence landscapes, etc.) from local connectivity data of point clouds. However, there is considerably less work on relating back these global features to the local scale of the points with a real-world meaning. This is surprising, considering that extracting information like cluster assignments on a point level is highly useful in application scenarios and a key goal of many areas of data analysis. Network Science, and especially signal processing on networks, on the other hand, deals with similar problems trying to connect the local connectivity information on network nodes to the global behaviour of the network and dynamics on it, and then relating this back to the individual nodes.

The overarching theme of my research is to combine perspectives of algebraic topology, network science, signal processing and clasical data anlysis to develop tools from extracting higher order information from point clouds and networks.

Complexity and Algorithms in Optimization under Uncertainty

Christoph Grüne (gruene@algo.rwth-aachen.de)
Supervisor: Prof. Dr. Gerhard Woeginer

Introduction. Optimization under uncertainty is a field in which problems are optimized against some form of uncertainty. For this, finding measures of robustness to find solutions that deal with the given form of uncertainty is of interest. The project will focus on different measures of robustness and different complexity viewpoints to analyze certain forms of problems. Among those may be problems with one player against an adversary (the "nature"), two players against each other or multiple player settings playing against or with each other. That is, the uncertainty is modelled by an adversary player playing against the agent. The complexity analysis may be based on classical complexity classes as well as parameterized complexity.

The first project is on Recoverable Robustness with a Hamming-distance measure which shall encounter combinatorial uncertainty scenarios. In this setting, a solution S is given and for every possible scenario, which may occur in this setting, we can choose another solution, S', which differs in at most only k elements from solution S'. That is, we can recover from a harmful scenario by choosing a different solution, which is not too far away from the first solution.

The project surveys the complexity of k-Hamming-distance recoverable robust version of problems that are in NP for different types of scenarios among a constant number of arbitrary scenarios, Gamma-scenarios, and general scenarios for elements of the universe. The analysis is primarily based on classical complexity measures such as the polynomial hierarchy. There are already results that have to be formulated into a paper. The results contain a hardness proof for the recoverable robust version of the undirected s-t-path problem, which may extend to a variety of other problems. The aim is to provide a structural theorem that captures this very variety of combinatorial problems that have this hardness structure. The second project, which is currently planned, may inspect parameterized complexity counterparts to the classical complexity analysis of the first paper. Instead of NP problems, $W[t]$-problems and other problems in parameterized hierarchies are considered; they may have a similar or the same hardness structure.

References:

1. Christoph Grüne. Dial-a-Ride for Railway Traffic. Master Thesis, RWTH Aachen University 2019
2. Jörg Flum, Martin Grohe. Parameterized Complexity Theory, Springer, 1998.
3. G. Rodney Downey, M. R. Fellows. Parameterized Complexity, Springer, 1999.
4. Marek Cygan, Fedor V. Fomin, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx. Parameterized Algorithms, Springer, 2015.
5. Raymond Greenlaw, James Hoover, Walter L. Ruzzo. Limits to Parallel Computation: P-Completeness Theory, Oxford University Press, 1995.

Approaches to Slicing of Probabilistic Programs

Darion Haase (darion.haase@informatik.rwth-aachen.de)
Supervisor: Prof. Dr. Joost-Pieter Katoen

Program slicing is a well-established method for analyzing and understanding the behavior of classical deterministic programs. It has a wide range of applications in software maintenance, program comprehension, debugging, and optimization. Slicing aims to extract parts of the original program that behave equivalently with respect to the property of interest by focusing on specific semantic aspects, such as the values of certain variables or a particular part of the computation. Depending on this property of interest, the so-called slicing criterion, there exists a wide-variety of applicable techniques.

Probabilistic programs build on classical deterministic programs and introduce concepts such as sampling and conditioning from probability theory to the programming language. Instead of manipulating a single state, the programs now manipulate a distribution of states.

The aim of this research project is to explore program slicing approaches for probabilistic programs. Existing work on syntactic slicing builds upon classical program slicing approaches by relaxing the considered notion of program equivalence. Specifically, requiring slices to produce equivalent instead of identical distributions on the relevant variables introduces potential to improve slicing precision through the exclusion of observe statements. The starting point of this research lied in capturing the different approaches in a unifying manner and reducing them to well-known formalizations using the Program Dependence Graph. This view reveals that ideas of d-separation from Bayesian networks play a crucial role at the heart of the existing approaches, and opens up the exploration of different syntactic conditions for capturing of stochastic independence as part of this project.

An important application of program slicing lies in the area of automated program verification. As reasoning about programs is hard, and only further complicated by the presence of probabilistic behaviour, slicing offers a potential way to simplify verification and help developers in understanding the faults in their programs as reported by a verifier. Extending slicing approaches to a language useful for deductive verification faces fundamental challenges. So far, extensions to semantic slicing using expectation-based reasoning are mostly unexplored in the literature. On a fundamental level, this research needs to investigate definitions of semantics for probabilistic programs that can be used to formally capture the notion of slicing correctness, while being useful for automated verification. Consequent integration of these results into Caesar [1], a deductive verifier for probabilistic programs, allows for direct practical application in probabilistic program verification.

Further generalization of the discovered results to weighted programming, a generalization of probabilistic programming, may be of interest for future work.

[1] Philipp Schröer, Kevin Batz, Benjamin Lucien Kaminski, Joost-Pieter Katoen, Christoph Matheja. A Deductive Verification Infrastructure for Probabilistic Programs. Proceedings of the ACM on Programming Languages 7, OOPSLA 2023. DOI: https://doi.org/10.1145/3622870.

Approximate Model-Predictive Control for Dynamic Multi-Robot Systems

Henrik Hose (henrik.hose@dsme.rwth-aachen.de)
Supervisor: Prof. Dr. Sebastian Trimpe

Fast feedback responses, stability, and constraint satisfaction are critical requirements for control in robotics to ensure safety. Model predictive control (MPC) achieves stability and constraint satisfaction, but is notoriously slow to evaluate. Approximation of such MPC controllers via (deep) neural networks (NNs) allows for fast online evaluation. However, the approximation introduces inaccuracies that can cause instabilities or constraint violations. In this project, novel methods for offline validation and safe online evaluation of approximations of MPC type controllers are developed. This work builds upon existing results in statistical offline validation, online safety certification in control, and explores the use of formal verification methods. Novel approximate MPC schemes with offline validation and safe online evaluation methods are evaluated in real-world problems from the robotics domain, such as the Wheelbot.

The Wheelbot, a small reaction wheel balancing robot, was originally developed at the DSME and MPI Stuttgart under the supervision of Prof. S. Trimpe. A video of the Wheelbot is available here [link]. The Wheelbot is a challenging robotics test bed for non-linear control when balancing, and even hybrid-systems with contact switches for stand-up maneuver. The next generation — the Mini Wheelbot — is engineered for production in small fleet quantities to serve as a hardware test bed at DSME.

Analyzing Termination and Expected Runtime Complexity for Probabilistic Term Rewriting

Jan-Christoph Kassing (kassing@cs.rwth-aachen.de)
Supervisor: Prof. Dr. Jürgen Giesl

Using random actions or selections is a very useful ingredient for the development of algorithms. It is typically used to change deterministic algorithms with bad worst-case behavior into efficient random algorithms which produce correct results with a high probability. The Rabin-Miller primality test, Freivalds' matrix multiplication, and the random pivot selection in Hoare's quicksort algorithm are prime examples. These kinds of algorithms can be elegantly expressed as probabilistic programs. Determining runtime and termination in the probabilistic case is a difficult problem with often unintuitive results. In the probabilistic case there are multiple notions of termination. Two of the most important ones are Almost Sure Termination (AST), i.e., the program terminates with probability 1, and (Strong) Positive Almost Sure Termination (PAST), i.e., the program terminates within a finite number of expected steps. Whereas in the deterministic case a single diverging infinite run leads to non-termination and infinite runtime, this is not the case for either notion of termination in the probabilistic case. There are approaches amenable to automation, but most current techniques only focus on programs on numbers and disregard programs operating on data structures such as lists or trees. In contrast, in the non-probabilistic setting, many powerful and automatic approaches have been developed to analyze termination and complexity of these types of programs using an automatic analysis of term rewriting systems.

This project deals with the challenging question of how to automatically determine the respective properties of probabilistic programs dealing with data structures. The focus of the project is on analyzing probabilistic term rewriting systems (PTRSs). There are some results for PTRSs which use polynomial and matrix orders to determine PAST, but as with the non-probabilistic case, these techniques alone are not very powerful. The key idea for termination analysis in the classical case was the introduction of dependency pairs and the resulting possibility of a modular analysis of a term rewriting system. Therefore, one of the goals of this project is the adaption of this analysis technique to the probabilistic case, to develop a fully automated technique for the termination analysis of PTRSs. We have already created an adaption of the dependency pair framework to automatically analyze innermost AST, where we only consider rewrite sequences that follow an innermost evaluation strategy. Currently, we are investigating different ideas from the non-probabilistic framework on how to increase the effectiveness and applicability of our newly developed framework.

Privacy Preserving Online Algorithms

Andreas Klinger (klinger@itsec.rwth-aachen.de)
Supervisor: Prof. Dr. Ulrike Meyer

Introduction: Secure multi-party computation (SMPC) allows parties to evaluate a function over their private inputs in a distributed fashion in such a way that each party only learns its prescribed output and anything it can deduce from combining its prescribed output with its own private input. An SMPC protocol describes the individual communication and computation steps each party follows during the distributed evaluation of the function it implements. Informally speaking, such a protocol is said to securely evaluate the function if it correctly computes the prescribed output of the function and if during the evaluation, the parties do not learn anything that goes beyond what they would learn if the function was evaluated centrally by a trusted third party (TTP) - even in the presence of an adversary. However, the general assumption underlying SMPC is that such a TTP that is incorruptable and trusted by everyone does not exist. Therefore, the parties execute an SMPC protocol to simulate the behavior of a TTP.

Secure function evaluation (SFE), a special case of SMPC, considers functions that could - if they were evaluated by a TTP - be evaluated by a single run of an (offline) algorithm that takes the inputs of a fixed number of parties and computes the desired output for each party. The security notions of SFE can also be extended to reactive SMPC to cover reactive algorithms [1] which allow an a priori known fixed number of parties to provide input over multiple rounds and obtain output in each round. In addition, the output can depend on a state, which itself depends on all previous inputs and outputs. However, the prominent set of functions or problems that can be solved with the help of online algorithms [2] has not yet been considered in SMPC: Online algorithms receive events one after another and for each event they have to decide immediately how to deal with it. The main difference between offline/reactive and online algorithms is that whereas the participating parties are fixed for offline/reactive algorithms, parties can dynamically join and leave in online algorithms.

Secure and privacy preserving protocols for offline (bipartite) matching is an important research field[3][4][5][6]. However, using online matching algorithms instead has the added benefit of not requiring to restart the complete computation as soon as a new party joins and thus reducing waiting times. Such online matching problems naturally arise for example when open job positions are to be filled by applicants, or students apply for

[1] Oded Goldreich, “Foundations of Cryptography: Basic Applications, Cambridge University Press, 2004

[2] Amos Fiat and Gerhard Woeginger,Online Algorithms: The State of the Art. Springer, 1998

[3] Philippe Golle. A Private Stable Matching Algorithm. In International Conference on Financial Cryptography and Data Security, pages 65–80. Springer, 2006

[4] Marina Blanton and Siddharth Saraph. Secure and Oblivious Maximum Bipartite Matching Size Algorithm with Applications to Secure Fingerprint Identification, 2014

[5] B. Anandan and C. Clifton. Secure minimum weighted bipartite matching. In 2017 IEEE Conference on Dependable and Secure Computing, pages 60–67, 2017

[6] Stefan Wüller, Michael Vu, Ulrike Meyer, and Susanne Wetzel. Using Secure Graph Algorithms for the Privacy-Preserving Identification of Optimal Bartering Opportunities. In Proceedings of the 2017 on Workshop on Privacy in the Electronic Society, WPES ’17, pages 123–132. Association for Computing Machinery, 2017

internship offers and need to be told immediately whether or not they are accepted. The aim of this dissertation project is to analyze how to securely evaluate online algorithms in a privacy-preserving fashion. We want to model the problem in a general fashion and find its limitations in regard to securely realizing such online functionalities. The main research questions are: How to provide privacy and security for online matching algorithms? What are the limitations? To what degree can privacy be provided?

Model: Ideally evaluating an online functionality such as online matching in a privacy-preserving way does not only entail protecting the inputs and outputs of all parties involved, and securely keeping a state over a changing set of parties, but also hiding the arrival and departure of parties from the other participating parties. This includes hiding the point in time when a party provides input. We therefore propose several new models for online TTPs that can evaluate online functionalities, i. e., the mapping of inputs and outputs of online algorithms.

Calculation of capacities in railway networks under the consideration of network effects

Maren Maus (maus@via.rwth-aachen.de)
Supervisor: Prof. Dr. Nils Nießen

Capacity calculations are essential for the long-term planning of railway infrastructure. The calculations are influenced by many different parameters such as the train mixture, the delay behavior and the intended operational quality. Many of the methods currently used in practice calculate characteristic capacity values separately for the single elements of a railway network, namely lines, track groups and route nodes. Due to the different calculation approaches, the determined capacity values are not directly comparable with each other and interactions with the surrounding infrastructure elements are not sufficiently considered.

In this project an optimization model that examines the possible capacity utilization of railway lines and nodes while taking network effects into account is developed. To quantify network effects, the resulting capacities are compared to the values of the single elements calculated via the state-of-the-art methods.

Automated Complexity Analysis of Probabilistic Programs

Eleanore Meyer (eleanore.meyer@cs.rwth-aachen.de)
Supervisor: Prof. Dr. Jürgen Giesl

In recent years, the study of probabilistic programs and methods to ensure their correctness has been an active field of research. Probabilistic programs are classical programs that are enriched with a notion of probabilistic choice. Such programs may then for instance branch on the outcome of a coin flip or assign a value that is sampled according to a probability distribution to a program variable.

One of the most important correctness properties of programs is their termination behaviour. When compared to classical programs the termination behaviour of their probabilistic counterparts is much more nuanced. One distinguishes between almost surely terminating (AST) programs, i.e., programs that terminate with probability 1 and positively almost surely terminating (PAST) programs that are characterised by the finiteness of their expected time to termination. In this project, we focus on the development of algorithms and techniques for the automated computation of (non-trivial) bounds on the expected time to termination. Such bounds can be interpreted as a measure of the efficiency of the analysed programs. Moreover, a finite bound guarantees the analysed program to satisfy PAST as well as AST (since PAST implies AST). Probabilistic ranking functions, a variant of ranking functions adapted for probabilistic programs based on the theory of ranking supermartingales[1](RSM), present a natural way to obtain bounds on the expected time to termination.

In recently published work[2], we introduced the concept of expected sizes of program variables. Moreover, we presented a novel modular approach for the computation of (upper) bounds on the expected time to termination in a fully automated fashion by combining bounds on the expected time to termination for parts of the program with bounds on the expected variable sizes.

In ongoing work we are looking at possible improvements to the expressiveness of probabilistic ranking functions. While lexicographic variants already exist[3][4], there is, to the best of our knowledge, no equivalent of multiphase-linear ranking functions (MΦRFs)[5]. In the classical setting the nested version of MΦRFs leads to linear bounds on a program's runtime. If this does transfer to the probabilistic setting it will be particularly useful due to the linearity of the expected value operator.

For deterministic programs there are classes of non-trivial loops for which termination is known to be decidable[6][7]. In further ongoing work we investigate whether similar classes, which would allow the decidability of AST, exist in the case of probabilistic programs.

[1] L.M.F. Fioriti, Hermanns, H.: Probabilistic termination: Soundness, completeness, and compositionality. In: Proc. POPL '15. pp. 489–501 (2015)

[2] F. Meyer, M. Hark, J. Giesl: Inferring Expected Runtimes of Probabilistic Integer Programs Using Expected Sizes. In: Proc. TACAS (2021).

[3] S. Agrawal, K. Chatterjee, P. Novotný: Lexicographic ranking supermartingales: An efficient approach to termination of probabilistic programs. In: Proc. ACM Program. Lang. 2(POPL) (2017)

[4] K. Chatterjee, E.K. Goharshady, P. Novotný, J. Zárevúcky, D. Zikelic: On Lexicographic Proof Rules for Probabilistic Termination. In: Proc. Formal Methods (2021)

[5] A.M. Ben-Amram, S. Genaim: On multiphase-linear ranking functions. In: Proc. CAV '17 (2017)

[6] M. Hosseini, J. Ouaknine, J. Worrell: Termination Linear Loops over the Integers. In: Proc. ICALP 2019 (2019)

[7] M. Hark, F. Frohn, J. Giesl: Polynomial Loops: Beyond Termination. In: Proc. LPAR23 (2020)

Optimization under Adversarial Uncertainty

Komal Dilip Muluk (muluk@algo.rwth-aachen.de)
Supervisor: Prof. Dr. Britta Peis

An optimization problem under adversarial uncertainty can be essentially formulated as a game between a player and an adversary: The player partially constructs a feasible solution for a given scenario, and then the adversary completes this to a full feasible solution. The goal of the player is to optimize some objective function and the goal of the adversary is to make the player perform as bad as possible. There are various types of adversarial problems. The PhD thesis of Berit Johannes (2011)[1] develops a machinery for deriving hardness results for large classes of the optimization problems with adversarial uncertainty. The thesis only discusses the negative aspects (hardness results) of the area.

