Felix Freiling (editor)

Proceedings of the 2021 Joint Workshop of the German Research Training Groups in Computer Science

May 31–June 1, 2021

DFG Deutsche Forschungsgemeinschaft

Bibliographic Data:

Felix Freiling (ed.): Proc. 2021 Joint Workshop German Research Training Groups in Computer Science. Erlangen, 2021.
DOI: 10.25593/opus4-fau-16426

Preface

Initiated in 1996 and run regularly since 2007, researchers of the German Research and Training Groups (RTGs) funded by the Deutsche Forschungsgemeinschaft (DFG) in the field of computer science meet annually at Schloss Dagstuhl — Leibniz Center for Informatics, one of the world's premier venues for computer science-related seminars. The goal of these workshops is to foster an interchange of ideas and experiences in order to strengthen the connection within the German computer science community.

Exactly 25 years after the first event in this series, this year's event was organized by RTG 2475 (Cybercrime and Forensic Computing). It took place on Monday, May 31, 2021 and Tuesday, June 1, 2021. Like the year before, the organization was affected by travel and contact restrictions to prevent the spread of the Corona virus. It was therefore decided to organize the meeting as a shortened online event. Still it featured presentations of the individual RTGs, short networking pitches by funded researchers, networking meetings for PIs and RTG coordinators, and — as a particular highlight — a live interview with Professor Otto Spaniol who had initiated the workshop series 25 years ago.

The editor wishes to thank Lena Voigt and Lena Reinfelder from FAU for their support in organizing and running the meeting and the production of the proceedings. Thanks also go to the organizers of the previous year, Gabriela Pipa, Toni Mattis, and Stefan Ramson, for their helpful advice in organizing this year's meeting.

Certainly, an online event can never match the real experience, and so we all hope to see each other soon again in wonderful Schloss Dagstuhl — one of the best places to meet, talk and interact!

Erlangen, May 31, 2021 Felix Freiling

Year	Organizers	Title	Location (event #)
2021	Freiling (Erlangen)	Joint Meeting of the German RTGs in Computer Science	online
2020	Ramson, Mathis (HPI), Pipa (Osnabrück)	14th Joint Meeting of the German RTGs in Computer Science	online
2019	Bolke-Hermanns, Katoun (Aachen), Mühlhäuser (Darmstadt)	Gemeinsamer Workshop der Graduiertenkollegs GRK 2050 und GRK 2236	Dagstuhl (19253)
2018	Aker, Fuhr (Duisburg-Essen), Rehof, Scholz, Tarkun, Winkels (Dortmund)	Gemeinsamer Workshop der Graduiertenkollegs GRK 2167 und GRK 2193	Dagstuhl (18223)
2017	Arnold, Grube, Gurevych, Heinzerling, Maaß, Mühlhäuser, Tauchmann, Wessels (Darmstadt)	Gemeinsamer Workshop der Graduiertenkollegs — 11th Joint Workshop of the German Research Training in Computer Science: GRKs 1994 AIPHES and 2050 PAT	Dagstuhl (17243)
2016	Herrmann, Lehner, Voigt, Weißbach (Dresden)	Gemeinsamer Workshop der Graduiertenkollegs: GRK 1780 Crossworlds der TU Chemnitz & GRK 1907 RoSI der TU Dresden	Dagstuhl (16213)
2015	Babari (Leipzig), Dotenco, Müller, Stamminger, Yupatova (Erlangen)	Gemeinsamer Workshop der Graduiertenkollegs — 9th Joint Workshop of the German Research Training Groups in Computer Science: GRK 1773 and GRK 1763	Dagstuhl (15232)
2014	Engelmann, Flick, Gao, Ody (Oldenburg), Hahn, Jentzsch, Pasewaldt (HPI)	Gemeinsamer Workshop der Graduiertenkollegs: Interdisciplinary scientific working principles	Dagstuhl (14252)
2013	Bader, Feinen, Kolb (Siegen), do Carmo, Scholl, von Stryk (Darmstadt)	Gemeinsamer Workshop der Graduiertenkollegs: GRK 1362 und GRK 1564	Dagstuhl (13222)
2012	Duske, Janusz, Reisig (Berlin)	Gemeinsamer Workshop der Graduiertenkollegs I: GK 1651 SOAMED et al.	Dagstuhl (12252)
2012	Gladisch, Kirste, Yordanova (Rostock)	Gemeinsamer Workshop der Graduiertenkollegs II: GK 1424 MUSAMA el al.	Dagstuhl (12253)
2011	Seidl (München)	Gemeinsamer Workshop der Graduiertenkollegs I	Dagstuhl (11252)
2011	Mitschele-Thiel (Ilmenau)	Gemeinsamer Workshop der Graduiertenkollegs II	Dagstuhl (11253)
2010	Thomas (Aachen)	Gemeinsamer Workshop der Graduiertenkollegs	Dagstuhl (10222)
2009	Fahland (Berlin)	Gemeinsamer Workshop der Graduiertenkollegs	Dagstuhl (09242)
2008	Meyer (Kaiserslautern)	Gemeinsamer Workshop der Graduiertenkollegs	Dagstuhl (08212)
2007	Spaniol (Aachen)	Gemeinsamer Workshop der Graduiertenkollegs	Dagstuhl (07232)
2005	Spaniol (Aachen)	Gemeinsamer Workshop der Graduiertenkollegs	Dagstuhl (05212)
2003	Spaniol (Aachen)	Gemeinsamer Workshop der Informatik Graduiertenkollegs	Dagstuhl (03252)
2000	Hommel (Berlin), Spaniol (Aachen), Vollmar (Karlsruhe)	Gemeinsamer Workshop der Graduiertenkollegs Aachen, Berlin, Karlsruhe	Dagstuhl (00222)
1996	Spaniol (Aachen)	Gemeinsamer Workshop der Graduiertenkollegs Aachen, Berlin, Darmstadt, Paderborn, Stuttgart	Dagstuhl (96232)

Contents

GRK 1907: Role-based Software Infrastructures for continuous-context-sensitive Systems

Prof. Dr.-Ing. Wolfgang Lehner
Email: wolfgang.lehner[at]tu-dresden.de
Technische Universität Dresden
Internet: https://rosi-project.org

Software with long life cycles is facing continuously changing contexts. New functionality has to be added, new platforms have to be addressed, and existing business rules have to be adjusted. In the available literature, the concept of role modeling has been introduced in different fields and at different times in order to model context-related information, including - above all - the dynamic change of contexts. However, often roles have only been used in an isolated way for context modeling in programming languages, in database modeling, or to specify access control mechanisms. Never have they been used consistently over all levels of abstraction in the software development process, i.e. over the modeling of concepts, languages, applications, and software systems.

The central research goal in this program is to deliver proof of the capability of consistent role modeling and its practical applicability. Consistency means that roles are used systematically for context modeling on all levels of the modeling process. This includes the concept modeling (in meta-languages), the language modeling, and the modeling on the application and software system level. The subsequent scientific elaboration of the role concept, in order to be able to model the change of context on different levels of abstraction, represents another research task in this program. Thus, consistency also means to systematically define relationships between the identified role concepts to allow for model transformations and synchronizations. Such consistency offers significant advantages in the field of software systems engineering because context changes are interrelated on different levels of abstraction; plus, they can be synchronously developed and maintained. Potential application fields are the future smart grid, natural energy based computing, cyber-physical systems in home, traffic, and factories, enterprise resource planning software, or context-sensitive search engines.

Currently, the research training group is being financed by DFG in its second phase (01.04.2018 until 30.09.2022). During the first phase of the research training group (01.10.2013 – 31.03.2018) 10 RoSI PhDs completed their doctorate.

By the end of the second funding phase, additionally, 22 RoSI financed PhDs plus a number of non-DFG-financed PhDs are expected to have completed their doctorate.

The research training group is run by ten Principle Investigators from TU Dresden plus a number of associate members. Regular thesis advisory board meetings, off-site

workshops, lecture series, seminars, invited talks, long-term stays abroad, and a soft skill program are essential elements of the program.

A Formal Foundation of Particle Methods

Johannes Bamme (johannes.bamme@tu-dresden.de)
Supervisor: Prof. Dr. sc. techn. Ivo F. Sbalzarini, Prof. Dr.-Ing. Wolfgang Lehner

Particle Methods are a classic and popular class of algorithms in scientific computing, from applications such as computational plasma physics to computational fluid dynamics. Some of the historically first computer simulations in these domains utilized particle methods and the field is still under active development today. A key advantage of particle methods is their versatility, as they can simulate both discrete (e.g. Discrete Element Methods) and continuous models (e.g. Particle Strength Exchange) either stochastically or deterministically. Particle Methods also include particle-based image processing methods, point-based computer graphics, and computational optimization algorithms using point samples — all of these rest on a common concept, which is so far not formally defined.

We propose a broader definition of particle methods to include all computational algorithms that use points. Therefore, points or particles are zero-dimensional computational objects that store positions and properties and are able to interact and evolve. We restrict particles only to interact pairwise. Higher-order interactions are not permitted — sequences of pairwise interactions using additional particle properties can circumvent this. These concepts lead to a generalized formal definition of particle methods across application domains, from simulation to graphics to computer vision and optimization.

We intentionally formulated the definition in the most general way in order for it to encompass everything that we call a particle method. Most practical instances, however, do not exploit the full generality of the definition. For example, one would frequently restrict a particle method to be order-independent to produce results independent of the particles' indexing order. The generality of our definition also prevents efficient concurrent formulations of particle methods unless it is suitably restricted. Another limitation of our definition of particle methods is its monolithic nature. An algorithm composed of many smaller algorithms, like, for example, a solver for the incompressible Navier-Stokes equation, will become very large and complex with several nested cases when explicitly formulated in our definition. Finally, our definition is not unique. Alternative, possibly more compact or elegant, but equivalent definitions are possible. We chose our formulation as we believe it to be close to practical implementations.

Notwithstanding these limitations, our definition establishes the so-far loose notion of particle methods as a rigorous algorithmic class. It paves the way for future research both in the theoretical, algorithmic foundations of particle methods and in the engineering of their software implementation.

Adaptive Heterogeneous Computation for Database Systems

Johannes Fett (Johannes.fett@tu-dresden.de)
Supervisor: Prof. Dr. Wolfgang Lehner

Nowadays, query processing in column-store database systems is highly tuned to the underlying (co-)processors. This approach works very well from a performance perspective, but has several shortcomings from a conceptual perspective. For example, this tuning introduces high implementation as well as maintenance cost and one implementation cannot be ported to other
(co-)processors. To overcome that, we developed a column-store specific abstraction layer for hardware-driven vectorization based on the Single Instruction Multiple Data (SIMD) parallel paradigm. Thus, we are able to implement vectorized query operators in a hardware-oblivious manner, which can be specialized to different SIMD instruction set extensions of modern x86-processors. To soften the limitation to x86-processors, our vision is to integrate GPUs in our abstraction layer by interpreting GPUs as virtual vector engines. To achieve this it is necessary to map a vector size to a CUDA configuration. Which consists of Threads per block, number of blocks and choice of asynchronous CUDA stream. By conduction a number of experiments, we have shown, that our approach is promising.While the observed throughput does not outperform CUDA native implementations, it allows developers to use CUDA-based fast GPU primitives without requiring knowledge of GPU implementations. Our vectorized approach still achieves reasonable performance that is not an order of magnitude slower than native CUDA implementations. Moreover, tuning the primitives and operators by choosing the right virtual vector size and configuration is critical for achieving good performance.

From Histograms to Sampling: Optimizing (R)SQL

Axel Hertzschuch (Axel.Hertzschuch@tu-dresden.de)
Supervisor: Prof. Dr. Wolfgang Lehner, Prof. Dr. Franz Baader

Modern software systems face a variety of demanding challenges. Emerging hardware, changing environments, and changing development requests drive traditional software modeling processes towards their limit. As pointed out by recent research, immutable runtime objects are likely to become the Achilles heel of modern software systems. To model long-living, highly adaptive, and seamlessly extensible systems, the concept of roles has been introduced. Like any traditional runtime object, role objects can be organized in a relational schema to be persisted in a database.

To gain business insights database systems need to answer complex queries as fast as possible. On existing infrastructures, this task essentially boils down to finding efficient query execution plans. Although being a core research topic for decades, this task is far from being solved, and designing a query optimizer tailored to role-based data yet again reveals the weak-spots of existing query optimizers: (i) frequently changing data distributions, (ii) arbitrary complex filter conditions, (iii) strong attribute correlations, (iv) and complex join paths. While the work of Jaeckel et al. [1] provides the necessary query language extension (RSQL), this thesis focuses on finding fast query execution plans.

For complex filters, that consist of various sub-expression, it is of paramount importance to apply the most selective sub-expressions first to reduce the scan cost of the remaining expressions. Sub-optimal filter execution plans are known to be the result of strong correlation misestimates over multiple attributes (e.g. salary and age). The thesis, therefore, develops a novel estimator that achieves precise estimates independent of the filter complexity, attribute correlation, used data types, and data updates. This novel estimator is implemented into the existing optimization infrastructure of SAP HANA/S4 and studied with real-life data input.

The thesis further studies the connection between precise filter selectivity estimates and good join-orderings. These insights enable us to tackle complex select-project-join queries with a novel enumeration scheme that leads to robust join-orderings and scales well within the number of joined tables. This join-order concept is —yet again— studied over a real-life workload that mimics core characteristics of role-based data.

As query response times are not only dictated by the order in which operators are applied but also the availability of indexes, the thesis develops a data structure that efficiently subsumes and analyses query execution plans to recommend indexes and appropriate physical join operators.

[1] Jaekel, Tobias, The Role Concept for Relational Database Management Systems, International Conference on Conceptual Modeling,277–286, 2013

Context Management in Database Systems with Word Embeddings

Michael Günther (Michael.Guenther@tu-dresden.de)
Supervisor: Prof. Dr.-Ing. Wolfgang Lehner

Natural language processing focuses recently on the development of learned language models called word embedding models. Those methods have shown their usefulness for a wide range of NLP and information retrieval tasks. In today's complex adaptive systems, automatic data integration and data exploration play an increasingly important role. Since word embeddings provide a deeper understanding of the domain of text they are trained on, they can be facilitated to gather useful domain information to integrate different information sources. Moreover, word embedding operations enable capabilities for semantic comparison. This has been utilized already by using word embedding models on text values in tabular data for table search, data discovery, and data integration tasks. However, word embedding models are rather inappropriate to model the context-specific meaning of short text values and infrequent named entities in database systems. Furthermore, trained on documents word embeddings do not fit perfectly for text values in tabular formats. This thesis thus focuses on two main aspects to support applications using word embeddings on tabular data:

On the one hand, we integrate additional functionality for text analysis and machine learning into database systems by utilizing the powerful abilities of word embedding models.

On the other hand, we improve the representation of text values in database systems. This can either be done by optimizing the embedding representations of text values in the database obtained by as word embedding models, or by training word embeddings directly on tabular data. To optimize existing embeddings, algorithms are developed to incorporate the context-specific meaning of the text values in database systems into their embeddings. To train embeddings on tabular data, a novel embedding model is introduced that learns from the alignment of text values in tables separate embeddings for text values occurring in the schema and in the instance data of a database.

A role-based architecture for distributed self-adaptive systems

Tim Kluge (Tim.Kluge1@tu-dresden.de)
Supervisor: Prof. Dr. Uwe Aßmann

Today's computing world features a growing number of connected distributed systems that require the cooperation of many physical devices. Examples include cyber-physical systems like autonomous cars and co-working robots, which are expected to appropriately adapt to any possible context they find themselves in (e.g. the presence of a nearby human).

However, the controlling software continues to be developed using established object-oriented modelling techniques like UML, which do not natively possess a notion of context and thus may introduce accidental complexity. With increasing complexity, the probability of the introduction of software errors rises, which can have fatal consequences in cyber-physical systems. To address this, we envision a model-driven architecture for self-adaptive distributed systems that explicitly models structured context using the compartment role object model (CROM)[1] developed in our research training group. Entities are modelled as message-passing parallel processes and can play roles in specific contexts, which dynamically alter their behaviour and relationships with other parts of the system. A possible concurrency model to use is a context-aware refined Hewitt Actor model. Since the planning of complex adaptations can be cumbersome in real-world scenarios, we envision an intuitive declaration of adaptations as graph rewriting rules on the context model. Rule-based graph rewriting on other model types has been proposed for self-adaptation by related work[2]. Therefore, our work approaches the following research questions:

- RQ1 How can role-based context-models be used to build distributed systems?
- RQ2 How can decentral adaptations of such a system be planned?

To show that the proposed architecture actually supports the development of self-adaptive cyber-physical systems, it is planned to qualitatively evaluate the system by conducting multiple case studies. Additionally, a quantitative evaluation will be provided. We plan to use the code complexity and adaptation performance as the main measures. As claimed, our approach is expected to reduce the accidental complexity of software by supporting the development of context-adaptive systems with cross-cutting concerns.

[1] Thomas Kühn, Stephan Böhme, Sebastian Götz, and Uwe Aßmann. 2015. A combined formal model for relational context-dependent roles. In Proceedings of the 2015 ACM SIGPLAN International Conference on Software Language Engineering (SLE 2015). Association for Computing Machinery, New York, NY, USA, 113–124.

[2] Basil Becker and Holger Giese. 2008. Modelling of correct self-adaptive systems: a graph transformation system based approach. In Proceedings of the 5th international conference on Soft computing as transdisciplinary science and technology (CSTST '08). Association for Computing Machinery, New York, NY, USA, 508–516.

Decision Making using Probabilistic Model Checking in Self-Adaptive Systems

Max Korn (Max.Korn@tu-dresden.de)
Supervisor: Prof. Dr. Christel Baier

The ever growing demand on our software systems creates a lot of difficult demands on software systems. Many of these systems need to properly function in a variety of scenarios, with often completely different outside parameters. This makes software systems that are able to properly function under different contexts a necessity.

Self-adaptive systems are systems with the ability to make context dependent decisions at run-time, depending on internal and external factors. This ability to handle themselves in a multitude of environments makes them a widely studied research topic.

One kind of system that is especially dependent on context information are the role based systems, where certain actors have different behavior depending on the role they play. The goal of this thesis is to utilize formal methods, in particular probabilistic model checking, to create a suitable run-time decider for role based models, allowing these role-based models to become self-adaptive.

For this I first started creating a efficient run-time decider for systems without roles. For this I create a framework for decision making, that takes the formal model of a system to adapt, and creates the decider.

To do this, we use several instances of the formal model, which represent various expected environments, and compute expected consequences of possible decisions using probabilistic model checking. These expectations are then saved in a database. The creation of this database takes place outside of the system run-time, though on of the future goals is to allow the filling of this database at run-time, at least partially.

The decider then uses this database, as well as a specification of its current goals, to decide on the decision with the best expectation regarding this goal.

We check the performance of the decider, by combining certain instances of the formal model with the decider, on using statistical probabilistic model checking to compute performance measures. The instances used can vary from the ones used to create the database, to gather the deciders reaction on unknown environments.

In future, this thesis aims to incorporate role based systems, knowledge creation at least partially in run-time, and the ability to change the objectives of the decider during run-time.

Towards Robust Decentralized Self-Adaptive Systems

Daniel Matusek (daniel.matusek@tu-dresden.de)
Supervisor: Prof. Dr. rer. nat. habil. Dr. h. c. Alexander Schill

Today's computer networks are largely distributed and therefore require steady maintenance. To tackle this problem, so-called context-aware self-adaptive systems could be used to change the behaviour depending on their environment without human interaction. However, most of the proposed systems use a central instance to control adaptation across multiple devices, which could lead to bottlenecks and reduce scalability.

Solving the problem of adapting systems to changing context and its environment would allow for nearly perpetual running systems. A decentralized solution for coordinating adaptation would increase robustness and allow for improving the scalability of those systems[1].

In the recent years, researchers proposed first approaches for the decentralized coordination of adaptations. By using adaptation managers on each device which are responsible for the administration of the nodes and developing a communication protocol to invoke adaptations decentrally, the need for a central instance was superseded. Nevertheless, those approaches need further investigation regarding the robustness of decentralized adaptation to make them more viable. Problems can still occur when it comes to node failures in big systems and a whole system gets partitioned into several subsystems. The resulting subsystems now perform adaptations by their own, which requires synchronisation of the resulting states afterwards. The same applies for systems of systems, which are partitioned on purpose. When they get aggregated to perform a common task or to dynamically higher performance, decentral adaptations need to get applied on subsystems which are originally in different states.

[1] Weyns, D., Bencomo, N., Calinescu, R., Camara, J., Ghezzi, C., Grassi, V., Grunske, L., Inverardi, P., Jezequel, J.-M., Malek, S., Mirandola, R., Mori, M., and Tamburrelli, G. (2017). Perpetual Assurances for Self-Adaptive Systems. In R. de Lemos, D. Garlan, C. Ghezzi, and H. Giese (Eds.), Software Engineering for Self-Adaptive Systems III. Assurances (Vol. 9640, pp. 31–63). Springer International Publishing. https://doi.org/10.1007/978-3-319-74183-3_2

Numerical Simulations of Topology-Driven Morphogenesis

Abhinav Singh (absingh@mpi-cbg.de)
Supervisor: Prof. Dr. Ivo F. Sbalzarini

Morphogenesis is one of the most intriguing phenomena in biology. Despite multiple theoretical and experimental approaches, it is still elusive how cells self-organize into complex shapes and how the process could be computationally modeled and simulated. The theory of incompressible viscous active polar gels provides a physically consistent model to explain patterning and shape emergence in morphogenesis with convincing simulation studies in 2D[1] and on effectively 1D surfaces[2]. However, no such study exists in general 3D geometries. In this project, we aim to develop a high-performance numerical simulation framework for simulating morphogenetic models in complex 3D geometries. Using these simulations, we aim to answer questions about active fluid flow, such as the existence of active turbulence in 3D. The eventual goal of the project is to enable a comprehensive computer simulation framework that copes with biological shape complexity. Building such a software system is challenging, as it needs to support various models and different numerical algorithms all while maintaining scalability on parallel computer architectures. To aid this process, we present a novel context-aware C++ expression system for scalable simulations using OpenFPM[3]. This generic expression system generates scalable code from mathematical equations models at compile time. We show that using this framework simplifies implementation and enables rapid testing of numerical codes without rewriting. It also cleanly separates the implementation of the model equations from that of the numerical algorithms. As a benchmark, we consider a parallel Discretization Corrected Particle Strength Exchange[4] solver for Poisson equations and couple it with a pressure-correction scheme[5] to enable Lagrangian simulations of active polar gels.

[1] R. Ramaswamy et al. "A hybrid particle-mesh method for incompressible active polar viscous gels," Journal of Computational Physics, vol. 291, pp. 334–361, Jun. 2015.

[2] A. Mietke et al. "Self-organized shape dynamics of active surfaces," PNAS, vol. 116, no. 1, pp. 29–34, Jan. 2019.

[3] P. Incardona et al. "OpenFPM: A scalable open framework for particle and particle-mesh codes on parallel computers" Computer Physics Communications, 241:155– 177, 2019.

[4] B. Schrader et al. "Discretization correction of general integral PSE Operators for particle methods" Journal of Computational Physics, vol. 229, no. 11, pp. 4159–4182, Jun. 2010.

[5] P. K. Papadopoulos "An auxiliary potential velocity method for incompressible viscous flow" Computers and Fluids, vol. 51, no. 1, pp. 60–67, Dec. 2011.

Role-based Context-aware Monitoring of Distributed Systems

Ilja Shmelkin (ilja.shmelkin@tu-dresden.de)
Supervisor: Prof. Dr. Alexander Schill

To create software systems that function as intended during run time is the main goal of every software developer. Nevertheless, unforeseen events and circumstances may lead to a software malfunction. In times where software systems are so big, that they cannot be surveilled solely by administrators, in turn, software is used to assist at that task. Such software is commonly referred to as monitoring software. There is an enormous amount of different monitoring solutions on the market, ranging from small monitoring solutions for local systems to big, industrial scale applications which monitor hundreds to thousands of servers. Thereby, the state-of-the-art approach to extract information out of a monitored system is to either query an API in cases where it allows direct access to the systems metrics or, which is the case predominantly, install probes on the monitored system which allow interaction with the monitoring system. Monitoring systems typically are very rigid regarding their supported use cases. The areas of deployment often only enclose one specific domain (e.g., monitoring different operating systems and their underlying hardware, extracting sensor data of specific hardware or creating time-series data by querying metrics on a regular basis). However, as software diversity and the number of possible deployment areas grow rapidly, the need for a general, flexible solution exists more then ever.
This thesis aims to approach this issue by, firstly, by proposing a model for a monitoring software that allows to monitor clients in arbitrary domains, as long as they support basis access to their properties, and secondly, by presenting a working prototype that implements that model. For this, the role-based approach is chosen as it provides high grade of dynamism allows for context-dependent dispatch of role-bound methods. The prototype is then evaluated in multiple scenarios spanning different problem domains, e.g. infrastructure auto-scaling, surveillance of unmanned aerial vehicles (UAV's). For this, it will be compared to state-of-the-art solutions of the according domain if such solution exists.

Adaptive Query Processing on Vectorized Hardware

Johannes Pietrzyk (johannes.pietrzyk@tu-dresden.de)
Supervisor: Prof. Dr.-Ing. Wolfgang Lehner

High performance is considered one of the main goals of modern query processing, particularly for analytical workloads. Therefore modern query engines need to facilitate the underlying hardware with all of its features to the maximum extent. One of these features is vectorization, based on the Single Instruction Multiple Data (SIMD) parallel paradigm. Over the last years, vectorized processing was established as a state-of-the-art approach for increasing single-query performance. However, since modern systems have to deal with workloads consisting of many concurrent analytical queries, single-query optimization may miss crucial optimization potential. Data accesses as well as computations, for instance, maybe accessed redundantly among different queries. Various techniques have been proposed to minimize redundant memory access within the last decades, ranging from cooperative scans to materialized views for common sub-queries.

Continuing this line of research, we investigate the opportunities offered by vectorized hardware to optimize analytical query workloads. We envision using vector registers as a resource for sharing data and SIMD instructions to share similar computations across concurrently running queries.

To prove the applicability of our idea, we developed a prototype where a vector register is used as a shared resource for multiple queries. Every single entry in a vector register, hereafter referred to as lane, can be assigned to a single query. Depending on the number of queries with common operators within a workload, the number of occupied lanes per query can vary. Thus, a reasonable degree of data parallelism can be ensured while redundant memory transfers can be minimized. We systematically evaluated this approach by investigating the impact of the used vector register size, the processed data value size, and the number of shareable operations and accessed data. Our approach improved the overall workload execution time if operations and data can be shared. Moreover, even in the worst-case scenario, where no data could be shared, the introduced overhead for maintaining the shared vector was negligible. We will develop a comprehensive multi-query processing model for complex analytical queries based on vector registers as a work-sharing resource for future work. Therefore, as a next step, we will develop a unified proxy model, enabling an on-the-fly adaptation of vector sharing to any given workload depending on the number of current concurrent or batched queries and the number of accessed columns. One of the most exciting questions regarding the input and output proxies is how intermediate data consumed from different operators should be organized.

Role-based integration of structural and behavioral modeling

Tarek Skouti (tarek.skouti@tu-dresden.de)
Supervisor: Prof. Dr. Susanne Strahringer; Prof. Dr. Frank J. Furrer

System development is complex and requires a shared understanding of concepts between stakeholders of different domains. Models reduce the complexity and support the development, but different domains need different models. Behavior models like process models describe the dynamic aspects of a system. Structural models like the UML class diagram describe the static aspects of a system. Both models are needed, but the gap between them leads to errors in developing a system. This work focuses on the role-based integration of structural and behavioral modeling.

Roles are the fundamental concept of the solution. While the Business-Role Object Specification (BROS)[1] used roles to introduce behavioral awareness to structural modeling, we use roles to introduce structural awareness to behavioral modeling. Since BPMN is the defacto-standard of process modeling, we developed a role-based BPMN extension called RBPMN. RBPMN uses roles to increase the adaptability and expressiveness of process models. We integrate both modeling approaches with roles as a shared concept between structural (BROS) and behavioral (RBPMN) modeling.

Furthermore, roles increase the ontological expressiveness of models. Modeling a wider variety of performers (physical and virtual) eventually increases the adaptability of the process at runtime. After developing the RBPMN language, evaluation of it is the next step. Evaluating a process modeling language can be based on workflow pattern coverage and ontological completeness. Workflow patterns support researchers in evaluating workflow systems on features a system can theoretically support. Evaluating ontological completeness can then show the expressiveness of RBPMN.

The first results of the evaluation show that RBPMN can express more workflow patterns than by BPMN and that the ontological completeness increases. This led to the new question that the combined ontological overlap might be too high. The theory of combined ontological completeness and overlap states a tipping point where the combined ontological completeness decreases if the ontological overlap of models is too high. The overlap between RBPMN and BROS is naturally higher than between BPMN and UML. Thus as a next step, we intend to apply and test the theory for the combined modeling approach of BROS and RBPMN.

For future work, one direction is the migration to role-oriented models from object-oriented systems. A second direction lies in organizational mining to learn the roles of physical and virtual performers participating in a system.

[1] Schön, H., Strahringer, S., Furrer, F.J., Kühn, T., "Business Role-Object Specification: A Language for Behavior-aware Structural Modeling of Business Objects," In: Ludwig, T. and Pipek, V. (eds.) Wirtschaftsinformatik 2019 Proceedings. pp. 244–258 (2019).

Role-Based Embedded Domain-Specific Language for Collaborative Multi-Agent Systems through Blockchain Technology

Orçun Oruç (orcun.oruc@tu-dresden.de)
Supervisor: Prof. Dr. Uwe Aßmann, Prof. Dr. Susanne Strahringer

Multi-agent systems have evolved with their complexities over the past few decades. To create multi-agent systems, developers should understand the design, analysis, and implementation together. Agent-oriented software engineering applies best practices through software agents with abstraction levels in multi-agent systems. However, abstraction levels take a considerable amount of time due to the design complexity and adversity of the analysis phase before implementing them. Furthermore, trust and security of multi-agent systems have never been detailed in the design and analysis phase event through the implementation of trust and security on the tamper-proof data are necessary for developers. Even though object-oriented programming is the right way to do implement complex software agents, there are some problems regarding it. one of the major problems is object-oriented programming approach still has a complex process-interaction and a burden of event-goal combination to represent actions by multi-agents.

In this research, we propose a domain-specific language called GASMASk [1], which is an embedded domain-specific language. We would like to generate codes from templates for agents and custom annotations for smart contract applications to reduce the boilerplate codebase for developers. After developing the language, the evaluation phase will be proceeded by quantitative and qualitative methods. One of the results from the research, a Turing-complete smart contract language that makes a value transaction between agents can improve the usability of agent-oriented applications. Another result is the implementation of compartments and roles needs advanced smart contract language such as Solidity; however, smart contract applications still far away from the realization of customized annotations at the compiler level.

In future work, a preview of the domain specific-language will be developed for a role-based multi-agent system. This tool includes an annotation processor and template-based code generator for agent behaviors. By selecting a general-purpose language, the system will be evaluated with qualitative and quantitative features. Smart contracts will be generated through annotation processing with customized annotations and agents will be generated with a template-based code generator for a specific practical agent framework.