The goals of my doctoral project are twofold: On the one hand, the project will derive new negative results, perhaps by extending and generalizing the machinery of Johannes to other families of optimization problems, such as problems in robust optimization. This should lead to new families of hardness and completeness results for the first or the second level of the polynomial hierarchy or for one of the intermediate complexity classes. On the other hand, the goal of the project is to develop positive results for the considered optimization problems. Major emphasis will be put on the investigation of crucial problem parameters, which will be done by applying the tool kit of parameterized complexity. A further goal is the development of fast exact algorithms with decent running times. Finally, the project will identify tractable special cases, for instance by constraining the combinatorics of underlying graph structures, or by imposing additional conditions on underlying cost matrices.

[1] B. Johannes, “New Classes of Complete Problems for the Second Level of the Polynomial Hierarchy,” Doctoral Thesis, TU Berlin, 2011

Algebraic Methods in SMT-Solving

Jasper Nalbach (Nalbach@cs.rwth-aachen.de)
Supervisor: Prof. Dr. Erika Ábrahám

Introduction. Algorithms and tools for checking the satisfiability of quantifier-free first-order logic formulas over different theories have many applications in e.g. verification, planning and numerous other fields and enjoy increasing interest. The theory of non-linear real arithmetic (also called real algebra), whose formulas are Boolean combinations of (in)equalities between polynomial expressions evaluated over the real numbers, admits a high expressive power at the cost of high computational costs for satisfiability checking. A subset of this theory, linear arithmetic, where the polynomial expressions are all linear, can be solved more efficiently. In particular, these theories are expressive enough for encoding complex properties about uncertainties. These could be safety properties of systems with linear and non-linear behaviour such as neural networks, and more generally non-linear probability distributions This project is about the general problem of solving (non-)linear arithmetic rather than specific applications. For this, several algorithms are developed and extended, which are implemented and evaluated in our SMT solver `SMT-RAT` [1,2] which builds on top of our computer algebra library `CArL`.

Non-linear arithmetic. Although Tarski [3] proved in 1948 that non-linear arithmetic is decidable, the cylindrical algebraic decomposition (CAD) method published in 1975 by Collins [4] was the first complete decision procedure for its solution. Recently, several novel approaches have been developed; namely the model-constructing satisfiability calculus (MCSAT) [5], the one-cell construction method [6] and the cylindrical algebraic coverings method (CAlC) [7]. MCSAT and the one-cell construction can be used in a symbiotic way to solve existential real-arithmetic problems. This new approach is still based on the CAD idea, but instead of a full decomposition it uses the CAD idea to generalize a non-satisfying sample point to a non-satisfying region. The cylindrical algebraic covering methods generates a covering of unsatisfying regions using similar ideas. We developed and implemented a more flexible variant of the original one-cell construction algorithm. This work allows future improvements of both theoretical as well as heuristic nature.

Currently, a publication with a formal proof of the one-cell algorithm algorithm and its experimental evaluation is in progress. In the future, we will develop further improvements of this method and will re-implement the cylindrical algebraic coverings to benefit from these ideas as well.

Linear arithmetic. Linear arithmetic is of interest as it is not only a subset of non-linear arithmetic but also (incomplete) reductions from non-linear arithmetic to linear arithmetic exist. Thus, improving our linear arithmetic solver also benefits the non-linear solver.

The general Simplex algorithm [8] is the most common method for solving linear arithmetic in SMT solving. Despite its exponential running time in worst case, it is efficient in practical instances, heavily depending on chosen heuristics. We are working on improving our Simplex implementation using state-of-the-art heuristics.

Furthermore, we are developing a novel approach that could be promising in the SMT solving context based on the Fourier-Motzkin variable elimination [9] procedure. Extensions of this novel method for learning combinatorial properties of the problem as well as deeper interleaving with the Boolean structure of formulas are conceivable.

While working on these problems, we proved the extension of the Simplex method and others for strict inequalities, which is currently under review. Although a proof already exists, we think that our publication provides more insights into the nature of the problem.

`SMT-RAT` *and* `CArL`. For several reasons, we maintain our own library for arithmetic operations. We are currently evaluating our library against other libraries with regards to efficiency and examine possible extensions or integrations of our library.

References:

1. Kremer, Gereon, and Erika Ábrahám. Modular strategic SMT solving with SMT-RAT. Acta Universitatis Sapientiae, Informatica 10.1: 5-25, 2018.
2. Kremer, Gereon. Cylindrical Algebraic Decomposition for Nonlinear Arithmetic Problems. Dissertation RWTH Aachen, 2020.
3. Tarski, Alfred. A decision method for elementary algebra and geometry. Quantifier elimination and cylindrical algebraic decomposition. Springer, pages 24–84, 1998.
4. Collins, George E. Quantifier elimination for real closed fields by cylindrical algebraic decompostion. Automata Theory and Formal Languages, pages 134–183, Springer, 1975.
5. Jovanović, Dejan, and Leonardo De Moura. Solving non-linear arithmetic. International Joint Conference on Automated Reasoning. Springer, 2012.
6. Brown, Christopher W., and Marek Kosta. Constructing a single cell in cylindrical algebraic decomposition Journal of Symbolic Computation 70: 14–48, 2015.
7. Ábrahám, Erika, et al. Deciding the consistency of non-linear real arithmetic constraints with a conflict driven search using cylindrical algebraic coverings. Journal of Logical and Algebraic Methods in Programming 119: 100633, 2021.
8. Dutertre, Bruno, and Leonardo De Moura. A fast linear-arithmetic solver for DPLL (T). Proc. of CAV. Springer, 2006.
9. Fourier, Jean Baptiste Joseph. Solution d'une question particuliere du calcul des inégalités. Nouveau Bulletin des Sciences par la Société Philomatique de Paris 99: 100, 1826.

Analysis of the expressivity of Graph Neural Networks (GNNs) and similar deep learning architectures for graphs

Eran Rosenbluth (rosenbluth@informatik.rwth-aachen.de)
Supervisor: Prof. Dr. Martin Grohe

Graph Neural Networks (GNNs) are a class of computation models operating on graphs, widely used in tasks of learning on graphs: Learning to classify graphs (or their vertices) and learning to regress features of graphs (or their vertices). In such tasks, it is desired that the computation model of choice is both isomorphism-invariant and inherently scalable to arbitrary graph sizes. Most GNN models comprise a sequence of message-passes interleaved with vertex-level computation blocks, each computation block consisting of a message-aggregation and a neural-network. The computation blocks are identical for every vertex, making GNNs isomorphism-invariant and inherently scalable algorithms. The use of neural networks makes GNNs deep architectures, potentially benefiting from the power of that paradigm. A key characteristic of any computation model is its expressivity. While there are significant results regarding the expressivity of GNNs, many interesting and important questions are yet to be answered. In my research I focus on resolving such questions, e.g. how does a certain hyperparameter of GNNs affect their expressivity – whether a specific configuration subsumes others. My work is mainly theoretical – formulating and proving theorems, with some projects including also the development of experiments that examine how the proven theory is manifested in practice.

Structural Network Analysis

Michael Scholkemper (scholkemper@cs.rwth-aachen.de)
Supervisor: Prof. Dr. Michael Schaub

Networks are a powerful abstraction to understand a range of complex systems such as

- protein interactions relating to drug efficacy,
- structural changes due to disease in respective tissue,
- the emergence of consensus or polarization through social interactions and epidemic spread, or
- the flow of traffic.

To comprehend such networks, we often seek patterns in their connections, e.g., densely interconnected communities or core-periphery structure, which facilitate a simpler or faster analysis of such systems. Communities are in this context typically envisioned as comparably tightly-knit nodes within a graph, though a range of different notions exists often defined in an algorithmic way, or by means of some cost function that is to be optimized. A complementary notion to „community“ is that of a role partition of the node. The concept of node roles – or node equivalences – originates in social network analysis, where a node‘s role – contrary to its community – is often related to symmetries or connectivity, rather than proximity. Both of these complementary approaches aim to simplify the network‘s complexity and provide a reduced view of the networks structure.

The intricacy of node role extraction derives from the lack of a clear definition of the role of a node in mathematical terms. Traditionally, exact equivalences such as automorphic or regular equivalence are considered. These approaches, while extremely expressive, yield a multitude of distinct roles that are incomparable to one another. More recently, node representation through embeddings aiming to capture these structural symmetries have been proposed. While these are clearly comparable by means of some distance measure in the vector space, the plethora of node embedding techniques yields no insight into what the role of a node is.

This research aims to derive a general definition of a node's role and the resulting reduced view of the network's structure. This, in turn, can then be applied to signal processing on graphs to provide a reduced view of a complex system of which only a process on the nodes but not the underlying network is observed. Apart from the structural analysis of a graph these node roles can also be used to obtain generative models that assert a certain global structure but are flexible locally. This is useful e.g. as a means of anonymization when sharing sensitive network data.

The Tournament Isomorphism Problem

Tim Frederik Seppelt (seppelt@informatik.rwth-aachen.de)
Supervisor: Prof. Dr. Martin Grohe

Introduction. The Graph Isomorphism Problem (GI), i.e. the computational problem of deciding whether two given graphs X and Y admit an isomorphism $X \leftrightarrow Y$, is of both theoretical and practical relevance in Computer Science and many adjacent fields [7]. For example, in chemistry it is desirable to determine whether two molecules encoded as graphs are structurally the same. The main interest from a theoretical viewpoint stems from the fact that despite intensive research efforts, the complexity of GI remains unknown. It is neither established that GI is NP-complete nor that it is in P. The best known algorithm, developed by Babai [2], runs in quasi-polynomial time in the number of vertices of the input graphs.

In order to resolve the complexity status of GI, restricted graph classes such as planar graphs and graphs with excluded minors have been considered in the past [6,8]. In each of these cases, researchers succeeded in showing that GI, when restricted to these classes, can be solved in polynomial time.

While the aforementioned graph classes have been eliminated as barriers for a potential polynomial time algorithm for GI, the class of tournaments persists in representing a bottleneck. Tournaments are directed graphs whose underlying undirected graphs are complete. The best known algorithm for the Tournament Isomorphism Problem (TI) from Babai and Luks [4].

Faster Algorithms for TI. Although decades-long research efforts have produced a variety of tools for variants of the GI, only a few methods tailored for the TI are known. TI fundamentally differs from other variants of GI in the sense that the automorphism group of a tournament is soluble which renders an efficient treatment of the occurring groups possible [9]. This in turn creates the need for refined combinatorial techniques. Subsequently, possible approaches for resolving the complexity status of TI are outlined.

Probabilistic Approaches. Probabilistic methods have been fruitfully used in the past in the context of TI. This includes randomized algorithms and reductions [11] but also probabilistic arguments used to derive structural insights into the involved combinatorial objects [1]. It is, therefore, desirable to further develop such probabilistic techniques in order to deepen the understanding of TI.

Exploiting Regularity. Whenever vertices of a graph can be distinguished, e.g. by their degrees, divide-and conquer techniques can be applied efficiently. These strategies fail if the graphs considered are regular. Looking at arcs instead of vertices gives rise to more powerful notions such as strong regularity. Especially in the realm of undirected graphs, the study of strongly regular graphs has led to a deep structural insights [5] and advanced algorithms [3,12]. This raises the question as to whether a structure theory for highly regular tournaments can be developed.

The Weisfeiler–Leman Algorithm. The Weisfeiler–Leman (WL) algorithm [13] is a ubiquitous tool in the context of the Graph Isomorphism Problem. Its k-dimensional version colors k-tuples of vertices according to their local structure. It is, hence, natural to identify levels of regularity with monochromaticity with respect to WL in certain dimensions. For example, graphs are strongly regular if and only if they are monochromatic with respect to 2-WL. Along these lines, the power of WL deserves further scrutiny. In [10], we studied the expressiveness of WL from a spectral perspective.

References:

1. László Babai. On the Order of Uniprimitive Permutation Groups. The Annals of Mathematics, 113(3):553, 1981.
2. László Babai. Graph Isomorphism in Quasipolynomial Time (Extended Abstract). In Proc. of STOC, pages 684—697, 2016.
3. László Babai, Xi Chen, Xiaorui Sun, Shang-Hua Teng, and John Wilmes. Faster Canonical Forms for Strongly Regular Graphs. I Proc. of FOCS, pages 157—166, 2013.
4. László Babai and Eugene M. Luks. Canonical labeling of graphs. Proc. of STOC, pages 171—183, 1983.
5. László Babai and John Wilmes. Asymptotic Delsarte cliques in distance-regular graphs. Journal of Algebraic Combinatorics, 43(4):771—782, 2016.
6. Martin Grohe and Dániel Marx. Structure Theorem and Isomorphism Test for Graphs with Excluded Topological Subgraphs. SIAM Journal on Computing, 44(1):114—159, 2015.
7. Martin Grohe and Pascal Schweitzer. The Graph Isomorphism Problem. Commun. ACM, 63(11):128–134, 2020.
8. Sandra Kiefer, Ilia Ponomarenko, and Pascal Schweitzer. The Weisfeiler-Leman dimension of planar graphs is at most 3. Proc. of LICS, pages 1—12, 2017
9. Eugene Luks. Permutation groups and polynomial-time computation. In Larray A. Finkelstein and William M. Kantor, editors, Groups and Computation, volume 11 of DIMACS Series in Discrete Mathematics and Theoretical Computer Science. American Mathematical Society, 1993.
10. Gaurav Rattan and Tim Seppelt. Weisfeiler–Leman, Graph Spectra, and Random Walks. 2021. Under Review for WG 2021.

11. Pascal Schweitzer. A polynomial-time randomized reduction from tournament isomorphism to tournament asymmetry. arXiv:1704.08529 2017.

12. Xiaorui Sun and John Wilmes. Faster Canonical Forms for Primitive Coherent Configurations: Extended Abstract. In Proc. of STOC, pages 693-–702, 2015.

13. Boris Weisfeiler. On Construction and Identification of Graphs, volume 558 of Lecture Notes in Mathematics. Springer, 1976.

Probabilites in Database Queries: Power and Complexity

Christoph Standke (standke@informatik.rwth-aachen.de)
Supervisor: Prof. Dr. Martin Grohe

Uncertainty plays a central role in modern database systems. It arises from many use cases, such as data integration, noisy data, data from unreliable sources or randomized processes. To obtain a quantitative framework to handle uncertainty, it is modelled via probability distributions.

Best studied in this area is the notion of finite *probabilistic databases (PDBs)*, which are finite probability spaces over database instances. For finite PDBs, the questions of representation and query evaluation under set semantics are well understood. We extend this knowledge in different directions:

Tuple-Independent Representations of Infinite Probabilistic Databases We systematically study the representability problem for infinite PDBs by means of tuple-independence and first-order views. Although first-order views over tuple-independent PDBs are not a complete representation system for infinite PDBs, they form a fairly robust class: Adding first-order constraints does not give them additional expressive power, and they cover many relevant special cases such as block-independent disjoint PDBs, and PDBs of bounded instance size. We identify criteria for representability (or non-representability) in this class and explore their limits.

Probabilistic Query Evaluation with Bag Semantics We study probabilistic query evaluation under *bag semantics* where tuples are allowed to be present with duplicates. Due to potentially unbounded multiplicities, the bag probabilistic databases we consider are no longer finite objects, which requires a treatment of representation mechanisms. Moreover, the answer to a Boolean query is a probability distribution over non-negative integers, rather than a probability distribution over $\{\mathsf{true}, \mathsf{false}\}$. Therefore, we explore two flavors of probabilistic query evaluation: computing expectations of answer tuple multiplicities, and computing the probability that a tuple is contained in the answer at most k times for some parameter k.

Furthermore, we also consider other sources of randomness than distributions over database instances:

The Importance of Parameter Values in Database Queries We propose and study a framework for quantifying the importance of the choices of parameter values to the result of a query over a database. These parameters occur as constants in logical queries such as conjunctive queries. This quantity is the SHAP score of the individual parameter values, as previously applied to feature values in machine-learning models. The application of the SHAP score requires two components in addition to the query and the database: (a) a probability distribution over the combination of parameter values, and (b) a utility function that measures the distance between the result for the original parameters and the result for hypothetical parameters. The main question we investigate is the complexity of calculating the SHAP score for different distributions and distances.