[1] https://www.researchgate.net/publication/350877450_Role-Based_Embedded_Domain-Specific_Language_for_Collaborative_Multi-Agent_Systems_through_Blockchain_Technology

Compilation and Interpretation Techniques for Role-based Programming Languages

Lars Schütze (lars.schuetze@tu-dresden.de)
Supervisor: Prof. Dr. Jeronimo Castrillon-Mazo

In object-oriented programming, the run-time type of objects defines the actual target of a method use, i.e., the dispatch of method calls on these objects. That is, the method call may not be resolved at compile-time as the static type may differ from the run-time type which is generally known as polymorphism.

Role-based programming introduces another dimension that influences the run-time type of an object. The role concept allows to enhance an object's interface which is determined by the roles currently played by that object. Roles being played by an object depend on current active contexts. Thus, available methods depend on the dynamic type of an object, as well as the role types an object is playing.

To realize these features, existing solutions incur in large overhead and suffer from inferior run-time performance[1]. That is because most role-based programming languages are translated into another programming language, while some provide a framework solution.

In recent years many role-based programming languages have been developed. Since there is no common understanding of what roles are, existing role-based programming languages provide different sets of role features[2]. The more features are supported by a role-based programming language the more complexity is added to the runtime.

To succeed in adopting role-based software infrastructures the performance must be on par with current mainstream programming languages. In this dissertation we will study the performance bottlenecks inherent to programming with roles and devise solutions, from the compiler and the language perspective, to circumvent them. If a program does not use specific role features, the run-time performance should not suffer. Profiling the usage of role features during run-time, dynamic compilation could provide performance dependent on the specific use of an application[3].

[1] L. Schütze and J. Castrillon, "Analyzing State-of-the-Art Role-based Programming Languages," in Proceedings of the International Conference on the Art, Science, and Engineering of Programming, 2017, pp. 1–6

[2] T. Kühn, M. Leuthäuser, S. Götz, C. Seidl, and U. Aßmann, "A metamodel family for role-based modeling and programming languages," in Software language engineering, vol. 8706, Springer, 2014, pp. 141-160.

[3] L. Schütze and J. Castrillon, "Efficient Dispatch of Multi-object Polymorphic Call Sites in Contextual Role-Oriented Programming Languages," in 17th International Conference on Managed Programming Languages and Runtimes, 2020, pp. 52–62

Context-Sensitive Description Logics in a Dynamic Setting

Satyadharma Tirtarasa (satyadharma.tirtarasa@tu-dresden.de)
Supervisor: Prof. Dr.-Ing. Franz Baader, Prof. Dr. Ivo F. Sbalzarini

Description Logics (DLs) are a family of knowledge representation formalism that cover a large number of application domains by choosing an instance with appropriate expressivity and complexity. However, the picture changes when the capability to represent meta-concepts, such as contextual knowledge is needed. It is not possible, or at least in an intuitive way, to describe context-sensitive information using classical DLs. This leads to the investigation of DL extensions with a context facet. One result of the research in this direction are contextualized DLs (ConDLs)[1], a family of two-dimensional DLs tailored for reasoning on a context-sensitive domain. By imposing the restriction that a signature in the object level can not access the meta level, the decidability is maintained even with the presence of rigid concept and role.

Unfortunately, the problems under consideration so far are static in some senses. We take a look at some common reasoning problems in a dynamic domain with considering context-sensitive DLs and investigate how they are intertwined. The expressiveness and computational complexity results will be very important, especially considering both formalisms are prone to undecidability.While the interaction between ConDLs and dynamic domains is relatively unexplored, the interaction between classical DLs and dynamic domain, such as action formalisms, has been studied[2]. Our study yields a ConDL-based action languages[3] which the projection problem with a basic ConDL is well-behaved, i.e., has the same complexity with the consistency problem of the underlying ConDL.

Furthermore, we inspect how to use such formalized framework as an underlying knowledge base for the application domain, and naturally in the means of context-sensitive languages and systems. Specifically, we will consider role-based modeling languages and formalize the problems that arise there to our framework. It has been shown that the aforementioned ConDLs are capable to represent role-based systems[4]. Some examples of the problems that can be tackled in a dynamic setting are building run-time system monitor or verifying temporal properties over the system. Intermediate result gives us a formalism to represent temporal properties over context-sensitive knowledge.[5]

[1] S. Böhme and M. Lippmann, "Decidable Description Logics of Context with Rigid Roles," Proceedings of the FroCoS 2015, p. 17–32, 2015.

[2] F. Baader, C. Lutz, M. Milicic, U. Sattler, and F. Wolter, "Integrating Description Logics and Action Formalisms: First Results," Proceedings of the AAAI 2005, p. 572–577, 2005.

[3] S. Tirtarasa and B. Zarrieß, "Projection in a Description Logic of Context with Actions," Proceedings of the GCAI 2019, p. 81–93, 2019.

[4] S. Böhme and T. Kühn, "Reasoning on Context-Dependent Domain Models," Proceedings of the JIST 2017, p. 69–85, 2017.

[5] S. Tirtarasa, "Temporal Properties over Contextualized Description Logics," Proceedings of the DL 2020, 2020.

Adaptive Routing in DTN for Public Transport Systems

Jose Irigon (jose.irigon@tu-dresden.de)
Supervisor: Prof. Dr. Alexander Schill

Disruption-Tolerant Networking (DTN) enables communication in environments that lack contemporaneous end-to-end connectivity. Devices decide autonomously based on information available locally or exchanged with peers over time, and therefore frequently outdated, making routing decisions a challenging topic that has been actively researched.

This work introduces an adaptive routing module for DTN, which sets the routing algorithm based on the context acquired from properties exchanged during contact opportunities. As a first step, we propose decoupling property exchange from the routing strategy to enable multi-routing support. Based on a classification of properties exchanged in the DTN literature, we chose a representative set to support multiple routing algorithms for the concrete use-case of DTN over Public Transport Systems (PTN). We evaluate the overhead caused by information exchange, the performance of different routing algorithms in multiple PTN scenarios, and the effect of different types of adaptations on routing performance.

Our results indicate that a parameterization suitable for the typical mobility may perform poorly or even preclude communication when mobility changes unexpectedly. Besides, adapting the routing algorithm led, in some contexts, to an improvement of the desired metric. A role-based design and prototyping demonstrated the feasibility of modeling the collaboration between multiple aspects that may be part of the routing decision.

Supporting Lecturers in Properly Using Digital Learning Environments

Tommy Kubica (Tommy.Kubica@tu-dresden.de)
Supervisor: Prof. Dr. rer. nat. habil. Dr. h. c. Alexander Schill

Digital learning environments, such as Audience Response or Backchannel Systems, provide a promising opportunity to address issues occurring in traditional or live-stream lectures, e.g., the lack of interaction, by allowing students to participate actively using their mobile devices. This can promote the students' attention, increase the communication between the lecturer and the students, and foster active thinking during class. In order to choose an appropriate digital learning environment, numerous surveys list and classify these systems according to different criteria.

However, the introduction of such systems leads to its own challenges: The lecturers have to adjust their preferred teaching strategy to the chosen system, as this usually relies on a single supported didactic scenario and therefore has a limited, fixed functional scope. Moreover, the lecturers have to select and use the system's functionality and interpret the received data by themselves – support or recommendations of a suitable functional scope are rarely provided[1]. Another issue becomes obvious by investigating different didactic scenarios: While collaboration with subsequent group discussions is an integral part of various scenarios, it is rarely or not even supported by these systems[2].

Using the means of adaptation, we target to overcome these limitations. The following research question arises: How can different levels of adaptation support the lecturer in properly using digital learning environments? In order to answer this question, three sub-theses have to be validated: (a) Modeling adaptation allows lecturers to create customized teaching scenarios that support their individual teaching strategies; (b) Runtime adaptation allows to adjust the teaching scenarios on-the-fly in order to respond to real-time results; (c) The concept of roles provides a promising extension to integrate the means of adaptation in digital learning environments.

Therefore, the concept of an adaptable collaborative learning environment that is supported by role concepts is proposed and implemented in an approach called scenario-tailored Audience Response System (stARS)[3]. Using stARS does not only enable lecturers to model and execute their individual teaching scenarios as customized workflows, but also to adapt those at runtime. This allows to evaluate the validity of the research theses and thus, to answer the overall research question.

[1] Kubica, T., Hara, T., Braun, I., Kapp, F., Schill, A., "Guided selection of IT-based education tools," 47th Frontiers in Education Conference (FIE), 2017.

[2] Shmelkin, I., "Untersuchung der Adaptierbarkeit webbasierter Audience Response Systeme," Main Seminar, Technische Universität Dresden, 2018.

[3] Kubica, T., Shmelkin, I., Peine, R., Roszko, L., Schill, A., "stARS: Proposing an Adaptable Collaborative Learning Environment to Support Communication in the Classroom.," 12th Int. Conf. on Computer Supported Education (CSEDU), 2020.

Adaptable programming models and compilers for 5G and beyond

Julian Robledo Mejia (julian.robledo@tu-dresden.de)
Supervisor: Prof. Dr. Jeronimo Castrillon; Prof. Dr. Uwe Aßmann

With the increasing number of wireless devices, as well as the wide range of applications they can offer, adaptivity in baseband processing systems becomes essential to accomplish optimized performance. The high-reliability required for mission-critical tasks such as a remote medical surgery, or the low-latency needed by a moving car downloading 4K streaming video, contrast with the demands of low-complexity applications computed by a massive volume of Internet of Things (IoT) devices. In 5G, the receiver base station has to deal not only with a huge number of users expecting connectivity, but also with a high workload heterogeneity. Adaptable systems appear as a promising solution to cope with the aforementioned promises.

The physical layer of a 5G base station, consists on a series of multiple computation kernels that process the modulated data from multiple users in order to decoding it and transmit it to the upper layers. There is a trend from the telecommunications industry to implement software-based solutions which can be easily updated with the changing standards, specially in the context of Cloud Radio Access Networks (Cloud RANs). Currently, heterogeneous multi-core Digital Signal Processor (DSP) platforms are widely used to implement such systems. Due to the dynamic nature of 5G requests to be processed in a base station, a baseband processing system would benefit from a software solution that allows the computation kernels to adapt their behavior to the specific context.

The overall goal of this dissertation is to implement adaptive methodologies to optimize 5G baseband processing in heterogeneous platforms. For this purpose, we want first to understand the variation of workloads in uplink communication and capture it into an execution model. We can then define an adaptive Model of Computation (MoC) to fully describe the uplink physical layer of a base station. Moreover, given the high number of concurrent requests to be processed in a base station, a clever scheduling strategy is a key point to meet the strict real-time deadlines imposed by the protocol while saving energy. Therefore, we also want to investigate on adaptive model-based optimizations to enhance multi-core scheduling performance.

Model-based Encodings

Juliana Hildebrandt (juliana.hildebrandt@tu-dresden.de)
Supervisor: Prof. Dr.-Ing. Wolfgang Lehner

Database systems are characterized by two important aspects: (1) consistent and permanent data storage and (2) efficient and isolated data processing. To ensure these two aspects, database systems have to meet three challenges nowadays, which are briefly outlined below.

The first challenge is the efficient use of increasing main memory capacity. In order to achieve this, numerous database systems utilize a memory-centric architecture which occupies an important role for data compression. Because of the reduction in data size, the transfer times from CPU and main memory are reduced as well. This leads to decreased processing time. For that, beside basis data also intermediate results are compressed. Different algorithms for compression and decompression are suitable for various data characteristics, query types and hardware capabilities. But all algorithms share the opportunity to process the compressed data.

The second challenge is the protection of data misuse. For this, encryption and anonymization can be used. According to data characteristics and query types different property preserving encryptions like order preserving or homomorphic encryptions are suitable. Property preserving encryptions enable an efficient query processing on encrypted data.

A third challenge is the fault-tolerance on unreliable hardware. Because of the reduction in circuit line widths more transistors fit on a single chip, but the error rate of these chips increases. This is marked by transient bit flips, that occur in main-memory and CPU as well as during the data transmission. With the focus on data consistency, different kinds of error-detecting and error-correcting codes are suitable to handle this challenge.

In summary one may say that it is necessary to integrate an individual data representation for the storage and processing of data regarding to data characteristics, query types, hardware capabilities and weight of processing efficiency, confidentiality and fault-tolerance. With a view to the GRK RoSI, the framework conditions represent a context, so that a role modeling is an adequate approach. This very aspect is intended to be examined. Moreover it shall be analyzed how to use role modeling and a corresponding model-driven approach to integrate the different encoding algorithms. The manual integration of such algorithms is possible, but time-consuming and error-prone. Furthermore a configurability of data representations and adaptivity according to the three challenges is limitedly feasible.

Managing Parallelization and Heterogeneity with Declarative Invasive Software Composition

Johannes Mey (johannes.mey@tu-dresden.de)
Supervisor: Prof. Dr. rer. nat. habil. Uwe Aßmann

Complex code transformations and compositions can be useful in a variety of application domains. One of these is the programming of wildly heterogeneous systems, which introduces new challenges to maintain programmability, because developers do not only have to produce parallel code, but also code for very different and potentially unanticipated target platforms.

Current parallelization and distribution strategies are able to handle those concerns for specific target platforms efficiently, but are both not easily exchangeable and require detailed knowledge of the application code, its performance, and the target platform. A common strategy to create a parallel program is to gradually extend a sequential program with the commands and pragmas, a process called progressive parallelization. This process, however, causes an entanglement of concerns that makes testing, benchmarking, and evolving both the core code and the additions difficult.

To overcome these problems, we introduce a novel code composition and transformation approach called Orchestration Style Sheets (OSS). OSS borrow the well-known concept of style sheets and transfer it to annotation and enrichment of source code, in this case for parallelization. Parallelization (and other secondary) concerns are externalized into styles that can be defined separately as well as exchanged and reused for multiple programs. This approach is similar to aspect-oriented programming (AOP), but goes further in some respects: e.g., the pragmas used in the Open* languages are heavily parametrized with elements of the core code like lists of private or shared variables. In principle, many of these parameters can be deduced automatically, which requires static analysis of the core program. OSS offer user-defined attributes that can perform this analysis, thus eliminating the error-prone process of deriving the parameters manually and allowing a reuse of styles.

While parallelization is the current main focus, the approach is both language-independent and applicable to different use cases. Currently, there are several systems based on this framework: OSS for Fortran, a language (still) frequently used in parallel programming has been used to evaluate the concept in realistic use cases like simulations used in mechanical engineering. A second, large system is Java-based and uses a complete attribute grammar-based compiler front-end. Additionally, the framework can easily be applied to smaller languages, i.e. domain-specific languages.

Domain Specific Language and Compiler Optimizations for Computational Biology Applications

Nesrine Khouzami (nesrine.khouzami@tu-dresden.de)
Supervisor: Prof. Dr. Jeronimo Castrillon-Mazo

Systems biology are trying to understand fundamental questions regarding the orchestration of the molecules in the human body. To answer these questions, they have to simulate complex models, solve large equation systems, and process immense amounts of data. Particles methods have been firmly established as a computational framework to process different kind of simulations. Such simulations rely on huge power of computing to give results on reasonable time thus the use of high performance computing (HPC) environments. As much as these technologies are evolving continuously, it is getting more difficult for scientists to use efficiently such architectures as they are not expert on low level and HPC paradigms.

Domain specific languages (DSLs) are a promising solution to ease writing high performance codes. In our project, we aim to help close the productivity and performance gap for HPC applications in systems biology using particle-mesh methods. To achieve this goal, we will develop a new DSL OpenPME (Open Particle Mesh Environment) for particle-mesh methods on top of OpenFPM [1] (Open Framework for Particles and Meshes) library and runtime system.

First, we formalize particle-mesh abstractions to reason about transformations and optimizations. Then we design and develop the particle-mesh DSL based on the conceived formal model. The DSL should hide the complexity of template meta-programming present in OpenFPM to ensure its usability and lower the entry barrier. All should be accompanied with a modern development environment where the language includes high-level information about the modeled phenomenon that helps developing new compiler optimizations for a given set of target architectures. We are interested particularly in the possibilities for compiler-controlled memory management and coherency. By defining the particles structure, how they interact and update positions, we can provide a clear sense of data locality that will be exploited by the compiler to insert data-prefetching calls to improve performance. Eventually, we will consider extending OpenFPM library and the runtime to get aware of the information available at the DSL level and extracted by the compiler. This will allow to implement auto-tuning approaches, especially for simulations with high dynamics, and improve dynamic load balancing. We also consider implementing the numerical solvers in OpenFPM-numerics required by the driver applications. We will expose the abstract operators corresponding to these solvers in the DSL and the development environment.

[1] Incardona, P., Leo, A., Zaluzhnyi, Y., Ramaswamy, R., Sbalzarini, I.F.: OpenFPM: A scalable open framework for particle and particle-mesh codes on parallel computers. arXiv:1804.07598 [physics] (April 2018)

Achieving situative privacy protection in a fog environment using situative privacy modeling of a DSPL of role-based pseudonym systems

Frank Rohde (frank.rohde@tu-dresden.de)
Supervisor: Prof. Dr. Uwe Aßmann

Problem The fog computing paradigm was proposed as an extension to the cloud computing paradigm enabling low-latency IoT applications that involve widely distributed nodes. Privacy protection is a challenge for those applications. It can be tackled by the use of pseudonyms [1]. While fog computing mitigates some of the disadvantages of cloud computing it creates a new challenge related to the heterogeneity of the nodes providing the fog service and the presence of highly dynamic client-server relations: situative privacy protection (scientifically described by Mann et al. .[2]), i.e., the challenge to deal with privacy threats that dynamically emerge from the current situation of the client within the fog-based computing environment (e.g., available computing power and hardware security measures of the service-providing fog nodes). Pseudonym systems that implement pseudonym specific behaviour show another problem with respect to their code quality: the respective code is scattered throughout the code base and is tangled with unrelated code as pseudonym specific behaviour is a cross-cutting concern.
Solution This thesis investigates the applicability of feature-oriented software development as an approach to design and implement a pseudonym system that is able to deliver situative privacy protection for fog-based applications. We propose a pseudonym system architecture that allows for the functional variability which is required for situative privacy protection and describe this functional variability by means of a feature model. Decisions with respect to runtime feature adaptation are taken based on a model of the current client situation. We analyze and discuss the applicability of petri nets for client situation modelling in fog computing. Moreover, we analyze and discuss the suitability of the Role-concept to mitigate scattering and tangling in implementations of pseudonym-specific behaviour.
Contributions The contributions of this thesis are threefold: (1) The suitability of feature oriented software development, more specifically software product line engineering, to achieve situative privacy protection for fog computing is analyzed and discussed by means of an exemplary privacy technology: pseudonym systems (2) The applicability of petri nets as a context model for situative privacy protection in fog computing is discussed; (3) The suitability of the role concept to tackle scattering and tangling of code that implements pseudonym-specific behaviour is discussed.

[1] Schaub et al. “Pseudonym schemes in vehicular networks: A survey”. IEEE communications surveys and tutorials. 17.1 (2014): 228-255.

[2] Mann et al. “Situativer Datenschutz im Fog-Computing”. Informatik Spektrum, 42.4 (2019): 236-243

GRK 2050: Privacy and Trust for Mobile Users

Prof. Dr. Max Mühlhäuser
Email: muehlhaeuser@privacy-trust.tu-darmstadt.de
Technische Universität Darmstadt
Internet: https://www.privacy-trust.tu-darmstadt.de

The RTG 2050 Privacy and Trust for Mobile Users is a highly interdisciplinary collaboration between Computer Science and the fields of Law, Sociology, Information Systems (in Economics), and Usability (in Psychology). We aim at improving the position of mobile users – think of smartphone users – vis-a-vis digital service networks, social networks in form of digital collectives, and sensor-augmented environments, i.e., "IoT" environments (all summarized in the follwing as 'networks').

In the mobile users' experience, these networks and the players therein are becoming increasingly opaque while the users themselves are becoming increasingly transparent. The term 'players' here refers to all kinds of digital 'counterparts' of mobile users and to the responsible people and organizations, such as service providers, social network providers and peers, smart environment operators, network operators, hard- and software vendors. In a multi-disciplinary effort, our RTG counters these 'paired trends' – transparent users and opaque networks – with the 'paired goals' privacy & trust: privacy is considered as the main instrument for limiting user transparency, while assessing the expected trustworthiness of players in the network is considered as the main instrument for countering the opaqueness of the network players.

Privacy and trust are not yet commonly perceived as paired, i.e., tightly interwoven necessities for making the Internet (and networks in general) a liveable digital habitat. This is in part due to a somewhat misleading use of the term trust in cybersecurity research: fields like trusted computing, trustworthy ICT, and trust management refer to issues of reliability-plus-security, tamper-free hard- and software, and digital identities, respectively – all quite remote from the primary meaning of the term trust. Our RTG fosters research into trust in its primary meaning: justified readiness to engage in a risky engagement, with risks including privacy violations and other negative experience with service provision. An important area of our trust research is computational trust, where trust is formalized as the probability of a trustee acting as expected; expectations in turn are justified from two categories of evidence: experience (own prior experience, reputation) and indicators (certified audit results, attestations, etc.). Since trust assessment relies on evidence, i.e., information about the trustee, there is a potential conflict: trust aims at revealing what privacy aims at concealing: information about an entity. This is relevant if trusters and trustees do not form two distinct sets (cf. social network participants and agents in peer-to-peer economies). In the RTG, privacy related research is (at least) as prominent as research on trust. Due to their interweaving, we are addressing both aspects jointly in our research areas, structuring our RTG according to the above-mentioned network categories: (social) collectives, service networks, and sensor networks in

form of the 'IoT' – with an additional focus area emphasizing novel mobile user support.

Outside the digial world, both trust and privacy were concerns since millenia. This mandates our interdisciplinary approach that involves Sociology, Psychology, Laws and Economics. Our experts from these fields contribute long standing experience in linking their disciplines to issues from the digital world, which greatly facilitates their cooperation with our computer scientists.

Trust in Artificial Intelligence

Mariska Fecho (fecho@is.tu-darmstadt.de)
Supervisor: Prof. Dr. Peter Buxmann

Recent advances in digitization and the availability of high volume data have led to higher interest and usage of artificial intelligence (AI). Thus, AI and machine learning methods are increasingly influencing many areas of our lives (e.g. medical diagnosis, autonomous driving, digital voice assistants). A key concern, which is often discussed in the context of intelligent systems, is their black-box behaviour. Due to their complexity, the results and functions of the algorithms are often not transparent for the user. Some users are even completely unaware of the intelligence of a system. This is particularly critical with regard to automated decision processes, where decisions are delegated completely or partially to a machine or system.

Trust is a multi-faced concept that has been examined in various disciplines. It helps to overcome perceived uncertainties and risks, especially in unfamiliar situations.[1] By investigating trust between human and technology, trust has been shown as a decisive factor for the usage and adoption of technologies.[2] As a result, the literature in the area of technology adoption has identified several factors that influence trust in technology.

This project aims to investigate relevant factors for the adoption and usage of AI-based technologies in specific contexts. Furthermore, concepts and dimensions of trust for the initial adoption of AI-based technologies shall be investigated.

[1] McKnight, D. H., Choudhury, V., and Kacmar, C. (2002). "Developing and validating trust measures for e-commerce: An integrative typology". Information systems research, 13(3), 334–359.

[2] Gefen, D., Karahanna, E., and Straub, D. W. (2003). "Trust and TAM in online shopping: An integrated model". MIS quarterly, 51–90.

Psychological Effects of Smartphone Usage

Julius Frankenbach (julius.frankenbach@tu-darmstadt.de)
Supervisor: Prof. Dr. Stephanie Pieschl

Are smartphones trustworthy daily companions that make life easier, or are they ultimate distraction machines that prevent us from achieving meaningful goals in life? The present dissertation project concerns the psychological effects of extensive smartphone usage, especially within the field of education.

This research question is motivated by the following observations. (a) Smartphones are now ubiquitous worldwide. At this scale, even small detrimental effects are potentially of great concern. (b) Smartphones and their software are designed by organisations that may have other goals for the users than users for themselves. In other words, smartphones may not be entirely "on the users' side".[1] (c) Psychological theory suggests clear pathways by which smartphones may disrupt learning or other mental activities that require deep concentration. (d) Finally, many people wish to use their smartphone less.

The dissertation project will be guided by the following principles. First, causality will be taken seriously. The dissertation project will therefore rely on interventions and experiments with random assignment, rather than correlational analyses that can never establish causal links. Second, the project will appreciate the complexity and diversity of behavior in the digital space. Rather than collapsing digital life into a singular (self-reported) measure of screen time, the project will examine the what, why, and when of smartphone usage. Third, the project will focus on psychological processes (e.g., how do adaptive versus maladaptive patterns of daily smartphone usage develop over time?), rather than surface level effects (e.g., what is the correlation between screen time and academic achievement?).

The first step toward these goals has been the development of a new smartphone application for the Android operating system. The objective of the application is to capture smartphone usage behavior as well as concurrent information about the users' mental state. To this end, three features have been implemented. First, the app logs usage directly. Second, the users' mental state can be probed in real time through notifications. Third, the app compiles usage sessions and presents them to the user. The users can then label these usage sessions and thereby report their mental state during usage retrospectively. This new app will be used in an intervention study with university students during the exam phase of the current summer term (2021). The study will evaluate, whether a 10-step program to reduce problematic smartphone usage[2] during exam preparation can improve exam outcomes.

[1] James Williams, "Stand out of our Light: Freedom and Resistance in the Attention Economy", Cambridge University Press, 2018

[2] Jay Olson, Dasha Sandra, Denis Chmoulevitch, Amir Raz, and Samuel P. L. Veissière, "A ten-step behavioural intervention to reduce screen time and problematic smartphone use", preprint, 2021

Information (In-)security of Human-Centric Sensor Data

Matthias Gazzari (mgazzari@seemoo.tu-darmstadt.de)
Supervisor: Prof. Dr. Matthias Hollick

In our world of ever increasing complex computing systems it becomes increasingly difficult to stay in control of the information flow between devices. An increasing number of more accurate sensors creates a lot of opportunities but also possibilities to violate the privacy and identity of users. In my current work I am focusing on analysing data from mobile sensors and input devices in respect to private information retrieval and user identification.

In the main part of my work, I am focusing on the analysis of side-channel attacks for reconstructing user inputs on input devices like keyboards or touchscreens. For doing so, I investigate device-targeted attacks, as well as user-targeted attacks by observing sensor values from devices measuring human actions. Similarly, I focus on (un-)intended user identification based on sensor data, looking into implications relevant to the privacy of the user as part of revealing or impersonating their identity.

As part of my first work, I am leveraging EMG and IMU data from the forearms to reconstruct whether and what has been typed on a keyboard. To do so, we have collected a data corpus containing about 318000 keystrokes from 38 participants typing predefined texts and passwords. Using end-to-end machine learning we could show that we are able to detect keystrokes, as well as reduce the search space for passwords.

Similarly, we did a preliminary study on a keyboard-targeted temporal side-channel by observing the wireless network traffic. For a qualified set of key-pairs and random passwords, we could show that is is possible to reduce the search space of a brute force attack.

As part of the second part of my work, we are trying to implement a photoplethysmography (PPG) based inter-sensor impersonation attack on an electocardiography (ECG) based authentication system. Using conditional generative adversarial networks, we transform PPG samples into impostor ECG samples in order to fool an ECG identification system. For evaluating the attack on an authentication system we are currently in the process of extending our data corpus of synchronized ECG/PPG data.

Leveraging the same type of sensor data, we are revisiting preliminary work for a user-targeted side-channel attack to reconstruct handwriting. As with all of the projects introduced above, data studies with multiple participants are planned in order to study whether such attacks could be applicable between subjects or only for a single target.

In the future, we will be looking into incorporating these findings from the studies above, in order to develop techniques for circumventing or reducing the impact of such attacks.

Building and Using Social Capital in Digital Collectives

Hendrik Jöntgen (joentgen@wiwi.uni-frankfurt.de)
Supervisor: Prof. Dr. Oliver Hinz

Digital Collectives are temporary unions of social media users who cooperate and share their resources and knowledge with each other in order to achieve a shared goal. Due to the ever-increasing usage of social media, these Digital Collectives are becoming progressively common and are exerting ever more influence.

These collectives can be very short-lived or sustain themselves over a longer period of time. Furthermore, they can be uncoordinated or have a single or multiple leaders. And finally, these collectives can result in positive or negative outcomes.

For example, although companies can use social media platforms as an effective channel to promote their products and services, they also become vulnerable to online firestorms where their community turns against them and consequently suffer damages to brand value or boycotts. On the other side, Open Source Software Communities allow the independent creation of software. The fundamental question here is how users are being motivated to participate in these projects. Regarding online firestorms, previous literature is mainly focused on giving overviews about the phenomenon and advice on how companies should react to them while previous literature on Open Source Software Communities has already addressed the motivation to participate in those communities but neglected the motivation to join a specific Open Source project. The goal of this research is therefore the examination of motivations to join a specific Digital Collective and the role of Social Capital in these decisions.

In order to do so, this research is using crawled data from Twitter and GitHub users as well as from their social networks. In addition to statistical tests on real-word data, surveys to further test the effects of Social Capital on the motivation to join a Digital Collective are being conducted.

The results of the analysis of different kinds of Digital collectives will contribute to a better understanding of them and their formation.

Research on the recursive construction of Social Networks as models of order

Florian Müller (florian.mueller@uni-kassel.de)
Supervisor: Prof. Dr. Jörn Lalma

More and more of social life takes place in digital spaces, especially in so-called Online Social Networks (OSN), which act as a communication platform, intermediary and business model at the same time. Additionally they are often interconnected with numerous other online services and applications. Therefore, making the platforms, through which social networks are formed, decidedly central and influential actors in societies. Nonetheless, the practices through which large platforms generate, process, analyze, and use data are highly controversial and problematic, especially with regard to the trustworthiness of these processes and actors. A few large Internet corporations such as Google or Facebook have access to and control a comparatively large amount of data. In general, users do not know what information has been generated about them, who has access to it, in what context and for what purpose the data has been used, or on the basis of which algorithms and correlations the users themselves have been classified (and possibly also evaluated).There seem to be clear discrepancies between a frontend, which is used to stage self-determination and optionality, and a backend, which ultimately pursues economic or political goals and serves predictivity. As a consequence, users in the digital world are confronted with great challenges, especially with regard to deciding whom or what they can trust and to what extent.

With regard to the subproject B1 of our RTG "Trust and Trust Assessment in Social Networks", this raises the central question of how socio-technical structures respectively orders can be created that counteract the imbalance between transparent users and opaque networks by creating problem-appropriate frameworks for a collective distribution of responsibilities and thus supporting users in the formation and evaluation of trust.

In my research, I want to contribute to the discussion and processing of this question by directing my attention — quite in a sociological manner — to the relations and contexts that are responsible for the formation and transformation of Social Networks as specific models of order. The central goal of my research is to investigate how different actors form, stabilize, and change specifically ordered Social Networks through recursive processes of negotiation and construction, in which trust and privacy appear as central reference variables and structuring elements. Thus, trust and privacy are both an essential component of models of order and themselves the subject of ongoing negotiation processes. Depending on how trust and privacy are constructed and staged, the relational structures and dependency relationships between the different actors or components of the network and thus the order of the network itself also change. In this context, OSNs are initially interpreted relatively general as platform-bound (or, in the case of decentralized networks, platform-independent) dynamic actor-networks in which the various actors of the network embed and position themselves reciprocally in a media-mediated manner. These can be classic Social Networks such as "Facebook" or "Twitter" as well as networks that are formed via a platform such as "ebay" or "Airbnb". Due to the variety of different Social Networks and the different definitions of what is considered a Social

Network, an essential part of my research will be to first differentiate and classify different forms of Social Networks by referring to or developing a taxonomy. Using the taxonomy, I will select at least two empirical cases, which will then be used to explore the question of this thesis in more depth.