Exploring the Expressivity and Generalization Power of expressive GNNs

Antonis Vasileiou (antonis.vasileiou@log.rwth-aachen.de)
Supervisor: Prof. Dr. Christopher Morris

Graph data is prevalent across various applications, leading to the rapid advancement of graph-based machine learning techniques. However, traditional machine learning algorithms designed for graphs have limitations in capturing complex relationships between nodes and higher-order patterns. Recent findings by Morris et al. [2020] and Chen et al. [2019] have provided valuable insights into the expressivity of Graph Neural Network (GNN) architectures. Their results show that the expressive power of all possible GNNs is restricted to the ability of the Weisfeiler Leman (WL) algorithm to distinguish non-isomorphic graphs. In their work, Morris et al. [2020] proposed higher-order neural networks that can capture more complex patterns in graph data based on the k-WL algorithm. However, these models suffer from scalability issues, making them challenging to apply to real-world datasets. In Morris et al. [2020], Zhao et al. [2022], Morris et al. [2022], more scalable algorithms than k-WL for detecting higher-order patterns in graph structured data are presented. These algorithm computes sets of nodes instead of tuples leading to significant less expensive models. However, the expressivity of this set-based algorithm was not investigated in detail. For example, it is not clear if increasing the number k of the elements of the sets leads to an increasing expressivity of the algorithm, or how the set-based WL algorithm is compared to the classical k-WL. In this work, we build upon the aforementioned studies by conducting a theoretical comparison of the different algorithms and neural architectures proposed. Additionally, we propose an extension of the above set based algorithm as well as a multiset based variant of the WL algorithm and we show that our algorithms strictly increasing by increasing k. These models aim to strike a balance between capturing higher-order patterns and maintaining scalability and sparsity.

Moreover, we aim to investigate the generalization power of our proposed higherorder GNNs. From our empirical results, we have observed that in many cases, more expressive architectures do not directly imply significantly higher performance on unseen data. Based on recent work in the area of the generalization power of GNNs [Morris et al., 2023, Franks et al., 2024], we aim to establish generalization bounds for our proposed architectures.

Combining Knowledge-based Reasoning with Data Driven Methods in Production Scenarios

Tarik Viehmann (viehmann@kbsg.rwth-aachen.de)
Supervisor: Prof. Dr. Gerhard Lakemeyer

In knowledge-based systems, the usage of rules, constraints, formal logic, and domain expertise is leveraged to make decisions on the base of explicit knowledge representations and predefined decision-making criteria. These systems rely on structured information and expert-defined rules to guide reasoning and inference processes, ensuring consistency and transparency in decision-making. However, knowledge-based systems may struggle to handle complex and uncertain situations which are not covered by the defined rules and constraints or where domain expertise is limited.

In contrast to that, data-driven algorithms operate by extracting patterns, trends, and insights directly from data, often without relying on predefined rules or domain expertise. This can be very beneficial in scenarios, where data is either readily available or obtainable through simulations and has the potential to uncover hidden relationships and patterns. However, they may suffer from issues like overfitting, lack of interpretability, and ethical concerns, particularly in sensitive domains.

To leverage the strength of both techniques In modern production scenarios, a combination of both knowledge-based and data-driven methods may useful in order to leverage the strength of both techniques. In this field a few different use-cases are studied: Firstly, automated flexible production lines with autonomous robots are considered. There one key objective is to maximize the efficiency of the production by scheduling and dispatching tasks on multiple autonomous robots while facing uncertainties in the requested order patterns. These considerations may be further constrained by company policies (such as maintenance cycles), safety concerns and hardware limitations. Using the domain of the RoboCup Logistics League, which replicates such challenges from smart factories, the CLIPS Executive, a knowledge-based goal reasoning framework, is extended to support reinforcement learning in order to aid the high-level decision process and to be used alongside of classical PDDL planning.

On the topic of robotics, data-driven methods are also promising good results in low-level control, especially in high-dimensional spaces such as when controlling multiple robotic arms in shared spaces or in order to process visual information into structured knowledge. These approaches can then be utilized by a knowledge-based system such as the CLIPS Executive, where additional execution monitoring may be deployed in order to ensure a robust execution.

Lastly, in scenarios where knowledge from past problems may be of value, Case-Based Reasoning (CBR) is explored as a technique to retrieve and reuse relevant historical data with the help of similarity metrics. These metrics may be partially defined by human experts and partially synthesized using machine learning techniques such as Siamese Neural Networks (SNNs). Therefore a modern CBR framework is being developed to allow hybrid similarity definitions based on user needs to provide decision support in complex scenarios such as the forging of a drive shaft for heavy-duty applications.

Programming and Verifying Uncertain Phenomena

Tobias Winkler (tobias.winkler@cs.rwth-aachen.de)
Supervisor: Prof. Dr. Joost-Pieter Katoen

In recent years, programming languages have been enhanced with probabilistic constructs, allowing programmers to write statements like "Flip a fair coin, if heads comes up then increment variable x by 1" or "If two processes A and B are in the same state, then process B crashes with probability 5%". It is important to understand that this extra randomness does not present a contradiction to the unambiguous nature of programming languages: Instead of yielding a predetermined output like classical programs, a probabilistic program typically results in a predetermined probability distribution over possible outputs. The following are a few of the most important use cases of probabilistic programs:

Randomized Algorithms are traditional algorithms extended with coin flips to increase performance or enable realizability of certain computational tasks. The latter is especially the case for computations distributed among several agents [1]. Such algorithms are meant to be actually implemented and run on a physical machine, often using a pseudo-random number generator.

Probabilistic Model Checking aims at verifying behavioural properties of processes involving randomness. The process under consideration is usually modelled by means of a probabilistic program. The purpose of the program is not to be actually executed but describe the process of interest in a precise mathematical manner. Application areas include verification of randomised—often distributed— algorithms (internal randomness), systems making decisions in an uncertain environment (external randomness), biological processes and many more. A distinguishing feature of model checking is that the program at hand is typically (but not always) finite-state. This enables exact algorithmic solvability (decidability) of almost all properties of interest by constructing a finite low-level model of the process such as a continuous- or discrete-time Markov chain, a Markov Decision Process, a stochastic game and others. See [2] for an overview of the field.

Probabilistic Programming (e.g. [3]) is a relatively new paradigm that aims to automate statistical inference. Similar to model checking, programs of the corresponding languages are not meant to be run directly but rather to describe a process in which unknown events may occur. The purpose of a Probabilistic Programming System is to automatically infer the likelihood of those events given observations about the outcome (or intermediate stages) of the process. It can, thus, be seen as an automated approach to Bayesian statistics, and it generalizes traditional graphical models such as Bayesian networks. Languages for Probabilistic Programming typically support continuous probability distributions and have additional primitives for observations. In general, inference can only be done approximately, using sampling based approaches.

Clearly, the three directions are closely related. Moreover, program verification is key in all of them: While this is obvious for Randomised Algorithms and Probabilistic Model Checking, it turns out that verification and inference mostly coincide in the case of Probabilistic Programming. The aim of my research is twofold:

A. To help foster a common theoretical basis for the three areas: More specifically I am interested in the development of new verification logics in the spirit of classical Hoare logic and weakest precondition transformers [4,5] to facilitate and systemize the verification tasks mentioned above. This is closely related to program semantics—mathematical definitions of the meaning of a program—as different approaches to semantics lead to different verification rules.

B. Contributions to the ample field of Probabilistic Model Checking, more concretely:

   1. Stochastic games. Such games arise from controlled stochastic processes, additionally faced with unquantifiable uncertain external events, i. e., events whose occurrence cannot be described be means of probabilities, e.g. because relevant statistical data is unavailable. I plan to investigate the two-player turn-based variant of such games under non-standard multi-objectives [6].
   2. Recursive stochastic processes. These are naturally described by imperative probabilistic languages allowing (mutually) recursive function calls. I plan to work on Model Checking finite-state versions of such programs [7]. Applications include self-reproducing stochastic processes.
   3. Program rewriting. Another direction I intend to pursue is to rewrite probabilistic programs prior to Model Checking with the aim of simplifying the latter task, in particular by decreasing the size of the resulting finite-state model.

References:

1. Ted Herman. Probabilistic Self-Stabilization. Inf. Process. Lett. 35(2), p. 63–67 (1990)
2. Joost-Pieter Katoen. The Probabilistic Model Checking Landscape. LICS 2016, p. 31–45 (2016)
3. Andrew D. Gordon et al. Probabilistic Programming. FOSE 2014, p. 167–181 (2014)
4. Annabelle McIver and Carroll Morgan. Abstraction, Refinement and Proof for Probabilistic Systems. Monographs in Computer Science, Springer (2005)
5. Benjamin L. Kaminski. Advanced weakest precondition calculi for probabilistic programs. PhD Thesis (RWTH Aachen University), Germany, (2019)
6. Taolue Chen et al. On Stochastic Games with Multiple Objectives. MFCS 2013, p. 266–277 (2013)
7. Javier Esparza et al. Model Checking Probabilistic Pushdown Automata. LICS 2004, p. 12–21 (2004)

GRK 2428: ConVeY: DFG Research Training Group on Continuous Verification of CYber-Physical Systems

Dirk Beyer
Email: dirk.beyer@sosy.ifi.lmu.de
LMU Munich and TU Munich
Internet: https://convey.ifi.lmu.de

Networks, computers, sensors, and actuators are increasingly integrated into cyber-physical systems, that is, software systems that interact with the physical world and must cope with its continuous behavior. An increasing number of cyber-physical systems operate in safety-critical domains, e.g., autonomous vehicles, robotic surgery, traffic control, human-robot collaboration, and smart grids. For this reason, their design, development, and deployment should ideally be accompanied by a formal check of correct behavior. A fundamental challenge in the verification of cyber-physical systems is the fact that they are subject to change. The physical environment changes continuously, at runtime, and in ways that cannot be completely foreseen at the design stage. At the same time, the requirements may change. Sought-after aspects include more functionality, lower power consumption, or faster response. In many cases, the system should be migrated to a different hardware platform.

To face this multi-level continuous change, ConVeY performs research in three areas:

Robust System Design We develop techniques that guarantee correct behavior under changes in plant parameters, under certain classes of perturbations including sensor measurement errors, and under uncertainties introduced by the implementation platform. In particular, we investigate the design of controllers that are, by construction, robust against these changes.

Evolving Systems We investigate novel construction and verification techniques that adapt to offline changes in the specification, the hardware, or the implementation of control software, and that reuse efforts from earlier stages as much as possible.

On-the-fly Synthesis and Verification For the online verification and synthesis of controllers, we develop techniques that operate—and provide a correctness guarantee—only within a given time horizon. Repeated execution of these techniques, combined with the availability of a fail-safe strategy, ensures safe operation.

Logical Foundations of Concurrent Cyber-Physical Systems

Marvin Brieger (marvin.brieger@sosy.ifi.lmu.de)
Supervisor: Prof. Dr. André Platzer

Cyber-physical systems (CPSs) are ubiquitous in our everyday lives. They occur as cars, robots, airplanes, etc., and are often safety critical. However, due to their interlocked discrete and continuous dynamics CPSs are difficult to get right making them a natural target of verification. Further, cars, robots, and airplanes often don't operate alone but populate the same part of the world, where they may meet or even interact by purpose. Analyzing these parallel CPSs by verification is a huge challenge as hybrid systems verification and concurrency verification are already nontrivial challenges each own their.

To tackle the CPS concurrency challenge, we study the dynamic logic of communicating hybrid programs $d\mathcal{L}_{\text{CHP}}$[1]. $d\mathcal{L}_{\text{CHP}}$ is an extension of differential dynamic logic $d\mathcal{L}$ to concurrency.[2] In $d\mathcal{L}$, CPSs are modeled using hybrid programs and verified using dynamic logic. We study communicating hybrid programs (CHPs), which extend hybrid programs with a parallel operator and communication primitives, as modeling language for parallel CPS behavior. For verification, $d\mathcal{L}_{\text{CHP}}$ adapts the assumption-commitment approach[3] to $d\mathcal{L}$, which allows for compositional reasoning about parallelism and communication behavior.

To enable large-scale applications of $d\mathcal{L}_{\text{CHP}}$, the logic is supposed be supported in the theorem prover KeYmaera X. Fundamental for this is the study of uniform substitution as the technique underlying KeYmaera X and enabling its parsimonious prover kernel.

[1] Marvin Brieger, Stefan Mitsch, André Platzer, "Uniform Substitution for Dynamic Logic with Communicating Hybrid Programs," CADE, 2023, http://dx.doi.org/10.1007/978-3-031-38499-8_6

[2] André Platzer, "Logical Foundations of Cyber-Physical Systems," Springer, 2018

[3] Jayadev Misra and K. Mani Chandy, "Proofs of Networks of Processes," IEEE Transactions on Software Engineering, vol. 7(4), p. 417 – 426, 1981

Bridging the Gap between Hardware and Software Verification

Po-Chun Chien (po-chun.chien@sosy.ifi.lmu.de)
Supervisor: Prof. Dr. Dirk Beyer

Computing systems are omnipresent in our daily lives. Assuring the correctness of these systems is crucial, as errors in their implementations could lead to catastrophic consequences. Formal verification provides correctness guarantees with mathematical rigor and has been successfully applied in practical scenarios. Verification tools and algorithms are often developed for a specific type of systems, e.g., software programs, hardware circuits, and cyber-physical devices, although they share common theoretical foundations. This creates a barrier between the research communities and makes it harder for a community to benefit from the advancement of the other. In this project, we aim to bridge the gap between hardware and software verification in two ways: by (1) translating verification tasks, and (2) transferring verification approaches.

In the first approach, we translate a verification task encoded as a hardware circuit into a software program, and the other way around. We develop Btor2-Cert, a certifying hardware-verification framework using software analyzers.[1] Btor2-Cert first translates a circuit into a program with the translator Btor2C[2], and then employs software verifiers to analyze the translated program. To enhance the explainability of verification results, the certificates provided by the software verifiers are additionally translated back to the circuit domain. Furthermore, we develop CPV, a circuit-based program verifier.[3], which translates a program into a circuit and verifies the translated circuit with a hardware model checker. This novel verifier achieved competitive results in the recent edition of SV-COMP, manifesting the feasibility of circuits as an intermediate representation for program analysis.

In the second approach, we adopt several interpolation-based hardware-verification algorithms for software model checking. Additionally, we combine one of the adopted algorithms with an invariant generator based on data-flow analysis[4] to increase the effectiveness of the verification process.[5] The cross-disciplinary algorithm adoption is beneficial, as the newly-adopted approaches were able to tackle tasks unsolvable by existing methods. Our study also helps consolidate and transfer knowledge across the two research communities.

[1] Zs. Ádám, D. Beyer, P.-C. Chien, N.-Z. Lee, and N. Sirrenberg, “Btor2-Cert: A Certifying Hardware-Verification Framework Using Software Analyzers,” Proc. TACAS, p. 129-149, 2024

[2] D. Beyer, P.-C. Chien, and N.-Z. Lee, “Bridging hardware and software analysis with Btor2C: A word-level-circuit-to-C translator,” Proc. TACAS, p. 152-172, 2023

[3] P.-C. Chien, and N.-Z. Lee, “CPV: A Circuit-Based Program Verifier (Competition Contribution),” Proc. TACAS, p. 365-370, 2024

[4] D. Beyer, P.-C. Chien, and N.-Z. Lee, “CPA-DF: A Tool for Configurable Interval Analysis to Boost Program Verification,” Proc. ASE, p. 2050-2053, 2023

[5] D. Beyer, P.-C. Chien, and N.-Z. Lee, “Augmenting Interpolation-Based Model Checking with Auxiliary Invariants,” Proc. SPIN, 2024, to appear

Modular and Efficient Creation of Function Summaries Using Abstract Interpretation

Julian Erhard (julian.erhard@tum.de)
Supervisor: Prof. Dr. Helmut Seidl

Abstract interpretation is an established approach for static program analysis, but analysis times still limit the applicability for large code bases. One approach to increase the scalability of abstract interpretation is to make an existing whole program analysis incremental, which was considered in previous work for C programs, for an analyzer using a top-down solver[1]. We also proposed efficient algorithms for efficient tracking of conjunctions of relations between two pointers [2] Another practical problem for scalable abstract interpretation is the treatment of calls to library functions. As library functions may constitute a large amount of the executed code, a complete analysis of these functions for each context in which they are called might cause prohibitive analysis run times. Optimistic assumptions about calls to library function, i.e. assuming that called functions do not have an effect on the program state abstracted by the analysis, will yield unsound results in many cases. For example, a function might write to memory that is accessible via a pointer passed as a parameter, which has to be accounted for by the analyzer. Without further knowledge of the called function, a sound static analyzer would have to make worst-case assumptions about the values of all memory blocks accessible to the callee for the program state after the call. One approach is to use a modeling language to summarize the effect of library functions for the use in abstract interpreters. While this allows to supply function summaries with precise information, it requires manual annotation, which is usually costly and error-prone. We aim at providing an efficient analysis for library code which delivers sound summaries of the effect of library functions. These summaries then can be used in the analysis of code using the summarized functions. In particular, summaries will contain information about which memory blocks reachable from global variables and parameters might have been written to, but may also include whether memory was allocated or de-allocated, which functions were called via passed function pointers and which threads were spawned. The analysis of the library functions is performed in a modular manner, that is, each function is analyzed only in one or a few canonical contexts. The representation of abstract values of formal parameters and global variables uses a symbolic representation for memory blocks, while not retaining any precise information about integer values. Our analysis takes possible aliasing into account, by distinguishing memory blocks that are pointed to by parameters or global variables only by their type. We integrate the analysis into the Goblint static analyzer and evaluate its efficiency on programs using libraries such as *sqlite3* and others.

[1] Seidl, H., Erhard, J., and Vogler, R. (2020). Incremental abstract interpretation.

[2] Seidl, H., Erhard, J., Schwarz, M., and Tilscher, S. (2024). 2-Pointer Logic.

An Unifying Framework for Transformation of Verification Properties

Marek Jankola (marek.jankola@sosy.ifi.lmu.de)
Supervisor: Prof. Dr. Dirk Beyer

The model checking problem is to determine whether a given program satisfies a given property. Since the range of possible properties that can be defined in numerous models, logics, or languages is wide, researchers usually focus on smaller subsets of specifications. One of the properties with high attention of software verification and hardware verification communities is called reachability or Reach Safety property. Thanks to the modularity of the property, there are many highly efficient algorithms and well-maintained tools that can verify programs against Reach Safety. In our research, we would like to observe whether the advantages can be leveraged for the verification of other practical properties.

There are some previous procedures on how the modeling transition system needs to be transformed so that the verification of liveness properties is reduced to the verification of Reach Safety. Furthermore, multiple tools reduce various properties to Reach Safety internally. However, these approaches lack modularity. They do not support the use of any Reach Safety algorithm for the reduced problem. Moreover, it requires an engineering effort to transform either a Reach Safety verification approach or the internal representation to be useful for other properties.

Our Unifying framework goals to mitigate the two problems by (1) doing transformation on the input level, i.e. transforming a given program to a new one that can be verified against Reach Safety, and (2) defining an input language, where one can conveniently define the instrumentation of the programs for the desired property. We have defined the transformation of three practical properties from SVCOMP in our framework - No Overflow, Memory Cleanup, and Termination. We ran experiments and evaluated the performance of Reach Safety algorithms applied to the transformed programs against the performance of state-of-the-art algorithms, for verification of the mentioned properties, applied to the original programs. Some of the Reach Safety analyzers were more effective and efficient on the transformed benchmarks than the best verifiers of the original properties.

The next step is to extend the framework to support more practical properties and features of C programs. After showing that the performance of Reach Safety analyzers can be leveraged to efficiently verify other properties, a new branch of research opens. In addition to a verdict of the model checking problem, the verifiers usually also provide a witness. The witness is a container of information that helps reconstruct the proof for the verdict. However, witnesses are bound to the input property and the input program. Therefore, a witness that we obtain after verifying the transformed program needs to be transformed back to match the original property and program.

Formalization and Verification of Post-Quantum Cryptography

Katharina Kreuzer (kreuzerk@in.tum.de)
Supervisor: Prof. Tobias Nipkow

Since communication is a key point in cyber-physical systems, it is important to ensure a safe communication. Especially since quantum computers come into reach more and more, the threat of breaking current, widely used crypto systems is imminent. Developing quantum resistant cryptography – and verifying it – is a major task of modern research. For the long term goal of my research, the focus lies on verifying post-quantum crypto algorithms applicable to cyber physical systems, using the proof assistant Isabelle.

For the main project, the post-quantum crypto system Kyber was formalized and its correctness and some security properties were verified in the theorem prover Isabelle [1]. An extended version [2] was accepted at CSF24. Kyber is a lattice-based post-quantum public-key encryption (PKE) scheme based on the module Learing-with-Errors problem and was chosen as the first winner of the NIST standardisation process for post-quantum PKE schemes. Since several error terms in the algorithms are chosen to blur the keys and the message, one has to make sure that the message is indeed decoded correctly. The goal was to verify this in Isabelle. During the formalization, an error in the original proof was uncovered. Moreover, my research could provide toy examples where this error breaks the security proof. One important new research question is therefore: Can we perform (statistical) analyses on various parameter systems whether this error also occurs in practical situations? Is the security proof broken or can we still fix it?

It is not only important to verify security proofs against classical computers, but also against quantum attackers. This is my next goal. The One-Way-to-Hiding (O2H) Lemma is the core theorem for security proofs against quantum adversaries for multiple post-quantum crypto systems. However, the O2H Lemma was axiomatized in the "qrhl-tool"[3], currently the only tool that allows formalization of security properties against quantum adversaries. A foundational formalization of the O2H Lemma would thus significantly improve the trustworthiness of security proofs against quantum adversaries. However, this project is more difficult to tackle as it requires thorough knowledge in quantum mechanics and the mathematical foundations thereof (like operator theory and infinite dimensional Hilbert spaces).

[1] https://isa-afp.org/entries/CRYSTALS-Kyber.html

[2] https://eprint.iacr.org/2023/087

[3] https://github.com/dominique-unruh/qrhl-tool

Software Verification Witnesses

Marian Lingsch-Rosenfeld (Marian.Lingsch@sosy.ifi.lmu.de)
Supervisor: Prof. Dr. Dirk Beyer

Automatic software verification methods attempt to proof the correctness of, or find bugs in, software systems with respect to a given specification. In practice, tools implementing these methods are usually massive software systems themselves, which are almost impossible to formally verify and difficult to test. This leads to the question of how to trust the results of these tools.

For this purpose software verification witnesses were invented [1]. They are exported by automatic software verification tools and contain information about the verification process. This information aids a validator in efficiently reconstructing the proof or replaying the violation.

The original witness format has the caveat that it is tied to the internal representation a tool has of the program it is analyzing, also called control-flow automaton (CFA). Due to this, the witnesses had unclear semantics and made it more difficult for other tools to understand them. In our recent work [2] we introduced a new format for witnesses that is based on the syntax of the input program. Due to this, the new witness have clear semantics, makes them more portable and easier to understand. Currently, we are working on extending the format to aid in the validation of more properties. For example, we are working on extending the support to express facts about memory safety and termination.

While validation was the original goal of witnesses, they have since found extended usage as artifacts for cooperative verification. For example in our recent work [3] we used witnesses to allow any verifier to make use of an abstraction refinement procedure based for the control-flow of a program. Currently, our work in this direction is focused on enabling the exchange of information with tools which are not automatic software verification tools. For example, deductive verification tools or interactive theorem provers. The goal is to leverage the benefits and reduce the drawbacks inherent to the different approaches to formal methods.

Software verification witnesses have shown to be a helpful way of exchanging information, not only to increase the trust in the results of automatic software verification tools, but also to enable cooperation between different verification approaches. The goal of our work is to further extend the format and the tools that work with it to make them useful in a wider variety of settings.

[1] Beyer, D., Dangl, M., Dietsch, D., Heizmann, M., Lemberger, T., Tautschnig, M., Verification witnesses. ACM Trans. Softw. Eng. Methodol. 31(4), 57:1-57:69 (2022)., https://doi.org/10.1145/3477579

[2] P. Ayaziová and D. Beyer and M. Lingsch Rosenfeld and M. Spiessl and J. Strejček. Software Verification Witnesses 2.0. Proc. SPIN, 2024, to appear

[3] D. Beyer and M. Lingsch-Rosenfeld and M. Spiessl. CEGAR-PT: A Tool for Abstraction by Program Transformation. Proc. ASE 2023, p. 2078-2081. https://doi.org/10.1109/ASE56229.2023.00215

Verified Solution Methods for Markov Decision Processes

Maximilian Schäffeler (maximilian.schaeffeler@tum.de)
Supervisor: Prof. Tobias Nipkow

Markov decision processes (MDPs) are a standard model for decision-making problems in probabilistic systems. In MDPs, a decision-maker selects actions with random outcomes to maximize long-term rewards. MDPs are widely used in reinforcement learning, planning, model checking and operations research. Since algorithms on MDPs have applications in safety-critical scenarios, we require a high level of trustworthiness from both the underlying theory and the implementation.

A methodology with notable success in developing provably correct software involves the use of interactive theorem provers (ITPs). In our project, we study the application of the ITP Isabelle/HOL to the development of trustworthy software for solving MDPs. Compared to developing other types of verified algorithms, algorithms on MDPs bring their own particular set of challenges. First, formal proofs of correctness of MDP solving algorithms in ITPs need a combination of non-trivial formal mathematical libraries and concepts. Second, at an implementation level, multiple challenges exist, what kind of implementation of numerics should be used. Last is the overall architecture of the verified system: should it be a verified implementation, or should we use an unverified system to produce a certificate, which is later validated using a formally verified certificate checker. Previous verification efforts in the realm of MDPs have all focused on the abstract mathematical challenges, while we also address the problem of deriving efficient executable code.

As a first step, we have already verified dynamic programming algorithms that can solve tabular MDPs optimally. For infinite-horizon problems, we model four fundamental iteration-based methods: value iteration, policy iteration, modified policy iteration, and splitting-based methods. We experimentally evaluate our implementations on standard probabilistic planning problems and show that they are practical. Finally, we experimentally show that combining our verified implementations with an unverified implementation yields significant performance improvements: one can use a fast floating-point implementation to perform all the iterations and then use the formally verified implementation for the last iteration. We have published our formalization efforts on MDPs in the Archive of Formal Proofs and at AAAI-23.

We are currently verifying an approximate policy algorithm that determines policies with formal guarantees for much larger MDPs. As part of this algorithm, we implement a verified certificate checker for linear programming solutions. Therefore, a much faster, unverified linear programming solver can be used in our implementation. Finally, we employ the technique of certificate checking to quantitative model checking. We are developing formally verified certificate checkers for the model checkers Storm and Prism.

Verification of Top-Down Solvers

Sarah Tilscher (sarah.tilscher@tum.de)
Supervisor: Helmut Seidl

For program analysis, fixpoint solvers are essential for computing a solution to the constraint system generated from the control flow of the program. The TD is a generic and demand-driven top-down solver that tracks the dependencies between unknowns on-the-fly. Several improvements to the initial TD have been proposed, such as the dynamic detection of widening/narrowing points[1]. While the improved TD is convenient to use for program analysis, the interplay of the advanced solving strategies is often difficult to understand, and every further extension makes it harder to reason about its correctness. This makes the implementation of the solver fragile and vulnerable to bugs. Therefore, we want to back up its correctness with machine-checked proofs written in the interactive theorem prover Isabelle.

We observe that the tracking of stable unknowns in the TD, as well as extensions with e.g. the dynamic detection of widening/narrowing points, can be viewed as an abstract interpretation of the abstract interpreter (A^2I) [2] [3] itself: the solver maintains a state that tracks an abstraction of its reaching left-context to implement an optimized evaluation strategy.

Based on this observation, we want to divide our proof and first only prove the partial correctness of the simpler TD_{plain}, a variant of the TD that simply re-evaluates unknowns whenever they are queried again. Secondly, we want to prove the equivalence of TD_{plain} and the original TD, i.e., they share the same termination behavior and also return the same result. Together with the partial correctness of the TD_{plain} this would then also imply the partial correctness of the TD.

In practice, widening and narrowing are often employed to guarantee termination even when a lattice has infinite ascending chains. Therefore, we want to investigate the termination behavior of the TD and variants of it which are extended with different forms of widening and narrowing. Beyond that, we are also interested in formalizing the TD with *side-effects*, an extension that allows to conveniently express e.g. analyses for multithreaded code or analyses that additionally collect flow-insensitive information about globals.

[1] Apinis, K., Seidl, H. and Vojdani, V., "Enhancing top-down solving with widening and narrowing", LNCS, vol. 9560, p. 272-288, 2016

[2] Cousout P. et al., "A^2I: abstract^2 interpretation", POPL, vol. 3, 2019

[3] Tilscher S. et al., "The Top-Down Solver—An Exercise in A2I.", In Intelligent Systems Reference Library (pp. 157-179), vol. 238, 2023

Infrastructure for Incremental and Cooperative Software Verification

Henrik Wachowitz (henrik.wachowitz@ifi.lmu.de)
Supervisor: Dirk Beyer

As software enriches more aspects of our daily lives, it becomes increasingly important to ensure its correctness. In safety-critical systems, such as cyber-physical systems (CPS), ensuring this correctness through mere testing is not enough. We instead need the strong guarantees provided by Software Verification. However, there are limitations: Software Verification techniques can consume large amounts of resources and time, making it often difficult to integrate Software Verification throughout the development process.

With Incremental Verification, we want to close this gap. Reusing knowledge gained from verifying earlier versions of a software system increases speed and lowers resource consumption of successive verification tasks. Embedding this incremental framework in a service infrastructure further improves its accessibility, creating an ideal fit for continuous integration.

To achieve this, I am working on solutions that enable incremental verification and ease with the integration of verification tools into modern development processes[1]. With the Project "CPA-Daemon" we showed how a Java-based verification tool can be lifted into a, potentially stateful, service. This new service framework serves as the foundation for incremental verification approaches built on top of the verification tool CPAchecker. CPA-Daemon is designed with local deployment in mind, although it could serve as a cloud service as well. To access verification tools as a service we built CoVeriTeam Service[1], a web service that can run a variety of verification tools.

To foster cooperation between verification tools I established a Python library, fm-actor, that consumes a standardized Yaml format describing tools. With this library, tools can be used in programs without much effort. A great showcase of this is another tool I developed, fm-weck. Fm-weck encapsulates the verification tool in a container. For rapid testing, fm-weck can download and execute the tool using the pre-configured arguments from the standardized Yaml as well as custom ones provided by users.

[1] Beyer, Kanav, Wachowitz, "CoVeriTeam service: Verification as a service", Proc. ICSE, companion, pp. 21–25, 2023

Multi-modal Machine Learning for Hardening Firmware Binaries

Yunru Wang (yunru.wang@ifi.lmu.de)
Supervisor: Prof. Dr. Johannes Kinder

The extensive dependence on third-party libraries in embedded firmware exacerbates the threat to software supply chain security. Reverse engineering serves as the initial step in identifying vulnerabilities within firmware; however, this process is impeded by the lack of semantic information in machine code. Besides, when known malicious code is present in either the source code or binary code, Binary Code Similarity Detection (BCSD) [1] emerges as a pivotal tool in recognizing malicious code patterns from binaries compiled across various domains; and aiding in the automation of firmware patching efforts.

Traditional BCSD methods, which heavily rely on manual feature selection and computationally intensive procedures, exhibit insufficient scalability. In contrast, deep learning-based BCSD techniques have garnered attention for their superior capability in feature representation and adaptability across diverse architectures, platforms, compilers, and compilation configurations. However, existing deep BCSD approaches solely focus on learning features from binary code, which may not suffice due to the absence of symbols and other high-level structures inherent in source code. Motivated by Multi-Modal Machine learning (MMML) [2], we propose leveraging source code and additional contextual information to enhance the binary representations. Furthermore, our project not only functions as a BCSD tool but also provides the reconstruction of valuable high-level structures, such as code summaries and pseudocode, assisting in other reverse engineering tasks.

In our methodology, we align binary encoding with function names in latent space using contrastive loss to incorporate function name semantics into binary encoding. Additionally, we employ a function name generation task as a proxy for further finetuning. With the trained binary encoder, we evaluate binary similarity by computing the cosine distance between their generated embeddings, thereby determining whether they originate from the same source code. Furthermore, we apply a similar approach to reconstruct high-level structures from binaries. During training, code summaries or pseudocode, served as captions, together with preprocessed binaries are input to finetune the binary representation model and train a decoder. During inference, high-level structures are reconstructed using the processed binaries alongside the finetuned binary encoder and the trained decoder.

[1] Xu, X., Liu, C., Feng, Q., Yin, H., Song, L. and Song, D., 2017, October. Neural network-based graph embedding for cross-platform binary code similarity detection. In Proceedings of the 2017 ACM SIGSAC conference on computer and communications security (pp. 363-376).

[2] Yu, J., Wang, Z., Vasudevan, V., Yeung, L., Seyedhosseini, M. and Wu, Y., 2022. Coca: Contrastive captioners are image-text foundation models. arXiv preprint arXiv:2205.01917.