Limits of Commercial Profiling in the European Law

Dirk Müllmann (muellmann@jur.uni-frankfurt.de)
Supervisor: Prof. Dr. Spiecker genannt Döhmann, LL.M. (Georgetown)

Due to the extensive inclusion of networked technologies into our daily routines it is possible to acquire a comprehensive picture of our activities, attitudes and interests by processing usage data. Users' information, which is often very sensitive, is collected and intertwined with other personal details in profiles allowing the analysis of the collected data, the deduction of metadata and the monetization of both. Profiles can hold whole daily routines, movement- or activity- profiles with the result that almost every operation in the real world finds a virtual equivalent. This situation can create an enormous "surveillance pressure" under which citizens might refrain from actions deviating from those of the majority population as they fear possible negative impacts of being different. Furthermore, such algorithm-based analysis of behavior can lead to a determination of human conduct. The algorithmic method, specifically, assumes that a person does not change and will always act in a similar way. Based on this premise the algorithm won't recommend action alternatives which don't conform to former decisions. The fundamental dangers of the technological advances have already been legally addressed in the 80s and have been adapted on the European level. The scientific investigation at hand aims to apply the findings of European data protection law on the technology of profiling in commercial contexts. In this regard, it attempts to reach a legal balance between economic chances of profiling and its dangers for democratic societies. Therefore, the thesis strives to answer two questions: Is there a quali- or quantitative limit for the acquisition, compilation, analysis and use of personal data in commercial profiles? And if so, is it possible to reproduce this limit in a practical system to ensure the protection of fundamental rights and to provide legal security for companies? In order to answer these questions it is necessary to analyze the technological aspects of profiling by examining how data for profiles is collected and exploited. Moreover, the methods depicted have to be legally classified in order to gain a legal understanding and definition of 'Profiling'. Starting with the European primary legislation it has to be examined if a quali- or quantitative limit for commercial profiling exists. For this purpose, it is crucial to find out if profiling is reconcilable with the essence of the conflicting fundamental rights. As those are predominantly defence rights against the state it is, furthermore, necessary to evaluate, if they develop a 'third-party-effect' under European Law. The legal consideration of secondary European legislation will, in addition, require to assess the impact of a granted consent on the creation of profiles. Finally, it is necessary to determine basic legal, sociological and psychological aspects for a quali- and quantitative scale to indicate the legality of a profile.

Privacy in user-based Bluetooth Protocols

Olga Sanina (sanina@privacy-trust.tu-darmstadt.de)
Supervisor: Prof. Dr. Marc Fischlin

The main goal of research area D is to develop a system, AlterEgo, that can be trusted in representing a user and his interests in digital networks. However, this is not possible without having the system being verified to be trustworthy enough by the means of identification, certification, and, if required, some other including new ways. Before deploying the infrastructure, it is important to analyse it in a theoretical sense, therefore, D2 project specifically supports research area D on the mathematical level.

Bluetooth is one of the private area networks (PANs) where AlterEgo can be used. Due to its short range for data transmission, application of Bluetooth is limited. However, developing of contact tracing, wearables and IoT has put Bluetooth back to be a popular solution for the need of communication between devices.

Bluetooth is being studied for a long time by researchers from different areas. Whereas some attacks and vulnerabilities of Bluetooth devices depend on the manufacturers, it is essential to prove security guarantees on the level of the standards used as building blocks to create devices. For instance, some pairing modes in Core Specification[1] involve humans into authentication process by confirming or entering the digits displayed on the screens. Manufacturers hence implies this approach when designing devices with Bluetooth. In this regard, the project aims to find the answers to the following questions:

- Is key exchange in Bluetooth protocol secured? Does it provide security guarantees?
- Is authentication sufficient? Does it rely on human being involved? What consequences can lead trust-on-the-first use approach to? Does it provide authentication of parties involved into communication at all?
- How security guarantees are dependent on the modes of device pairing? Is it possible to maliciously change the mode to the one with lower guarantees? Do some modes indeed protect from man-in-the-middle attacks?
- Does Bluetooth provide privacy for users? What actions can be taken to protect user's privacy from invading when user's device comes within the range of other Bluetooth device?

Since privacy is a focus of this research, it is important to understand whether Bluetooth devices may disclosure data of users and in a what way. Project implementation will allow users to make sure that Bluetooth is trustworthy and standard developers to improve it.

[1] Bluetooth SIG Proprietary "Bluetooth Core Specification Version 5.2", P. 3256, 2019.

Privacy and Trust in Value Related Fields of Tension

Enno Steinbrink (steinbrink@peasec.tu-darmstadt.de)
Supervisor: Prof. Dr. Christian Reuter

Privacy is an issue in the online world that often arises in conflict with other interests and values. These may be social connectedness online, financial interests in free-to-use services where people pay with their data, personal health by the exchange of healthcare data, and more. Consequently, users have to make complex decisions in which they evaluate many factors, possibly resulting in the so-called 'privacy paradox'. This paradox describes the willingness of people to give up privacy for little incentives while at the same time stating to consider privacy as an important factor. This project tries to explore the interplay between different factors that impact privacy behavior in different situations, especially focusing on the factor of trust and on crisis situations when personal stakes are high. Trust has been previously identified by research as a key factor, because it impacts the behavior towards the person or entity that is receiving the user data, as well as it impacts the acceptance of technological solutions as most users lack technical expertise.

The project approaches the subject from three perspectives: First, from a perspective in which specific situations are examined where user privacy collides with other values in a complex way (one example is the smartphone use of asylum seekers during their journey). Second, on an abstracted level in order to aggregate the results and to identify common denominators of different contexts. Last but not least, it aims to address the issues identified by the development of privacy enhancing technologies (PETs) as well as the guidelines for trust-building.

Human Factors in Privacy

Alina Stöver (stoever@privacy-trust.tu-darmstadt.de)
Supervisor: Prof. Joachim Vogt

Recent studies prove again, that people still have privacy concerns[1]. At the same time, they are confused regarding privacy policies and express a lack of control over their personal information[2]. This leads to the question: How can we support users in protecting their privacy? This question is addressed in two parts. Part I deals with the measurement of users' intentions regarding privacy, using a persona approach. Part II aims to investigate the communication of the users' intentions to a so-called privacy assistant.

If we design privacy support solutions for users, we face two issues: (1) Users differ in terms of privacy issues (e.g. privacy concerns, motivation, knowledge)[3] and (2) we can find evidence that one solution might not fit all users[4]. One approach that addresses these issues, is to cluster users into privacy personas. Already existing clustering approaches, cluster users in terms of concerns[5] or their knowledge and motivation to protect their privacy[6]. But the existing instruments that assign users to a cluster lack quality (e.g. Westins 3-item-scale shows low validity). The goal of part I is to develop an instrument (questionnaire) that meets quality criteria such as objectivity, reliability, and validity and allows to cluster people into privacy personas. Therefore first cluster criteria will be identified, second, the instrument will be developed and validated.

Part II applies the results from part I in the context of a so called privacy assistant that supports users by enforcing their privacy intentions. The idea here is to develop a prototype of a privacy assistant that uses the privacy personas and to investigate the communication of users intentions to privacy assistants.

[1] Braun, M. and Trepte, S. (2017). "Privatheit und informationelle Selbstbestimmung: Trendmonitor zu den Einstellungen, Meinungen und Perspektiven der Deutschen". Stuttgart: Universität Hohenheim.

[2] Pew Research Center. (2019, December 31). "Americans and Privacy: Concerned, Confused and Feeling Lack of Control Over Their Personal Information". Pew Research Center: Internet, Science and Tech.

[3] Baruh, L., and Cemalcılar, Z. (2014). "It is more than personal: Development and validation of a multidimensional privacy orientation scale". Personality and Individual Differences, 70, 165–170.

[4] Rudolph, M., Polst, S., and Doerr, J. (2019, March). "Enabling Users to Specify Correct Privacy Requirements". In International Working Conference on Requirements Engineering: Foundation for Software Quality (pp. 39–54). Springer, Cham.

[5] Westin, A., and Harris Louis and Associates (1991). Harris-Equifax "Consumer Privacy Survey". Tech. rep., Conducted for Equifax Inc. 1,255 adults of the U.S. public.

[6] Dupree, J. L., Devries, R., Berry, D. M., and Lank, E. (2016, May). "Privacy personas: Clustering users via attitudes and behaviors toward security practices". In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (pp. 5228–5239). ACM.

Mechanisms for Protecting Privacy in Applications

Amos Treiber (treiber@privacy-trust.tu-darmstadt.de)
Supervisor: Prof. Dr.–Ing. Thomas Schneider

Today, mobile applications are central to our lives. Driven by the goal of personalized user experience through machine learning (ML) techniques, operators collect large quantities of individual user data. As a result, user data has become essential to them, raising the need for privacy protection and spawning legislation like the General Data Protection Regulation (GDPR).

The usage of privacy-enhancing technologies from applied cryptography such as secure computation (SC) has been shown to be a promising approach to preserve privacy while still allowing an application to process user data. Recently, research has been focused on making machine learning techniques privacy-preserving. However, using these techniques usually requires large-scale computations even without privacy in mind. Existing solutions with optimal leakage do not scale well and require expert knowledge for deployment, which disincentivizes privacy protection in real-world applications. While some privacy-preserving solutions gain efficiency by leaking some information, this approach leaves open the real-world impact on privacy, partly because attacks exploiting leakage have only been studied in artificial environments.

In this work, we evaluate and build mechanisms for protecting privacy, focused on large-scale applications from the domain of machine learning. Our goal is for these mechanisms to enable practical ways for effectively preserving privacy in real-world applications that can even be used by non-experts.

To achieve this, we develop methods from SC for efficient, privacy-preserving applications at large scale. Building on existing private ML work that was solely focused on privacy-preserving neural networks and decision trees, we show how to practically protect privacy in crucial upcoming variants from machine learning. As an important use case that requires the protection of biometric information due to international standardization efforts, we demonstrate how to apply SC techniques to allow for highly efficient, privacy-preserving speaker recognition. Further, in collaboration with legal experts we design a novel system building on SC technologies that allows security agencies to exchange suspect information in a manner that satisfies European data protection laws, thereby moving towards solving the problem that data protection hinders modern law enforcement. Our developed tools are published as open source and are targeted to be usable by non-experts.

Additionally, we examine the practical security of existing solutions. We prove the insecurity of a protocol central to a line of prior privacy-preserving ML research and show how to learn private inputs. We also provide a first understanding of the practical impact of information leakage by searchable encryption schemes, an SC mechanism for querying databases used in private ML. For this, we evaluate existing attacks in scenarios surveyed by real-world data, laying out in which use cases common leakage profiles violate privacy.

Distributed Private Analytics in Online Social Networks

Aidmar Wainakh (wainakh@tk.tu-darmstadt.de)
Supervisor: Prof. Dr. Max Mühlhäuser

Online Social Networks (OSNs) became essential means of communication in our modern society. People increasingly use the services provided by OSNs in their daily life. Currently, the dominant OSNs (e.g., Facebook and Twitter) are functioning in a centralized fashion. The service providers have full control of the user data. They collect and process the data to make revenue in several ways. Unfortunately, the providers show consistently insufficient commitment to the privacy of the users. The user data oftentimes is used without informed consent or even misused in different ways. It is disclosed to third parties (e.g., data miners companies). The data is prone to hacking activities (e.g., Facebook tokens hack 2018). In addition, some parties violate the usage policy of the OSNs and harvest the user data for suspicious purposes (e.g., Cambridge Analytica). That is, the users' privacy under the centralized OSN scheme is seriously and continuously violated.

Within the subproject B.2 of RTG 2050, we focus on enhancing the privacy aspect in OSNs by giving the users the ability to control their own data. For that, we propose the concept of hybrid OSN (HOSN), where users can still use the centralized OSNs but with additional means of privacy control. The HOSN is based on three objectives. First, providing users with techniques for distributed anonymous communication. Hence, users are able to communicate and exchange data privately and efficiently. Second, increasing the users privacy awareness by providing users with measures to quantify their privacy level. Third, putting the data access control in the hands of users. Thus, users control what data and in which accuracy is accessible by the providers.

Realizing the concept of HOSN requires to consider the financial sustainability of provider companies. The main source of these companies' revenue is the targeted advertisements. Thus, in order to keep their business model functioning, the providers need to obtain sufficient data to run advertisements. Therefore, I focus in my research on the data exchange between users and the providers, i.e., the data access control objective. Users can deliver data models to the providers instead of the raw data. These models can be built by the users collaboratively in a privacy-preserving manner. To achieve that, I investigate several methods, e.g., federated machine learning.

AlterEgo as Trustworthy Device Collective

Dr. Ephraim Zimmer (zimmer@privacy-trust.tu-darmstadt.de)
Supervisor: Prof. Dr. Max Mühlhäuser

Smartphones have become a common and ubiquitous device for handling our personal data as well as for interacting with services and devices—mobile devices are becoming our digital counterpart, i.e., are becoming our AlterEgo. Rather than protecting our privacy, today's mobile devices on the contrary distribute personal data. More worrying, our options to assess their trustworthiness are slim to non-existing. Finally, today's mobile devices lack the ability to prove our trustworthiness to others, and, in return, to allow us to quantify the trust in services, online social networks (OSNs), and devices. The goal of this project is to evolve mobile devices towards a true digital counterpart—an AlterEgo. Users should be able to assess the trustworthiness of their digital counterparts and to control their personal data. Further, users, services, OSNs, and devices quantify the trust in each other. Ultimately, AlterEgo should not only be capable of supporting the user but also of acting autonomously on the user's behalf.

To achieve this goal, this subproject D.4 of our Research Training Group (RTG) follows a multi-layered approach: (1) deep and collaborative activity monitoring architecture, (2) distributed AlterEgo architecture, (3) proactive user assistance, and (4) mechanisms to protect privacy and assess trust according to dynamic constraints.

At the lowest layer 1, the lack of possibilities to assess the trust in a personal device is addressed. No matter what functionality, data, and even security or privacy mechanisms are deployed on the devices of users, if the hardware and software of those devices cannot be trusted, than the personal data is at stake. Mechanisms for deep as well as inter-device collaborative monitoring are able to assess the nature of ongoing device and network activities. As a result, among others, hidden functionalities in hardware and software can be identified and a level of trustworthiness can be established.

On layer 2 the privacy interests and rights of a user are provided and enforced accross different devices by the functionality of an AlterEgo, which acts in a distributed manner. Either agent-based or by means of an even stronger interconnection of lower device levels than it is known nowadays, the AlterEgo is providing access to private data for third parties but at the same time keeps the control over this private data fully in the hands of the respective owner.

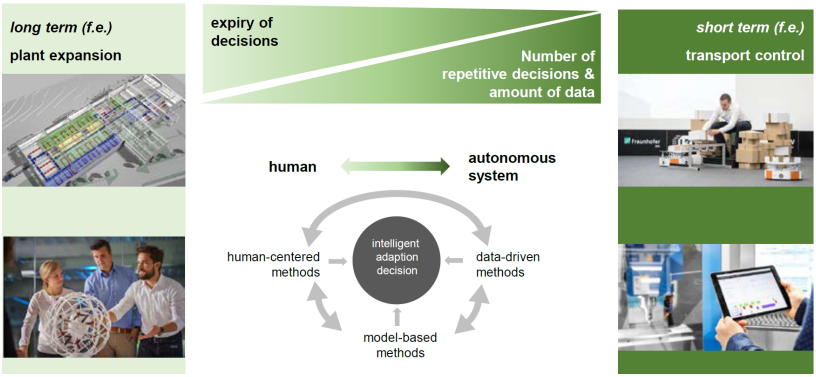
Layer 3 extends AlterEgo with device-local proactive assistance functionality. This functionality intends to increase the users' understanding of their actions' implications on privacy and trust and even aims at interacting on the users' behalf.

Layer 4 links this research project to the other research areas of our RTG 2050 by combining their research into a holistic AlterEgo with additional measures to flexibly assess privacy and trust, and enabling respective enforcement measurements.

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

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

Current research in the field of factory planning is not characterized by an adequate consideration of multidisciplinary procedures. These, however, are indispensable in order to grasp the complexity of factories, to consider their interdependencies, and to achieve shorter reaction time and adaption efficiency. The primary intention of the Research Training Group 2193 "Adaption Intelligence of Factories in a Dynamic and Complex Environment" is to strengthen an interdisciplinary education of doctoral researchers in integrated factory planning in order to reach a higher performance of collaborative planning and adaption in practice.



Based on this interdisciplinary approach, the PhD students of the GRK2193 are dealing from different perspectives with the central working hypothesis that

adaptation-intelligent factory systems can only be realized if human-centered methods, model-based methods and data-driven methods are combined. For the application domain, solutions for a target-oriented cooperation of human and algorithmic intelligence are therefore developed and evaluated in order to adapt factory systems more efficiently and safely, i.e. more intelligently. The focus is on the development and evaluation of AI-based and autonomous planning and control methods for factory systems, the application of machine learning methods within human-technology interactions, and the generation of a permanent planning readiness by digital and virtual factory models based on online data processing and automated software generation using component-based synthesis. Further research focuses on the effects of autonomous decisions on the sense of responsibility and traceability for humans and the role as well as the areas and limits of the irreplaceability of human action in autonomous systems.

Component-based Software Synthesis of Manufacturing Simulation Models

Fadil Kallat (fadil.kallat@tu-dortmund.de)
Supervisor: Prof. Dr. Jakob Rehof

In the factory planning and adoption process, a planning team chooses between a wide range of system configurations. Usually, simulation supports the decision-making process. Often, only selected configurations can be evaluated by simulation due to project budget and time restrictions. However, omitted solutions may be the most promising ones.[1]

A promising solution to this problem is the Automatic Simulation Model Generation (ASMG). A challenging area in this field is the structural variance of simulation models, which is important in terms of factory systems. For instance, a factory system may differ in the number and positions of machines. In addition, there is still a need for the adaption of complex control strategies, which often are functions implemented in higher programming languages. Moreover, most approaches in the field of ASMG are tailored to specific use-cases and can not be transferred inherently on other applications.[2]

In this PhD project, I develop an approach that uses component-based software synthesis to generate a set of simulation models. The simulation models are not built from scratch but are composed of building blocks held in a repository. The framework Combinatory Logic Synthesizer CLS, which implements a type inhabitation algorithm based on combinatory logic with intersection types, performs the synthesis. Given a target type, CLS generates all possible variants that meet the target type.[3] In addition, constraint solving allows filtering of not proper models considering domain-specific constraints. Besides, the CLS framework is also used to synthesize the control strategies in the simulation model.

This project aims to migrate discrete-event simulation models to enable the synthesis of a simulation model product line. Therefore, the repository is automatically built up by extracting components from an existing AnyLogic 8 simulation model. The factory planner configures the synthesis by adjusting the target type and modifying the components in a user-friendly web front-end. We tested this approach within different use cases of our RTG such as a sheet metal box production[1].

[1] F. Kallat, C. Mieth, J. Rehof, and A. Meyer, "Using Component-based Software Synthesis and Constraint Solving to generate Sets of Manufacturing Simulation Models," Procedia CIRP, vol. 93, pp. 556–561, 2020

[2] S. Wenzel, J. Rehof, J. Stolipin, and J. Winkels, "Trends In Automatic Composition Of Structures For Simulation Models In Production And Logistics," 2019 Winter Simulation Conference, pp. 2190-2200, 2019.

[3] J. Bessai, A. Dudenhefner, B. Düdder, M. Martens, and J. Rehof, "Combinatory Logic Synthesizer", 6th ISoLA, pp. 26-40, 2014

Control of decentralized systems under uncertainty

Alexander Puzicha (alexander.puzicha@cs.tu-dortmund.de)
Supervisor: Prof. Dr. Peter Buchholz

On the one hand, single intelligent and complex autonomous robots are developed for a wide variety of missions, for example, in disaster areas. Autonomous robots of this type are expensive; thus only a few of them are created during development of the control software. This leads to a system design where each robot is constructed to solve missions on its own and where the communication and collaboration with other agents is mostly neglected. On the other hand, there are small low cost swarm robots that focus on communication and collaboration but suffer from insufficient resources and capabilities to solve complex missions. The combination of complex and expensive agents in swarms to tackle the challenging problems at larger scale has not be completely considered in research, because the benefits and feasibility of these intelligent swarms have to be proven. Unfortunately, field tests are expensive and not always possible. Thus, the only viable alternatives are simulations. However, the test of software with real-time requirements has to be done in a real-time environment which puts very strict demands on the simulation software. The control software has to be part of the simulator, the dynamics of the robots has to be simulated realistically. The focus here is on the question how to describe dezentralized missions as mathematic formulas, which can be extracted by model predictive controllers to directly obtain control signals without creating waypoints by global planners. The mission functions have to react on events and messages of unrealiable networks. In addition, central servers, load distributions and mission planners are unwanted, because the form a singlepoint of failure and they hinder the advantages of fast local broadcast communication like LCM[1]. Sample based model predicitve control seems to be a promising starting point[2], because of its relaxation on possible cost functions since it does not require steadiness. This leads to the question which methods and optimization strategies shall be used and how to choose the sample points. To achieve compettitive results on trajectory smoothness and optimality explicit model knowledge should be considered in the optimizer. Another topic is how to model larger dynamic unknown environments to cover on as much entities as possible, but still guarantee real time behavior. As a result the handling of the unreliable network is important[3].

[1] Huang, A. S.; Olson, E.; Moore, D. C. (2010 - 2010): LCM: Lightweight Communications and Marshalling. In: 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems

[2] Hämäläinen, P., J. Rajamäki und C. K. Liu (2015): „Online Control of Simulated Humanoids Using Particle Belief Propagation“. In: Proc. SIGGRAPH '15. New York, NY, USA: ACM

[3] Puzicha, Alexander; Buchholz, Peter (2020): Real-Time Simulation of Robot Swarms with Restricted Communication Skills. In: 2020 IEEE/ACM 24th International Symposium on Distributed Simulation and Real Time Applications (DS-RT)

Stochastic Production Scheduling Using AlphaZero

Alexandru Rinciog (alexandru.rinciog@tu-dortmund.de)
Supervisor: Jun.-Prof. Dr.-Ing. Anne Meyer

This work investigates the applicability of AlphaGo Zero (AZ), one of the most prominent advances in reinforcement learning (RL) today, for optimizing large scale dynamic stochastic production scheduling problems. Production scheduling is the task of assigning operations grouped into jobs to processing resources such that a target goal, e.g. tardiness or makespan is optimized. Constraints on the job structure and resource capabilities define individual problems, most of which are NP-complete.

While many established scheduling problems, e.g. the Job-Shop Scheduling Problems, are deterministic, dynamic and stochastic problem variants better capture the nature of many real-world production systems: Jobs arrive in a continuous often stochastic fashion, operation processing times are subject to noise, due dates may change due to unforeseen events in the supply chain and the availability of production resources may change because of machine breakdowns, for instance.

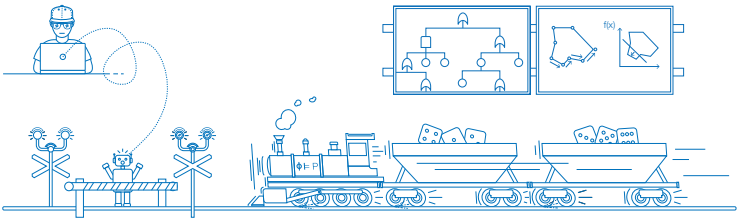
Traditionally, (near-)optimal schedules are found using exact methods, e.g. constraint programming, if the production instances are small, or through search, e.g. evolutionary algorithms, for larger instances. For production problems where the instance size makes exact planning and search too computationally taxing, simple heuristics, e.g. least processing time, are employed.

Given the rise in production data availability due to more exact production tracking systems, learning methods, particularly RL, are increasingly considered for scheduling. RL algorithms have been applied to both deterministic and stochastic scheduling problems. In the deterministic case, while outperforming simple heuristics, RL fails to outperform the state of the art optimization approaches. In stochstic settings the state of the art is difficult to establish owing to experiment reproducibility issues. For non-deterministic problems, RL may provide a competitive approach, since RL solvers are adaptive and potentially robust in highly stochastic settings.

Our contribution is threefold: Firstly, we develop an RL and planning method compatible benchmarking simulation framework guaranteeing experiment reproducibility. Secondly, we assess various production scheduling problems with respect to stochasticity, flexibility and resource workload to quantify the situations where simple heuristics perform well relative to (near-)optimal solutions, thus warranting RL application. Thirdly, we evaluate different RL modeling approaches and benchmark AlphaZero against heuristics, other popular RL methods, e.g. Deep Q-Networks, and search methods in the aforementioned cases.

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 lots 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. Systems have to deal with unpredictable and sometimes hostile environments. A different, also inevitable, kind of uncertainty arises from abstractions in system models focusing on the control of events. Probabilistic modelling and randomization are key techniques for dealing with uncertainty. Many trends witness this. Real-world modelling in planning is advancing by probabilistic programs describing complex Bayesian networks. In security, hostile environments are often captured by probabilistic adversaries. Probabilistic databases deal with uncertain data by associating probabilities to the possible worlds. In systems verification, probabilistic model checking has emerged as a key technique allowing for correctness checking and performance analysis. Similar developments take place in logic and game theory. The pervasiveness of uncertainty urges to make substantial enhancements in probabilistic modelling and reasoning so as to understand, reason about, and master uncertainty. The focus of the interdisciplinary RTG UNRAVEL is to significantly advance probabilistic modelling and analysis for uncertainty by developing new theories, algorithms, and tool-supported verification techniques, and to apply them to core problems from security (e.g., probabilistic protocols), planning (robotics and railway engineering), and safety and performance analysis (railway systems). To tackle these research challenges, theoretical computer scientists from computer-aided verification, logic and games, algorithms and complexity, together with experts from management science (robust optimization), applied computer

science (robotics and security), and railway engineering form the core of this RTG. 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. The curriculum consists of bi-weekly research seminars, soft-skill courses, reading groups, annual workshops, a summer school in the first Ph.D. year, and advanced (guest) lectures.

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

Robust Primary Care Systems

Martin Comis (comis@math2.rwth-aachen.de)
Supervisor: Prof. Dr. Christina Büsing

Introduction. Primary care systems are generally considered to be the backbone of universal health care. However, as the population ages and the number of primary care physicians declines, this foundation is starting to crumble. There result increasing access distances, waiting times, and workloads up to the point where the system's functioning can no longer be guaranteed. To counteract these developments, representatives from the government, insurances, and associations discuss an array of novel supply concepts and policy changes. This project aims to advance this discussion by providing suitable decision support tools, algorithms, and theoretic results. Special attention is thereby put on rural primary care systems, as these are particularly vulnerable due to their geographic-demographic facts. The resulting contributions can be categorized into three main groups and we summarize them hereinafter.

The first part of this project addresses the fundamental question of how the quality of primary care systems can be quantified. Due to the inherent complexity and micro-level detail of primary care systems, this turns out to be a highly non-trivial problem and the predominant method of choice is therefore still an assessment of the physician-to-population ratio. To facilitate a more refined analysis, this project introduces the hybrid agent-based simulation model `SiM-Care`. `SiM-Care` models and tracks the micro-interactions of patients and primary care physicians on an individual level. The model thereby enables decision makers to access several performance indicators such as patient waiting times and physician utilization that can serve as a sound basis for the assessment and comparison of primary care systems. Furthermore, it becomes possible to evaluate changes in the infrastructure, patient behavior, and service design which is otherwise impossible with purely ratio-based assessments.

The second part of this project examines mobile medical units (MMUs) for the supply of primary care services in rural environments. MMUs are customized vehicles fitted with medical equipment that are easy to relocate and therefore enable a demand-oriented and local provision of health services. Prior to their operation, MMUs require a complex prelaunch strategy to ensure their effectiveness and sustainability. To devise such strategies, this project contributes an integrated multi-phased optimization framework. Novel to this framework is the consideration of two types of patient demands; namely, patients who seek health services through a centralized appointment system as well as walk-ins who do not announce their visits. Moreover, the framework allows for the incorporation of uncertainties in both types of patient demands which was previously unconsidered.

The third part of this project studies two matching problems that derive from the application of MMUs in primary care. It is shown that very restricted variants of these matching problems are already strongly NP-hard. Consequently, this project focuses on restricted graph classes and contributes a range of polynomial and pseudo-polynomial algorithms.

Part I: Agent-based Modeling for Primary Care. The planning, analysis, and adaptation of primary care systems is a highly non-trivial problem due to the systems'

inherent complexity, unforeseen future events, and scarcity of data. To support the search for solutions, Part I introduces the hybrid agent-based simulation model SiM-Care. SiM-Care models and tracks the micro-interactions of patients and primary care physicians on an individual level. At the same time, it models the progression of time via the discrete-event paradigm. Thereby, it enables modelers to analyze multiple performance indicators such as patient waiting times and physician utilization to assess and compare primary care systems. Moreover, SiM-Care can evaluate changes in the infrastructure, patient behavior, and service design. To showcase the strengths of SiM-Care and its validation through expert input and empirical data, we present a case study for a primary care system in the northern Eifel region of Germany. Specifically, we study the immanent implications of demographic change on rural primary care and investigate the effects of an aging population and a decrease in the number of physicians, as well as their combined effects.

Part II: Operational Planning for Mobile Medical Units. Mobile medical units (MMUs) are customized vehicles fitted with medical equipment that are used in the provision of primary care in rural environments. As MMUs can be easily relocated, they enable a demand-oriented, flexible, and local provision of health services. In Part II of this project, we investigate the operational planning of an MMU service in three sequential phases to which we refer as Phase 1, Phase 2, and Phase 3. Phase 1 considers the strategic planning problem for MMUs (SMMU) in which we decide where MMU operation sites are set up and how often these are serviced in the course of one week. To that end, we model the problem as a capacitated set covering problem that includes existing practices and two distinct types of patient demands: i) steerable demands representing patients who seek health services through a centralized appointment system and can be steered to any treatment facility within a given consideration set and ii) unsteerable demands representing walk-ins who always visit the closest available treatment facility. We propose an integer linear program for the SMMU that can be solved via Benders decomposition and constraint generation. Starting from this formulation, we focus on the uncertain version of the problem in which steerable and unsteerable demands are modeled as random variables that may vary within a given interval. Using methods from robust optimization and duality theory, we devise exact constraint generation methods to solve the robust counterparts for interval and budgeted uncertainty sets. In Phase 2, we address the planning of MMUs at the tactical level. To that end, we investigate a bottleneck partitioning variant of the k-center problem that we call the tactical partitioning problem for MMUs (TPMMU). We show that the metric TPMMU is NP-hard to approximate within a constant approximation factor $1 < \alpha < 2$ and subsequently derive a MILP formulation. Moreover, we show that all our results from Phase 1 for the SMMU translate to a session-specific problem extension that combines strategic and tactical planning and thereby enables for a joint consideration of Phases 1 and 2. The final Phase 3 is devoted to the vehicle routing of MMUs. For a single depot, we reduce the problem to a minimum weight perfect matching problem in a bipartite graph which can be solved in polynomial time. In the multi-depot setting, we show that the vehicle routing of MMUs is a special case of the so-called budgeted colored bipartite perfect matching problem which we subsequently prove to be strongly NP-hard. To solve the vehicle routing problem for MMUs with multiple depots, we derive a compact integer linear programming formulation. Finally, we evaluate the

entire three-phased optimization framework in a computational study based on a set of instances that we generate from the rural primary care system in Germany that we considered in Part I of the project.