GRK 2475: Cybercrime and Forensic Computing

Prof. Dr.-Ing. Felix Freiling
Email: felix.freiling@fau.de
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)
Internet: https://cybercrime.fau.de

Cybercrime and
Forensic Computing
Research Training Group 2475

Information technology has caused a new form of crime to emerge: cybercrime. It is incurring an increasing cost on modern society and is arguably threatening the stability of our economic system. Traditional law enforcement approaches appear to struggle with this new development. However, with new technologies also come new forms of criminal investigation, like large-scale data analysis and police trojans for covert surveillance. The effectiveness of such methods routinely raises questions regarding their impact on the constitutional rights of affected citizens. The inherent bounds of national law complicate matters further.

Funded since October 2019, this Research Training Group aims to disentangle the many open ends of this research area arising from the interaction between computer science and criminal law by bringing together established scientists from both areas. Computer science is represented through the areas of cryptography (Dominique Schröder), theoretical computer science (Lutz Schröder, Stefan Milius), multimedia security (Christian Riess), hardware-software-co-design (Jürgen Teich, Stefan Wildermann) and computer security (Felix Freiling). Colleagues from law represent criminal law (Hans Kudlich), criminal procedural law (Christoph Safferling) and criminology (Gabriele Kett-Straub). Starting with the second funding phase from April 2024, this expertise is complemented by three additional scientists: Zinaida Benenson (human factors in security), Paul Rösler (cryptography, provable security, secure messaging) and Paulina Pesch (data protection in law enforcement, criminal law and criminal procedure).

After almost four years of joint interdisciplinary research in this research training group, we have taken an analytical-empirical look at the topic of cybercrime and the criminal liability and prosecution of cybercrime. In doing so, the project succeeded both in the area of technical analysis of digital evidence as well as in the legal and social understanding of the phenomenon of cybercrime and the challenges of its prosecution (like the work of IT expert witnesses). Progress has been made, for example, in the fundamental assessment of the falsifiability of digital traces or the reliability of artificial intelligence methods in criminal prosecution. While many research topics naturally have until now explored the limits of forensic computing, the second funding phase aims more towards applicability of results.

Overall, the research field of forensic computing is undergoing a national and international phase of “demythologization”. Thus, the involved scientific disciplines are realigning themselves and each other. Within this context, the continued goal of this research training group is to slowly but systematically work towards establishing new methodological standards in handling digital evidence, interpreting and developing national and international law in the years to come.

The individual research and training programme of funded researchers is undertaken in cooperation with an interdisciplinary advisory committee and supported by a joint lecture series, a research seminar and interaction with international guests. During the annual cybercrime workshop, funded researchers interact by solving selected cybercrime cases involving forensic analysis of digital evidence and its presentation in front of an expert panel consisting of computer security professionals, public prosecutors and judges.

Graded Semantics and Logics and their Applications in Digital Forensics and Security

Üsame Cengiz (uesame.cengiz@fau.de)
Supervisor: Prof. Dr. Lutz Schröder

State-based systems are a useful formal method for modeling process behaviour. In its most basic form, a system consists of states and transitions between them. Over the years many different system types have emerged, such as labelled, weighted, probabilistic, and game-based systems. The category-theoretic framework of universal coalgebra provides a generalization of all these settings.

But even when one restricts to just a simple system type, there is a wide variety of pertinent semantics. Two systems which behave differently under a finer semantics may be indistinguishable under a coarser semantics. The desired semantics always depends on the use case. Processes can thus be compared in different ways, ranging from just their outputs to their exact internal behaviour. All of these types of process equivalences can be used in e.g. specifying the degree of strength of an attacker model for verifying security features of protocols. This is done by modeling implementation and specification and checking for their equivalence under the appropriate semantics.

In recent years, the idea to use graded monads in the study of process semantics has emerged. An advantage of taking this approach, in addition to its coalgebraic generality, is that in some cases one can systematically extract the so called characteristic logic of a semantics. Simply put, a logic consists of certain kinds of formulas and a definition of when they are satisfied by, in our case, a process. The characteristic logic for a semantics defines the types of formulas in such a way that whenever two processes are semantically equivalent, there is no formula, i.e. property, that distinguishes them from each other, e.g. is satisfied by one and not by the other. In the case of inequivalence, distinguishing formulas serve as an easily verifiable certificate of this fact.

The first goal is to extend the framework of graded semantics to include temporal logics, which enable reasoning over unspecified depth of steps. Temporal logics are used in model checking and can be useful in e.g. checking for properties of unspecified depth in data analysis such as “Was program A executed before program B?”.

The study of these semantics and extending the theoretical foundation for their applications in digital forensics and cybersecurity is the main aim of the thesis.

Coalgebraic Automata and Learning Algorithms and their Application in Forensics

Hans-Peter Deifel (hans-peter.deifel@fau.de)
Supervisor: Prof. Dr. Stefan Milius

The study of dynamic systems has a long and rich history in computer science, spanning fields such as classical automata theory, concurrency theory, and IT security. Such systems include deterministic automata, (labeled) transition systems, and probabilistic systems. Historically, algorithms developed for one type of system had to be adapted or reinvented for another one. In contrast, the theory of *Universal Coalgebra* aims to provide a generic framework for systems that encompasses the instances mentioned above and many others.

The use of coalgebraic techniques has recently facilitated the development of a generic *partition refinement algorithm*, which we implemented in a tool that can efficiently minimize a wide array of state based systems. In fact, for many of the studied system types, the generic algorithm matches the run-time complexity of the best known specialized algorithm and for some system types even surpasses it. Genericity is achieved by varying the coalgebraic type functor, but the base category is assumed to be the category of sets.

In the above algorithm, partition refinement is used to compute the state space of a minimal system w.r.t. behavioral equivalence. In this thesis we will, as a first step, extend the algorithm into a fully fledged minimization procedure. This entails moving from computing the state space to also computing the transition structure of the minimal system, while retaining the full genericity.

We will also add support for data automata to the algorithm, by porting it to another base category. Data automata deal with infinite alphabets that are accessible only by a limited API. They arise e.g. when dealing with user data in XML processing.

Another class of algorithms that has recently seen the introduction of coalgebraic techniques is active automata learning, which allows to infer automata models by querying a black-box system. E.g. Angluin's original learning algorithm reconstructs a deterministic finite automaton by posting a series of questions to an adequate *teacher*. Since this pioneering work, similar learning algorithms have been developed for a variety of different systems, motivating the search for a generic method. Advances in this direction were made using coalgebraic methods by Silva et al. with their Coalgebraic Automata Learning Framework and more recently, by Barlocco et al. The latest development is an algebraic approach by Schröder and Urbat. All of these approaches have various shortcomings, in particular, they do not yield a concrete ready-to-use generic algorithm. In this thesis, we will investigate the applicability of those approaches and hope to devise a concrete algorithm with a high level of genericity.

As a case study, we will apply active learning techniques in the field of digital forensics, e.g. by constructing accurate models of black-box systems in digital evidence.

Viktimologie Cybercrime

Julia Drafz (julia.drafz@fau.de)
Supervisor: Prof. Dr. Gabriele Kett-Straub

Die Digitalisierung schreitet in unserer Gesellschaft immer weiter voran und bringt neue Technologien hervor. Nach der JIM-Studie 2019 des Medienpädagogischen Forschungsverbunds Südwest ist von einer flächendeckenden Vollausstattung sowohl mit dem Internet als auch Smartphone in den deutschen Haushalten auszugehen. Das Internet ist somit nicht mehr aus dem Berufsleben und privaten Alltag wegzudenken. Doch die technischen Errungenschaften gehen jedoch nicht nur mit positiven Aspekten einher, da auch Kriminelle das Potential des Internets zum Missbrauch für ihre eigenen Zwecke entdeckt haben. Während das Schadensausmaß enorm ist, ist das Aufdeckungsrisiko aufgrund der Anonymität des Internets und fortschreitenden technischen Entwicklungen gering. Im Gegensatz zu anderen Kriminalitätsbereichen steht die Forschung im Gebiet der Internetkriminalität noch am Anfang. Insbesondere in der (Cyber-)Viktimologie, einem Teilbereich der Kriminologie, das sich mit verschiedenen Facetten der Kriminalitätsopfer beschäftigt, besteht ein großes Forschungsdesiderat.

Bisherige Studien beliefen sich bisher relativ erfolglos auf die ausschließliche Anwendung quantitativer Methoden zur Identifizierung von Risikofaktoren bei Opfern von Cyberkriminalität in der allgemeinen Bevölkerung. Um den Opfern nach einer Viktimisierung zu helfen und Taten im Vorfeld zu verhindern, erfordert es eine Forschung, die sich nicht alleine auf statistische Analysen beschränkt und sich neben der Aufdeckung von Risikofaktoren auch anderen viktimologischen Themenfeldern widmet.

Im Rahmen der Dissertation steht deshalb neben der Aufarbeitung des aktuellen Stands der Opferforschung sowie einer grundlegenden Darstellung des Phänomens Cyberkriminalität die Durchführung einer eigenen empirischen Studie im Vordergrund. Mithilife eines standardisierten Fragebogens sollen Opfererfahrungen von Privatnutzer*innen im Internet und ihr Online-Verhalten erfasst und Daten für eine statistische Analyse gewonnen werden. Im anschließenden qualitativen Forschungsteil werden ausgewählte Cybercrime-Opfer zu ihrem Umgang mit der Tat und der Bewältigung der Tatfolgen interviewt. Diese kombinierte Vorgehensweise schafft zum einen die Möglichkeit, statistische Kennwerte zu erhalten und zum anderen mittels einer qualitativen Inhaltsanalyse nach Mayring Wissen über den Umgang mit der Tat von Cybercrime-Opfer zu generieren, welche dann in die Opferhilfe und Präventionsarbeit einfließen können.

Digital Stratigraphy: Chronological Dating for Digital Forensics

Lisa Marie Dreier (lisa.dreier@fau.de)
Supervisor: Prof. Dr.-Ing. Felix Freiling

In criminal investigations, understanding temporal relationships often is crucial to interrelate evidence or exonerate suspects. In the case of digital evidence such temporal relationships are usually established by collecting and interrelating timestamps in file systems or log files. Although this can be a good approach in general, such an analysis entirely depends on the existence of reliable timestamps. But these do not necessarily exist, as illustrated in the case of file recovery in file systems: Even though deleted files often can be recovered fully or partially, associated metadata (and thus corresponding timestamps) might already be overwritten. Besides, some recovery methods (e.g., file carving) do not recover the associated timestamps by design.

But temporal relationships can not only be established by interpreting a set of timestamps. Instead, Casey proposed a method to estimate a time frame for the creation date of a deleted file based on the file's location on disk combined with time specifications of neighboring files.[1]. It is part of a bundle of methods and observations summarized under the term digital stratigraphy, with their concepts transferred from archaeology and geology. In these two disciplines, stratigraphy is a well-known method for establishing a relative chronology (and thus temporal relationships) between different sediment or rock layers (including the objects they contain). As establishing temporal relationships and assigning dates to objects is an essential objective in both disciplines, they have a wide range of such methods summarized under the term chronological dating.

Thus, this thesis will explore different ways of chronological dating in different disciplines and investigate which of these methods can be transferred to digital forensics. For this purpose, it will focus on the inherent concepts of the methods and constitute an overview of their characteristics before investigating how they can be applied to digital forensics. Nevertheless, one main focus of the work will still be digital stratigraphy, with the aim of improving and extending its capabilities to establish temporal relationships: this thesis will extend the method in a way that makes it applicable for further digital media, e.g., databases, and refine it with additional concepts transferred from other disciplines, such as the Harris-Matrix taken from archeology.

[1] Eoghan Casey, "Digital Stratigraphy: Contextual Analysis of File System Traces in Forensic Science", Journal of Forensic Sciences, vol. 63, p. 1383-1391, 2018.

Investigations into Automata for Data Languages and their Applications in Forensics

Florian Frank (florian.ff.frank@fau.de)
Supervisor: Prof. Dr. Stefan Milius

Infinite alphabets are used to model the communication of values from infinite data types such as nonces, channel names, process identifiers, URLs, data values in XML documents, object identities, or abstract resources. Automata models for *data languages*, which are languages over an infinite alphabet, have received considerable attention in recent research. Typically, these languages are accepted by register automata, first introduced by Kaminski and Francez[1] and extended by Kaminski and Zeitlin[2]. These automata have a finite state description, which then generates an infinite configuration space by use of the infinite alphabet. Another model for data languages are *nominal automata*, first introduced by Bojańczyk, Klin, and Lasota[3]. These automata have infinitely many states but are often required to be *orbit-finite*: they have only finitely many states up to renaming implicitly stored data values. Both types of automata have shown to be equi-expressive, yet nominal automata enjoy many properties convenient for talking about them in an abstract manner.

In this thesis we shall explore different topics arising at the interface of automata theory, logic, and (co-)algebras: We will look at new automata models for data languages and work out the details of their semantics and equivalent descriptions. In addition, we will describe equivalent logics for the class of data languages accepted by these automata models. In particular, we have started an investigation of *presheaf automata*. During this, we found a class of languages accepted by specific presheaf automata, which had wonderful properties for reasoning about them. We found a characterization of these by using monadic second order logic with equality tests and also equivalences to other automata types already studied before. Later we will also work out the coalgebraic semantics for presheaf automata, similarly to our previous work where this was done for non-deterministic orbit-finite nominial automata (NOFAs) and regular non-deterministic nominal automata (RNNAs).[4] With these presheaf automata we will use the internal logics of toposes.

Finally, we will discuss applications of these automata in different fields of digital forensics. Nominal automata have previously been used for model checking, where an automaton model is exhaustively verified against a property specified in an expressive logic. We aim to extend this towards other automata models and to properties of interest in forensics.

[1] M. Kaminski, N. Francez, ?Finite-memory automata?, *Theoretical Computer Science*, Volume 134, Issue 2, 1994, Pages 329-363, ISSN 0304-3975

[2] M. Kaminski, D. Zeitlin, ?Finite-Memory Automata with Non-Deterministic Reassignment?, *Int. J. Found. Comput. Sci.*, Volume 21, Issue 5, 2010, Pages 741-760

[3] M. Bojańczyk, B. Klin, S. Lasota, ?Automata theory in nominal sets?, *Logical Methods in Computer Science*, Volume 10, Issue 3, 2014

[4] F. Frank, S. Milius, H. Urbat, ?Coalgebraic Semantics for Nominal Automata?, In Helle Hansen, Fabio Zanasi, eds.: *Coalgebraic Methods in Computer Science. CMCS 2022.* Lecture Notes in Computer Science, vol 13225. Springer, Cham.

Foundations of Adaptor Signatures

Paul Gerhart (paul.gerhart@fau.de)
Supervisor: Prof. Dr. Dominique Schröder

Adaptor signatures are a novel cryptographic primitive with numerous applications to payment channels, blind conditional signatures, and verifiable witness encryption. On a high level, this primitive allows the signer to compute pre-signatures on messages for statements of NP relations. Pre-signatures are publicly verifiable that simultaneously hide and commit to a signature of an underlying signature scheme on that message. Anybody possessing a corresponding witness for the statement can adapt the pre-signature to obtain the “regular” signature. These properties allow building some sort of smart contracts on blockchains that only allow the storage of transactions on chain. Unfortunately, the formal security notions of adaptor signatures are not well understood. The security of many works building on top of adaptor signatures relies on stronger security assumptions not necessarily covered by adaptor signatures. Therefore, we try to build a new formal security model for adaptor signatures that allows building solid proofs. Furthermore, we want to build protocols that do fair exchange based on these new definitions and explore applications in digital forensics.

Forensic Application of Side-Channel Analysis

Paul Krüger (paul.krueger@fau.de)
Supervisor: Prof. Dr. Jürgen Teich

With the ever-increasing spread of embedded systems used in, for example, smart devices and the Internet of Things, and the public's heightened consciousness regarding privacy and data security, the need for secure computing and communication was never higher. While current standards for cryptographic algorithms can be considered mostly secure the issue of side-channel vulnerabilities is omnipresent. Side-channel vulnerabilities may occur when an algorithm implemented on a physical platform is executed generating traces possibly containing information related to a confidential part of the system. This issue is exacerbated when considering the implementation of cryptographic algorithms on embedded systems, where side-channel security is often traded for computational efficiency, further increasing the risk of side-channel attacks. However, these vulnerabilities also present an opportunity for forensic investigation, as systems that might be otherwise unbreakable may still be susceptible to side-channel attacks, enabling investigators to access possibly critical information.

In order to apply side-channel attacks in forensic investigations it is necessary to handle emerging issues due to discrepancies between side-channel attack research and forensic practice: While research usually focuses on often simplistic target platforms, the devices encountered in practical forensic investigations are usually much more complex. Device-specific attacks from the research literature therefore regularly have to be adapted.