Part III: Variations of the Matching Problem. Assignment problems are among the most famous combinatorial optimization problems and have been studied in many variations. In Part III of this thesis, we consider two such variations which are motivated by the vehicle routing and staff assignment for MMUs. The first variation is a weighted matching problem with k independent edge cost functions called the multi-budgeted matching problem mBM. The total cost of a matching with respect to each cost function must not exceed a corresponding budget. We show that the mBM is strongly NP-hard on paths with uniform edge weights and budgets. Subsequently, we propose a dynamic program for series-parallel graphs with pseudo-polynomial running time for a fixed number of budget constraints. As an extension, we show how this algorithm can be used to solve the mBM on trees using a graph transformation. Realizing that both these graph classes have a bounded treewidth in common, we introduce a dynamic program based on tree decompositions. This approach leads to a pseudo-polynomial algorithm for the mBM with fixed number of budget constraints on graphs of bounded treewidth. The second matching problem that we study is the minimum color-degree perfect b-matching problem Col-BM) which represents a new extension of the perfect b-matching problem to edge-colored graphs. The objective of the Col-BM is to minimize the maximum number of differently colored edges in a perfect b-matching that are incident to the same node. We show that the Col-BM is strongly NP-hard on two-colored bipartite graphs and that there exists no α -approximation algorithm for $1 < \alpha < 2$ unless P=NP. Still, we identify a class of two-colored complete bipartite graphs on which we can solve the Col-BM in polynomial time. Furthermore, we use dynamic programming to devise polynomial-time algorithms solving the Col-BM with a fixed number of colors on series-parallel graphs and simple graphs with bounded treewidth.

Design and Analysis of Algorithms for Combinatorial Optimization Problems under Uncertainties

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

Introduction. 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. Combinatorial algorithms to find equilibrium prices often use the concept of duality. For example, in two-side matching markets with valuations on both sides, equilibrium prices can be found by iteratively applying a primal-dual algorithm to find a max-cardinality matching in a bipartite graph [1].

Another example are markets in which trade occurs via intermediaries. Sellers and buyers have valuations for trading with each other and only trade if they profit. It can be shown that welfare maximizing equilibrium prices exist and can be computed efficiently via a primal-dual algorithm for solving a max-weight matching in a bipartite graph and adapting the dual prices so that they form equilibrium prices [2].

[1] Gabrielle Demange, David Gale and Marilda Sotomayor, "Multi-item auctions," Journal of Political Economy, vol. 94.4, p. 863–872, 1986

[2] Lawrence E. Blume, David A. Easley, Jon Kleinberg and Éva Tardos, "Trading networks with price-setting agents," Games and Economic Behavior, vol. 67.1, p. 36–50, 2009

Optimization under Uncertainty

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

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

[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

from the follower 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

Robust Infrastructure

Nadine Friesen (friesen@via.rwth-aachen.de)
Supervisor: Prof. Dr. Nils Nießen

The extended planning periods and the long life cycle of railway infrastructure require a lengthy planning horizon. Thus, bottlenecks in the infrastructure have to be recognized at an early stage to initiate adequate measures. At the present, the infrastructure is planned while only little about the intended operation is known. Hence, the timetable and the operation are adjusted to the infrastructure. Since space, time and money for extension measures of railway infrastructure are limited, each modification has to be done carefully and and in a long lasting manner. To meet the customers' future needs, infrastructural projects have to planned such that the infrastructure will be appropriate for future unknown demand.

For the long-term service life of the planned infrastructure, it makes sense to include timetable scenarios in the planning in order to be able to expand the railroad infrastructure, which is already reaching its capacity limits on some lines, in a targeted manner.

The aim of the project is to provide a procedure for timetable-based, robust infrastructure planning to complement the previous infrastructure-based timetable construction. In doing so, the idea of a long-term timetable and the infrastructure adaptation based on it will be precisely defined and further developed. For this purpose, infrastructure planning is modeled as a network design problem under uncertainties. Subsequently, a solution is to be found by means of robust optimization.

The term "robustness" is generally understood as the ability of a method to find a correct solution even under uncertain input data. In the context of this project, the timetables are not yet fixed until the end of the infrastructure's service life and are, thus, still uncertain at the time of infrastructure planning. For example, it is not a realistic scenario to run only one timetable throughout, but likewise infrastructure cannot be held in reserve for every scenario, no matter how unlikely, for both financial and spatial reasons. In the context of this project, both the various, potential timetable scenarios as well as the actual operational scenarios are to be included in the considerations.

For this purpose, the uncertain input data will be modeled in a first step. In the context of the project, it is first determined which criteria for a "similarity" of timetable scenarios have to be fulfilled to which extent.

Subsequently, infrastructure planning is modeled as a network design problem. Here, a solution has to be found for which new edges, i.e. track sections, are needed and for which edges the capacity should be increased. This solution should ensure the feasibility of all planned timetable scenarios at the lowest possible cost.

Special Online Problems with Advice

Janosch Fuchs (fuchs@algo.rwth-aachen.de)
Supervisor: Prof. Dr. Gerhard Woeginger

The graph exploration problem demands the shortest tour that visits every vertex at least once in an unknown graph. Regarding the decisions of the algorithm, its knowledge is limited by the perception of the explorer. There are different models regarding the perception of the explorer. For our research, we used the fixed graph scenario, proposed by Kalyanasundaram and Pruhs (Proc. of ICALP, 1993), where the explorer starts at a vertex of the network and sees all reachable vertices, their unique names and their distance from the current position. Thus, the algorithm recognizes already seen vertices and can adapt its strategy during exploring as it does not forget anything.

Since the algorithm learns the structure of the graph during computation, it cannot deterministically compute an optimal tour that visits every vertex at least once without prior knowledge. Therefore, we are interested in the amount of crucial a-priori information needed to solve the problem optimally, which we measure in terms of the well-studied model of advice complexity.

There are different variations of the graph exploration problem and they can be differentiated between directed or undirected edges, cyclic or non-cyclic solutions, unit costs or individual costs for the edges and different amounts of a-priori structural knowledge of the explorer.

For general graphs, it is known that $n \log n$ bits of advice suffice to compute an optimal solution. We found algorithms with an advice complexity of $O(m+n)$, thus improving the classical bound for sparse graphs. Our algorithms solve the problem on directed or undirected graphs. We can solve the cyclic graph exploration problem, where the explorer has to end at the same vertex where it started, and can be used to also solve the non-cyclic version of the problem by adding $\log n$ bits of advice.

The basic idea of the algorithms is to use the structure of the graph that is induced by the edges that are used in an optimal solution. There are edges that are used: never, precisely once or more than once in the optimal solution. By giving this information to the algorithm, mistakes can be avoided. But to find the shortest tour that visits every vertex at least once more information is necessary.

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.

Satisfiability Checking for Optimisation of Timetables in Railway Engineering under Consideration of Uncertainties

Rebecca Haehn (haehn@cs.rwth-aachen.de)
Supervisor: Prof. Dr. Erika Ábrahám

Introduction. In many application areas of logic in computer science the aspect of uncertainty plays an increasingly important role. For example in railway systems the high utilization of the existing infrastructure often causes a train's delay to impact several other trains as it is not possible to reserve the infrastructure for an extensive time after each train. When scheduling additional trains to a given timetable, the train's travel time in practice does not only depend on the distance, but also on external influences such as weather conditions, traffic and passengers. An additional uncertainty results from unreliable railway networks. Some of the tracks might even fail completely, for example due to construction work. Especially for long forecast periods the models of railway traffic include various uncertainties.

Despite the uncertainty in the available data long-term decisions on the design of railway networks, train timetables, and construction periods have to be made. These decisions are required to be robust, i.e. to deliver (almost) optimal solutions even for uncertain input data.

In an approach to estimate the remaining infrastructure capacity of railway systems by determining how many additional trains can be scheduled is presented. This approach can be used to compare different solutions to the above mentioned problems, however, uncertainty is completely neglected.

In [1] other approaches where uncertainty is considered in the form of varying primary delays this is often done by implementing Monte Carlo simulation, e.g. in [2] or [3]. There is one approach that implements analytic procedures to compute delay propagation instead of Monte Carlo simulation, presented in [4].

The aim of this project is to create an efficient system that can cope with uncertainty in railway traffic while making reliable statements about the network capacity. Thus, Christian Meirichs work, presented in his dissertation [1], is extended to take uncertainty into account.

Current project status. As a basis for our work, we had to mathematically model railway systems at a suitable level of detail. In order to consider uncertainty in a meaningful way, we modeled time discretely but at least in minutes, in contrast to the model in [1], where just one time period is considered. In order to cope with the considerable size of the problems, we modeled the infrastructure slightly less detailed. This should not be problematic as we do not require the individual infrastructure elements capacity due to considering individual time steps. This model is presented in [5], as well as an algorithm to schedule additional trains without disturbing a given timetable.

In order to consider uncertainty when scheduling additional trains, respectively computing how many additional trains might be scheduled, we developed a probabilistic simulation algorithm. This algorithm can be used to compute, for example, the expected utilization of infrastructure elements over time, which can then be considered instead of the planned utilization. The first approach for this algorithm, where stochastic dependencies are neglected, is presented in [6].

Currently, we modify the analytical simulation to take stochastic dependencies into account. Additionally, we are working on the visualization of the simulation results in order to make them better understandable and more accessible for human users to take into account for decision making. Therefore, we also work on the computation of the causes of delays. It is also planned to integrate these simulation results in the computation of additional trains.

References

1. Christian Meirich. Berechnung und Bewertung der Gesamtleistungsfähigkeit von Eisenbahnnetzen. Dissertation. RWTH Aachen University,
2. Janecek, D. , Weymann, F. LUKS - Analysis of lines and junctions 12th World Conference on Transport Research, 2010.
3. Schneider, Walter, Nießen, Oetting, Andreas. Moses / Wizug - Strategic modelling and simulation tool for rail freight transportation ETC (European Transport Conference) 2003.
4. Thorsten Büker, Bernhard Seybold. Stochastic modelling of delay propagation in large networks. J. Rail Transp. Plan. Manag. 2(1-2): 34-50 (2012)
5. Rebecca Haehn, Erika Ábrahám, Nils Nießen. Freight Train Scheduling in Railway Systems. MMB 2020: 225–241, 2020.
6. Rebecca Haehn, Erika Ábrahám, Nils Nießen. Probabilistic Simulation of a Railway Timetable. ATMOS 2020: 16:1–16:14, 2020.

Automated Runtime Analysis of Probabilistic Programs

Marcel Hark (marcel.hark@cs.rwth-aachen.de)
Supervisor: Prof. Dr. Jürgen Giesl

Analyzing the correctness of a program has become one of the most important steps in program development. In order to derive total correctness, reasoning about termination of a program is crucial. Usually, termination is not the only property of interest but one would like to know bounds on the asymptotic complexity in order to exclude the inefficient use of resources and security leaks such as possible denial of service attacks. Therefore, tools to infer the asymptotic runtime complexity of programs fully automatically, such as AProVE [9], have been developed and have shown good results in practice.

Probabilistic programs have become more and more popular over the years. Whereas in deterministic programs the simulation of non-deterministic behavior is limited, introducing probabilistic behavior enables a more fine-grained approximation of real-world systems and has shown great results in improving the efficiency of existing algorithms such as primality testing and sorting data.

Nevertheless, when introducing probabilistic actions, the concept of runtime changes from ordinary functions to random processes and random variables. To ease analysis, the exact behavior of a program is approximated by its expected behavior, such as the expected runtime of a program, yielding different concepts of termination known as positive almost sure termination (expected runtime is finite) and almost sure termination (termination with probability one). Although these concepts coincide in deterministic systems, there are probabilistic programs terminating almost surely with infinite expected runtime [13]. Furthermore, deciding termination with respect to these concepts becomes even more involved in the sense of the arithmetic hierarchy than in the deterministic case [14], limiting the chance of techniques for inferring the exact expected runtime of a probabilistic program.

However, in applications it is usually enough to know the asymptotic behavior of the expected runtime, e.g., linear, polynomial or even exponential. Thus, developing techniques for reasoning about bounds on the expected runtime is sufficient for real-world scenarios. The objective of this project is the development of effective algorithms for reasoning about the different termination concepts fully automatically.

In [15], we generalized the approach of [4] for inferring upper bounds on the runtime of non-probabilistic programs to probabilistic programs. Here, we used the already existing generalization of ranking functions to the probabilistic setting, called ranking supermartingales (see, e.g., [1, 3]). Additionally, we developed a concept of expected size for probabilistic systems. We combined these concepts in an alternating way to a powerful fully automatic technique for inferring upper bounds on the expected runtimes of probabilistic programs. However, this is a non-trivial extension of [4] since for probabilistic programs the concepts to prove the soundness of our approach differ significantly from the concepts for deterministic programs considered in [4]. In our experimental evaluation, our approach outperformed existing tools [2, 17] capable of a fully automatic inference for upper bounds on expected runtimes.

While some fully automatic approaches for inferring upper bounds on expected runtimes of probabilistic programs exist, only very few results on inferring lower

bounds on expected runtimes or expected outcomes exist [8, 16]. But having an upper bound is only an advantage if the upper bound already excludes inefficient effects such as memory usage or runtime. However, having an exponential upper bound is not useful because it means that the measured effect is at most very bad, whereas an exponential lower runtime bound would express that a possible adversary could make the system run very long by choosing an appropriate input. So, being able to infer lower bounds is crucial for a thorough program analysis.

Unfortunately, lower bounds for probabilistic programs are more intricate than upper bounds. This is due to the fact that in probabilistic systems the expected runtime or the expected outcome can be described as the minimum of a certain set of values. In order to find an upper bound it is then enough to give an upper bound for some of the values. For inferring a lower bound one has to bound every value from below. Hence, the straightforward generalization of the related deterministic concept (metering functions [6]) is not sound for analyzing probabilistic runtimes.

In [11], we presented the first inductive rules for lower bounds on expected outcomes and expected runtimes of probabilistic programs. Our rules are compositional, and thus,capable of analyzing complex programs also containing nested loops. Moreover, our results generalize existing techniques such as metering functions for lower bounds, and thus combine them to a more general concept.

Finally, for both non-probabilistic and probabilistic programs complete approaches for inferring bounds on the (expected) runtime or analyzing termination are desirable. However, as the underlying problems are undecidable such approaches only exist for limited classes of (probabilistic) programs. For non-probabilistic programs, we studied a class of so-called polynomial loops. We showed that for these loops termination is decidable over the real numbers and semi-decidable over the rationals and the integers. Additionally, we gave tight bounds on the complexity of deciding termination for these loops [7]. Furthermore, we showed that for these loops an upper bound on the runtime can always be computed and all witnesses for non-termination can be enumerated [12]. Our results can be integrated into existing tools for complexity and termination analysis to be applied to sub-programs of larger programs which fall into the class of loops studied by us.

For probabilistic programs, we determined a class of loops where there is a simple decision procedure for (positive) almost sure termination and both upper and lower bounds on the expected runtime can always be computed. Moreover, we developed an algorithm to compute the exact expected runtime for these loops [10]. In general, even if an asymptotic bound is often sufficient in applications, whenever it is possible, inferring the exact runtimes of programs is preferable to asymptotic runtimes.

References

1. Sheshansh Agrawal, Krishnendu Chatterjee, Petr Novotný: Lexicographic Ranking Supermartingales: An Efficient Approach to Termination of Probabilistic Programs. Proc. ACM Program. Lang. 2(POPL), 34:1-34:32 (2018)
2. Martin Avanzini, Georg Moser, Michael Schaper: A Modular Cost Analysis for Probabilistic Programs. Proc. ACM Program. Lang. 4(OOPSLA), 172:1-172:30 (2020)
3. Olivier Bournez, Florent Garnier: Proving Positive Almost-Sure Termination. Proc. of RTA, 323-337 (2005)

4. Marc Brockschmidt, Fabian Emmes, Stephan Falke, Carsten Fuhs, Jürgen Giesl: Analyzing Runtime and Size Complexity of Integer Programs. ACM Trans. Program. Lang. Syst. 38(4): 13:1-13:50 (2016)
5. Florian Frohn, Jürgen Giesl: Complexity Analysis for Java with AProVE. Proc. of IFM, 85-101 (2017)
6. Florian Frohn, Matthias Naaf, Marc Brockschmidt, Jürgen Giesl: Inferring Lower Runtime Bounds for Integer Programs. ACM Trans. Program. Lang. Syst. 42(3), 13:1-13:50 (2020)
7. Florian Frohn, Marcel Hark, Jürgen Giesl: Termination of Polynomial Loops. Proc. of SAS, 89-112, (2021)
8. Hongfei Fu, Krishnendu Chatterjee: Termination of Nondeterministic Probabilistic Programs, Proc. of VMCAI, pp. 468-490 (2019)
9. Jürgen Giesl, Cornelius Aschermann, Marc Brockschmidt, Fabian Emmes, Florian Frohn, Carsten Fuhs, Jera Hensel, Carsten Otto, Martin Plücker, Peter Schneider-Kamp, Thomas Ströder, Stephanie Swiderski, René Thiemann: Analyzing Program Termination and Complexity Automatically with AProVE. J. Autom. Reason. 58(1), 3-31 (2017)
10. Jürgen Giesl, Peter Giesl, Marcel Hark: Computing Expected Runtimes for Constant Probability Programs. Proc. of CADE 2019, 269-286 (2019)
11. Marcel Hark, Benjamin Lucien Kaminski, Jürgen Giesl, Joost-Pieter Katoen: Aiming Low Is Harder: Induction for Lower Bounds in Probabilistic Program Verification. Proc. ACM Program. Lang. 4(POPL), 37:1-37:28 (2020)
12. Marcel Hark, Florian Frohn, Jürgen Giesl: Polynomial Loops: Beyond Termination. Proc. of LPAR: 279-297 (2020)
13. Benjamin Lucien Kaminski, Joost-Pieter Katoen, Christoph Matheja, Federico Olmedo: Weakest Precondition Reasoning for Expected Runtimes of Randomized Algorithms. J. ACM 65(5), 30:1-30:68 (2018)
14. Benjamin Lucien Kaminski, Joost-Pieter Katoen, Christoph Matheja: On the Hardness of Analyzing Probabilistic Programs. Acta Informatica 56(3): 255-285 (2019)
15. Fabian Meyer, Marcel Hark, Jürgen Giesl: Inferring Expected Runtimes of Probabilistic Integer Programs Using Expected Sizes. Proc. of TACAS, 250-269 (2021).
16. Annabelle McIver, Carroll Morgan: Abstraction, Refinement and Proof for Probabilistic Systems. Monographs in Computer Science, Springer 2005, ISBN 978-0-387-40115-7, pp. 1-310
17. Van Chan Ngo, Quentin Carbonneaux, Jan Hoffmann: Bounded Expectations: Resource Analysis for Probabilistic Programs. Proc. of PLDI, 496-512 (2018)

Robust Execution of Abstract Task Plans on Mobile Robots

Till Hofmann (hofmann@kbsg.rwth-aachen.de)
Supervisor: Prof. Dr. Gerhard Lakemeyer

Introduction. Formalisms such as Golog [6] and PDDL [7] allow the specification of a robot's behavior in an abstract manner. Based on a logical model of the environment, the agent's actions are specified with preconditions and effects. This allows for determining the course of action by searching for an appropriate action sequence (PDDL), possibly intertwined with agent programs specified by the user (Golog). However, when deploying such a system on a real robot, one often faces additional challenges, such as the need to calibrate a robot arm before its usage. Those issues are intentionally ignored when specifying the abstract behavior, as it would impair the reasoner performance. This research project aims to close the gap between high-level reasoning and low-level robot platform [5]. Instead of specifying all the low-level details in the reasoning domain, we instead model the platform components separately as timed automata. Then, we specify constraints that connect a high-level program with the platform, e.g., by requiring that the arm needs to be calibrated five seconds before the robot picks up an object. We then need to transform the high-level program into a sequence of actions that satisfies all those constraints, resulting in a task specification that follows the high-level program while dealing with the low-level platform details.

A logic for specifying metric temporal constraints for Golog programs. In the first step, we extended ESG [3], a modal variant of the Situation Calculus that allows temporal constraints, with metric temporal constraints [4]. The resulting logic retains most of the properties of ESG and thus allows the specification of basic action theories and Golog programs extended with metric constraints.

Plan Transformation based on Timed Automata Reachability Analysis. In a first approach to solve the temporal platform constraints, we looked at timed automata reachability analysis. In a first step, we transform a high-level action sequence into a timed automaton such that each action is one location in the resulting automaton. This automaton is then combined with the platform model such that in the product automaton, all edges that violate a constraint are removed. Finally, we apply reachability analysis using the model checking tool UPPAAL [1]. The resulting path describes the transitions of the platform models such that all constraints are satisfied during the execution of the original plan. Despite the combinatorial blowup due to the automata product, this approach performs well and we were able to transform plans with 50 actions and several platform components in a few seconds.

Controller Synthesis for Golog Programs. The first approach, however, poses some limitations: For one, it only works on pre-determined plans. Thus, it cannot be used with any formalism that uses online sensing, as this would require online decision making. Also, it does not distinguish between controllable actions (e.g., starting to pick up an object) and actions that are controlled by the environment (e.g., the arm going into an error state, or even the end of an action). To tackle those limitations, we used a different approach based on MTL synthesis. Instead of applying reachability analysis, we build on top of results on controller synthesis for MTL specifications [1]. We first convert a given Golog program into a timed

automaton, apply MTL controller synthesis on the automaton, the platform model, and the platform constraints, and then use the resulting controller to guide the Golog executor. We presented the theory of the approach in [8] and we are currently working on an implementation in cooperation with the former UnRAVeL researcher Stefan Schupp and UnRAVeL supervisor Erika Ábrahám. This cooperation, which allows us to combine expertise in robotics with expertise in hybrid systems, directly resulted from an UnRAVeL workshop in April 2020, where we presented preliminary results for the synthesis approach.

References

1. Behrmann, G., David, A., and Larsen, K. G. A Tutorial on Uppaal. In: Formal Methods for the Design of Real-Time Systems: International School on Formal Methods for the Design of Computer, Communication, and Software Systems, Revised Lectures (pp. 200-–236). Springer, 2004.
2. Bouyer, P., Bozzelli, L., and Chevalier, F. Controller Synthesis for MTL Specifications. In Proceedings of the 17th International Conference on Concurrency Theory (CONCUR) (pp. 450–-464). Springer, 2006.
3. Claßen, J. and Lakemeyer, G. A Logic for Non-Terminating Golog Programs. Proceedings of the 11th International Conference on Principles of Knowledge Representation and Reasoning (KR), 589–-599, 2008.
4. Hofmann, T. and Lakemeyer, G. A logic for specifying metric temporal constraints for Golog programs. Proceedings of the 11th Cognitive Robotics Workshop (CogRob), 2018.
5. Hofmann, T., Mataré, V., Schiffer, S., Ferrein, A., and Lakemeyer, G. Constraint-based online transformation of abstract plans into executable robot actions. AAAI Spring Symposium: Integrating Representation, Reasoning, Learning, and Execution for Goal Directed Autonomy, 2018.
6. Levesque, H. J., Reiter, R., Lesperance, Y., Lin, F., and Scherl, R. B. GOLOG: a logic programming language for dynamic domains. Journal of Logic Programming, 31(1-3), 1997.
7. McDermott, D., Ghallab, M., Howe, A., Knoblock, C., Ram, A., Veloso, M., Weld, D., and Wilkins, D. PDDL - The Planning Domain Definition Language. The AIPS-98 Planning Competition Committee, 1998.
8. Hofmann, T. and Lakemeyer, G. Controller Synthesis for Golog Programs over Finite Domains with Metric Temporal Constraints. In: 17th International Conference on Principles of Knowledge Representation and Reasoning (Poster), 2020.

Privacy Preserving Online Algorithms

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

In secure multi-party computation a number of parties wants to compute a function over their inputs such that their inputs are kept private. The participating parties shall only learn their prescribed output without learning anything beyond that. The output can be either the same for all parties or each party obtains a different output. A trusted third party can be used to perform these computations. However, in some settings the parties want to keep their inputs private, e.g., if it is confidential or private information the parties are not willing to share with anyone. In order to keep the inputs private the parties avoid the trusted third party by computing the function in a distributed fashion, i.e., they jointly execute a secure protocol to simulate the trusted third party. In addition, such a protocol shall provide privacy and security in the presence of adversaries, i.e., a malicious party that wants to learn more than intended or deviates from the protocol specification arbitrarily.

For the most common secure multi-party computation settings it is assumed that everything is known prior to the protocol execution, i.e., the parties know their personal input and the set of parties participating in the protocol execution is somehow known. For such a determined setting there exist already a variety of protocols for different requirements. However, there are cases where the scenario is more uncertain and might change over time.

The aim of this dissertation project is to analyze these scenarios in more detail and provide a framework to define security and privacy in these settings. We will focus our research on online algorithms and develop protocols that can deal with different types of uncertainty.

Robust Hospital Management

Tabea Krabs (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.

The Theory of Infinite Probabilistic Databases

Peter Lindner (lindner@cs.rwth-aachen.de)
Supervisor: Prof. Dr. Martin Grohe

Introduction. Probabilistic databases (PDBs) are a formalism extending database theory towards probability spaces over database instances. Their application is the storage, processing, and manipulation of (large amounts of) uncertain data in the framework of databases. Typical situations and settings in which such data arise include information extraction, data integration, and in handling data from sensors or data obtained by machine learning methods. The semantics of probabilistic (relational) databases are centered around the notion of possible worlds referring to the possible outcomes of the database. Theoretical work on probabilistic databases has so far mostly been restricted to a very simple setting, where the probability space contains only finitely many possible worlds. Probabilistic databases bear the inherent problem of their representation. Database instances are typically large, even more so spaces of sets of database instances. This issue is usually approached by imposing simplifying structural assumptions on the probabilistic databases. For example, a tuple-independent probabilistic database may be specified by just listing all of its entries together with their marginal probability, and probabilities of instances are simply calculated as products of these. Then, the representation is not much larger than that of a traditional database.

Unfortunately, the semantics of probabilistic databases underlying the sketched means of representation follow a closed-world assumption [13]. This means that any information about the probability space that is not explicitly specified, is treated as an impossible event. It has been pointed out by Ceylan et al. [3] that this has undesirable consequences conflicting the intuitive behaviour of query answering and updating mechanisms. They proceed to describe a model of "open-world probabilistic databases" in an attempt to overcome this flaw. Yet, their model still has the drawback of considering finite probability spaces only.

The starting point of this research project was the idea to take the considerations of Ceylan et al. to the level of infinite probability spaces—after all, data records typically involve infinite domains like integers or real numbers. Some existing work proposed systems and models that can even handle continuous distributions [1,11,12]. Yet, the research field lacks a unifying framework for infinite probabilistic databases, rooted in possible worlds semantics.

Independence in Infinite PDBs and the Open-World Assumption. As a first step towards infinite probabilistic databases, we give a formal definition of infinite probabilistic databases and queries following the possible worlds semantics, the introduction of which has been requested in [5] (and to our knowledge has not been resolved in the meantime). In order to extend the ideas of [3], we investigate independence assumptions in this infinite setting. We characterize the existence of probabilistic databases with prescribed marginal probabilities and develop a notion of completions similar to [3] in order to perform open-world query evaluation in infinite probabilistic databases. As it turns out, obtaining multiplicative approximations of query results is impossible. Yet, the known methods from the finite setting can be used to obtain additive approximations for query answering. Resorting to independence assumptions immediately raises the question of expressiveness of the

model, and we investigate the expressive power for different kinds of independence assumptions, and using different classes of views.

This subsection describes the content of [7] and [10]. An extended version of [7] (which was invited for submission to a special issue of the JACM) is currently under review.

Standard Probabilistic Databases. The model from [7] is quite powerful, but it lacked a thorough discussion of measurability issues. Due the finite setting, much of previous theoretical work eluded such analysis, yet it is inevitable as soon as uncountable attribute domains are involved in the data model. We present a thorough construction of measurable spaces for infinite probabilistic databases that is rooted in the theory of finite point processes [4]. This completes the picture of [7] towards a unifying framework for infinite probabilistic databases, yielding the notion of standard PDBs. The technical assumptions that have to be made for the development are quite mild and amply cover the typical applications. We show that for standard PDBs, all typical kinds of database queries are measurable functions, which means that they have a well-defined semantics that can be lifted from the single instance semantics in the natural way.

This subsection describes the content of [9]. An extended version of this paper (which was invited for submission to a special issue of LMCS) is currently under review.

Generative and Probabilistic Programming Datalog. As an application of the previously described developments, we use our model of standard PDBs to extend the semantics of the probabilistic Datalog language PPDL [2] towards the support of continuous distributions. As a side effect, we obtain more general semantics, rendering the language compositional and robust with respect to different modes of evaluation.

This subsection describes the content of [8]. An extended version of this paper (which was invited for submission to a special issue of JACM) is currently under review.

References

1. Parag Agrawal, Jennifer Widom: Continuous Uncertainty in Trio. MUD 2009: 17–32, 2009
2. Vince Bárány, Balder ten Cate, Benny Kimelfeld, Dan Olteanu, Zografoula Vagena: Declarative Probabilistic Programming with Datalog. ACM Trans. Database Syst. 42(4): 22:1-22:35 (2017)
3. Ismail Ilkan Ceylan, Adnan Darwiche, Guy Van den Broeck: Open-World Probabilistic Databases. KR 2016: 339–348, 2016
4. D.J. Daley and D. Vere-Jones. An Introduction to the Theory of Point Processes–Volume I: Elementary Theory and Methods, Springer, 2003.
5. Nilesh N. Dalvi, Christopher Ré, Dan Suciu: Probabilistic databases: diamonds in the dirt. Commun. ACM 52(7): 86–94 (2009)
6. Christopher De Sa, Ihab F. Ilyas, Benny Kimelfeld, Christopher Ré, Theodoros Rekatsinas: A Formal Framework for Probabilistic Unclean Databases. ICDT 2019: 6:1–6:18, 2019
7. Martin Grohe, Peter Lindner: Probabilistic Databases with an Infinite Open-World Assumption. PODS 2019: 17-31, 2019
8. Martin Grohe, Benjamin Lucien Kaminski, Joost-Pieter Katoen, Peter Lindner: Generative Datalog with Continuous Distributions. PODS 2020: 347–360, 2020

9. Martin Grohe, Peter Lindner: Infinite Probabilistic Databases. ICDT 2020: 16:1–16:20, 2020
10. Nofar Carmeli, Martin Grohe, Peter Lindner, Christoph Standke: Tuple-Independent Representations of Infinite Probabilistic Databases. CoRR abs/2008.09511 (2020)
11. Ravi Jampani, Fei Xu, Mingxi Wu, Luis Leopoldo Perez, Chris Jermaine, Peter J. Haas: The Monte Carlo database system: Stochastic analysis close to the data. ACM Trans. Database Syst. 36(3): 18:1-18:41 (2011)
12. Oliver Kennedy, Christoph Koch: PIP: A database system for great and small expectations. ICDE 2010: 157–168, 2010
13. Raymond Reiter. On Closed World Data Bases. Readings in Artificial Intelligence. 119–140, 1981.

Probabilistic Action Formalisms with Applications to Robotics

Daxin Liu (liu@kbsg.rwth-aachen.de)
Supervisor: Prof. Dr. Gerhard Lakemeyer

Introduction. A belief program [1], a member of GOLOG program family, is a type of probabilistic program where test conditions inside the program refer to the agent's subjective belief and the agent's effectors and sensing could be noisy. Due to the action-centered feature and subjective feature, it is extremely suitable for high level robot control. An important step before deploying such a program is to verify whether the program satisfies certain properties. There is related work that focuses on the verification problem of programs where the agent's knowledge is categorical and sensing is accurate [2]. Also, efforts have been made to verify programs where tests refer directly to the real world [3,4]. All of them rely to some degree on the hypothesis of full accessibility to the world. Yet, such a hypothesis no longer holds in the context of a belief program, where effectors and sensing could be noisy and the agent's knowledge is quantitative belief.

Before considering verification, formal semantics of the belief program are required, more concretely, one must design a probabilistic logic of belief that incorporates noisy sensors and stochastic actions. Perhaps the most successful work in doing this is the BHL model which combines probability theory and situation calculus [6]. This is the formalism adopted in the semantics of belief programs in [1]. However, everything in BHL's model and its variant are axiomized. Such axiomatic approaches would suffer when expressing temporal property.

For instance, temporal properties such as Eventually and Globally have to resort to second-order logic and the fix-point calculus. Based on possible world semantics, a modal variant of BHL, i.e. the logic DS, seems to be more suitable for our purpose by which the usual temporal operators G and F can be defined smoothly. Nevertheless, the DS has its disadvantages as well. For example, it lacks the expressiveness of specifying belief distribution and also it lacks the projection reasoning mechanism. Projection reasoning is a mechanism with which one can infer whether a given formula is entailed by an initial knowledge base. The task is accomplished by either translating the initial knowledge base to a knowledge base about the future (progression) or translating the formula about the future to another about the initial state (regression). Such a mechanism is the foundation of verification.

My research begins with lifting the expressiveness of the Logic DS and investigating projection therein. Afterwards, I will explore the verification of belief programs. So far, a variant formalism of BHL has been proposed and we are investigating verifying belief programs. The following section details the results achieved.

Reasoning about beliefs and beta-beliefs in an expressive probabilistic action logic. In this work, we overcome the downsides of the logic DS by first introducing the notation rigidity to ensure fixed interpretation on rigid terms. Among other things, this fixes the drawback of DS that truth of belief formula might rely on the real world which is counter-intuitive. Additionally, such a fixed interpretation enables us to express arbitrary mathematical probabilistic distribution as belief. In terms of regression, we propose a regress operator that works on not only beliefs about the world but also beliefs about beliefs, namely meta-beliefs. The paper arising from this work is under review.

Reasoning about beliefs by progression. Progression in general is much more difficult than regression since, as proved by [7], progression might not be first-order definable. In this work, we show that progression in the nullary fluent fragment of DS is first-order definable, roughly speaking, this amounts to assuming the system only has finite random variables. Subsequently, based on the notion of progressed distribution, we provide a semantic account of progression. We are preparing a submission for this work.

Verification of belief programs. In terms of verification, based on our improved logic, we provide a semantics of a belief program, where specification of property can be expressed by a logic similar to the probabilistic computational tree logic. As observed by [5], the decidability of GOLOG program verification depends on the underlying logic, the program constructs and the domain specifications. Roughly speaking, domain specifications are a set of axioms describing rules the world obeys. Perhaps the dimension of domain specification is less well-known compared to the others. We studied the decidability regarding domain specification. As it turns out, even under very restricted domain specification setting, the problem is still undecidable. Therefore, we restrict the properties needed to verify and investigate fragments where the verification problem is decidable. This work is ongoing.

References

1. Vaishak Belle, Hector J. Levesque. ALLEGRO: Belief-Based Programming in Stochastic Dynamical Domains. IJCAI 2015: 2762–2769, 2015.
2. Benjamin Zarrieß, Jens Claßen. Verification of Knowledge-Based Programs over Description Logic Actions. IJCAI 2015: 3278–3284, 2015.
3. Benjamin Zarrieß, Jens Claßen. Decidable Verification of Golog Programs over Non-Local Effect Actions. AAAI 2016: 1109–1115, 2016.
4. Jens Claßen, Benjamin Zarrieß.: Decidable Verification of Decision-Theoretic Golog. FroCoS 2017: 227–243, 2017.
5. Jens Claßen, Martin Liebenberg, Gerhard Lakemeyer, Benjamin Zarrieß. Exploring the Boundaries of Decidable Verification of Non-Terminating Golog Programs. AAAI 2014: 1012–1019, 2014.
6. Fahiem Bacchus, Joseph Y. Halpern, Hector J. Levesque. Reasoning about Noisy Sensors and Effectors in the Situation Calculus. Artif. Intell. 111(1-2): 171–208, 1999.
7. Fangzhen Lin, Raymond Reiter: How to Progress a Database. Artif. Intell. 92(1-2): 131–167, 1997.

Termination and Complexity Analysis of Probabilistic Programs

Dominik Meier (dominik.meier@rwth-aachen.de)
Supervisor: Prof. Dr. Jürgen Giesl

Introduction. 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 behaviour into efficient random algorithms which produce correct results with a high probability. The Rabin-Miller primality test, Freivalds' matrix multiplication, andthe random pivot selection in Hoare's quicksort algorithm are prime examples. These kinds of algorithms can be elegantly expressed as probabilistic programs. Determining runtimeand 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 with a finite number of expected steps.

While PAST is theoretically harder, AST is often considered more difficult to prove. 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.

Details. The project "Termination and Complexity Analysis of Probabilistic Programs" deals with the challenging question of how to automatically determine the respective properties of probabilistic programs. There are approaches amenable to automation, but current techniques usually only focus on programs on numbers and disregard programs operating on data structures such as lists or trees. In contrast, in the non-probabilistic case, many powerful approaches have been developed to analyze termination and complexity of term rewriting systems automatically. Therefore, the focus of the project is on analyzing probabilistic term rewriting systems (PTRSs). Due to better properties regarding modularity, AST will be considered as the main notion of termination.

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 adaptation of this analysis technique to the probabilistic case, in order to develop a fully automated technique for the termination analysis of PTRSs. Currently, in the project after having reviewed the state-of-the-art literature key challenges and requirements have been identified. Further, there has been significant progress in adapting the dependency pair technique to the probabilistic case.

Optimization under Adversarial Uncertainty

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

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.

Learning Definable Relations in Graphs

Martin Ritzert (ritzert@informatik.rwth-aachen.de)
Supervisor: Prof. Dr. Martin Grohe

Introduction. This work is embedded in the general setting of learning logical formulas over some background structure or background graph. We consider the algorithmic problem in which the background structure and a set of positive and negative examples are given and the learning algorithm has to return a consistent model consisting of a formula and a set of parameters. For a formula Φ and a set of parameters $\overline{v}$, the corresponding classifier accepts all those $\overline{u}$ such that the formula $\Phi(\overline{u}, \overline{v})$ is satisfied and rejects otherwise. The rationale behind this setting is explained in more detail in [1]. We generally are interested in the complexity of this learning problem for different types of graph (bounded degree, strings, trees, nowhere dense graphs) and logics (first-order logic, monadic second-order logic). All our results can easily be extended from graphs to general structures.

We found that on graphs of bounded degree, learning first-order formulas with bounded quantifier rank and number of variables is possible. The runtime of such an algorithm can be bounded in terms of the highest degree of a node in the graph and the restrictions on the formula. Since both are bounded by constants, the resulting algorithm runs in constant time [1]. Restricting the quantifier rank and number of variables in the retuned formula enforces that it cannot simply memorize the training set and the learning algorithm instead must identify a distinguishing factor between positive and negative examples.

On strings and trees it is easy to prove that linear time is necessary, but we showed that there is an algorithm that pre-processes the background structure (in linear time) and, then, can learn in sub-linear time[2][3]. Both results hold for learning first-order and monadic second-order formulas, but only for learning unary relations. Learning a unary relation corresponds to a classification task on the nodes of the graph, but not on (possible) edges or hyperedges.

Recently, we showed that, in general, the learning problem is always at least as hard as the corresponding model checking problem, which for full first-order logic implies that on the class of all graphs, the learning problem is (under common complexity theoretic assumptions) not in FPT and, thus, not tractable[4]. For nowhere dense graphs, one of the largest classes on which first-order model checking is in FPT, we presented a learning algorithm that works for higher-order learning tasks but needs to increase the number of parameters and the quantifier-rank of the returned formula.

[1] Martin Grohe, Martin Ritzert, "Learning first-order definable concepts over structures of small degree," LICS, 2017

[2] Martin Grohe, Christof Löding, Martin Ritzert, "Learning MSO-definable hypotheses on strings," ALT, 2017

[3] Emily Grienenberger, Martin Ritzert, "Learning Definable Hypotheses on Trees," ICDT, 2019

[4] Steffen van Bergerem, Martin Grohe, Martin Ritzert, "On the parameterized complexity of learning logic," under review at ICALP, 2021

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.

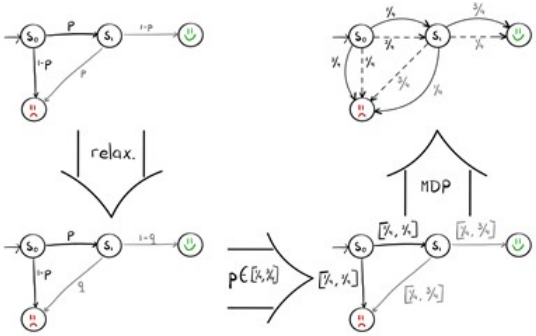
References

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.

Monotonicity in Parametric Markov Chains

Jip Spel (jip.spel@cs.rwth-aachen.de)
Supervisor: Prof. Dr. Joost-Pieter Katoen

In several kinds of systems probabilistic behaviour occurs. For instance unreliable or unpredictable behaviour in computer networks can be seen as probabilistic behaviour. Also, in a communication protocols, messages might not be received with a given probability, this yields a probabilistic state change.



Research has been done on formal methods for the specification and verification of probabilistic systems. Questions such as: "What is the probability that the file is transferred correctly if messages are lost with a probability 0.05?" could be analyzed through formal methods. One way to describe these probabilistic systems is through Markov chains. In a subset of these Markov chains all state changes are probabilistic and in discrete time.

However, the probabilities of these state changes are not always known in advance. Therefore, parametric Markov chains have been developed. They allow the use of parameters in the probabilities. For instance, in a biochemical reaction network, the rates of reactions might not be exactly known. In the past, they were then estimated. However, parametric Markov chains allow the analysis of them more precisely. Also in the case of transferring a file, the probability that a message is lost might not be known in advance. Instead of estimating this probability, we can now — based on the parametric Markov chain and a requirement, for instance "the probability that the file is transferred correctly should be at least 99%" — obtain parameter values for which the requirement holds.

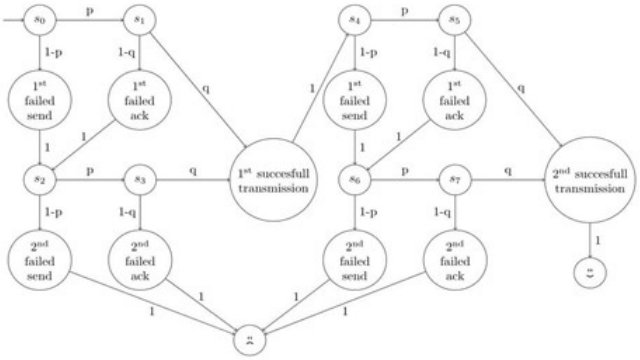


Figure 1: Parametric Markov chain

I want to investigate the effect of changing these parameter values on the probability that a requirement holds. In particular, parameters might have a monotone effect on the probability that a given system state is reached. I want to find this monotonicity in parameters and exploit this to improve the analysis on the behaviour of systems. During my Master's thesis I began work on this by providing a framework to determine monotonicity based on the probabilistic program describing a system.

My first two publications focus on finding monotonicity in parametric Markov chains and integrating this approach into existing techniques. The next goal is to extend the search for monotonicity to Markov decision processes, and possibly also to other Markov models. Furthermore, together with colleagues, I'm looking at other methods to improve the existing techniques.

Logics with Multiteam Semantics

Richard Wilke (richard.wilke@rwth-aachen.de)
Supervisor: Prof. Dr. Erich Grädel

Classical logics model queries structures such as graphs, groups, databases and so on. They are usually evaluated via so-called "Tarski-style" semantics. This means that during the evaluation process we use an assignment, i. e. a function mapping each free variable of the current sub-formula to its value, as an element of the target structure. There may be multiple assignments which one can use in this process, but there is no way of letting these communicate with each other, as the evaluation processes are independent of each other. In the game semantic sense, one can think of such an assignment as a play in the model-checking game, i.e. the path taken by both players when they choose which value a variable should be assigned in order to prove that the formula is satisfied, or not, depending on the player.

Classically, one is only interested in the question as to whether a winning strategy for one of the players exists, because this shows whether the given formula is satisfied by the structure at hand. By analysing these games one can find dependencies between variables that cannot be expressed in classical logics. For example, we could find out that one player can only win if she adapts her move depending on the previous move by her opponent. As mentioned said before, a move can be thought of as selecting a value for a variable. Historically, many attempts have been made to define logics that are able to speak about dependencies between variables. Such logics are sometimes called logics of imperfect information, and originated in the work of Henkin, Enderton and Walkoe, among others. Initially, the semantics of these logics were defined in a top-down manner, meaning that one cannot infer anything about the formula just from looking at its sub-formulae. One example is independence friendly logic (IF), as proposed by Hintikka and Sandu, where quantifiers are restricted in such a way that only the knowledge about certain variables can be used to determine next value. In the model-checking game this means a player must be able to choose a value without knowing which values the other player has chosen for certain variables. Another example are Henkin quantifiers, which can be interpreted as parallel classical quantifiers.

It was conjectured that the semantics of these logics cannot be defined in a compositional fashion, but in 1997 Hodges disproved this informal conjecture by providing a model-theoretic semantics for IF-logic in terms of what he called "trumps". Today, this semantics is called "team semantics" and it enabled Väänänen to propose a new logic called "dependence logic". The main idea is that dependencies are treated as atomic properties. Therefore, one has to collect all information about the variables, resulting in a set of assignments, in contrast to the classical case. Such a set is called a "team".

On the atomic level we can evaluate a dependence or independence statement by looking at the team arriving at it, hence providing bottom-up semantics. A quantifier no longer provides a single value to a variable, but rather, induces a set of values. Grädel and Väänänen introduced independence logic which, interestingly, corresponds precisely to the complexity class NP in terms of expressive power (over finite structures), i. e. existential second order logic. There have been many extensions of logics with team semantics, most of which share a single weakness: A

team is viewed as a set of assignments, hence ignoring the number of occurrences of each assignment. While this is appealing in many theoretical settings, it fails to give a good description of real-world scenarios. One can easily think of examples where not only the existence, but rather the number of assignments matters, e.g. in voting. The goal of this project is to augment dependence logics with the ability to handle multiplicities. Our approach is to consider multisets instead of set of assignments, resulting in logics with multiteam semantics. These will, for example, be able to express statistical dependencies between data. One can also think of a multiteam as a way to incorporate uncertainty in logic, since it allows us to make statements on which a portion of the (input) multiteam satisfies a certain criterion.

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)

Probabilistic Operating Concepts for Highly Automated and Autonomous Rail Vehicles in Rural Areas

Stephan Zieger (zieger@via.rwth-aachen.de)
Supervisor: Prof. Dr. Nils Nießen

Introduction. Mobility is an important component of various human activities, ensuring the access of citizens to exercise their social rights and the capacity to partake in productive activities. In an urban environment, with higher population densities and levels of economic activity, mobility drives economic development and contributes to social equity.

In large parts of Europe, it can be observed that the population is drawn to metropolitan areas. As a result, the population in the rural areas declines which in turn leads to a closure of several rural railway lines as regular services became unprofitable.

Current technology, i.e., highly automated vehicles, modern communication technology as well as intelligent vehicles aware of their environment, and flexible transportation concepts such as Demand-Responsive Transport can help to overcome current challenges and enable rural local rail transport again. Therefore, an on-demand rail service with small highly automated vehicles is proposed and its feasibility is researched.

Mathematical Programming Approach. A mixed-integer program is derived from existing approaches to formulate general scheduling tasks and Dial-a-Ride problems. Modelling railway operations carries several pitfalls, e.g., the inclusion of technical minimum headway times which are especially relevant on single-track sections. The modelling suffers greatly from its large number of interdependent constraints and variables. Although approaches for drastically reducing the number of constraints were implemented, the problem can only be solved for small instances due to the strongly increasing number of variables and constraints.

Simulation Approach. Another possible means to investigate the feasibility and behaviour of such an on-demand railway service is simulation. All passenger requests are gathered centrally and then the vehicles act on these requests according to some strategy profile, e.g., the closest vehicle in terms of distance serves the request, if possible. Several such strategies will be simulated, compared and ideally the gap to an optimal behaviour can be investigated.

GRK 2428: ConVeY — Continous Verification of Cyber-Physical Systems

Prof. Dr. Helmut Seidl
Email: seidl@in.tum.de
TU München and Ludwig Maximilian Universität München
Internet: https://convey.in.tum.de/

Networks, computers, sensors, and actuators are being increasingly integrated into cyber-physical systems, i.e., 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 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, we propose to

- develop verification and synthesis technology for robust system design, i.e., for the design of systems that maintain correct behavior under change
- develop verification and synthesis technology able to cope with frequent or even continuous change in the specification and the environment.

Areas of Research

Robust System Design. We will develop techniques to 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 will investigate the design of controllers that are robust by construction against those changes.

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

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

Formal Analysis of Large-Scale Stochastic Systems against Temporal Logic (Hyper)Properties

Mahathi Anand (mahathi.anand@lmu.de)
Supervisor: Prof. Dr. Majid Zamani

Cyber-physical systems (CPS), i.e., systems with interacting physical and software components, have achieved significant attentions in the past two decades. They model many applications such as power grids, air traffic networks, medical equipment, etc., and are often required to perform complex logic tasks. Examples of such tasks include those expressed as linear temporal logic or (in)finite strings over automata. Due to the large system size and the presence of random disturbances and uncertainities, the development and verification of safe and secure CPSs is a challenging problem. Therefore, formal analysis of large-scale stochastic control systems against temporal logic specifications has received significant attentions in past few years. Traditionally, such systems have been analyzed using discretization-based methods [1]. These approaches suffer from the curse of dimensionality since the computational complexity grows exponentially with the number of state variables. More recently, discretization-free techniques using barrier certificates [2] have been developed to potentially alleviate this computational burden. However, the search for suitable barrier certificates is still difficult for large-scale systems. Our research mainly aims at handling these limitations by proposing a compositional framework to construct control barrier certificates (CBCs) for large-scale stochastic control systems. Utilizing CBCs, the goal is to synthesize hybrid controllers enforcing specifications that are expressed by automata over (in)finite time horizons, while providing a (potentially tight) lower bound on the probability that the system satisfies the given specifications. Though the specifications considered describe a wide range of linear-time properties that take into account individual trajectories of the system such as safety, liveness, etc., many important security properties like opacity and non-interference require quantifying the relationship between multiple trajectories and cannot be expressed by standard temporal logic. Such properties are called as hyperproperties [3] and have received significant attentions in the past few years. However, the techniques presented in literature are suitable for finite-state systems and are not applicable to real-world CPSs that evolve over continuous state sets. Therefore, our research also investigates the verification of stochastic control systems against hyperproperties via barrier certificates.

[1] A. Lavaei, S. Soudjani, and M. Zamani, "Compositional (in)finite abstractions for large-scale interconnected stochastic systems," IEEE Transactions on Automatic Control, vol. 65, no. 12, p. 5280–5295,2020.

[2] S. Prajna, A. Jadbabaie, and G. J. Pappas, "A framework for worst-case and stochastic safety verification using barrier certificates", IEEE Transactions on Automatic Control, vol. 52, no. 8, p. 1415–1428,2007

[3] M.R.Clarkson, B. Finkbeiner, M. Koleini, K.K. Micinski, M.N. Rabe. and C. Sánchez, "Temporal logics for hyperproperties", Principles of Security and Trust, Lecture Notes in Computer Science, p. 265-284, 2014

Study of Weak Models of Distributed Computing

Philipp Czerner (czerner@in.tum.de)
Supervisor: Prof. Dr. Javier Esparza

Many natural or artificial distributed systems, such as molecules, cells, microorganisms or nano-robots, consist of parts with limited computational capacities. These parts (called agents) can, for example, store only a small amount of information, have no identities, and interact stochastically. Various weak models of distributed computing have already been researched intensely in the literature, such as population protocols or chemical reaction networks.

The goal of this area of research is to find efficient protocols to perform distributed computations in these models and analyse their characteristics. Additionally, we want to develop automated procedures which can prove properties of specific protocols, such as their correctness or running-time.

I focus both on extending the existing theoretical knowledge on models such as population protocols, and on considering variants of known models.

Population protocols are a weak model of distributed computing, where agents have only finitely many states. They interact pairwise and stochastically – an agent has no knowledge about the global state of the population. Despite these limitations, they can compute global properties of the initial configuration, e.g. whether initially more "red" than "blue" agents exist in the population.

A natural topic of inquiry is the succinctness of population protocols: how many states does a protocol need to implement certain properties? Here, we showed the first elementary lower bound, by proving that protocols for properties of the form $x \geq k$, where k is a constant and x the number of agents in the population, have at least $\Omega(\log \log \log k)$ states.

Continuing a line of research investigating extensions to population protocols, where additional types of transitions are added, we consider broadcast consensus protocols (BCPs). Here we add an atomic broadcast to the model, which allows an agent to send a signal to the rest of the population. We characterise the set of predicates which BCPs can decide efficiently (i.e. in polynomial time). Based on these theoretical results, it is possible to implement efficient protocols in this model, or rule out that certain tasks can be done within limited time.

Chemical reactions are often assumed to be well-stirred, meaning that an agent does not have a fixed location and could interact with any other agents. For other, e.g. biological, systems this assumption does not hold, however, and an agent can communicate only with its neighbours in some fixed structure (a communication graph). We investigate protocols operating on these structures, where an agents perceives only the states of its neighbours. Based on a previous classification, we determine the expressive power for a number of classes. This enables comparisons between different models, and helps answering basic questions such as: "How important is randomness in this model?" or "How much information do agents need about their neighbours to decide certain properties?"

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 practical problem for static analysis employing 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.

Ouadjaout and Miné[1] propose 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 that the abstract interpreter could not provide itself, it requires manual annotation, which is usually costly and error-prone.

We aim at providing an efficient analysis for library code which shall deliver 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 aim to integrate the analysis into the Goblint static analyzer and evaluate its efficiency on implementations of the C standard library, such as musl.

[1] Abdelraouf Ouadjaout and Antoine Miné, "A library modeling language for the static analysis of C programs", vol. 12389 of Lecture Notes in Computer Science, p. 223 - 247, 2020

Synthesizing Controllers With Guarantees

Kush Grover (grover@in.tum.de)
Supervisor: Prof. Dr. Jan Křetínský

My research aims to develop techniques for synthesizing controllers with guarantees for safety critical systems. A safety-critical system is a system whose failure or malfunction may result in death/serious injury to people or loss of a lot of money. That is why it is necessary to prove correctness of such systems.

Markov decision processes (MDP) are widely used formalism for modeling non-deterministic and probabilistic behaviors of systems like a robot, warehouse storage management etc. Usually, for any continuous system, a model is generated by some sort of abstraction which discretizes the actual continuous space. Although, there exist sound analysis techniques which gives guarantees on the discrete models, these abstractions can be a source of errors in the final result. Hence, to tackle this problem, we have developed an algorithm to solve the rechability problem in a continuous MDP directly while preserving the guarantees[1]. This algorithm also generates an optimal controller for which the error is bounded by some given precision.

In contrast to this, I also worked on synthesizing controllers for discrete MDPs which satisfies some specification. We modeled a robotic arm using a discrete MDP which only incorporates the high level tasks the arm can perform. We use PRISM to generate the controller and dtControl to store and use it efficiently. So far, the planners that roboticists use generate a sequence of actions to be executed next but that plan may become invalid because of some changes or faults in the environment and the planner will have to do the planning again. The controller that we synthesize is "universal" which only depends on the current state of the system. Hence, it is much faster for the robot to find what to do next and have a fail-safe mechanism as well. This approach can be extended to work with more complicated models and also to find the best strategy w.r.t some reward structure.

Motion planning is the problem of finding a path (usually for a robot) from some starting point which satisfies some specification. There exists algorithms which does this quite efficiently. People have also wondered how can we solve this problem in an unknown environment i.e. the robot does not know the map of the environment, instead it can sense things within a certain radius. In this case, the robot has to explore the environment while searching for the path. We gave an algorithm to find a path which satisfies some specification in an unknown environment. We compare our approach to "first explore, then plan" algorithm on 100 randomly generated environments and our approach performs significantly better. This work can be extended by also considering the dynamics of the robot.

[1] Kush Grover, Jan Křetínský, Tobias Meggendorfer, and Maximilian Weininger. An anytime algorithm for reachability on uncountable mdp, arXiv:2008.04824, 2020.

Population Protocols and Chemical Reaction Networks

Martin Helfrich (helfrich@in.tum.de)
Supervisor: Prof. Esparza

Population protocols are a model of distributed computation where a constant but unknown number of finite-state agents interact to decide a property. For example, in a majority protocol, there are initially agents that vote for "yes" and agents that vote for "no". The agents need to decide if there is a majority for "yes" by interacting in pairs. All agents follow the same protocol that determines how two agents in a rendez-vous interaction change their state. By interacting in a stochastic manner, the agents need to stabilize to the correct consensus in order to answer the property in question for every possible number of agents.

Population protocols are widely studied in the distributed computation community. Research areas are for example their computational power, their computation speed, their succinctness as well as automatic verification and synthesis procedures. Extensions of the model like broadcasts or are also investigated. Another interesting and related model are graph automata, where the communication graph of the agents is not complete, like in population protocols, but arbitrary.

While population protocols are a theoretical model, the closely-related model of chemical reaction networks has more practical applications such as modelling and analysis of biochemical systems, high-level programming of molecular devices and synthetic biology. In a chemical reaction network, molecules interact in reactions with different speeds that correspond to actual chemical reactions.

An important research topic is the efficient analysis of these complex and possibly infinite-state systems to accurately predict the evolution of a mixture of molecules without the need for an expensive and potentially dangerous wet lab. Because chemical reaction networks have an infinite state space, abstractions are used to make the system more tractable while preserving its global behaviour. To allow chemists to answer high-level questions like "Is DNA produced after 3 days?" or "Does the system oscillate between acidic and alkaline?", a specification language for model checking needs to be developed. This framework could also help to find unknown reactions in chemical reactions or even allow the synthesis of systems with desired properties.

Formal Synthesis of Controllers for Interconnected Stochastic Control Systems with Partial Information

Niloofar Jahanshahi (niloofar.jahanshahi@lmu.de)
Supervisor: Prof. Dr. Majid Zamani

This research is motivated by the challenges arising in the synthesis of controllers for stochastic systems enforcing complex logic specifications. Stochastic control systems are becoming ubiquitous and an integral part of our daily lives. Examples of such systems range from robots and medical devices to smart grids and automative networks. In many real-world applications, these systems are expected to do complex logic tasks. Such tasks can usually be expressed using temporal logic formulae or as (in)finite strings over finite automata. For this reason, formal synthesis of controllers enforcing complex logic specifications has attracted significant attentions from both academic and industrial communities. In the past few years, abstraction-based techniques have been very promising for formal synthesis of controllers for stochastic control systems [1]. Since these techniques are based on discretization of state and input sets, when dealing with large-scale systems, unfortunately, they suffer severely from the curse of dimensionality (i.e., the computational complexity grows exponentially with the dimension of the state set). In order to overcome the large computational burden, a discretization-free approach, based on control barrier functions [2] has shown potential to solve the formal synthesis problems. In our research, we provide a systematic approach to synthesize a hybrid control policy for stochastic control systems without discretizing the state sets. Furthermore, since in many real-world applications access to full state information is not available, this research is considering partially-observed stochastic control systems. Our goal is to utilize the notion of control barrier functions to synthesize control policies providing (potentially maximizing) a lower bound on the probability that the trajectories of the partially observed stochastic system satisfy some complex specifications (usually expressed by temporal logic formulae[3]). Though synthesis of controllers for lower-dimensional systems is challenging itself, the task is much more computationally expensive (if not impossible) for large-scale interconnected system. Driven by this challenge, in this research we further extend our results for networks of partially-observed stochastic control systems by proposing a compositionality approach for the construction of control barrier functions, resulting in a compositional controller synthesis scheme.

[1] A. Lavaei, S. Soudjani, and M. Zamani, "Compositional (in)finite abstractions for large-scale interconnected stochastic systems," IEEE Transactions on Automatic Control, vol. 65, no. 12, pp. 5280–5295,2020.

[2] Prajna S and Rantzer A, "On the necessity of barrier certificates," IFAC Proceedings Volumes, vol. 38, no. 1, pp. 526–531, 2005.

[3] Baier C and Katoen JP,"Principles of model checking," The MIT Press, 2008.

Provable Safe Reinforcement Learning for Motion Planning of Autonomous Vehicles

Hanna Krasowski (hanna.krasowski@tum.de)
Supervisor: Prof. Dr. Ing. Matthias Althoff

The recent development of motion planning techniques for autonomous vehicles has become more data driven due to the advance in computation power and the increasing amount of available traffic data. Compared to rule-based methods, data-driven learning approaches require much less expert knowledge based on their ability to ascertain complex dependencies from data. Motion planning tasks can be modeled as Markov decision processes, for which reinforcement learning offers solutions.

For real-world motion planning tasks of autonomous cars, vessels, drones or other mobile robots, it is desirable to provide guarantees that ensure specified safety bounds for the system behavior. As the reinforcement learning agent usually explores at random, unsafe actions are possibly executed and this impedes the applicability for real-world tasks. Thus, safe reinforcement learning researchers have recently begun to develop methods to decrease the randomness in reinforcement learning such that less or no unsafe states are explored. Still only some of the safe reinforcement learning research can prove safety during learning and deployment.

To achieve provable safety for reinforcement learning, our approach constrains the action space to the subspace of safe actions by determining safe actions with formal methods. For motion planning of autonomous vehicles on highways, we implemented reinforcement learning for a discrete action space while the safe actions are verified through set-based prediction of the surrounding traffic participants. We showed that our approach is provably safe during training and deployment. Furthermore, the safety layer leads to a good agent performance from the beginning of the training as policies causing collisions instead of reaching the goal are not explored.

In the future, we will transfer our concept to other motion planning applications and extend it with interaction rules. The goal is to generalize the concepts in a framework which can handle all types of action spaces, additional interaction rules, and different types of learning algorithms for different motion planning tasks while preserving verifiability. Finally, this framework should be easily adaptable for motion planning application in order to facilitate research in provable safe reinforcement learning.

Verification of Quantum Resitant Cryptography

Katharina Kreuzer (k.kreuzer@tum.de)
Supervisor: Prof. Tobias Nipkow

As quantum computers come into reach more and more, the threat of their implications becomes more imminent. With quantum computers, most of the existing crypto algorithms like RSA, Diffie-Hellmann and elliptic curve Diffie-Hellmann algorithms could be broken easily. Developing quantum resistant cryptography – and verifying it – is a major task of modern research. For the long term goal of my research, the focus will lie on verifying such post-quantum crypto algorithms, in particular lattice-based algorithms, using the proof assistant Isabelle. In order to get familiar with Isabelle, the short term goal is to formalize a proof of the Van der Waerden Theorem. This theorem from combinatorics states that for integers k and l there exists a number N which guarantees that if the integers up to N are coloured with k colours, there will always be an arithmetic progression of length l of the same colour. This result has not yet been formalised. Indeed, finding the smallest number such that the Van der Waerden Theorem holds is still an open problem for most integers k and l.