In particular, the issue of *Systemic Noise* in power side-channel attacks is investigated. In this context the Systemic Noise of a device is an encapsulation of all noise directly originating from the device's hard- or software components. These noise components may negatively affect side-channel attacks as they can introduce for an observer non-deterministic execution behavior and may invalidate the attacker's assumptions about the system behavior. This thesis mainly investigates the effects of Systemic Noise on the current standards for cryptographic algorithms, namely the Advanced Encryption Standard (AES) and its related modes of operation. The overall goal of this thesis is to enable efficient cross-device application of side-channel attacks by providing generic solutions to Systemic Noise issues that can be applied to different classes of target devices. For this the following three approaches are pursued:

First, we adapt current side-channel attack techniques to enable them to detect, interpret and adapt to Systemic Noise. Second, we investigate side-channel attack techniques that approach Systemic Noise as an additional side channel to attack and obtain further information about the system. And third, we develop side-channel attacks completely circumventing the issue of Systemic Noise by relying on sources of information unaffected by Systemic Noise.

Cyberangriffe auf kritische Infrastrukturen

Mathis Ohlig (mathis.ohlig@fau.de)
Supervisor: Prof. Dr. Hans Kudlich

Immer wieder gibt es Schlagzeilen, dass Infrastrukturen und deren IT-Systeme von Akteuren jedweder Art angegriffen wurden. Eine ausführliche Untersuchung des strafrechtlichen Umgangs mit diesem Phänomen unter Berücksichtigung des geltenden Rechts ist deshalb geboten, indes ist sie bislang nicht erfolgt.

Im Fokus der Analyse der aktuellen Rechtslage stehen zunächst die rund um das Phänomen der Hackingan auf kritische Infrastrukturen relevanten Straftatbestände –Computerdelikte im engeren und im weiteren Sinne. Dabei werden auch Fragen des Allgemeinen Teils des Strafgesetzbuchs beleuchtet. Daran schließt sich die Untersuchung an, ob die bestehenden Straftatbestände und die bestehende strafrechtliche Dogmatik Hackingangriffe auf kritische Infrastrukturen angemessen erfassen. Hierfür ist freilich vorweg die Frage zu beantworten, was überhaupt „kritische Infrastrukturen“ nach der aktuellen Rechtslage sein können. Auch wenn es bisher keine strafrechtliche Legaldefinition gibt und noch keine Straftatbestände besondere Folgen an die Kritikalität von angegriffenen Infrastrukturen knüpfen, ist die Begriffsdefinition doch deshalb relevant, weil es denkbar ist, dass sich aus der Einstufung als kritische Infrastruktur im Rahmen normativer Kriterien besondere dogmatische Folgen ergeben. Es ist das Ziel der Dissertation, die gefundenen Probleme in der gegebenen Rechtslage unter Berücksichtigung der Besonderheiten der Fallgruppe dogmatisch zu lösen.

Auch mögliche Rechtsentwicklungen bedürfen aufgrund aktueller Gesetzesvorhaben verschiedener Akteure schon jetzt ausführlicher Würdigung. Besonders interessant zu sein, verspricht der Vorstoß des Bundesjustizministeriums in einem Eckpunktepapier, insbesondere der §§ 202a ff. StGB zu reformieren. So soll IT-Sicherheitsforschung in Zukunft nicht mehr strafrechtlich verboten sein. Vor dem Hintergrund des bayerischen Antrages im Bundesrat „Für einen effektiven strafrechtlichen Schutz von kritischen Infrastrukturen gegen Cyberangriffe“ stellt sich die Frage, welche Delikte oder Deliktstypen einzuführen im Kontext der kritischen Infrastrukturen sinnvoll und mit der bestehenden Dogmatik – unter Berücksichtigung der erörterten Besonderheiten – und Strafzwecken vereinbar wäre. Nicht nur, aber gerade auch im Hinblick auf die beiden genannten Ansätze, ist zu erforschen, ob und welcher Regelungsbedarf mit Blick auf Hackingangriffe auf kritische Infrastrukturen tatsächlich besteht.

Bringing Science to Mobile Device Forensics

Jenny Ottmann (jenny.ottmann@fau.de)
Supervisor: Prof. Dr.-Ing. Felix Freiling

Mobile devices like smartphones have become an irreplaceable companion for many people today and are used for a multitude of activities such as navigation, communication and entertainment. Therefore, the data stored on a smartphone can serve as a valuable source of evidence during a criminal investigation. Accessing data on a mobile device can be a technical challenge. In smartphones, for example, many measures are employed to protect the users' data, and various methods, some hardware-based, others software-based, have been developed to facilitate data extraction and subsequent analysis[1]. When such extraction methods are used it is important to know if they make any changes to the device under investigation and how reliable their results are as this influences their usability in court proceedings.

To establish under which circumstances an extraction method is reliable and in what cases the produced data could contain errors, testing needs to be performed. With regard to the ever evolving hard- and software, it is important that testing is continously performed under the new circumstances. Because of the possible importance of digital evidence it is not enough to rely upon the word of the tool vendors that they are performing this continuous testing. And, in some cases, the number of different settings in which a method could be used is too big to rely upon one entity to perform testing for all of them. Therefore, it is necessary that practitioners and researchers also perform tool and method validation. However, there are obstacles to this like limited resources and a lack of reference data. More testing could also help to establish the limitations of tools better[2].

As testing is an important factor for the extracted data to serve as reliable evidence, in this thesis possibilities to validate data extraction methods used in the context of mobile device forensics are explored. First, an extensive overview of methods that could be used in digital investigations will be given and the methods classified according to the maximal data access they can provide. Then quality criteria need to be defined and different possibilities for a validation setup assessed. Finally validation methods should be implemented and evaluated regarding their usability and the transferability of results between different devices.

[1] Maxim Chernyshev, Sherali Zeadally, Zubair Baig and Andrew Woodward, "Mobile Forensics: Advances, Challenges, and Research Opportunities", IEEE Security and Privacy, vol. 6_15, p. 42-51, 2017

[2] Graeme Horsman. ""I couldn't find it your honour, it mustn't be there!" – Tool errors, tool limitations and user error in digital forensics". In: Science and Justice 58.6 (Nov. 2018), pp. 433–440.

„Der IT-Sachverständige im Strafverfahren“ — Heuristik und Beweiswürdigung

Nicole Scheler (nicole.scheler@fau.de)
Supervisor: Prof. Dr. Christoph Safferling

Nicht nur viele unserer Lebensinhalte spielen sich nunmehr digital ab, auch die Beweismittel haben längst die analoge Welt verlassen („eEvidence“). Durch die Allgegenwärtigkeit der Informationstechnik in unserem Alltag (Smartphones, Laptops, Wearables, Navigationsgeräte, Sprachassistenten, etc.), können anhand der dabei entstehenden Daten umfassende Persönlichkeits- und Aktivitätsprofile erstellt und digitale Abbilder gespeichert werden. Diese Daten können umfangreiche Spuren enthalten, die auf Sachverhalte aus der körperlichen Welt schließen lassen und menschliches Verhalten nachweisbar machen. Sie zu finden, zu sichern und gerichtsverwertbar auszuwerten ist Gegenstand der IT-Forensik. Diese digitalen Spuren müssen als gerichtsfestes Beweismittel in die Hauptverhandlung eines Strafverfahrens eingeführt werden. Neben den Herausforderungen der Massendatenauswertung, der Heterogenität von Daten sowie der Verschlüsselung der Kommunikation und von Festplatten, stellt sich u.a. auch der „Übersetzungsvorgang“ von digitalen Beweismitteln durch IT-Sachverständige für die anderen Prozessbeteiligten vor Gericht als problematisch dar. Die Gerichte können in vielen Verfahren nicht mehr auf die Hilfe von IT-Sachverständigen verzichten. Aufgrund der steigenden Komplexität informationstechnischer Systeme ist hierfür — neben der reinen Übersetzungstätigkeit in eine menschenlesbare Form durch Software — in zunehmendem Maße auch eine tiefgehende Erläuterung der Ergebnisse von Datenverarbeitungsvorgängen durch menschliche IT-Forensik-Expertinnen und Experten notwendig. Bei mangelnder Kompetenz der Gerichte im Bereich der IT-Forensik besteht die ernstzunehmende Gefahr, dass nicht mehr die Richterinnen und Richter (allein) über Schuld oder Unschuld befinden (§261 StPO), sondern die IT-Sachverständigen in weiten Teilen das Ergebnis hinsichtlich der Schuldfrage determinieren. Um dieser Gefahr vorzubeugen, sollen verschiedene Lösungsansätze entwickelt werden. Zum einen soll ein Vergleich zu den Anfängen anderer forensischer Wissenschaften vor Gericht hergestellt (u.a. DNA-Analysen, Rechtsmedizin, Glaubwürdigkeitsgutachten) und ggf. die dabei entwickelten Regeln auf die IT-Forensik übertragen werden. Standardisierte Verfahren sowohl in der IT-Forensik als auch bei der Bewertung und Würdigung digitaler Beweise sind dringend notwendig für eine vertrauenswürdige und nachvollziehbare Tatsachenqualität, die juristischen und grundrechtseinschränkenden Entscheidungen (wie Ermittlungsmaßnahmen und Verurteilungen) zugrundeliegen. Zum anderen könnte eine präzisere Kommunikation zwischen verfahrensbeteiligten Juristinnen und Juristen und IT-Sachverständigen notwendig sein, sowie Grundkenntnisse aller Verfahrensbeteiligten hinsichtlich der Besonderheit der IT-Forensik und Daten als Beweismittel, um die Ergebnisse der Gutachten im Rahmen der Beweiswürdigung auf Plausibilität überprüfen zu können.

Tempting Bytes: Vergleiche der Neigung zu Cyberkriminalität und herkömmlicher Kriminalität

Laurin Schwemer (laurin.schwemer@fau.de)
Supervisor: Prof. Dr. Gabriele Kett-Straub

Im November 2022 veröffentlichte das Bundeskriminalamt (BKA) die Ergebnisse der Dunkelfeldstudie „Sicherheit und Kriminalität in Deutschland" (SKiD). Der Studie zufolge sind die Menschen in Deutschland am häufigsten von Straftaten betroffen, die sich dem Bereich der Cyberkriminalität zuordnen lassen. Gleichzeitig sind die Anzeigequoten in diesem Bereich recht gering und das Dunkelfeld daher vergleichsweise groß. Mit der fortschreitenden Digitalisierung in der Gesellschaft nimmt die Bedeutung digital verübter Kriminalität offenbar zu, während herkömmliche Kriminalität vom Volumen her etwas abnimmt. Obwohl Cyberkriminalität häufig grenzüberschreitend ist und die hohe Zahl an Betroffenen in Deutschland nicht unbedingt auf eine hohe Zahl an Tatbegehenden aus Deutschland schließen lässt, stellt sich doch die Frage, ob in einer digitalisierten Gesellschaft die Kriminalitätsneigung zu Cyberkriminalität womöglich höher ist, als zu herkömmlicher Kriminalität.

Das Dissertationsvorhaben, will also herausfinden, inwiefern in der deutschen Bevölkerung eine Kriminalitätsneigung zu Cyberkriminalität besteht und in welchem Maße sich diese von der Kriminalitätsneigung zu herkömmlicher Kriminalität unterscheidet.

Noch zu füllende Forschungslücken betreffen die Tatgelegenheitsstruktur, mögliche Tatmittel und notwendige Kompetenzen im „Tatort" Internet. Des Weiteren wurden bisher fast nie kriminologische Theorien eingesetzt, um das unterschiedliche Ausmaß an Kriminalität im Cyberspace im Vergleich zur Realität beziehungsweise die unterschiedliche Kriminalitätsneigung zu Cyberkriminalität und herkömmlicher Kriminalität zu erklären. Dies möchte das Dissertationsvorhaben ändern, indem es sich den Fragen widmet: Welche technischen Faktoren und Merkmale des digitalen Raumes lassen die Begehung von Cyberkriminalität im engeren Sinne rational oder attraktiv erscheinen? Und: Besteht in Deutschland (als Beispiel für eine digitalisierte Gesellschaft) eine höhere Neigung zu Cyberkriminalität im engeren Sinne als zu herkömmlicher Kriminalität?

Zur Beantwortung der Fragen sollen mehrere Experteninterviews – unter anderem mit überführten Straftätern von Cyberkriminalität – geführt werden, um die tatsächliche Tatgelegenheitsstruktur im Internet besser einschätzen zu können und zu erfahren, welche Mittel und Kompetenzen für die Begehung von Cyberkriminalität im engeren Sinne vonnöten sind. Daran anschließend soll ein Fragebogen für eine allgemeine Bevölkerungsbefragung entworfen und angewendet werden, um die deliktspezifische Kriminalitätsneigung in der Gesellschaft zu erheben. Mit den Ergebnissen lassen sich die Erkenntnisse von SKiD weiter komplementieren und es wird ein Beitrag zur theoretischen Weiterentwicklung geleistet.

Open Source Ermittlungen im Strafverfahren

Tabea Seum (tabea.seum@fau.de)
Supervisor: Prof. Dr. Christoph Safferling

Das Internet beeinflusst und erleichtert nicht nur die Recherche im Privaten, sondern unterstützt auch die Ermittlungen der Polizei und Staatsanwaltschaft. Zur Nachforschung und Durchsuchung im Internet werden Suchmaschinen, wie Google, Yahoo oder duck-duck-go, genutzt. Dabei gibt es neben den genannten kommerziellen Suchdiensten auch spezielle für die Ermittlungen im Netz. Im Rahmen dieser Recherchen sind sensible Daten frei zugänglich und können auf den verschiedenen Plattformen, wie beispielsweise Facebook, Instagram, TikTok usw. gefunden werden. Diese privaten Daten können dann mit anderen Ermittlungsergebnissen verknüpft und abgeglichen werden. Es besteht somit eine Fülle an neuer, möglicher Beweisergebnisse und Indizien im Rahmen der Ermittlungen. Als Kehrseite muss jedoch ein strenger Blick auf das Vorgehen der Datengewinnung durch die Ermittler geworfen werden. Aufgrund von Fakeprofilen und Fakeinformationen kann die Gefahr bestehen, dass diese nicht authentisch und integer sind. Das vorliegende Promotionsverfahren befasst sich mit den aufkommenden, rechtlichen Problematiken bei der Verwendung solcher Suchmaschinen durch die Ermittlungsbehörden, die sich sowohl im Ermittlungsverfahren als auch in der Hauptverhandlung ergeben können.

Primär stellt sich die Frage nach einer angemessenen Rechtsgrundlage. Aktuell wird § 161 StPO, die Allgemeine Ermittlungsbefugnis, verwendet. Dabei ist fraglich, ob die Norm das Spannungsfeld der schutzwürdigen Daten zum einen und der Eingriffsintensität zum anderen erfasst und diese somit als ausreichend erachtet werden kann oder ob zur Gewährleistung der ausgewogenen Beachtung der Individualrechte des Beschuldigten/Angeklagten und einer effizienten Strafverfolgung der Gesetzgeber tätig werden müsste.

Weiterhin kann die Verknüpfung der digitalen Ermittlungsergebnisse mit anderen Ergebnissen aus den Ermittlungen zu einem gesonderten Eingriff in die Privatsphäre führen. Daraus folgt das Bedürfnis einer genaueren Untersuchung, ob ein solcher Eingriff vorliegen kann und wenn dies zu bejahen ist, wie der Eingriff rechtlich zu behandeln ist.

Auch die einzelnen Suchmaschinen, die die Ermittler verwenden, müssen genauer betrachtet werden. Je nach Ausgestaltung beinhalten diese KI-Anteile. Insoweit ist es notwendig, die Funktionsweise genauer zu untersuchen und zu verstehen. Aufgrund des gewonnenen Verständnisses können dann etwaige rechtliche Probleme erkannt, behandelt und gelöst werden.

Im Rahmen der Hauptverhandlung muss sich das Gericht dann kritisch mit der Herkunft und dem Vorgang der Beweisgewinnung auseinandersetzen. Nur durch eine sorgfältige Analyse kann eine Fehleinschätzung des Beweiswertes vermieden werden. Es wird somit die Frage untersucht, welche Anforderungen und Kriterien an die Vorlage der Open-Source Beweismittel zu stellen sind.

Detection of AI-Generated Images

Lea Uhlenbrock (lea.uhlenbrock@fau.de)
Supervisor: PD Dr. Christian Riess

In forensic investigations, images can show valuable leads or they can serve as evidence in court. A variety of digital tools has been developed to validate the origin and authenticity of images. Conversely, there also exist tools for detecting image tampering. This includes, for example, image splicing, object removal or copied image regions. One new challenge for image forensics analysis is the emergence of images generated by artificial intelligence (AI). With only minimal human interaction, such AI-based generators are able to create image content that is visually highly plausible. The detection of such generated content is an open problem. Currently, the most promising approach for forensic detectors is to use machine learning (ML).