Neural Network Abstraction for Accelerating Verification

Stefanie Mohr (mohr@in.tum.de)
Supervisor: Prof. Dr. Jan Krestinsky

Neural Networks (NN) are successfully used to solve many hard problems reasonably well in practice. However, there is an increasing desire to use them also in safety-critical settings, such as perception in autonomous cars, where reliability has to be on a very high level and that level has to be guaranteed, preferably by a rigorous proof. This is a great challenge, in particular, since NN are naturally very susceptible to adversarial attacks, as many works have demonstrated in the recent years [1]. Consequently, various verification techniques for NN are being developed these days. Most verification techniques focus on proving robustness of the neural networks, i.e. for a classification task, when the input is perturbed by a small ε, the resulting output should be labeled the same as the output of the original input. Unfortunately, verification tools struggle to scale when faced with real-world neural networks. Reducing the size of a NN by abstraction leads to several possibilities. Firstly, since the abstracted NN is smaller, it may be preferred in practice because generally smaller networks are often more robust, smoother, and obviously less resource-demanding to run. Note that there is a large body of work on distilling smaller NN from larger ones, e.g. re naturally very susceptible to adversarial attacks, as many works have demonstrated in the recent years [2], i.e. training a smaller NN based on the output of a bigger one. Secondly, and more interestingly in the safety-critical context, we can use the smaller abstract NN to obtain a guaranteed solution (robust or satisfying other properties) to the original problem: We can analyze the abstract NN more easily as it is smaller and then transfer the results to the original one, provided the differences are small enough.

We already developed an abstraction framework for NN. In contrast to syntactic similarities, such as having similar weights on the edges from the previous layer [3], our aim is to provide a behavioral, semantic notion of similarity, such as those of predicate abstraction, since such notions are more powerful. In future, we want to extend the tool for application to more complex settings and create a CEGAR-loop.

[1] Akhtar, Naveed and Mian, Ajmal, "Threat of adversarial attacks on deep learning in computer vision: A survey," IEEE, 2018

[2] Hinton, et al., "Distilling the Knowledge in a Neural Network," 2015

[3] Guoqiang Zhong et al., "Merging Neurons for Structure Compression of Deep Networks," ICPR, 2018

A Verified Proof Checker for Isabelle

Simon Roßkopf (rosskops@in.tum.de)
Supervisor: Tobias Nipkow

The results obtained by using an interactive theorem prover are only as trustworthy as the used system itself. The proof assistant Isabelle aims to achieve such trust by following the LCF-principle(theorems can only be

created by inference rules). However, its infrastructure is optimized for performance(for example supporting multithreading), complicating the system and making it harder to trust or verify.

We hope to further increase trust, not by verifying the (complicated) system itself, but by checking its generated proofs using a formally verified, external tool instead. Berghofer and Nipkow[1] have already extended Isabelle with functionality to export proof terms, which our tool can then check.

For our verification, we need a notion of what constitutes a correct proof. For this we have developed the first complete formalization of Isabelle's metalogic. Then we defined a (executable) proof checker and verified it with respect to our formalization of the meta logic. All this work was done in Isabelle/HOL and ML code was extracted using code generation. Lastly, we integrated the checker with the Isabelle system and used it to check some existing proofs, most importantly about 12% of the most commonly used HOL `Main` library. More details can be found in our paper[2] .

In the future we plan to scale up the checker to handle more and larger proofs, for example present in the Isabelle distribution or the Archive of Formal Proofs[3]. Therefore it is necessary to optimize the performance of our tool. For larger formalizations we also anticipate the size of the proof terms in memory becoming a problem. Isabelle can already emit compressed proof terms, to which we can adapt our formalization. Lastly, erroneous definitions make proofs based on them meaningless. Our tool can be extended to ensure for example the absence of circular dependencies in type or term definitions.

[1] S. Berghofer, T. Nipkow, "Proof terms for simply typed higher order logic," In J. Harrison, M. Aagaard, editors, Theorem Proving in Higher Order Logics, 13th International Conference, TPHOLs 2000, LNCS 1869 p. 38–52, 2000

[2] T. Nipkow, S. Roßkopf. "Isabelle's Metalogic: Formalization and Proof Checker", In Platzer, A., Sutcliffe, G. (eds.) 28th International Conference on Automated Deduction (CADE-28), LNCS, Springer, 2021

[3] https://www.isa-afp.org/

Thread-Modular Abstract Interpretation for Multi-Threaded Code

Michael Schwarz (m.schwarz@tum.de)
Supervisor: Prof. Dr. Helmut Seidl

Larger software systems tend to be multi-threaded where their correctness depends on the possible ways in which different threads can interact with each other. In particular, correctness may depend on the set of possible values of global variables.

However, analyzing all possible interleavings of different threads of larger programs is expensive in analysis time — in some cases even prohibitively so. Ideally, such analyses should be thread-modular, implying that their complexity does not increase exponentially with the number of threads.

As a reference semantics, we rely on a local trace semantics that is formulated by means of side-effecting constraint systems.[1] Local here means that each thread has only a local view of the system, i.e., it only knows things about its own past and those actions of different threads that are observable by it, but not about other, non-observable, actions of different threads.

Based on this setting, we will provide thread-modular non-relational value analyses and show that a generalization of the analysis provided by the static analyzer Goblint[2] as well as a natural improvement of Antoine Miné's approach [3] can be obtained as instances of this general scheme.

We will also investigate how to extend the semantics beyond the notion of local to express information about program points possibly-running-in-parallel. Such an extension is, e.g., convenient to reason about other synchronization primitives such as signaling and waiting that are used in multi-threaded programs in addition to mutexes.

Based on this semantic foundation, we will then design new thread-modular analyses such as relational analyses of the values of global variables and analyses of signaling and waiting in multi-threaded programs.

[1] Apinis K., Seidl H., Vojdani V., "Side-Effecting Constraint Systems: A Swiss Army Knife for Program Analysis.", APLAS, vol. 7705, p. 157-172, 2012

[2] https://goblint.in.tum.de/

[3] Miné A., "Static Analysis of Run-Time Errors in Embedded Real-Time Parallel C Programs", LMCS, vol. 8, 2012

Incremental Automatic Software Verification

Martin Spiessl (spiessl@sosy.ifi.lmu.de)
Supervisor: Prof. Dr. Dirk Beyer

Automatic Software Verification has become more and more powerful over the recent years, but still there are easy ways to generate verification tasks that cannot be solved by any of the currenty state-of-the-art tools. One of the reasons for this is that different analyses often have orthogonal weaknesses, and specific combination of techniques would be needed to proof a certain program correct. This lead to the development of Conditional Model Checking [1] [2], which we try to improve upon by further increasing the ways via which different tools, approaches, and the users can interact with each other.

One obvious way is to leverage the information exchange of invariants that are contained in the verification witnesses. Currently the main purpose of these witnesses is to validate the results of verification, and their usefulness in exchange between tools is limited.

As a first step we enable verifiers to directly reuse this information by encoding the information in the witnesses into a new verification problem that is potentially easier to solve. [3].

To better understand the information that is really important for a particular verification approach, a next step is to enable automatic verifiers to be used like interactive verifiers, i.e., provide easy ways for the users to add annotations and proof hints that can be transparently translated into verification tasks. Of course one can also use the information generated by the verifiers to automatically generate annotations. This can make the verification results more clear to the user, and help tool developers improve the quality of the exported information. For example, currently there is no way to make quantified invariants available in the verification witnesses, and more features like this might be revealed as necessary to further improve the state-of-the-art.

Lastly we envision a way for a precision-based parametric analysis that can choose between different verification approaches either automatically via CEGAR or interactively via user-provided annotations (very similar to how interactive proof assistants work). The goal is to use the insights gained in the previous steps to create new ways of designing powerful analyses that can apply working strategies for different subproblems in a larger verification task.

[1] D. Beyer and T. A. Henzinger and M. E. Keremoglu and P. Wendler, Conditional Model Checking: A Technique to Pass Information between Verifiers, Proc. FSE 2012, article no. 57, https://doi.org/10.1145/2393596.2393664

[2] D. Beyer and M.-C. Jakobs and T. Lemberger and H. Wehrheim, Reducer-Based Construction of Conditional Verifiers, Proc. ICSE 2018, pp. 1182-1193, https://doi.org/10.1145/3180155.3180259

[3] D. Beyer and M. Spiessl, Witness Validation via Verification, Proc. CAV 2020, pp. 165-177,https://doi.org/10.1007/978-3-030-53291-8_10

A Store for Software Invariants

Nico Weise (nico.weise@sosy.ifi.lmu.de)
Supervisor: Prof. Dr. Dirk Beyer

We consider program invariants as first-class citizens in software verification. Exchanging invariants as intermediate or final results is a crucial step towards truly incremental and modular software verification.

The value of reusable verification results is already recognized by the community of the competition on software verification (SV-COMP)[1]: Participating tools receive score points only if, in addition to a correct result, they produce a correctness witness that can be confirmed by a result validator. However, the current format for correctness witnesses[2] is not flexible enough for our use case. Also, the software-verification community desires a new witness format[3]. Therefore, we develop a new format for storing and exchanging invariants.

In our store, an invariant record contains not only an invariant, and references to the program, line number, and offset, but also meta data such as producer, time stamp, and specification. The idea is that verifiers add invariants to the invariant store, and use invariants from the invariant store in order to cooperate during the verification process. A verifier that uses an invariant from the store and has proved its correctness, will sign off the invariant by adding a new entry to the invariant's record and store it back into the invariant store. Each "signing off" will increase the trust that the invariant is correct.

We plan to propose the new format to be used in SV-COMP. With the large benchmark set from SV-COMP and tools supporting our format, we can develop a portfolio solver that uses the invariant store as central data structure for exchange. Our approach can then itself be evaluated in the SV-COMP setting against state-of-the-art verifiers.

With the results we want to understand requirements for useful invariants, limitations and potential for improvement of invariant generators and consumers, and applications of our approach to incremental software verification.

[1] Dirk Beyer, "Software Verification: 10th Comparative Evaluation (SV-COMP 2021)," Proc. TACAS 2021, p. 401–422, 2021

[2] Beyer, D., Dangl, M., Dietsch, D., Heizmann, M., "Correctness witnesses: Exchanging verification results between verifiers," Proc. FSE 2016, p. 326-–337, 2016

[3] https://github.com/sosy-lab/sv-witnesses/issues/17

Adaptive Reachability Analysis: Near-Optimal Effortless Safety Verification

Mark Wetzlinger (m.wetzlinger@tum.de)
Supervisor: Prof. Matthias Althoff

Applying cyber-physical systems in safety-critical environments requires formal verification techniques to ensure correct functionality. A contemporary example is the launch of a rover to another planet, where even small failures are critical as they might lead to severe consequences. One of the main techniques to provide safety guarantees is reachability analysis where all possible system behaviors over time are computed under the influence of uncertainty in the initial state and external input or disturbances. If the reachable set does not intersect an unsafe set determined by unwanted system behavior, safety is formally guaranteed.

In general, reachable sets cannot be computed exactly. The tightness of the reachable sets as well as the computational efficiency of the reachability algorithm heavily depends on the tuning of algorithm parameters, such as the time step size or the accuracy of the set representation. This entails the main research question: How can we measure the tightness of the reachable sets and tune the algorithm parameters accordingly? In the literature, a wide variety of reachability algorithms have been proposed, but many have to be manually tuned to yield good results. This requires expert knowledge about the intricacies of the algorithm, thereby hindering widespread application in practice and even among researchers.

Our research aims to develop methods to adaptively tune the algorithm parameters for reachability analysis of linear and nonlinear continuous systems. The effects contributing to the over-approximation of the reachable sets are analyzed and the induced errors are measured in an (over-)approximative way. This information is used to enhance the underlying reachability algorithm by an automated parameter tuning approach ensuring tight reachable sets.

For linear systems, we have proven that any prescribed error bound can be satisfied by a novel adaptive parameter tuning approach. This allows any practitioner to effortlessly yield reachable sets of desired accuracy. For nonlinear systems, we have devised an algorithm which balances the main sources of over-approximation by a formulation as an optimization problem. In both cases, the performance was evaluated on benchmark systems and shows to be competitive compared to expert manual tuning in both accuracy and computation time.

Future research questions will aim to increase the efficiency while still enabling the computation of tight reachable sets. Investigated methods include the abstraction of dynamics either by decomposition or order reduction methods, which both aim to find a simpler description of the original system which is easier to compute and therefore greatly enhances the scope of system for which the reachable sets can be efficiently computed.

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

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.

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). Our goal is to slowly but systematically work towards establishing new methodological standards in handling digital evidence, intepreting and developing national and international law in the years to come. At the same time, we attempt to (at least partially) remedy the lack of scientifically trained experts in this area.

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.

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.

Cryptocurrency Anonymity

Dominic Deuber (dominic.deuber@fau.de)
Supervisor: Prof. Dr. Dominique Schröder

Cryptocurrencies are digital means of payments that are based on the blockchain technology and cryptographic primitives such as digital signatures. In contrast to traditional currencies, cryptocurrencies do neither require a central bank to issue new units nor a central point to monitor transactions. These unique properties are the reason why cryptocurrencies increasingly change how payments are made worldwide.

In most cryptocurrencies, transactions use public keys as part of a digital signature scheme to specify senders and recipients of the payments. A person can generate an arbitrary number of public keys on the fly and the keys themselves do not reveal the identity of the person. Therefore cryptocurrencies working like this are often mistakenly considered anonymous. However, multiple public keys belonging to the same person can be grouped by linking heuristics.[1] For this reason, such cryptocurrencies only achieve pseudonymity and are thus non-privacy-preserving. However, two main techniques have been developed to realize anonymity. On the one hand, there are so-called overlays[2] that can be used on top of non-privacy-preserving cryptocurrencies to add anonymity by complicating linkage. On the other hand, there are privacy-preserving cryptocurrencies aiming for anonymity by design. The three largest privacy-preserving cryptocurrencies by market capitalization are Monero, Zcash, and Dash. While the privacy measures of Monero and Zcash have been extensively studied,[3] Dash has not yet been subject to analyses. Therefore the first part of my work is to understand and formalize Dash.

Cryptocurrencies, especially the aforementioned Monero, Zcash and Dash are more and more used by criminals[4] and thus gain the attention of law enforcement agencies. While the results of deanonymization attacks might be sufficient to start investigations, it is not yet clear what their meaning in a criminal trial might be. The reason is that deanonymization attacks are based on heuristics and thus might lead to false positives. This may raise problems given the standard of evidence required to find a defendant guilty. Thus, the second part of my work is to study how results based on those heuristics can be used in criminal procedures, especially how they should be interpreted.

[1] D. Ron and A. Shamir. Quantitative analysis of the full Bitcoin transaction graph. In FC 2013, pages 6–24, 2013. S. Meiklejohn et al. A fistful of Bitcoins: characterizing payments among men with no names. In Internet Measurement Conference, pages 127–140, 2013.

[2] S. Meiklejohn and C. Orlandi. Privacy-enhancing overlays in bitcoin. In FC 2015 Workshops, pages 127–141, 2015.

[3] M. Möser et al. An empirical analysis of traceability in the Monero blockchain. PoPETs, 2018(3):143–163, 2018. G. Kappos et al. An empirical analysis of anonymity in Zcash. In USENIX Security 2018, pages 463–477, 2018.

[4] G. Tziakouris. Cryptocurrencies——a forensic challenge or opportunity for law enforcement? An Interpol perspective. IEEE Security and Privacy, 16(4):92–94, 2018.

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.

Spectra of Behavioural Semantics via Graded Monads

Chase Ford (chase.ford@fau.de)
Supervisor: Prof. Dr. Lutz Schröder

The behaviour of concurrent processes is a topic of central importance in computer science. Time has shown that the sea of behavioural semantics arising in practice is vast, with variations in the system (labelled/probabilistic transition systems) and semantic (equivalence, preorder, metric) types. Each choice of these parameters further induces a spectrum of concrete behavioural semantics, e.g., the linear time–branching time spectrum[1]. Fitting these orthogonal dimensions into a common framework is a non-trivial task with prospects in both theory and practice.

To this end, Milius et al.[2] proposed a framework of graded semantics where behavioural equivalences are abstracted as graded monads with an embedding into the coalgebraic system type[3]. A strength of this framework is that it permits the extraction of characteristic modal logics[4]. The goal of this thesis is to contribute to the theory of (systems of) graded semantics and their associated graded logics.

A top priority is to extend graded semantics to semantic types beyond equivalences. We have established the theory of graded behavioural preorders, where the key ingredient was a characterisation of graded monads on Pos in terms of a notion of graded ordered algebraic theories introduced here. A further goal is the treatment of graded behavioural metrics. The development of a notion of "bisimulation game" for graded semantics is a further ambition. Finally, we aim to supply a generic axiomatic treatment of graded logics, and to develop a suitable notion of graded fixpoint logic.

As an application of this framework, we propose a definition of digital evidence based on modelling user-machine interactions as coalgebras under (systems of) graded semantics. In this direction, we aim to develop an implementable algorithm for coalgebra learning under a given graded semantics with the ambition of developing a novel approach to learning user-specific behaviour within a variety of cybercriminal contexts.

[1] R. van Glabeek, "The linear time–branching time spectrum I; the semantics of concrete sequential processes", Handbook of Process Algebra, p. 3-99, 2001.

[2] S. Milius, D. Pattinson, and L. Schröder, "Generic trace semantics and graded monads", Proceedings of the 6^{th} Conference on Algebra and Coalgebra in Computer Science (CALCO 2015), p. 253–269, 2015.

[3] J. Rutten, "Universal coalgebra: A theory of systems", Theor. Comput. Sci., p. 249:3–80, 2000.

[4] U. Dorsch, S. Milius, and L. Schröder, "Graded Monads and Graded Logics for the Linear Time–Branching Time Spectrum", Proceedings of the 30^{th} International Conference on Concurrency Theory (CONCUR19), p. 36:1–36:16, 2019.

Logic and Argumentation in Social Media

Merlin Göttlinger (merlin.goettlinger@fau.de)
Supervisor: Prof. Dr. Lutz Schröder

Social media hosts vast amounts of discussions about any topic. This abundance of data offers valuable information hidden from the human reader, due to its sheer amount and ranking algorithms. In this setting, short sentences and abbreviations are common or even enforced by character limits. This leads to argumentation being further obscured by incomplete arguments, where large parts of the argumentation are left to be filled from common knowledge.

We approach representing these arguments via a bespoke formalism that incorporates aspects from inference anchoring theory[1], the argument interchange format[2], and other structured argumentation frameworks. We propose a multi-layered graph where the main layers represent a dialogue game, and structured arguments, respectively. The argument layer has nodes annotated with formulæ in a suitable modal logic connected with conflict and deduction hyper-edges possibly labelled with argument schemes.

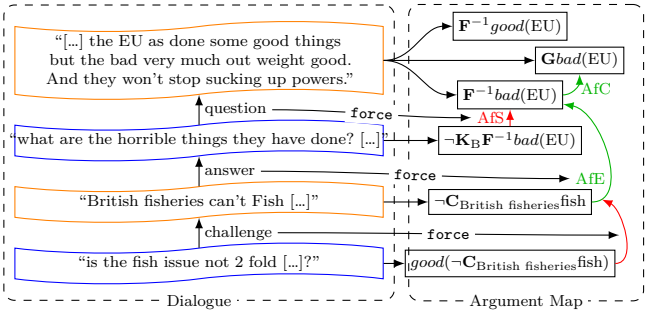


Fig. 1: Exemplary Twitter conversation about the EU. Af{C, E, S} = Argument from {Commitment, Evidence, Source}. Arguments are represented as modal formulæ, in this example featuring modalities $\mathbf{F}^{-1}$ at some past point, $\mathbf{G}$ always in the future, $\mathbf{K}_a$ agent a knows, and $\mathbf{C}_a$ agent a can.

We extract the formulæ from the written arguments via linguistic queries similar to regular expressions[3]. The framework of coalgebraic modal logic[4] enables recovering implicit deductions and conflicts from our annotated modal formulæ via a suitable logic reasoner.

[1] Budzynska, and Reed. Speech acts of argumentation. CMNA 2011
[2] Rahwan, and Reed. The Argument Interchange Format. AAI 2009
[3] Dykes, Evert, Göttlinger, Heinrich, and Schröder. (2020). Reconstructing arguments from noisy text. Datenbank-Spektrum, 20
[4] Cîrstea, Kurz, Pattinson, Schröder, and Venema. (2011). Modal logics are coalgebraic. Comput. J., 54

Reliable Models for Authenticating Multimedia Content as Forensic Evidence

Benedikt Lorch (benedikt.lorch@fau.de)
Supervisor: PD Dr. Christian Riess

Criminal investigations often need to handle photo and video recordings that may serve as forensic trace or be probative in a legal setting. Such recordings are found on a seized device or hard disk, or on social media platforms. In most cases, little is known about the origin of the recording, such as its processing history or the camera that captured it. To validate the authenticity and to identify the source of the recording, a broad set of so-called blind techniques has been developed that require little to no explicit prior knowledge about the image under analysis. Most of these blind techniques identify relevant forensic traces by learning statistical models from large sets of examples. Due to their generality and wide applicability in practice, blind techniques offer a good starting point for forensic analysis.

Despite achieving state-of-the-art performance, these learning-based methods implicitly assume that the training data is representative for the test image under analysis. If a test image differs too much from the training data, learning-based methods are prone to fail silently. This problem is known as the training-test mismatch and poses a severe challenge to blind multimedia forensics in practice. Since blind techniques are by definition concerned with images from unknown origins, these techniques must take care of mitigating the training-test mismatch. The lack of knowledge about the camera manufacturing and post-processing related to the distribution channel, however, make it difficult to completely avoid mismatches. While data augmentation helps to reduce this gap, it is virtually impossible to cover all possible image variations in the training data.

In recent works, we proposed Bayesian methods for mitigating the training-test mismatch. Along with their predictions, Bayesian methods calculate a confidence interval, which enables a forensic analyst to quantify when to trust in the prediction. If an image under analysis differs too much from the training data, this can be detected based on the confidence interval, such that potential failure cases can be anticipated. We recently explored the detection of JPEG double compression and open-set camera model identification as application scenarios.

Besides providing probabilistic predictions, forensic techniques should be robust to smaller variations in the test data. In future work, we aim to improve our method's robustness while maintaining its awareness of unfamiliar inputs.

For use in criminal investigations, forensic methods must meet high requirements regarding precision and reliability. We argue that uncertainty-aware methods provide a path towards reliable forensic tools that can be used in practical applications.

Die strafprozessualen Ermittlungs- und Eingriffsmaßnahmen im Lichte der Cyberkriminalität unter besonderer Berücksichtigung des Internets der Dinge

Florian Nicolai (florian.nicolai@fau.de)
Supervisor: Prof. Dr. Hans Kudlich

Cyberkriminalität nimmt stetig zu. Im Lichte dieser Entwicklung sind die im Rahmen der Strafprozessordnung (StPO) geregelten Eingriffsbefugnisse der Ermittlungsbehörden überarbeitungs- und reformbedürftig.

Im Fokus der Betrachtung steht das „Internet of Things“. In diesem Zusammenhang ist nicht nur von Interesse, inwiefern Daten aus diesem Bereich überhaupt für einen Strafprozess von Relevanz sein können. Vielmehr ist auch begutachtungswürdig, auf welche Weise diese Daten gerichtsfest und für den Strafprozess verwertbar gewonnen werden können.

Es stellt sich zudem die Frage, inwiefern vernetzte Geräte neue Herausforderungen für das materielle Strafrecht darstellen. So bilden sich aufgrund der Vernetzung verschiedener Geräte neue Kriminalitätsfelder und Phänomene, auf die das Strafrecht keine passende, das begangene Unrecht vollständig abbildende Antwort hat. Die materiellrechtliche Einordnung steht häufig in einer Wechselbeziehung zum Strafprozessrecht.

Die vorhandenen Eingriffsbefugnisse sind sowohl in rechtsdogmatischer Hinsicht als auch unter praktischen Gesichtspunkten nicht auf dem neuesten Stand. Zwar werden — teils unter Billigung höchstrichterlicher Rechtsprechung — bestehende Normen (analog) auf neue, die IT betreffende Sachverhalte angewandt. Jedoch bestehen hiergegen zum Teil (schwerwiegende) rechtsdogmatische Einwände. Bedenken bestehen bisweilen auch bzgl. der Normenklarheit und -bestimmtheit, denen mit Blick auf den Eingriffscharakter strafprozessualer Maßnahmen nicht immer mit einem bloßen Verweis auf die Entwicklungsoffenheit der StPO begegnet werden kann.

Damit einher gehen Probleme bei der Rechtsanwendung, insbesondere bei den Ermittlungsbehörden, für die die Unsicherheiten über den Anwendungsbereich der Ermittlungsmaßnahmen im Alltag der Strafverfolgung Schwierigkeiten bereiten. Ferner steht zu befürchten, dass mit weiterem Fortschreiten der Technik die existierenden Normen den neuen Anforderungen, denen sich Staatsanwaltschaft und Polizeibehörden ausgesetzt sehen, nicht gerecht werden.

Diese Betrachtungen können nur unter enger Bezugnahme auf technische Neuerungen und Grundverständnis der technischen Aspekte erfolgen. Unerlässlich ist ebenfalls eine dezidierte Auseinandersetzung mit den technischen Möglichkeiten, auf Grundlage derer den Ermittlungsbehörden neue Befugnisse und Ermittlungsmethoden an die Hand zu geben sind.

Es gilt, die Schwierigkeit zu meistern, die dabei besteht, eine möglichst für neue Technik offene Regelungen zu schaffen und diese dennoch hinreichend bestimmt in ihrer Anwendbarkeit zu gestalten.

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]. Today, these methods often rely on manipulating the device's software. This clashes with the requirement on forensic methods to preserve the evidence's integrity during an investigation. When changes are inevitable it should be clear what was changed and to what extent this could influence the extracted data as this affects its legal relevance.

Currently a lot of research in the area of mobile device forensics revolves around the development of novel methods for data extraction and analysis. Less focus is placed on the scientific evaluation of these methods. In fact, a thorough evaluation appears to be missing for many established methods too. Comparing categorizations of memory acquisition methods[2] it also becomes evident that differences in the understanding and naming in the academic field exist. Different tool vendors also use their own terminology for which definitions are not always easily accessible, complicating tool comparison.

This thesis aims at improving the possibilities to determine if a method used in the context of mobile device forensics is forensically sound. A first step towards this goal is to establish a clear terminology regarding data extraction methods. Additionally, concepts like forensic soundness need to be reexamined in the context of today's technology and their definitions updated, if need be. We further plan to develop methods that enable tool validation in mobile device forensics.

[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] For example, see Konstantia Barmpatsalou, Tiago Cruz, Edmundo Monteiro, and Paulo Simoes, "Current and Future Trends in Mobile Device Forensics: A Survey", ACM Computing Surveys (CSUR), vol. 3_51, p. 1-31, 2018

Understanding Privacy in Cryptocurrencies

Viktoria Ronge (vikoria.ronge@fau.de)
Supervisor: Prof. Dr. Dominique Schröder

Cryptocurrencies are digital currencies normally not issued by a government or other central authority relying on cryptographic tools. They enable users to transfer money all over the world in a secure way, where there is no need for intermediaries like banks or exchange the money into different currency. Thereby, no user can be prevented from transferring money, no one can spend money they do not own or spend it twice and money can only be created under rules everyone agrees to. They further provide different nuances of privacy, where somewhat fully private ones are rare. The largest two are Monero[1] and Zcash[2]. They pursue different approaches, which are, with our current knowledge about privacy, at least partly incomparable.

This research project focusses on the foundations of anonymous cryptocurrencies from different angles. One is to understand the theory behind different anonymous cryptocurrencies and to formalize them. This is necessary as without formalizing no security can be proven and no statements about actual privacy for users can be done. Another one is to extend our knowledge and comprehension of different anonymity measures and to use them for comparison of currencies. This would help us to answer simple questions like which currency offers better anonymity, but this research is also important from a legal perspective, because anonymous cryptocurrencies often are used by criminals. Understanding privacy of different systems might lead to ideas on how to attack a system. This raises the fundamental question if this is proportional in relation to the violation of privacy of honest users. Moreover, when using results from such attacks in prosecution, we need an understanding of the results' quality. For genetic tests we know well about the accuracy based on past experiences. For deanonymising we are lacking such a ground truth that exists in other areas used for evidence. Therefore it is urgent to gain confidence in the accuracy of deanonymisation to make sure no innocent is falsely accused.

We hope to help giving an overview of these issues to provide the community with a better understanding of what privacy means in this subfield and how reliable we can talk about it. A first step was already done in formalizing Monero as a whole[3]. We further gave a better understanding of choosing anonymity sets in Monero[4] and are working on a followup.

[1] The Monero Project, https://www.getmonero.org/, last visited March, 27^{th}, 2020

[2] Electric Coin Company, https://z.cash/, last visited March, 27^{th}, 2020

[3] Omniring: Scaling Private Payments Without Trusted Setup, R. W. F. Lai and V. Ronge and T. Ruffing and D. Schröder and S. A. K. Thyagarajan and J. Wang, Proceedings of the 2019 ACM SIGSAC CCS 2019

[4] Foundations of Ring Sampling, V. Ronge, C. Egger, R. W. F. Lai, D. Schröder, and H. H. F. Yin, Proceedings of PETS 2021

Digitale Daten als Beweismittel im Strafverfahren

Dr. Christian Rückert (christian.rueckert@fau.de)
Supervisor: Prof. Dr. Christoph Safferling

Durch die Durchdringung der Arbeitswelt und des Privatlebens durch Computertechnik und das Internet werden in zunehmendem Maße Daten über Aktivitäten, Beziehungen und Bewegungen von Personen erzeugt und gespeichert. Die erzeugten und gespeicherten Daten sind auch für das Strafverfahren interessant und relevant. Dies gilt nicht nur für den Bereich des sog. Cybercrime, sondern für alle Deliktsbereiche.

Das deutsche Strafverfahrensrecht bzw. die Auslegung seiner Normen ist derzeit nicht an die sich schnell entwickelnde IT-Technologie angepasst. Das Habilitationsvorhaben adressiert dabei die beiden aus Sicht des Verfassers dringlichsten Problemkreise.

Zunächst stellt sich im Bereich der Datenerhebung im Ermittlungsverfahren das Problem, dass die Regulierung von Eingriffsgrundlagen zur Datenerhebung nicht mit der technischen Entwicklung Schritt halten kann. Dies führt dazu, dass der Gesetzgeber in den letzten Jahren eine Vielzahl von einzelnen Eingriffsgrundlagen für jeweils spezifische Technologien geschaffen hat. Diese Eingriffsgrundlagen sind häufig eng zugeschnitten und lassen sich wegen des Vorbehalts des Gesetzes nicht auf die Anwendung neuer Technologien übertragen. Dennoch werden die Grenzen spezieller Eingriffsbefugnisse überschritten oder es werden neue Eingriffsgrundlagen unter Verstoß gegen den Vorbehalt des Gesetzes und die Wesentlichkeitstheorie des BVerfG durch Kombination verschiedener bestehender Befugnisnormen geschaffen.

Das Habilitationsprojekt möchte diese Problemstellung adressieren, indem aus höherrangigen Normen (Verfassungs- und Europarecht) allgemeine Leitlinien zur Auslegung bestehender und Schaffung neuer Eingriffsbefugnisse entwickelt werden. Als Ergebnis sollen Vorschläge zur technikneutralen Reform der Befugnisnormen für Datenerhebungen auf Grundlage der notwendigen Schutzmechanismen in Abhängigkeit von der Eingriffsintensität unterbreitet werden. Ein besonderer Schwerpunkt liegt dabei auf der Erarbeitung eines Kriterienkatalogs für Verhältnismäßigkeitsprüfungen bei strafprozessualen Dateneingriffen.

Der zweite Problemkreis betrifft die Würdigung von digitalen Daten als Beweismittel in der Hauptverhandlung. Hier geht es vor allem um die Schaffung und Bewahrung eines möglichst großen Beweiswerts. Da Daten flüchtig und leicht manipulierbar sind, müssen Regeln zur Sicherung der Authentizität und Integrität in das Beweisrecht der StPO intergriert werden. Weiterhin stellt sich das Problem, dass das Tatgericht die Daten nicht selbst auswerten kann. Die Richterinnen und Richter müssen sich daher auf die Auswertung durch IT-Forensiker/innen verlassen. Hierfür muss es daher einheitliche Regeln hinsichtlich Methodik und Qualifikation der herangezogenen Sachverständigen geben. Das Vorhaben entwickelt diese Regeln sowohl aus dem Stand der Wissenschaft und Technik der IT-Forensik als durch Auslegung der Normen des Beweisrechts der StPO.

„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.

Automated Side-Channel Evaluation of Embedded Devices

Jens Schlumberger (jens.schlumberger@fau.de)
Supervisor: Dr. Stefan Wildermann

The always increasing abundance of embedded devices dealing with sensitive or security critical data should incentivize side-channel security evaluations not only for vendors but also forensic investigators. Hereby, side-channels like electromagnetic radiation can compromise mathematically safe cryptography by leaking information about the key. This is of special interest, as smart home devices and the Internet of Things are on the rise and many devices can provide valuable information when their cryptographic key is revealed. To analyze the side-channel information of a specific device, emissions of several cryptographic operations need to be recorded, synchronized, and compared to detect leakage. In order to enable easier and faster ways to evaluate generic embedded devices, new approaches have to be developed.

Forensic investigations have specific requirements for side-channel analysis, as they should not modify or tamper with evidence during the task. Therefore, electromagnetic radiation probes can be used to measure the emissions. However, current side-channel evaluation techniques use highly device-specific training or information which is not feasible due to the diversity of embedded systems. Therefore, expensive experts and a lot of time and effort would be needed to retrieve side-channel information at a crime scene. As this is not feasible for every crime scene, valuable information may be lost.

To tackle these problems, this thesis investigates new approaches which will enable highly automated side-channel evaluation of embedded devices. With a main focus on the Advanced Encryption Standard (AES) as it is widely spread for embedded systems as a symmetric, round based block cipher. The goal is a system that automatically evaluates a device which uses AES without preliminary knowledge about the device. Furthermore, other block ciphers are evaluated as well as other cryptographic routines. Specifically, the following challenges are faced:

First, detecting and characterizing of cryptographic operations on a power trace with multiple recorded cryptographic operations without device-specific knowledge. Second, the approach shall be independent of the measuring setup as well as independent of the specific implementation of the cryptographic algorithm in software or hardware. A final goal is to build a framework that can do a live side-channel evaluation of a target device without modifying it physically.

Tools and Techniques for Structured Analysis of Digital Evidence

Janine Schneider (janine.schneider@fau.de)
Supervisor: Prof. Dr.-Ing. Felix Freiling

Digital evidence is an increasingly important form of evidence in courts of law today and it comes in many different forms, be it pictures stored on a hard disk, documents in a cloud or passwords stored in a computer's main memory. This form of evidence constantly introduces new challenges, changing with new technologies and applications. For example, because solid-state drives (SSDs) operate in an entirely different way as classical hard discs (HDDs) it is questionable whether classical techniques to recover deleted files (file carving) can still be applied. Furthermore, the increased risk of bit errors could lead to integrity check failures while using cryptographic hashes. Another example is the complex handling of cloud storage and shared documents. Vassil Roussev and Shane McCulley already did some extensive research on API-based data acquisition and analysis and developed a tool called kumodocs[1] which is able to extract artifacts of Google documents and slides. Besides, in contrast to other forms of evidence, the sheer quantity of digital evidence is actually a problem. Therefore, it needs new ways to acquire and analyze digital evidence efficiently, to ensure integrity and to combine already existing forensic approaches. Brian Carrier already observed that the task to reconstruct evidence on higher levels of abstraction from low level evidence is non-trivial[2] since it involves decoding the mapping between pieces of data on both layers and to bridge the semantic gap[3]. Within this PhD thesis we will develop a model of storage abstraction layers to formalize the problem of reconstructing evidence on higher levels from lower levels of abstraction. The model will make use of heuristics to formalizes different analysis and reconstruction problems and to create a generalized interface for enabling the combination of different solving approaches. Through the generic combination of various techniques results could be strengthen or the result quantity could be decreased. To demonstrate the applicability of the approach, a forensic analysis and reconstruction tool will be implemented. The tool will be an open-source C++ framework whose architecture will be directly derived from the model.

[1] Vassil Roussev and Shane McCulley, Forensic analysis of cloud-native artifacts, Digital Investigation, 16, S104-S113, 2016

[2] Brian Carrier, Defining digital forensic examination and analysis tools using abstraction layers, International Journal of Digital Evidence, 1, 2003

[3] Jain et al., SoK: Introspections on Trust and the Semantic Gap, IEEE Symposium on Security and Privacy, SP 2014, pp. 605-620, 2014

GRK 2535: Knowledge- and data-based personalization of medicine at the point of care

Prof. Dr. rer. nat. Britta Böckmann
Email: Britta.Boeckmann@uk-essen.de
Universität Duisburg-Essen
Internet: www.wispermed.org

Due to the increasing digitization in medicine, ever more data is being generated and made available, for example in electronic patient records, laboratory analyses or clinical guidelines. It is a challenge to make the knowledge contained in these different classes of data available and usable at the Point of Care (PoC) for concrete individual therapy decisions. Existing clinical information systems allow the collection and storage of important information, but relatively unstructured and without individual, context-related compilation of the relevant facts for treatment decisions. 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 can gain a holistic overview of the state of research on knowledge- and data-based personalisation of medical decision-making processes and learn to design new interdisciplinary methods, as well as implement prototypes such as malignant melanoma.
In novel ways, methods from the fields of information extraction, knowledge representation are combined with machine learning methods and findings on user interaction at the PoC. Interdisciplinary measures, in particular hospitations at the dermatological clinic, will reduce the barriers of understanding between these disciplines. A unique characteristic of the Research Training Group is the inter-institutional cooperation between the University of Applied Sciences Dortmund, the University Duisburg-Essen and the University Medicine Essen. This is based on an already existing cooperation through a joint study course Medical Informatics. The applicants represent a broad range of expertise in the fields of medical informatics, bioinformatics, epidemiology, artificial intelligence, psychology, radiology and melanoma research. Graduates of our program will be able to assume leading roles in the digitization process of the health care system and to further improve treatment methods with the help of artificial intelligence methods, taking into account the direct feedback and experience of the attending physicians.

Context modelling and mapping of guidelines and SOPs

Catharina Lena Beckmann (catharina.beckmann@fh-dortmund.de)
Supervisor: Prof. Dr. Britta Böckmann

Clinical guidelines (e.g. the national S3 guideline for the diagnosis, therapy and follow-up of melanoma[1]) and standard operating procedures (SOPs) provide useful recommended actions for evidence-based care. However, since these are available as an unstructured information base, corresponding guideline information on the patient context currently has to be searched for and compared with hospital-internal SOPs in a time-consuming manner by physi-cians[2]. Currently existing modeling and mapping procedures for guidelines do not consider the specific patient or user context.

In order to precisely represent the unstructured knowledge using context modelling and mapping and thereby reduce the time physicians spend treating patients, this research project aims to make a contribution. The overarching research question of this work is to investigate how relevant text passages, that support at the point of care, can be identified in unstructured documents (such as guidelines and SOPs) and summarized simultaneously. For the identification of these passages, the patient-specific comorbidities, comedications and the patient's general condition (ECOG) are relevant, as well as the inclusion of user-specific expertise.

To establish a semantic mapping between guideline and patient context, semantic analysis techniques are used to identify information modules in the guideline. Subsequently, these modules are mapped to a suitable ontology such as SNOMED CT. The resulting computer-interpretable guidelines and SOPs can then be queried relative to a contextual model incorporating the patient's position in the treatment pathway and their medical records.

[1] Deutsche Krebsgesellschaft, Deutsche Krebshilfe, AWMF, "Diagnostik, Therapie und Nachsorge des Melanoms", Langversion 3.3, 2020, http://www.leitlinienprogrammonkologie.de/leitlinien/melanom/, retrieved: 06.04.2021

[2] Becker M., Kasper S., Böckmann B., Jöckel KH., Virchow I., "Natural language processing of German clinical colorectal cancer notes for guideline-based treatment evaluation", International Journal of Medical Informatics, 127, 141-146, 2019

Extraction of argumentation structures

Marie Bexte (marie.bexte@uni-due.de)
Supervisor: Prof. Dr-Ing. Torsten Zesch

When consulting medical documentation to decide between different treatment options, finding documents a search term occurs in may not be sufficient in identifying the ones relevant to a certain information need. When for example searching for the drug Dabrafenib, one may find the following two statements:

- Constant findings with Dabrafenib, therefore continuation.
- Constant findings with Dabrafenib. Discontinuation due to side effects.

While both of them mention constant findings, one of them continues on to say that the side effects were severe enough to stop using it. We therefore aim to extract these argument structures to allow for a more fine-grained distinction of what kinds of arguments are made.

This task poses a number of challenges, starting with the nature of the documents. Their creation may be influenced by not only the personal style of the medical personnel writing them but also time pressure, which may lead to spelling mistakes or shortened forms. A further difficulty may arise from the use of implicit language. While state-of-the-art approaches are able to extract explicit argument structures from argumentative essays[1], social media text[2] and medical diagnostics[3], implicit reasoning remains difficult[4]. On top of that, the documents at hand include specific medical terminology, which will be addressed with the help of another research project in the graduate school.

We will create a reference corpus that covers all relevant document classes that arise in treating patients and annotate the argument structures within them. We intend to publish this corpus as freely as possible. Based on the corpus we will then develop and evaluate algorithms aiming to automatically extract these argument structures.

[1] Christian Stab and Iryna Gurevych, "Parsing argumentation structures in persuasive essays.", Computational Linguistics 43(3), p. 619–659, 2017

[2] Michael Wojatzki and Torsten Zesch, "SemEval-2016 Task 6: Stance detection in social media using stacked classifiers.", 10^{th} International Workshop on Semantic Evaluation, p. 428–433, 2016

[3] Claudia Schulz, Christian M Meyer, Jan Kiesewetter, Michael Sailer, Elisabeth Bauer, Martin R Fischer, Frank Fischer, and Iryna Gurevych, "Analysis of automatic annotation suggestions for hard discourse-level tasks in expert domains.", 57^{th} Annual Meeting of the Association for Computer Linguistics, 2019

[4] Michael Wojatzki and Torsten Zesch, "Stance-based argument mining–modeling implicit argumentation using stance.", 13^{th} Conference on Natural Language Processing, p. 313–322, 2016

Context-sensitive, personalized search at the Point of Care

Sameh Frihat (Sameh.Frihat@uni-due.de)
Supervisor: Prof. Dr.-Ing. Norbert Fuhr

In the last few years, new data science opportunities started to appear for more effective medical decisions due to the increasing amount of data and knowledge derived from patients and clinical experiments, which allows to retrieve similar cases and related treatments. For this purpose, we need a context dependent integrated search across multiple information sources[1], taking into account both the patient and user context. Thus, it is necessary to determine which patient- and case-specific information is relevant for a contextual search. It must be investigated which of these can be extracted by automatic procedures and which, if necessary, must be explicitly specified by the practitioner. Subsequently, procedures must include this context in the search, whereby interactive and iterative procedures are to be considered in particular. New user guidance procedures such as scaffolding[2] and the Interactive Probability Ranking Principle[3] should be considered to achieve an effective user interface. So far, no uniform context model for user interaction at the point of care is known, especially for the medical domain to support users effectively. This context orientation is missing in previous approaches in this area, as is the lack of interactivity and consideration of user acceptance.

Within this project, we focus on developing a user context that includes the current user as well as the specific application. Depending on this, the search functionality should be configurable (e.g., the selection of sources to be searched should depend on the physician's prior knowledge), and the presentation of results should be adaptable (e.g., more concise vs more detailed presentation depending on the situation). Besides, the system should allow physicians to comment on and annotate useful knowledge, as well as to set bookmarks and their own links. If there is a need for further information at a later stage of treatment because the selected therapy does not work or complications occur, this information can be accessed immediately.

In the end, it is planned to evaluate all results from different aspects like the user interface, procedures quality and user interactivity using user studies and Information retrieval evaluation techniques.

[1] Marcel Martin, Algorithms and tools for the analysis of high throughput DNA sequencing data, Dortmund: TU Dortmund, 2013.

[2] Kriewel S., Fuhr N., An evaluation of an adaptive search suggestion system,32nd European Conference on Information Retrieval Research, Springer, ECIR, 2010.

[3] Fuhr, N., A Probability Ranking Principle for Interactive Information Retrieval. Information Retrieval 11(3), 2008.

Treatment decision for melanoma patients: Identification of similar patients at the PoC

Wolfgang Galetzka (wolfgang.galetzka@uk-essen.de)
Supervisor: Prof. Dr. A. Stang

In the treatment of melanoma patients unusual constellations occur frequently. For those cases, making a treatment decision can be difficult. One aim of the PhD project is to develop methods to find patients similar to a currently treated patient among previously treated melanoma patients. The physician at point of care can then compare the suggested similar patients, their treatments and the corresponding outcomes, to draw conclusions for the treatment.

Different methods to measure similarity will be investigated, from learning an inner product for optimizing k-nearest neighbour classification[1] to more recent procedures like convolutional neural networks[2, 3] in combination with graph-based models to include ontologies[4]. Furthermore, the use of methods enabling counterfactual predictions[5] will be assessed.

Another important task is to integrate expert knowledge as well as the different types of data, e.g. data from genomic analysis, images or physicians notes, which requires close collaboration with the corresponding research projects of the RTG.

[1] K. Weinberger, S. Lawrence, "Distance Metric Learning for Large Margin Nearest Neighbor Classification," Journal of Machine Learning Research, vol. 10, p. 207-244, 2009

[2] Z. Zhu, C Yin, B. Qian, Y. Cheng, J. Wei and F. Wang, "Measuring Patient Similarities via a Deep Architecture with Medical Concept Embedding," 2016 IEEE 16th International Conference on Data Mining (ICDM), p. 749-758, 2016

[3] Q. Suo, F. Ma, Y. Yuan, M. Huai, W. Zhong, J. Gao and A. Zhang, "Deep Patient Similarity Learning for Personalized Healthcare," IEEE Transactions on NanoBioscience, vol. 17 no. 3, p. 219-227, 2018

[4] E. Choi, M. T. Bahadori, L. Song, W. F. Stewart and J. Sun, "GRAM: Graph-Based Attention Model for Healthcare Representation Learning," Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, p. 787-795, 2017

[5] A. Ahmed, M. van der Schaar "Bayesian inference of individualized treatment effects using multi-task Gaussian processes," Proceedings of the 31st International Conference on Neural Information Processing Systems, p. 3427-3435,2017

Context modelling for the point of care

Eva Maria Hartmann (eva.hartmann@fh-dortmund.de)
Supervisor: Prof. Dr. rer. nat. Sabine Sachweh

To support work processes and decisions, it is necessary to adapt the software used to its respective area of application. The field of medicine in particular poses the challenge of reflecting the dynamic knowledge base and displaying the data currently relevant to the user for patient care. From this problem many open questions about the provision, the need and the presentation of relevant data arise.
In this research project these challenges and questions are investigated for the context of medical care of malignant melanoma patients. The goal is to enable support for individual, knowledge-based and efficient decision making at the point of care. As a basis, existing context models from other domains, as well as existing models of conceptual modeling for the integration of treatment guidelines and clinical pathways will be considered. Further, it will be explored which additional influencing factors can be identified in general and for the specific use case. Taking this prior knowledge into account, a context model will be generated and a concept for an adaptive user interface developed. The conception will be based on a user-centered approach, which includes the clinicians in the development. This is done on the basis of the software-technical methods of Domain-Driven Design[1]. In addition, the concept must take into account the guidelines of the German government's data ethics committee[2], as well as data protection and information security aspects. The evaluation of the concept is to be carried out iteratively on the basis of a microfrontend and -service-based prototype with the stakeholders involved.

[1] Eric Evans, "Domain-driven design:Tackling complexity in the heart of software", 2003
[2] Datenethikkommission der Bundesregierung, "Abschlussbericht der Datenethikkommission" 2019

Evaluation and Proposal System for Current and Relevant Literature at the PoC

Ahmad Idrissi-Yaghir (ahmad.idrissi-yaghir@fh-dortmund.de)
Supervisor: Prof. Dr.-Ing. Christoph M. Friedrich

Clinical practice guidelines (CPGs) play a fundamental role in supporting medical decision-making in clinical practice for specific diseases such as melanoma. They describe the diagnosis and treatments based on the best evidence available in biomedical publications. Therefore, such guidelines require regular updates[1] to be able to reflect the current research.

However, the published guidelines are often outdated due to the constant and rapid changes in the evidence used in clinical practice. Keeping CPGs up-to-date is challenging as this requires continuous monitoring and review of the relevant literature by experts, which is time-consuming and labor-intensive. Therefore, developing automated approaches could play an essential role in maintaining the clinical practice guidelines.

This research project aims to leverage machine learning and natural language processing approaches to build models for literature recommendation and develop a personalised evaluation model for them. It also aims to investigate how the relatively new and promising transformer[2] architectures could contribute to improving such works. Moreover, based on the obtained models, a literature recommendation system at the point of care is to be developed as an application for the Smart Hospital Information Platform (SHIP) at the University Hospital Essen.

During this dissertation project, the following research questions will be investigated: First, is a suggestion system for current personalised literature recommendations possible? Second, what criteria does a rating system for literature recommendations on skin melanoma need? Furthermore, how can such a system be evaluated quantitatively and qualitatively? Last, can word embeddings, and transformer architectures improve these systems?

[1] Becker M, Neugebauer EA, Eikermann M., "Partial updating of clinical practice guidelines often makes more sense than full updating: a systematic review on methods and the development of an updating procedure", J Clin Epidemiol, vol 67(1), p 33-45, 2014

[2] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L. and Polosukhin, I., "Attention is all you need", Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17), 6000–6010, 2017

Giving information to counteract wrong conclusions - Empirical study on acceptance

Alisa Küper (alisa.kueper@uni-due.de)
Supervisor: Prof. Dr. phil. Nicole Krämer

Decision support systems have the potential to optimize the clinical decision-making process.[1]. This project aims to investigate how decision support systems can best support physicians by providing information, beyond the first own impression, to counteract wrong conclusions. Human decisions are subject to human error like availability bias, anchoring bias or confirmation bias.[2]. These cognitive biases can influence which diagnoses come readily to mind or whether all possible diagnoses are considered. Overall, unchecked biases have an impact on the quality of decision. Therefore, one overarching questions of this project is if and how a decision support system can counteract these biases? Furthermore, how can additional errors, such as automation bias or reactance, introduced through the technical support be avoided?

An important factor in avoiding reactance is technology acceptance.[3]. Another pivotal question of this research project will be under what circumstances a support system is accepted by physicians. Acceptance of a system is crucial to guarantee uptake by clinicians.[4].

These questions will be researched in four experimental-psychological studies with about 100 participants each and an additional long-term study with 30 participants. There, we examine how and how much information needs to be presented in order to support good decisions without leading to defensive reactions. The long-term study will focus on the acceptance of the decision support system in the field by investigating factors like perceived usefulness and trust in the system. Additionally, in collaboration with other research projects, this project aims to contribute to explainable AI research by investigating how much information needs to be given to make users trust algorithmic decision systems while avoiding overreliance on the system.[5].

[1] Cresswell K., Majeed A., Bates D., Sheikh A., "Computerised decision support systems for healthcare professionals: An interpretative review," Informatics in Primary Care, vol. 20, p. 115-128, 2012

[2] Croskerry P., "From Mindless to Mindful Practice -Cognitive Bias and Clinical Decision Making," New England Journal of Medicine, vol. 368, p. 2445-2448, 2013

[3] Khairat S., Marc D., Crosby W., Al Sanousi A., "Reasons for physicians not adopting clinical decision support systems: Critical analysis," Journal of Medical Internet Research, vol. 20, 2018

[4] Jun S., Plint A., Campbell S., Curtis S., Sabir K., Newton A., "Point-of-Care Cognitive Support Technology in Ermergency Departments: A Scoping Review of Technology Acceptance by Clinicans," Academic Emergency Medicine, vol. 25, p. 494-507, 2018

[5] Bussone A., Stumpf S., O'Sullivan D., "The Role of Explanations on Trust and Reliance in Clinical Decision Support Systems," International Conference on Healthcare Informatics, 2015

Analysis of Preclinical Image Data Including Additional Clinical Data

Daniel Sauter (daniel.sauter@fh-dortmund.de)
Supervisor: Prof. Dr. rer. nat. Markus Kukuk

The incidence of melanoma in Germany has continuously increased during the last decades. Furthermore, the five-year survival rate of type IV melanoma (UICC) in 2015 and 2016 was as low as 23% for women and 15% for men[1]. Therefore, reliable early detection and proper treatment are important. Nowadays, therapeutical options for melanoma include surgical resection, biochemotherapy, radiotherapy, photodynamic therapy, immunotherapy, and targeted therapy[2].

In the field of deep learning (DL), convolutional neural networks continue experiencing an upswing due to technological advancements[3]. These include improvements in architecture[4] as well as in training procedures[5]. Such modern machine learning algorithms combined with the increasing amount of medical data available are expected to improve medical diagnosis and treatment[6]. DL has been used in the context of cancer care for different purposes. These include both diagnosis[7] as well as prediction of the disease-specific survival rates[8]. For the specific case of melanoma, research has focused on dermatoscopy[9] and histopathology[10] image data, combined with several new DL techniques.

This thesis aims at investigating whether DL can be used for the prediction of clinical endpoints from photographs or histological images. Among others, different image types, data augmentation methods, hyperparameters, and transfer learning techniques will be compared. We hope to further improve AI-driven medical assessment in terms of validity and reliability. Thereby helping practitioners to better recognize and treat melanoma patients.

[1] Robert Koch-Institut, "Krebs in Deutschland für 2015/2016," Berlin, 2019, Accessed: Apr. 14 2021, Available: https://www.krebsdaten.de/Krebs/DE/Content/Publikationen/Krebs_in_Deutschland/kid_2019/krebs_in_deutschland_2019.pdf

[2] B. Domingues, J. M. Lopes, P. Soares, and H. Pópulo, "Melanoma treatment in review," ImmunoTargets Ther., vol. 7, pp. 35–49, 2018

[3] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015

[4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," CVPR 2016 Proc., Las Vegas, NV, pp. 770–778, 2016

[5] F. Zhuang et al., "A Comprehensive Survey on Transfer Learning," Proc. IEEE, vol. 109, no. 1, pp. 43–76, 2021

[6] A. Rajkomar, J. Dean, and I. Kohane, "Machine Learning in Medicine," N. Engl. J. Med., vol. 380, no. 14, pp. 1347–1358, 2019

[7] F. F. Ting, Y. J. Tan, and K. S. Sim, "Convolutional neural network improvement for breast cancer classification," Expert Syst. Appl., vol. 120, pp. 103–115, 2019

[8] J. N. Kather et al., "Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study," PLOS Med., vol. 16, no. 1, paper e1002730, 2019

[9] A. Esteva et al., "Dermatologist-level classification of skin cancer with deep neural networks," Nature, vol. 542, no. 7639, pp. 115–118, 2017

[10] A. Hekler et al., "Deep learning outperformed 11 pathologists in the classification of histopathological melanoma images," Eur. J. Cancer, vol. 118, pp. 91–96, 2019

Analysis of unstructured texts from publications

Henning Schäfer (henning.schaefer@uk-essen.de)
Supervisor: Prof. Dr. med. Peter Horn

Terminologies can be used to enrich clinical texts with contextual knowledge and normalized clinical concepts, such as drugs, measurement units, anatomical locations, mutations, diseases or genes. At present, there is no German-language terminology for melanoma. Therefore, the research question of this dissertation project is whether a semi-automatic creation of a terminology for malignant melanoma is possible. This is accompanied by the question of which tools and resources can be used to create a German-language terminology and how the terminology can be evaluated.

One way to leverage knowledge is to enrich clinical texts using machine learning models and mapping normalized concepts to existing terminologies or ontologies such as the Unified Medical Language System (UMLS)[1] or RadLex[2]. Normalization not only enables more detailed searching capabilities based on specific, uniquely defined concept IDs, but also allows for statistical analysis and filtering. Normalization addresses the problem of naming variations, such as typographical variants like the alternating use of hyphens, parentheses, spacing, abbreviations, and short formulas.

A promising approach for creating a terminology is named entity recognition through the use of transformer architectures[3]. A practical application of text enrichment is planned through the development of a terminology server that can be used by the Smart Hospital Information Platform (SHIP) at the University Hospital Essen. The server is planned for direct use at the point of care, e.g. to present personalized summaries of inclusion/exclusion criteria for clinical trials or physician letters enriched with clinical concepts. If successful, the architecture for creating a terminology on melanoma could also be extended to other diseases.

[1] Bodenreider, O, "The unified medical language system (UMLS): integrating biomedical terminology.", Nucleic acids research, vol. 32, D267-D270, 2004

[2] Langlotz, C. P., "RadLex: a new method for indexing online educational materials.", RadioGraphics, vol. 26, no. 6, 2006

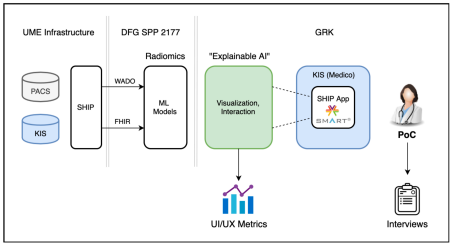
[3] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L. and Polosukhin, I., "Attention is all you need", Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17), 6000–6010, 2017

Analysis of clinical image data incorporating further clinical data - Explainable Radiomics

Yasmin Schmitt (yasmin.schmitt@uk-essen.de)
Supervisor: PD Dr. Felix Nensa

Machine learning techniques have experienced a huge development in the last decades, gaining high interest in both research and industry. In the field of image processing, for example, Deep Learning approaches have demonstrated highly predictive accuracies. Explicitly the medical factor can greatly benefit from it, however, the programs are very often so-called black-box models whose inner decision-making process cannot be accessed. When it comes to decisions with high stakes, such as a person's health, transparency is crucial. Physicians must be able to understand the predictions in order to be able to link their own expertise and to explain subsequent procedures to patients, to gain their trust, and to be able to identify possible mistakes.

Therefore, the question, how can these models be made more understandable? Explainable AI (XAI) is a research field that addresses exactly this question.[1] So far, mainly post-hoc methods have been developed to achieve interpretability, but intrinsic approaches can also be helpful, especially for end-users who have no prior knowledge of machine learning at all. Within the scope of this project, methods from the field of XAI will be implemented to provide explanations like summary reports, visualizations or counterfactuals and will be subsequently evaluated for use at the point of care.[2][3] Accuracy, user experience and interface as well as user trust will be examined in cooperation with physicians at the University Hospital Essen and other associated projects.



[1] Huff DT, Weisman AJ, Jeraj R. Interpretation and visualization techniques for deep learning models in medical imaging. Phys Med Biol. 2021 Feb 2;66(4):04TR01. doi: 10.1088/1361-6560/abcd17

[2] Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med. 2019 Oct 29;17(1):195. doi: 10.1186/s12916-019-1426-2

[3] Demircioglu A, Koitka S, Nensa F. Big Imaging Data: Klinische Bildanalyse mit Radiomics und Deep Learning. Nuklearmediziner 2019;42(02):97-111. doi: 10.1055/a-0838-8135

Predictive modeling for patient similarity based on an openEHR model of melanoma

Jessica Swoboda (jessica.swoboda@uk-essen.de)
Supervisor: Prof. Dr. Britta Böckmann

In the management and treatment of melanoma, there are currently no reliable predictive biomarkers related to primary resistance, development of resistance to therapy, or risk for the occurrence of (severe) adverse events, especially immunotherapy-related ones. Although several vector- and network-based approaches for generating similarity metrics[1] [2] [3] and for identifying similar patients for predictive modeling already exist in various medical disciplines[1] [4], AI (artificial intelligence) continues to hold great potential and challenges for personalized medicine[5]. The use of AI could identify previously hidden associations and generate new insights in the treatment of patients with minimal human intervention[5].

The goal of this work is the AI-supported creation of a knowledge-based prediction model to support personalized medicine using patient similarities and to make individualized predictions for each patient regarding their best possible treatment. Unlike many related work[4], this approach is intended to use longitudinal information, identify a variety of individual predictor variables, incorporate treatment context, and take biases into account.

A comprehensive openEHR model of melanoma will be created. With the help of 2000 retrospective data sets of the skin tumor center from SHIP (Smart Hospital Information Platform) of the Universitätsmedizin Essen, this openEHR model will build the basis for a predictive, knowledge-based prediction model. International standards will be applied in this approach. The clinical terminology SNOMED CT will be used for semantic annotation and HL7 FHIR (Fast Healthcare Interoperability Resources) for data persistence. Expert knowledge and AI-supported algorithms will be used to identify relevant predictors for an index patient and thus find similar patients, in order to make statements about (possibly dynamic) questions regarding the management, treatment, and possible outcomes of the individual patient. The results will be evaluated using expert knowledge and compared with etsablished algorithms.

[1] Sharafoddini A., Dubin J. A., Lee J., "Patient Similarity in Prediction Models Based on Health Data: A Scoping Review", JMIR Med Informatics, vol. 5, p. e7, 2017

[2] Zhu Z., Yin C., Qian B., Cheng Y., Wei J., Wang F., "Measuring Patient Similarities via a Deep Architecture with Medical Concept Embedding", 2016 IEEE 16th International Conference on Data Mining (ICDM), p. 749-758, 2016

[3] Pai S., Bader G. D., "Patient Similarity Networks for Precision Medicine", J MolBiol, vol. 430, p. 2924–2938, 2018

[4] Goldstein B. A., Navar A. M., Pencina M. J., Ioannidis J. P. A., "Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review", Journal of the American Medical Informatics Association, vol. 24, p. 198-208, 2017

[5] Wang F., Preininger, A., "AI in Health: State of the Art, Challenges, and Future Directions", Yearbook of medical informatics, vol. 28, p. 16-26, 2019

Uncertainty aware Bayesian methods for precision oncology

Hamdiye Uzuner (hamdiye.uzuner@uni-due.de)
Supervisor: Dr. Johannes Köster

Varlociraptor[1] is a variant calling tool comprising an underlying statistical model that is aware of uncertainties such as mapping and tumor heterogeneity as well as various biases. The goal of this project is to extend Varlociraptor to further broaden its applications in the treatment and diagnosis of cancer.