The subject of this thesis is to add to the state of the art in ML-based image forensics with neural networks. The pursued research addresses the following questions: Which traces are left by AI-based image generators? Which neural network architecture is best suited for detecting generated images? How can a learning-based detector framework be constructed, such that the results of the image analysis provide as much interpretability and explainability to an analyst as possible?

Towards answering these questions, first results were found for images that are generated from Generative Adversarial Networks (GANs): GANs exhibit specific traces in transform domains, and it was even found that specific model architectures leave fingerprints in the images. However, these insights only form the beginning of understanding the specific representations of generated images. For example, the current state-of-the-art Diffusion Models are considerably more difficult to detect by established forensic techniques. This shows that we need a broader and deeper understanding of the representation of synthetic images in order to be able to reliably analyze them.

Two aspects are considered particularly important for the pursued research on forensic detectors. First, the technology behind forensic detectors has to be sufficiently flexible to be able to keep up with the rapid progress in the state of the art in synthetic image generation. This aspect encourages research in features and representations that capture inherent properties of generated content. Second, forensic detectors have to be designed with interpretability and explainability in mind. This is a critical component for an analyst to understand whether a detection result is reliable and justified. Such understanding is on one hand important when considering that also pristine images are created nowadays with an increasingly large amount of advanced image processing. Hence, an analyst must distinguish between such "expected" image processing and post-hoc artificially generated content. On the other hand, forensic detection methods for generated content almost exclusively rely on black-box learning-based methods, which make it challenging for an analyst to recognize the circumstances when she can trust the results.

To tackle these challenges, we investigate a learning framework that consists of multiple neural networks. Each network is specialized in extracting one type of trace. One system containing the different networks then adapts their combined output to create a final result. This ensures generalizability to different scenarios and generator models. As each sub-network is specialized on finding specific traces, this framework also implies a certain level of explainability and interpretability of the end result. Our research analyzes and discusses the conceptional details of this framework and its implementation to allow valuable insights into the detection of synthetic images.

Forensic Disk Image Generation Revisited

Lena Lucia Voigt (lena.lucia.voigt@fau.de)
Supervisor: Prof. Dr.-Ing. Felix Freiling

Various usage scenarios that necessitate realistic forensic data have been identified in the past. Some of the most prominent ones are education and training, tool testing, malware analysis, and research and development. In most cases, the deployment of synthetic data is inevitable as the use of real-world data is severely restricted due to privacy concerns as well as non-disclosure obligations. However, synthetic data often contains traces of its *artificial* creation and lacks *realistic* background noise or wear-and-tear artifacts that are irrelevant to the case under consideration but contribute to the comparability of synthetic data to real-world data.

For more than a decade, separate approaches have been introduced by the scientific community – with various concrete objectives in mind – to enhance the creation of forensic data, resulting in frameworks like Forensig^2, EviPlant, TraceGen, or ForTrace. Nevertheless, the authors of previous work uniformly acknowledge that there still does not seem to be an adequate solution.

In this work, we build upon existing strategies in the field of forensic data generation, but take a novel viewpoint by integrating ideas known from other research areas, such as malware sandbox detection and Large Language Models. However, the first step towards tackling the problem of creating realistic data, is to precisely define and formalize the concept of *realistic* data, contrasting them to *non-realistic* data. Subsequently, ways to evaluate the realism of synthetic data by means of quantitative as well as qualitative measures will be studied. Moreover, we will distinguish the concrete requirements on generated data that are implied by different usage scenarios.

While different starting points are conceivable, in this thesis we will first focus our attention on the generation of realistic disk images for education and training purposes. An extension of the approach and evaluation of its suitability for other applications or data types is intended in the future.

GRK 2535: Knowledge- and Data-Driven Personalization of Medicine at the Point of Care (WisPerMed)

Prof. Dr. Britta Böckmann
Email: britta.boeckmann@uk-essen.de
University of Duisburg-Essen & Fachhochschule Dortmund
Internet: https://wispermed.com/

WisPerMed
Wissens- und datenbasierte
Personalisierung von Medizin
am Point of Care

Thanks to increasing digitization in medicine, more and more data is becoming available, for example in electronic patient records, through laboratory analyses, or even in treatment guidelines. One challenge is to make the knowledge contained in this very diverse data available and usable at the point of treatment for concrete individual therapy decisions. Existing clinical information systems allow the collection and storage of important information, but usually in a relatively unstructured way and without an individual, context-related compilation of the facts relevant for a treatment decision. The aim of the research training group is to train young researchers from the fields of medical informatics, computer science, statistics, epidemiology, and psychology so that they obtain a holistic overview of the state of research on knowledge- and data-based personalization of medical decision-making processes and learn to design new methods on an interdisciplinary basis and implement them prototypically using the example of malignant melanoma. For this purpose, methods from the fields of information extraction, knowledge representation with machine learning methods, and insights into user interaction at the point of care will be combined in a novel way. Through interdisciplinary measures, in particular through job shadowing in the dermatology clinic, barriers to understanding between the disciplines are broken down. Unique for a research training group is the cross-institutional cooperation between the Dortmund University of Applied Sciences and Arts, the University of Duisburg-Essen, and the University Hospital Essen, which is based on an already existing cooperation through a joint study program in medical informatics. Together, the applicants represent broad expertise in the fields of medical informatics, bioinformatics, epidemiology, artificial intelligence, psychology, radiology, and melanoma research. Graduates of our program will be able to take leading roles in the digitization process of healthcare and further improve treatment pathways using artificial intelligence techniques, taking into account the direct feedback and experience of the treating physicians.

Extracting prognostic indicators of patient outcome from pre-clinical histopathology image data and additional clinical data

Mohamed Albahri (mohamed.albahri@fh-dortmund.de)
Supervisor: Prof. Dr. Markus Kukuk

For malignant melanoma, an aggressive melanocyte-derived tumor, we are observing a global rise in cases[1]. Although it represents a small fraction of skin cancer diagnoses, it leads to the majority of skin cancer-related deaths[2]. This severity is heightened by genetic alterations, notably BRAF mutations.
Acknowledging the complexities clinicians face in making treatment decisions, deep learning emerges as a critical support tool by bridging the gap between advanced technology and personalized care. By deeply analyzing medical images and uncovering complex patterns, deep learning anables clinicians to consider more nuanced and effective treatment decisions.

The project focuses on leveraging pre-clinical histopathological images and additional clinical data to identify prognostic indicators for patient outcomes using deep learning techniques for skin cancer. It builds on the analysis of melanoma skin cancer whole slide images (WSI), employing both public datasets - such as the Cancer Genome Atlas (TCGA) Skin Cutaneous Melanoma (SKCM) - and our newly digitized image data.

Focusing on an end-to-end learning approach for gigapixel images without requiring pixel or patch-level labels, our research aims to enhance BRAF mutation predictions across entire WSIs. Additionally, our goal to explore prognostic variables, such as the presence of tumor-infiltrating lymphocytes, mitotic rates, and tumor necrosis factors, emphasizes a comprehensive strategy for understanding factors critical to melanoma patient survival rates. A further objective is to investigate how to integrate Dermato-Histopathology specific foundation models into our research and assess the feasibility of such integration.

The project adopts a self-supervised learning approach for initial feature representation enhancement, followed by fine-tuning for specific tasks, and investigates the effectiveness of transferring features from ImageNet-trained models to the medical domain, which showed in the last years very promising results compared to traditional supervised machine learning algorithms.

[1] Siegel R, Ward E, Brawley O, Jemal A. Cancer statistics, 2011: the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. CA Cancer J Clin. 2011;61:212–36.

[2] Dinnes J, Ferrante di Ruffano L, Takwoingi Y, Cheung ST, Nathan P, Matin RN, et al. Ultrasound, CT, MRI, or PET-CT for staging and re-staging of adults with cutaneous melanoma. Cochrane Database Syst Rev. 2019;7(7):CD012806. https://doi.org/10.1002/ 14651858.CD012806.pub2.

Application of Graph Representation Learning for Analysis of FHIR Graphs in Patient Data

Mikel Bahn (mikel.bahn@uk-essen.de)
Supervisor: Prof. Dr. med. Felix Nensa

Many current medical data models, algorithms, and applications restrict the data driven perspective on patients by a gridded or sequential nature of representation. Such approaches in current digital treatment programmatically don't take rich interconnected and multi-modal facets of the true nature of patients and their relation to medical entities such as diagnoses, treatments into account and can only hardly provide unified and holistic insight into data at the point of care. At the same time, standardized and graph-based knowledge representation like FHIR continue to gain in popularity in the medical domain leading to an increasing graph-based representation of medical entities and an increasing gap between existing knowledge and corresponding knowledge exploitation. Moreover, data-based representations of medical entities suffer from incomplete and incorrect data and more over are fundamentally limited in their representation qualities by the current understanding of the subject matter at hand. Especially complex medical operating procedures or complex diseases can hinder clear and effective treatment because of overwhelming or vanishingly small amounts of relevant data and knowledge negatively influencing the medical operator and algorithmic approaches in their results.

Graph-based analytical models promise to help to uncover previously undiscovered structure, patterns, and relationships in ever growing semantic medical data that may be critical for improved patient stratification, personalized therapies, diagnostics support, treatment gap or error resolution, and more accurate prediction of disease progression by allowing to exploit contextualized information derived from highly interconnected FHIR graphs. Due to their statistical holistic nature such methods allow to enhance current perspectives on knowledge representation, enable contextualized information extraction and retrieval, and provide a basis to overcome factual barriers of classical gridded or sequential data-based models.

This work proposes to use heterogeneous and multi-modal graphs as a more natural and unrestricted representation to model and analyze complex medical entities like a patient that may be composed of features like medication, care plan and so forth simultaneously. Upon this representation, Graph Representation Learning techniques such as Graph Neural Networks along with generative Large Language Models are at the center of this work and will be used in a two folded way. On one hand, along proper data modeling techniques that adhere the logical structure of FHIR formatted data, they'll be used to model rich interactions relevant at the point of care to allow for a wide class of use cases such as patient stratification, data augmentation or synthetic data generation. On the other hand, such learned representations will be used as a factual but also machine-readable basis to provide medical professionals new capabilities of interacting with these complex entities transparently. This is done by exposing natural language-based interfaces in interactive ways bridging factual and probabilistic knowledge but also investigating upon new ways to make contextualized and personalized information explicitly available within current applications at the point of care.

Multimodal Foundation Models for Medical Images and Text

Helmut Becker (helmut.becker@uk-essen.de)
Supervisor: Prof. Dr. Dr. Jens Kleesiek

When treating patients, clinicians rely on integrating multiple sources of information for their diagnosis and prognosis. The treatment is based on a wide variety of data modalities including medical images from pathology, radiology and nuclear medicine, as well as clinical notes and tabular data from laboratory tests. Adapting this multimodal approach to machine learning using foundation models as a starting point shows great potential to leverage the performance of predicting clinical endpoints for melanoma patients.

Foundation models have shown remarkable success in solving computer vision tasks such as semantic or instance segmentation[1]. These generalist models are typically trained on extensive datasets using self-supervised or unsupervised methods to learn valuable representations for later downstream-tasks. Successively, these base models can be fine-tuned utilizing few-shot or domain-specific learning to create a highly adaptable framework with increased performances on clinical tasks. This strategy already exhibits great potential for highly capable models in the medical domain without the use of multimodal data.[2]. The incorporation of multiple clinical data modalities could increase this potential even further.

The RTG WisPerMed strives for knowledge and data-driven personalization of medicine and clinical solutions at the Point of Care based on the use case of melanoma patients. In this context, this project aims to use state-of-the-art foundation models and extend their capabilities for multimodal data. We start by evaluating the fine-tuning of already existing foundation models and utilizing fusing approaches to merge multiple information streams for solving a set of advanced clinical tasks ranging from the outcome of cancer progression to survival prediction. Based on these results, a potential next step would be the development of unified training strategies to create a new multimodal foundation model. This model will be able to predict clinical endpoints by analyzing whole slide images (WSI) from pathology, CT scans from radiology, unstructured data like electronic health records (EHR) and structured data such as genomics data. The existing knowledge and experiences of previous work such as cell segmentation and detection in WSIs[3], whole-body CT segmentation[4] and text processing[5] can be used to find new solutions for the aims of this project and the processing of the above called data modalities within our foundation model-based architecture. The required computational resources for training and inference will be provided by our newly integrated KITE (KI Translation Essen)

[1] Kirillov, Alexander, et al. "Segment anything." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023.

[2] Ma, Jun, et al. "Segment anything in medical images." Nature Communications 15.1 (2024): 654.

[3] Hörst, Fabian, et al. "Cellvit: Vision transformers for precise cell segmentation and classification." Medical Image Analysis (2024): 103143.

[4] Jaus, Alexander, et al. "Towards unifying anatomy segmentation: automated generation of a full-body CT dataset via knowledge aggregation and anatomical guidelines." arXiv preprint arXiv:2307.13375 (2023).

[5] Dada, Amin, et al. "Information extraction from weakly structured radiological reports with natural language queries." European Radiology 34.1 (2024): 330-337.

infrastructure to leverage the outcome of this project.[6]. The results of our work will be used at the Point of Care by integrating the model into an interactive dashboard developed by the previous cohort.

[6] https://kite.ikim.nrw/

Where does the AI look? Influence of convergence of human attention foci and AI attention foci on reliance on a clinical decision support system

Noëlle Bender (noelle.bender@uni-due.de)
Supervisor: Prof. Dr. Nicole Krämer

Clinical decision support systems that employ AI-generated insights allow practitioners to reduce their mental workload and make elaborate and highly accurate decisions [1]. Deep learning, such as Convolutional Neural Networks (CNN), can support human decision-making, especially in medical imaging. However, accepting and relying on AI advice in medical decision-making does not necessarily reflect this trend [2]. The rise of contemporary technology has developed from tool-based usage to more human-like interaction [3], which requires an extension of the human mentalizing capabilities with a "theory of artificial minds" [4]. To accept and trust the support of another party, one needs to be aware of one's own perceptual uncertainty and confidence [5] as well as the inner workings and uncertainty of the AI.

Therefore, this research aims to bridge findings about explainable AI, transparency, interpretability, and high-performing medically trained human decision-making and information processing. The most influential processes for highly sensitive and consequential scenarios are aimed to be identified, mainly applying psychophysical measures to find suitable interventions. With experimental eye-tracking studies, potentially implicit criteria during decision-making will be examined to make evidence-based inferences about trust calibrations. This is central to providing optimal conditions to pave the way toward hybrid intelligence in a medical context [6]

[1] Antoniadi, A. M., Du, Y., Guendouz, Y., Wei, L., Mazo, C., Becker, B. A., and Mooney, C. (2021). Current challenges and future opportunities for XAI in machine learning-based clinical decision support systems: A systematic review. Applied Sciences, 11(11), 5088. https://doi.org/10.3390/app11115088

[2] Küper et al., in press

[3] Köbis, N., Bonnefon, J. F., and Rahwan, I. (2021). Bad machines corrupt good morals. Nature Human Behaviour, 5(6), 679–685. https://doi.org/10.1038/s41562-021-01128-2

[4] Krämer, N., Wischnewski, M., and Müller, E. (2023). Interacting with Autonomous Systems and Intelligent Algorithms – New Theoretical Considerations on the Relation of Understanding and Trust. https://doi.org/10.31234/osf.io/h32ze

[5] Fleming, S. M. (2024). Metacognition and confidence: A review and synthesis. Annual Review of Psychology, 75(1), 241–268. https://doi.org/10.1146/annurev-psych-022423-032425

[6] Akata, Z., Balliet, D., de Rijke, M., Dignum, F., Dignum, V., Eiben, G., Fokkens, A., Grossi, D., Hindriks, K., Hoos, H., Hung, H., Jonker, C., Monz, C., Neerincx, M., Oliehoek, F., Prakken, H., Schlobach, S., van der Gaag, L., van Harmelen, F., . . . Welling, M. (2020). A research agenda for hybrid intelligence: Augmenting human intellect with collaborative, adaptive, responsible, and explainable artificial intelligence. Computer, 53(8), 18–28. https://doi.org/10.1109/MC.2020.2996587.

Interactive retrieval methods for contradictory evidence

Bohao Chu (bohao.chu@uni-due.de)
Supervisor: Prof. Dr.-Ing.Norbert Fuhr

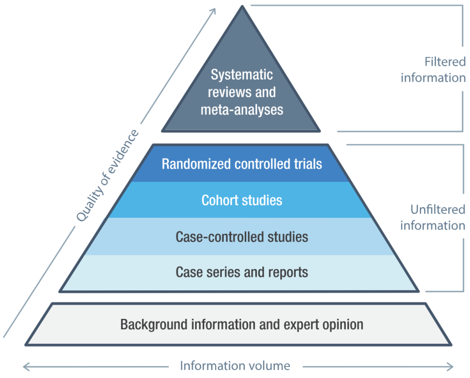


Figure 1: Levels of Evidence Pyramids [1]

For innovative treatments, there are often contradictory publications on their efficacy. Current retrieval methods offer little or no support for this situation. In this project, based on medical entity extraction methods, suitable literature will be searched for diagnosis-therapy pairs specified by the user and classified concerning their respective evidence and significance[2].

The main task is to extract the respective statements on efficacy from the publications found. Then these are aggregated under consideration of the publication's evidence (see Figure 1) [3]. For this purpose, we will explore novel methods which can be applied to retrieve contradictory information. This research will be roughly split into the following tasks:

1. Deploy a medical search engine to retrieve relevant publications based on user (i.e. medical professionals or patients) queries. **2.** Extract and summarize claims regarding the efficacy of therapies from publications or published clinical trial reports searched by users. **3.** Suitable classification schemes will be defined to describe the therapy's efficacy (and side effects) and the corresponding automatic classifiers will be implemented. **4.** Based on classification results, a systematic review related to the therapy will be generated automatically to find conflicting evidence. **5.** An appropriate user interface will be developed in collaboration with the intended end users, including defining interactive functionality. **6.** User studies will finally be conducted to evaluate the new system in comparison with the status.

Furthermore, as part of this research, the following questions will also be addressed: How can users influence the arrangement of the documents found? How to communicate the inherent uncertainty of system decisions? How useful are summaries generated by Large Language Models (LLMs)?

[1] Illustration adapted from model displayed in "Evidence-Based Practice in Health"

[2] Frihat, S. and Fuhr, N. (2024). Determining the Evidence Level of Studies Described in Medical Publications.

[3] Sabbah, F., and Fuhr, N. (2021). A Transparent Logical Framework for Aspect-Oriented Product Ranking Based on User Reviews. ECIR 2021, Part I (pp. 558-571). 2021

Recommending scientific literature for revising clinical guidelines with knowledge graphs

Hendrik Damm (hendrik.damm@fh-dortmund.de)
Supervisor: Prof. Dr.-Ing. Christoph M. Friedrich

Building upon prior efforts to automate the recommendation of scientific literature for clinical guideline revisions, this project narrows its focus on the S3 guideline for melanoma. It innovates by harnessing bibliometrics and transformer-based Natural Language Processing (NLP) methods alongside the integration of knowledge graphs. These graphs, derived from established melanoma ontologies and terminologies, aim to improve the transparency and accuracy of literature recommendations, thereby enhancing the explainability of these processes. This initiative directly addresses challenges identified in previous research, such as the accuracy of literature recommendations and the occurrence of "hallucinations" in outputs from large language models like ChatGPT. By leveraging knowledge graphs, the project proposes a novel solution to mitigate these issues, offering a more reliable basis for recommendations[1].

An evaluation study with guideline developers forms a core part of this project, aiming to validate the effectiveness of the proposed method. Furthermore, the potential of this methodology to be applied to other oncology guidelines is explored, seeking to broadly enhance the clinical guideline revision process across the field.

This approach signifies a significant leap forward in the application of knowledge graphs and computational techniques within medical informatics. It promises substantial contributions to the evolution of clinical practice guidelines, emphasizing the creation of a more robust and scalable framework for integrating the latest scientific evidence into clinical practices. By fostering a deeper understanding and utilization of knowledge graphs, the project not only aims to streamline the guideline revision process but also to establish a more robust and scalable framework for incorporating the latest scientific evidence into clinical practice.

[1] Michalowski, M. Rao, M.; Wilk, S.; Michalowski, W.; Carrier, M. (2023), „Using graph rewriting to operationalize medical knowledge for the revision of concurrently applied clinical practice guidelines“, in Artificial Intelligence in Medicine, Volume 140, 102550, https://doi.org/10.1016/j.artmed.2023.102550.

Digital sovereignty of medical practitioners

Enis Dogru (enis.dogru@uni-due.de)
Supervisor: Prof. Dr. Nicole Krämer

For technological developments to succeed in a clinical environment, medical practitioners must be willing to engage with the technology and consider its advice and support. Previous research on algorithm aversion and appreciation has shown that users may not trust AI-based support due to previous attitudes or negative experiences[1] [2]. The current project aims to explore how medical practitioners can be empowered to better judge when to trust AI-generated advice. The work will be grounded in assumptions on calibrated trust[3], algorithm literacy, and explainable AI[4], as well as assumptions on the balance of understanding and trust [5].

As a first step in this research project, it is planned to conduct a qualitative study – interviewing dermatologists at the University Clinic Essen to gain insight into the decision-making process of a physician and to ask about their current attitudes towards working with and relying on artificial intelligence.

Simultaneously, a quantitative online study will be developed to test cognitive forcing functions (CFFs) – design elements of decision support systems that are hypothesized to mitigate cognitive biases in the context of explainable AI[6]. If integrated at the right point in the decision-making process, CFFs could help build calibrated trust by encouraging the user to engage more deeply with the explanation provided by the AI. This could mitigate over-reliance and therefore reduce the number of errors made by the human-AI team.

Altogether, by means of qualitative interviews, subsequent quantitative surveys and experimental designs, this project aims to uncover which factors influence whether medical practitioners trust in the specific system is actually warranted and how to develop digital sovereignty.

[1] Dietvorst, B. J., Simmons, J. P., and Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. Journal of Experimental Psychology: General, 144(1), 114–126. https://doi.org/10.1037/xge0000033

[2] Logg, J. M., Minson, J. A., and Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. Organizational Behavior and Human Decision Processes, 151, 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

[3] Lee, J. D., and See, K. A. (2004). Trust in automation: Designing for appropriate reliance. Human Factors, 46(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

[4] Leichtmann, B., Humer, C., Hinterreiter, A., Streit, M. and Mara, M. (2023). Effects of Explainable Artificial Intelligence on trust and human behavior in a high-risk decision task. Computers in Human Behavior, 139, 107539. doi.org/10.1016/j.chb.2022.107539

[5] Krämer, N., Wischnewski, M., and Müller, E. (2023, May 7). Interacting with autonomous systems and intelligent algorithms – new theoretical considerations on the relation of understanding and trust. https://doi.org/10.31234/osf.io/h32ze

[6] Buçinca, Z., et al. (2021). "To Trust or to Think." Proceedings of the ACM on Human-Computer Interaction 5(CSCW1): 1-21.

Empirical evaluation of developed medical informatics tools at the point-of-care

Kilian Elfert (kilian.elfert@uk-essen.de)
Supervisor: Prof. Dr. Andreas Stang

During the first three years of the WisPerMed research training group, various medical informatics tools have been developed to assist in the clinical care of patients with malignant melanoma at the point-of-care.

The evaluation of software tools includes the systematic investigation of the usability of an object. The evaluation results and conclusions reached must be justified on the basis of empirical data (qualitative as well as quantitative). Subjective and objective methods can be applied. Objective methods include, for example, observational interviews, log file recording, video recordings, keystroke level method. Subjective methods mainly include surveys involving both users and experts (according to e.g. cognitive walkthrough). Usability tests simulate the real-life case and use problem-oriented tasks. Usability attributes include learnability, efficiency, memorability, errors, and satisfaction [1].

The evaluation approaches to be developed should take into account value criteria (task accomplishment, functionality, interaction effort), suitability for the task, self-decriptiveness, conformity with user expectations, controllability, error tolerance (e.g. undo dialog steps, understandable error messages), suitability of individualization, and suitability for learning. After the development of the evaluation approaches, the evaluations take place and are evaluated after completion and discussed in a group of experts. In addition to the usability assessment, the evaluation should also be formative, which means that it uncovers any remaining weaknesses that can subsequently be remedied.

[1] Nielsen J. Usability engineering. San Francisco, California: Morgan Kaufmann; 1993.

Personalization of Medical Decision-making at the Point of Care using Large Language Models and Chatbots

Bahadır Eryılmaz (bahadir.eryilmaz@uk-essen.de)
Supervisor: Prof. Dr. med. Felix Nensa

The introduction of Large Language Models (LLMs) into the medical sector has gained significant traction as the applications for clinical use cases seem limitless[1][2]. Unlike their Natural Language Processing (NLP) predecessors, LLMs have the ability to build a deeper understanding of language, and therefore a deeper understanding of complex domains such as medicine, which, if harnessed correctly, could significantly improve personalised and context-aware medical care at the point of care.

Before personalised medicine can be achieved, one of the major challenges facing healthcare is extracting relevant information from unstructured clinical documents and ensuring a high level of structured information about patients. Clinical notes and reports, among other sources, contain vast amounts of valuable data that are not easily accessible or interpretable by traditional data management systems[3]. Transforming this unstructured data into a recognised and interoperable data standard, such as Fast Healthcare Interoperability Resources (FHIR), is seen as a promising way to effectively leverage this detailed clinical insight. However, the process of extracting structured FHIR resources - such as patient health conditions, medical procedures and medication information - from unstructured text is dominated by Named Entity Approaches (NER), which require sophisticated mapping of documents.

To address this issue and leverage the capabilities of Large Language Models (LLMs), the proposed project focuses on the efficient and scalable transformation of unstructured clinical documents into structured FHIR resources using open-source LLMs such as Mistral[4] or Mixtral[5]. To achieve this, a dataset of pairs of clinical notes and corresponding FHIR resources should be curated and used for fine-tuning runs using a mix of manual and automated data processing steps. In addition, the training process should include the mapping of widely used coding systems like SNOMED, LOINC, or ICD10.

In this work, LLMs will be used to accurately extract relevant data from unstructured text. Key to this process is the quality of the training data, which can be assembled semi-automatically through the use of another LLM or manually curated with human input. Once a robust training dataset is established, a open-source generative LLM will be employed to create structured data. By establishing an automated pipeline for extracting

[1] Thirunavukarasu, A.J., Ting, D.S.J., Elangovan, K. et al. Large language models in medicine. Nat Med 29, 1930–1940 (2023). https://doi.org/10.1038/s41591-023-02448-8

[2] Yu, Ping et al. "Leveraging Generative AI and Large Language Models: A Comprehensive Roadmap for Healthcare Integration." Healthcare 11 (2023): n. pag.

[3] Adnan, K., Akbar, R., Khor, S.W., Ali, A.B.A. (2020). Role and Challenges of Unstructured Big Data in Healthcare. In: Sharma, N., Chakrabarti, A., Balas, V. (eds) Data Management, Analytics and Innovation. Advances in Intelligent Systems and Computing, vol 1042. Springer, Singapore. https://doi.org/10.1007/978-981-32-9949-8_22

[4] Jiang, Albert Qiaochu et al. "Mistral 7B." ArXiv abs/2310.06825 (2023): n. pag.

[5] Jiang, Albert Q. et al. "Mixtral of Experts." ArXiv abs/2401.04088 (2024): n. pag.

relevant patient information from unstructured documents, the first step is taken for subsequent use cases such as personalised medicine and healthcare-based chatbots.

Clinical pathway discovery based on electronic health record data

Praveen Jagan Nath (praveen.nath@fh-dortmund.de)
Supervisor: Prof. Dr. Britta Böckmann

When treating patients with melanoma, clinical guidelines and standard operating procedures (SOP's) are essential sources of evidence-based knowledge that support well-informed medical decision-making. Despite their significance, effective electronic patient documentation that is suited to clinical requirements, is frequently lacking from today's hospital information systems. This Deficit makes it difficult for doctors to make prompt treatment decisions, since they have to sift through a lot of data to find pertinent patient information.

In order to overcome these obstacles, our study improves the usability of Electronic Health Records (EHRs) for individualized melanoma treatment by utilizing previous work on integrating patient data in clinical pathways. Our goal is to simulate and dynamically learn patient-specific clinical pathways using real-time EHR data by utilizing machine learning techniques. Techniques for process mining will be used to extract event records from EHRs that show the course of a patient's illness and course of treatment. These event logs will be shown as clinical pathways unique to each patient, allowing for a direct comparison with SOP's to ascertain the best course of action for each patient.

To find paths from data, our method evaluates deep-learning based techniques like Variational- and Predictive-Auto-Encoders as well as stochastic techniques like Hidden-Markov-Models. The results of this study may serve as a foundation for further research aimed at modelling treatment outcomes according to the particular circumstances and medical background of each patient.

Summaries and lay translations of clinical documents with transformers

Tabea M. G. Pakull (tabeamargaretagrace.pakull@uk-essen.de)
Supervisor: Prof. Dr.-Ing. Christoph M. Friedrich

During the course of a patient's treatment, a variety of clinical documents accumulate that are often difficult for patients to understand. Providing lay summaries of the contents of this documents can improve the patient-doctor communication and can lead to better informed patients[1]. However, the manual preparation of these summaries is time-consuming and in most cases not feasible for doctors in their busy day-to-day clinical work. To address this problem, this thesis investigates the synthetic generation of summaries of German clinical texts and their lay translations with transformers.

Pre-trained large language models (LLMs), such as the Generative Pretrained Transformer (GPT)[2], have significantly improved Natural Language Processing (NLP), across a wide range of applications. Self-supervised pre-training on large amounts of unstructured text enables these models to capture semantics, lexical information and knowledge. LLMs outperform previous methods even in few-shot settings, where no labeled data is required. The use of LLMs for summarizing clinical texts at the point of care presents challenges related to data protection, peculiarities of German clinical texts[3], and ensuring the accuracy of the summaries.

To comply with data protection regulations, network-based models cannot be used. Instead, this work will be based on a large clinical language model developed by the members of the graduate school, together with previous work on semi-automatic ontology/ terminology extraction. As part of the project, a synthetic corpus consisting of pairs of findings and corresponding lay translations or summaries will be created. In addition to text-based summaries, the thesis examines alternative presentation formats, including dashboards. Furthermore, an explainability component will be developed and evaluated to ensure the credibility of the generated summaries and translations.

[1] Crucefix, Anna L et al. "Sharing a written medical summary with patients on the post-admission ward round: A qualitative study of clinician and patient experience." Journal of evaluation in clinical practice vol. 27,6 (2021): 1235-1242. doi:10.1111/jep.13574

[2] Brown, Tom et al. "Language Models are Few-Shot Learners" Advances in Neural Information Processing Systems vol. 33 (2020): 1877–1901.

[3] Starlinger, Johannes et al. "How to improve information extraction from German medical records" it - Information Technology vol. 59,4 (2017): 171-179. doi:10.1515/itit-2016-0027

Uncertainty aware haplotype based genomic variant effect prediction

Felix Wiegand (felix.wiegand@udo.edu)
Supervisor: Prof. Dr. Johannes Köster

The prediction of effects of genomic variants is a crucial step in the analysis of genomic data, in particular in precision oncology and precision medicine in general. So far, this has been done by considering each individual variant alone by predicting its effect on a regulatory element or a protein (e.g. causing the early termination of the protein or, less severe, a changed amino acid), estimating the severity of this effect, and annotating public knowledge about the variant (like allele frequencies, pathogenicity, and disease associations). Various tools that accomplish this task have been developed, for example VEP [1] and SNPeff [2].

The consideration of each variant alone can lead to incomplete or in the extreme case even wrong information. For example, on the one hand, a deletion at the beginning of a protein that shifts the reading frame invalidates any downstream predictions of amino acid changes. On the other hand, if this first deletion is followed by an insertion that restores the reading frame, the impact on the protein might be less severe, or at least completely different. Another example is the collaboration of multiple amino acid changing variants that only together generate an effect severe enough to impact a certain binding domain of a protein. Finally, a severe variant inside a promoter or enhancer does only play a role for the targeted genes on the same haplotype or subclone, which needs to be taken into account when interpreting the interplay with potential additional variants in the targeted gene.

To address these challenges, we will develop a completely new approach for variant effect prediction. Instead of considering each variant individually, we will develop a graph structure (impact graph) that represents haplotypes as paths, while representing individual variants as nodes. This way, the impact of a variant can be annotated and interpreted in the context of the surrounding haplotype. Uncertainty in haplotype assignment is captured by alternative paths. Moreover, we will leverage the Bayesian model of Varlociraptor [3] to annotate both edges and nodes with posterior probabilities for haplotype and variant calls. We will further develop an interactive visual representation of the impact graph that allows us to comprehensively assess the impact, any kind of common annotations and the uncertainty of the presented information.

[1] William McLaren et al. "The Ensembl Variant Effect Predictor". In: Genome Biology 17.1 (June 6, 2016), p. 122. doi: 10.1186/s13059-016-0974-4.

[2] Pablo Cingolani et al. "A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff". In: Fly 6.2 (Apr. 1, 2012), pp. 80–92. doi: 10.4161/fly.19695.

[3] Johannes Köster et al. "Varlociraptor: enhancing sensitivity and controlling false discovery rate in somatic indel discovery". In: Genome Biology 21.1 (Apr. 28, 2020), p. 98. doi: 10.1186/s13059-020-01993-6.