This will first happen in the form of generating a novel model for subclonal HLA typing. The human leukocyte antigen (HLA) system, also known as major histocompatibility complex (MHC), consists of proteins that serve immune system and its regulation against pathogens and certain diseases, including cancer[2]. In a population consisting of individuals that have different genetic backgrounds, the response of the immune system to certain diseases may not be the same, and variances at HLA loci may be one of the reasons. HLA alleles are known to contain different variants which make HLA genes highly polymorphic. For subclonal HLA typing, we plan a Bayesian model that integrates haplotype fraction likelihoods derived from Kallisto[3] with per-locus allele frequency likelihoods calculated by Varlociraptor.

Secondly, the Varlociraptor model will be extended to enable joint calling of genomic variants and CpG methylation. In order to achieve this, the model will be extended by a latent variable for the methylation state. This will allow to capture the uncertainty between genomic C to T substitutions and not methylated CpGs on bisulfite sequenced samples.

Utilization of Varlociraptor in clinical applications, coupled with the Snakemake[4] workflow management system is highly promising to offer sophisticated reports to doctors via the Smart Hospital Information Platform (SHIP) at the point of care. Therefore, the third part of the project entails the representation of Varlociraptor results via the Fast Healthcare Interoperability Resource (FHIR) and the presentation of results for every patient to the clinics. At this point, it is important to be involved in Intercollaborative Working Groups (IWGs) with other projects of the GRK. An example collaboration could concern the enhancement of Snakemake reports in a way that they become more serviceable and accelerate decision-making, e.g via a tight integration into SHIP and an optimization of the displayed information depending on the role of the viewer.

[1] Köster, J., Dijkstra, L.J., Marschall, T. et al. Varlociraptor: enhancing sensitivity and controlling false discovery rate in somatic indel discovery. Genome Biol 21, 98 (2020)

[2] Turner D. The human leucocyte antigen (HLA) system. Vox sanguinis, 87 Suppl1, 87–90. (2004)

[3] Bray, N. L., Pimentel, H., Melsted, P., Pachter, L. Near-optimal probabilistic RNA-seq quantification. Nature biotechnology, 34(5), 525–527. (2016)

[4] Mölder F, Jablonski KP, Letcher B et al. Sustainable data analysis with Snakemake. F1000Research. (2021)

HPI Research Schools on Data Science and Engineering and Service-Oriented Systems Engineering

Prof. Dr. Felix Naumann, Prof. Dr. Tilmann Rabl,
Prof. Dr. Andreas Polze, and Prof. Dr. Robert Hirschfeld
Email: {felix.naumann,tilmann.rabl,andreas.polze,robert.hirschfeld}@hpi.uni-potsdam.de
Hasso Plattner Institute, University of Potsdam
Internet: https://hpi.de/forschung/research-schools.html

The HPI research schools explore topics in the fields of data science and IT systems engineering that are of interest to academics and practitioners.

The research school on Data Science and Engineering unites top PhD students and researchers 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, digital health, and more.

The research school on Service-Oriented Systems Engineering is active in research areas such as system design, analysis and modeling; adaptability; component-based development and application integration; business process management; cyber security; software engineering; and programming technology.

Graph Immersions

Aikaterini Niklanovits (aikaterini.niklanovits@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Tobias Friedrich

Understanding the structure of a graph has been proven useful when it comes to dealing with algorithmically hard problems. Through the existence or obstruction of certain structures, graph classes with interesting properties arise.

One of the most well known, non-trivial such structures are graph minors. We say that a graph H is a minor of a graph G, if it can be obtained from it through a series of vertex or edge deletions and edge contractions. Graph minors where first introduced and studied by Robertson and Seymour in a series of thirteen papers under the name Graph Minors I-XIII. One of the most popular graph classes characterized through forbidding the existence of both K_5 and $K_{3,3}$ as minors, are planar graphs. Lipton and Tarjan, for example, provided a polynomial time algorithm that computes a balanced separator of size at most $\sqrt{n}$, of a planar graph G of order n.

Another interesting structure which was again introduced by Robertson and Seymour are graph immersions. We say that a graph H is an immersion of a graph G if it can be obtained from it through a series of edge or vertex deletions and edge lifts. Robertson and Seymour also proved that immersions, like minors are well quasi ordered. It has also been proven that immersions are closely related to the minimum degree of a graph while minors to its average degree. Immersions have gained more interest over the last decade but haven't been studied as deeply as minors.

A usual strategy to approach problems related to immersions is deriving inspiration from similar results for minors. Such a result comes from Giannopoulou, Kaminski and Thilikos who gave a structural characterization for graphs forbidding Kuratowski graphs as immersions. It has also been proven efficient to add restrictions to problems related with immersions like bounding the independence number of a graph. A significant structural theorem was proven by Wollan, that is that every graph either has bounded tree-cut width or admits a large wall as an immersion. This result has been used to obtain FPT algorithms for problems like Capacitated vertex cover and capacitated dominating set. Lastly, I believe and work on gaining a better understanding of this theory, and exploiting already useful structural results like Wollan's theorem, to develop efficient algorithms.

The Effect of Sparsity on Learning Disentangled Representations

Alexander Rakowski (alexander.rakowski@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Christoph Lippert

The goal of representation learning is to obtain latent space representations, which should map complex, usually high-dimensional data (e.g., images) into lower-dimensional feature spaces. These mapping are usually learned via a machine learning algorithm of choice. Ideally, these features should then characterize the data points on higher levels of abstraction - e.g., describing objects located in an image instead of the "raw" pixel values. However, these representations are usually not easy to interpret. This is especially true in unsupervised settings, where no labels are available. It is postulated that disentanglement is an important property to be achieved for such representations[1].

The majority of current state-of-the-art approaches to learning disentangled representations in deep neural networks relies on priors and regularizations imposed on the latent space, focusing mainly on obtaining an uncorrelated posterior. It has been shown however that such approaches are bound to be futile[2], due to the rotational invariance of the typically used priors.

Instead, in our work we investigate the effect of regularizing the network architecture itself, by reducing the number of connections between the hidden units. Dense connectivity increases the number of parameters in a model, and thus its potential expressive power. However, recent works have shown that in practice a large percentage of weights remains unused and can be removed without sacrificing performance. We argue that sparsely connected features might be desirable when learning disentangled representations. This is motivated by the assumption that it is not necessary for a feature (represented by a hidden unit) to be constructed using all available features from the subsequent layer - instead, it should suffice for it to be a combination of only a subset. We further propose techniques for obtaining such sparse models - Random Pre-Defined Masking and Learnable L1 Masking. These are then evaluated on standard datasets for measuring disentanglement, along with ablation tests to identify the factors crucial for improving performance.

[1] Y. Bengio, A. Courville, and P. Vincent, "Representation Learning: A Review and New Perspectives,"IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35.8, p. 1798–1828, 2013

[2] F. Locatello, S. Bauer, M. Lucic, G. Raetsch, S. Gelly, B. Schoelkopf, and O. Bachem, "Challenging Common Assumptions in the Unsupervised Learning of Disentangled Representations" arXiv preprint arXiv:1811.12359, 2018

Processing multi-dimensional geodata towards a virtual spatial model

Andreas Fricke (andreas.fricke@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Jürgen Döllner, Prof. Dr. Hartmut Asche

A virtual spatial model (VSM) can represent a complex amalgamation of various spatial multidimensional data. However, modelling with insufficiently valid base data limits its general applicability. Through functionalities such as spatial and semantic queries or analyses, a VSM can not only provide visual representations but also generate informative added value. This doctoral project deals with a question that is relevant both for science and for everyday reality: How can functionally rich virtual multidimensional city models be generated, derived, visualised and made usable in an application-oriented way considering insufficient basic data (reference or training data). On the one hand, questions of acquisition, processing (in a geometric as well as semantic context), function assignment and visualisation are addressed. On the other hand, the partial aspects mentioned are integrated into a generic, service-oriented entire process that runs largely automatically.

The approach covers a comprehensive workflow from the creation and management of valid base data to a usable, feature-rich virtual spatial model. It synergistically combines methods and techniques from geoinformatics (mostly semantic schematisation) and computer graphics (usually geometric reconstruction). The approach is exemplified using the real urban agglomeration area, East-Jersualem, in the Middle East, which lacks reliable base data (such as reference and training data). The workflow focuses on 1) the generation and handling of an individual reference data set, 2) the object-specific extraction from invariant point clouds as well as the semantic enrichment of extracted objects in a common database and 3) the application-related visualisation of those data and objects using developed visualisation functionalities in a virtual spatial model.

This approach is primarily characterised by comprehensively addressing all subareas in one workflow, from the data to the final model. Comparable work of this kind and scope had not previously been carried out. Another new aspect of this approach is the ability to derive a virtual spatial model from unorganised, unstructured point clouds, which serve as a highly precise geometric reference data set for reconstruction. Based on this, objects (for example buildings) are semantically enriched, to be abstracted and modelled according to the application. This allows (in the medium term) for the possibility of an application-related and ad hoc derived virtual spatial model and constitutes a clear contrast to established modelling methods such as the discrete level-of-detail concept.

Automatic conformity and interoperability tests for railway infrastructure

Arne Boockmeyer (arne.boockmeyer@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Andreas Polze

Germany has a widely spread and modern railway system, but the number of infrastructure manufacturers is really small. This is because of the usage of proprietary communication standards which makes it hard for new manufacturers. A new European standard from the EULYNX-consortium should solve this problem by defining an open standard for all infrastructure elements. The downside of the idea is the need for stronger conformity and interoperability testings. Right now, the testing happens manually by employees of the DB Netz. This approach does not scale enough for future challenges.

To close this gap, we are suggesting simulation-based conformity and interoperability tests. Therefore we started together with other researchers with the development of the simulation-based test framework marvis. Marvis combines techniques like co-simulation, virtualization, and the integration of hardware to execute realistic tests of distributed applications. The network simulator ns-3 and the traffic simulator SUMO are both integrated to constitute a simulation-based test environment. As hosts for the distributed application virtual machines or containers or real infrastructure elements can be used. To have a large set of available devices, we are going to connect our IoT-Lab, which already contains real railway infrastructure devices, with other laboratories in research institutes (like the DLR) and at the manufacturers. Besides the environment and the test-execution environment provided by marvis, also digitalized test case catalogs and evaluation techniques are necessary.

One big challenge is the configuration of the devices. Right now, most of the devices are configured in some static way (like configuration files on memory cards). To be able to run automated tests, this configuration must be exchangeable without manual reconfiguration. Together with students, we are working on solving this challenge with programmable memory cards. Besides the configuration also the test case catalog is a big challenge. Right now, all test cases are written in a test case catalog, but mostly in a natural language. Another step is to define a domain-specific language to describe these tests in a machine-readable format. In the end, the results must be evaluated.

In a first experiment, we showed that an axle counter product from Frauscher can communicate with an interlocking even having an artificial delay of 80ms on the connection. This dependability test is required by the standard (50ms are allowed). This experiment used marvis as the test framework.

Right now, we at the HPI and the colleagues from DB Netz are at the beginning of the realization of this vision. As a first step, we are going to research-related work, like current approaches to have simulation-based conformity and interoperability tests, to improve our approach. In the end, we will create a software architecture for the test suite.

Time-Series Analysis and Machine Learning for Read-level Analysis of NGS Sequencing Cycles

Athar Khodabakhsh (athar.khodabakhsh@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Bernhard Renard

Next-Generation Sequencing (NGS) is a technology used for medical genomics diagnostics from clinical samples in order to detect known organisms or discover unknown genomics[1]. However, NGS-based approaches for sequencing and analysis have a high turnaround time from sample arrival to final diagnostics[2]. Recent researches invested a lot in speeding up genomics diagnostics since it is crucial to provide a clinical application with diagnostics before the sequencing has finished and reduce decision-making time. This research aims to apply Time-Series modeling and Machine Learning/Deep Learning techniques[3] specifically, Long-Short Term Memory (LSTM) for earlier detection of present pathogens on sequencing data acquired from NGS cycles in read-level analysis. The results will be used to provide earlier evaluation of medical samples with high accuracy and find the earlier time for genomics diagnostics.

[1] Simon H. Tausch , et al., "PathoLive – Real-time pathogen identification from metagenomic Illumina datasets," BioRxiv (2020): 402370.

[2] Loka, Tobias P., et al., "Reliable variant calling during runtime of Illumina sequencing," Scientific reports 9.1 (2019): 1-8.

[3] Lipton, Zachary C., et al., "Learning to diagnose with LSTM recurrent neural networks," nternational Conference on Learning Representations (ICLR), 2015.

Causal Models of Software Fault Understanding for Sequential Decision Making during Code Inspection Tasks

Christian M. Adriano (christian.adriano@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Holger Giese

Software programmers spend from 20% to 40% of their time searching for the causes of software failures. To alleviate that, automated debugging techniques were developed to reduce the search space to a list of a few suspicious program statements. However, these debugging techniques assume "perfect fault understanding", i.e., that the programmer will always recognize the software fault among the list of suspicious program statements. Because inaccurate fault understanding inevitably happens, programmers waste time generating invalid bug fixes, which in turn undermines the programmers' trust on the debugging techniques.

The goal of this work is to analyze code inspection tasks for the purpose of understanding which factors can predict if a software fault was correctly identified from the perspective of the programmer in the context of software debugging.

We performed two large experiments with respectively 777 and 654 anonymous programmers who executed small, self-contained, independent code inspection tasks. These tasks consisted of answering automatically generated questions about possible relationships between a suspicious program statement and one out of 18 real software failures from various popular open source software projects. The independent variables were the programming ability of the participant and two types of program statements, faulty (experimental condition) or not faulty (control condition). Participants with different programming abilities were randomly assigned to one of the two conditions.

We uncovered a set of factors that can predict the accuracy of fault understanding. Factors combine both programmers' attributes (coding ability, profession, years of experience) and the outcomes of their code inspection tasks (perceived difficulty, confidence, duration, explanations provided). To confirm and refine the prediction factors, we built two causal models: programmer qualification model and task inspection model. This two-stage causal model guarantees that the inferences about participants programming ability are understood and explained before we use this information as input to the second causal model, which in turn explains the accuracy of the code inspection tasks. We applied these causal models to build an algorithm the minimizes the number of tasks needed to identify bugs high accuracy. The algorithm extends a contextual multi-armed bandit method to make sequential decisions about which tasks to generate next based on the answers of previous tasks. Our results allowed to locate all faults with more than 90% precision while requiring only 20% of total available tasks.

Voice-based Interactions for Editing Text On The Go

Debjyoti Ghosh (debjyoti.ghosh@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Patrick Baudisch

My research explores voice-based text interaction techniques for mobile (on-the-go) computing. Mobile text-based tasks are typically performed by typing on mobile phones which consume the users' full visual attention, limiting their awareness of the surrounding. Further, typing while walking is a dual-task walking scenario, hence, difficult to perform as typing requires precise motor movements. My research is motivated by the vision of an interaction paradigm where computing is more seamlessly integrated with the users' everyday mobility. In this interaction paradigm, speech input is likely to play a pivotal role as the interaction is hands- and eyes-free, leaving the user free to engage in other tasks. Additionally, speaking is a natural form of human communication and offers an untethered and device-independent channel of communication with a computing interface. Towards this end, my research investigates speech input as the primary modality for performing text input as a representative everyday mobile computing task, with the goal of lowering the interaction burden so that it can be efficiently and safely embedded into the users' everyday mobility and activities. In particular, the scope of my research is text editing — text entry is convenient to perform using speech (speech being natural and at least 3 times as fast as typing on a physical/virtual keyboard), however, editing the dictated text using speech alone is a difficult task to solve as it requires spatial referencing within the text to delimit where and how much of the text needs to be changed. Speech as a linear input modality is not naturally capable of doing spatial tasks. My research unlocks such capabilities and reimgaines voice-based text input. As a primary contribution, my research establishes a robust voice-based interaction technique for editing typed or dictated text on the go. The proposed technique is based on a methodical exploration of the characteristics of speech input and a detailed understanding of the natural user behavior of voice-based text editing. I propose a voice-based interaction technique that allows the user to correct either by using: (1) commands that might be simple and canonical like "Change jumped to jumps", or conversational-styled like "Could you please change the word jumped to jumps?"; or, (2) by re-dictating over an erroneous portion of the text to correct it, for example, to change 'jumped' to 'jumps' in the text snippet, 'the quick brown fox jumped over the [...]', the user can say, "quick brown fox jumps", or "fox jumps" or simply "jumps". The user can use either of the two techniques interchangeably without the need for changing modes. The scope of my proposed interaction technique extends to both eyes-free and visual interfaces, thereby, opening up a whole new research avenue in voice-based user interfaces where text-based tasks can be feasibly performed using voice as the primary input modality, and hence, easier and more convenient to perform on-the-go.

Discovering Business Process Architectures from Event Logs

Dorina Bano (dorina.bano@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Mathias Weske

The business process management [1] lifecycle deals with the discovery, modeling, executing and analysis of business processes in a perpetual repetition in order to accommodate the ever-changing business requirements. An important artifact in this context is a process model repository, which often captures hundreds of models. With the increase in size and complexity of process model repositories it gets harder to manage them. The area of BPM addressing such challenge is called Business Process Architecture (BPA) [2].

Typically, BPAs are designed using the relationships between process models in a given repository, i.e., BPA design is purely based on process models. In this paper we argue that these relationships might not be the ones that actually occur during business process executions and, hence, do not reflect their real-world relationships properly.

The introduced approach is based on two types of process relations, trigger flow and information flow. Trigger flow represents situations when a process triggers the instantiation of another process. Information flow captures data exchange between processes. The resulting approach is able to discover complex process interdependencies that have occurred in process executions; we extend an existing BPA graphical representation to accommodate these relationships properly. The approach is applied on two real-life set of event logs to prove its feasibility and effectiveness.

[1] Weske, M., "Business Process Management - Concepts, Languages, Architectures" Third Edition, Springer, 2019

[2] R.M. Dijkman and I.T.P. Vanderfeesten and H.A. Reijer. "The road to a business process architecture : an overview of approaches and their use", Technische Universiteit Eindhoven. BETA Working Paper, 2011

Detecting Layout Templates in Multiregion Spreadsheets

Gerardo Vitagliano (gerardo.vitagliano@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Felix Naumann

Data is the new oil: as more and more data is available and has been openly released, for example by governmental sources, a large number of structured files is available for public consumption. However, large collections of raw data files are often lacking an appropriate structure, requiring cumbersome, hand-crafted, and error-prone data preparation.

In our work, we focus on a specific class of data files, "multiregion" spreadsheets.

These files contain useful data, but rather than containing one single, properly formatted table, their data is arranged in a canvas-like fashion with a custom layout and no single proper tabular format. What is more, often multiregion layouts repeat themselves following the same template, which are not trivial to recognize as tables may have different shapes due to different number of rows, different missing values, etc. To assist practitioners with the large-scale data preparation of spreadsheet datasets,

We designed an automated approach, called Mondrian, to tackle region detection and layout inference in multiregion spreadsheets. First, the content of a spreadsheet is mapped to the image domain, by converting its cells into pixels, according to their content. If two files are found to have at least a pair of regions with the same fingerprint, they are selected as possible instances of the same template.

To compare file layouts for template recognition, each layout is encoded as a complete graph, with nodes encoding the information about regions and edges representing their connectivity. Layout graphs are finally compared with a graph similarity measure based on the similarity-flooding algorithm, and files whose graphs show a high similarity score are assigned to the same template. An experimental evaluation conducted on different real-world spreadsheet datasets with above 1,500 files, shows that Mondrian is very effective in dealing with region extraction and template recognition in multiregion spreadsheets.

Towards Joint Design-Time and Run-time Verification of the Complex System

He Xu (he.xu@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Holger Giese

When we concern about designing a complex system, especially those who interact with the environment and other systems a lot, e.g. a self-driving car system, the most crucial problem is how to guarantee assurances during operation. Run-time internal failures and external hazards may lead to severe consequences and some of them may lead to fatal accidents.

Indeed, for some systems, we can prove their correctness at design time, but for complicated systems, that would be extremely expensive, and sometimes it's impractical. For instance, acquiring an accurate model at design time for some systems is infeasible. Specific to the self-driving car system, we cannot anticipate all uncertainties in the environment and system at design time, and the system may also change its behavior (e.g., update its model) at run-time, which means in most cases there is no way to design a absolute safe system.

In this approach, we develop a joint verification method, which combines advantages of both design time and run-time verification. That is at design time the verification process will do as much as possible to ensure the safety, reliability, and robustness of the system. These works will cover the most resource and time consumption parts of the verification of the system. At run-time, the system will monitor the system itself and the context. When potential risks have been detected or the prediction of hazards occurs, the system will invoke the coping mechanism.

There are two checking methods correspond to run-time and design time verification. The backward bounded model checking will be used at design time to establish the unsafe areas around the inevitable unsafe states and to help engineers pick the proper countermeasures for each of these failures. This checking process can also help to analyze the uncertain and rare adverse events and conditions that may or may not happen in the operation of the system. Unexpected situations can be analyzed offline using backward checking, and build specific unsafe areas that include these unsafe states and related system states. After that, these unsafe areas can be updated to the operating systems, and systems can react to these new threats.

At run-time, the forward checking process only searches to states that are reachable from the current state within a fixed number of transitions. Once it detects the boundary of the unsafe area, it can execute the emergent mechanism to cope with the adverse situation. Combining with results from design time backward checking, the run-time forward checking can hence give the system enough time to prepare for potential failures or hazards. At the same time, it will reduce the time and resource consumption of time and resource at run-time.

Reactive and Proactive methods for failure analysis in Microservices Architecture

Iqra Zafar (iqra.zafar@uni-potsdam.de)
Supervisor: Prof. Dr. Holger Giese

Microservices (or microservices architecture) are cloud native architectural style that composes many loosely coupled and independently deployable smaller components, or services that are separated via bounded context in a single application. These services have their own database and communications with each other over a combination of REST APIs, event streaming, and message brokers. Microservice Architecture appears poised to replace SOA as the dominant industry architecture. Number of internet applications are using this approach due to its flexibility and clear logic. The stability of microservice is thus vitally important for these applications' quality of service.[1] The performance quality of microservice is of vital importance to the Internet company, because a microservice failure can degrade the user experience and bring economic loss. Therefore, an efficient failure detection is needed, which enables rapid service recovery and loss mitigation, becomes increasingly more important for microservices.[2]

This thesis aims for failure analysis in microservices based systems.Reactive and proactive failure analysis method are the major techniques for handling open challenges in these systems. Our aim to design an automotive framework to robustly localize the anomalous behavior in a microservice and mitigate the effect of failure using reactive and proactive methods. This research focuses on proposing a mechanism that can continuously observe and monitor the microservices architecture and be able to detect anomalous behaviour with high accuracy and generate low rate of false alarms. At the same time, this mechanism should be able to respond to true positive alarms by suggesting a set of adaptation policies (adaptation strategy), that can be deployed in the cluster to achieve high level of self-healing in response to changes in its operating environment. The envisioned property of this mechanism is that it can be easily deployed with fewer and smaller footprints on the limited resources found in the tiny containers running in a microservices cluster.

[1] A. Samir and C. Pahl. , "Dla: Detecting and localizing anomalies in containerized microservice architectures using markov models," 7th International Conference on Future Internet of Things and Cloud (FiCloud). IEEE,pages 205–213, 2019

[2] M. Ma, J. Xu, Y. Wang, P. Chen, Z. Zhang, and P. Wang. "AutoMAP: Diagnose Your Microservice-based Web Applications Automatically". In: Proceedings of The Web Conference 2020, pages 246–258.,2020

Process Mining in Healthcare

Jonas Cremerius (jonas.cremerius@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Mathias Weske

Today, healthcare organizations face the challenge of increased care needs and tightening budgets. This problem motivates these organizations to look at their processes and identify improvement possibilities to save money and provide better outcomes for patients. However, healthcare processes tend to be loosely framed and knowledge-intensive, including lots of stakeholders. This makes a manual recreation of processes difficult, resulting in an individual process instead of a general one for all patients.

Therefore, process mining uses real-life process execution data from health information systems to create process models. As process mining takes real-life data into account, it can provide evidence-based insights about a process of a specific patient group. This enables the confirmation of expected behavior and reveals novel insights about the treatment of patients.

So far, process mining creates process models showing the control flow and activities performed during the process. However, getting insights about a process is not only based on that, but also on the data generated and manipulated during process execution. Especially in healthcare, a tremendous amount of data is generated within one hospital stay, such as lab values, image analyzes results, vital sign recordings etc. Thus, we researched ways to include data in process models mined by process mining and created data-enhanced process models[1]. These allow attaching domain data to process activities, enabling monitoring of the behavior of domain data directly in the process model. Furthermore, activities can be compared based on their domain data, such as their resource efficiency.

Further work in this PhD includes ways to improve process mining in healthcare by looking at the interaction between domain experts and process analysts when analyzing a process together. Additionally, we will build upon the idea of data-enhanced process models and create a method to recommend attributes that are interesting to observe within a process model.

[1] Jonas Cremerius, "Towards a Framework for Data-Enhanced Process Models in Process Mining", ZEUS Workshop, ZEUS 2021, 2021

Estimation of Subjective Ratings of Perceived Exertion using Inertial Measurement Units, Electocardiogram and Computer Vision

Justin Amadeus Albert (justin.albert@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Bert Arnrich

Intensity quantification is an essential aspect of weight training to adjust training routines and prevent injuries. Usually, training intensity can be assessed objectively, such as measuring the lifted weight. On the other hand, subjective measurements such as questionnaires or ratings of perceived exertion (RPE) can also determine important information from athletes. The Borg scale is a standard RPE scale, ranging from 6 to 20. Six refers to minimal exertion, and 20 refers to maximal exertion. Usually, athletes report their RPE values after an entire training session (session RPE). There is, however, interest in getting RPEs on the fly. A study published in 2015 has applied Support Vector Machines (SVM) to estimate RPEs during upper-body resistance training exercises[1]. They have used a full-body Inertial Measurement Units (IMU) setup. A more recent study [2] aimed to predict the subjective fatigue levels during three exercises. Subjects have performed these exercises until exhaustion while being recorded with IMU sensors and force plates and reported an RPE value. They have utilized random forest (RF) regression and convolutional neural networks (CNN) to predict the experienced fatigue scores.

To estimate RPE values from athletes performing physical exercise, we record data using IMU sensors, electrocardiography (ECG), and RGB-D cameras. Subjects perform squats for a given number of sets while being recorded with six IMU sensors placed on the lower- and upper body. Further, we calculate heart-rate variability (HRV) from the ECG signal. We obtain hand-crafted data features in the time and frequency domains. After obtaining those features, we utilize machine learning methods such as SVM or K-Nearest-Neighbors (KNN) to estimate the subjects' RPE values. Primarily we will identify the most meaningful features from IMU and Kinect sensors and investigate the impact of HRV features. In addition, we will utilize Deep Learning for RPE prediction. Therefore, we will investigate multi-branch network architectures to feed in data from our three sensor modalities and spectrogram images as input for a CNN. Overall, this will allow for a comparison between Deep Learning versus traditional machine learning approaches.

[1] Pernek, I.; Kurillo G.; Stiglic G.; Bajcsy R., "Recognizing the Intensity of Strength Training Exercises with Wearable Sensors." Journal of Biomedical Informatics, 11., 2015

[2] Jiang, Y.; Hernandez, V.; Venture, G.; Kulić, D.; K. Chen, B. "A Data-Driven Approach to Predict Fatigue in Exercise Based on Motion Data from Wearable Sensors or Force Plate." Sensors 21, 1499, 2021

Multi-Tenancy in FPGA Accelerator Designs

Lukas Wenzel (lukas.wenzel@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Andreas Polze

With diminishing performance gains in classic general-purpose compute devices, heterogeneous system architectures are receiving an increased amount of attention. A heterogeneous system combines general-purpose compute devices with with accelerators, which are special-purpose devices designed to execute a limited set of workloads more efficiently than a purely general-purpose system could. Their specialization allows accelerators to reduce overheads and accomodate the particular computational structures of a workload class, but it is also the root of some practical problems. Development and production of custom compute devices is a very costly enterprise, so an accelerator must suit a sufficiently broad workload class to be it economically feasible. Field Programmable Gate Arrays (FPGAs) offer a compromise solution between customization and mass-production, by implementing a fabric of reconfigurable logic primitives, that can be adapted to realize arbitrary compute structures. Thus FPGAs offer a high degree of specialization and thus superior performance and energy consumption for suitable workloads.

These characteristics make FPGAs valuable components in heterogeneous system architectures. Even large scale compute environments like datacenters or clouds follow the trend towards a heterogeneous mix of compute resources, but in these particular scenarios FPGAs impose some challenges. As the specific hardware microarchitecture of an FPGA is not defined by system designers but application developers and might even change at runtime due to reconfiguration, many common assumptions about resource sharing and isolation can not be guaranteed. Therefore current approaches usually allocate the FPGA exclusively to a single workload at a time, sacrificing some optimization potential on the FPGA and limiting the overall throughput of the system which might process a large number of concurrent workloads. Consequently it would be desirable to enable the FPGA resources for multi-tenant operation.

In my current work on FPGA multi-tenancy, I approach the question from a throughput point of view. The alternative security-focused perspective has also received some recent interest. Works in this field show, that due to the possibility of side-channel attacks, the required complex isolation mechanisms are likely to outweigh any multi-tenancy benefits. Instead I concentrate on a scenario where tasks share the FPGA without malicious intent.

To build an FPGA design accomodating more than one task requires special care from application developers, as it requires efficient mechanisms to manage state between different task contexts and schedule access to FPGA resources. In this situation I propose an FPGA overlay which encapsulates the required isolation logic and allows developers to focus on the actual application design.

A Transformer based approach for Entity Linking

Manoj Prabhakar Kannan Ravi (manoj.prabhakar@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Gerard de Melo

Entity Linking (EL) is a compelling task in natural language processing for understanding the semantics and extracting information from text. The task is to link the entity mentions in an unstructured text to their corresponding entities in a knowledge base/knowledge graph. Wikidata is one such knowledge graph which is mainly crowd sourced and has several challenges due to its noisy, non-standard entity titles. Currently, there are very few approaches that try to solve the EL problem over wikidata which has a few shortcomings due to their outdated methods. We propose a transformer-based model to target end to end EL over wikidata. It is a context aware approach that feeds global context from an external knowledge source. We empirically illustrate our model on datasets, where it significantly outperforms on various baselines and also the role of external knowledge in the transformer-based models for EL over wikidata.

Efficient Approximation of Partition Functions in Statistical Physics

Marcus Pappik (marcus.pappik@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Tobias Friedrich

Spin systems are the common mathematical model for particle systems in statistical physics. Roughly, such spin systems can be divided into two categories: discrete and continuous systems.

Discrete spin systems are usually defined based on an undirected graph $G = (V, E)$ and a finite non-empty set of spins Q. Vertices model spatial positions that might be occupied by particles, edges represent possible interactions between those positions and the spins are used to represent possible states of the vertices in the graph. The set of possible states of the spin system consist of all assignments of vertices to spins $\sigma : V \to Q$. Further, every such state σ gets a weight $w(\sigma)$, encoding the energy of the system in the respective state. The so called Gibbs distribution now assigns each state σ a probability that is proportional to its weight $w(\sigma)$. Another important quantity of such a spin system is the partition function which sums the weights of all states of the systems (i.e., the normalizing constant of the Gibbs distribution).

Analogously to those discrete models, continuous spin systems are defined based on some continuous domain, for example a suitable region in d-dimensional Euclidean space $\mathbb{V} \subseteq \mathbb{R}^d$, and a set of spins Q. Possible states of the system are all assignments of the points in $\mathbb{V}$ to spins in Q. Similar to the discrete case, the Gibbs distribution and the partition function are now defined based on a weight function.

Whereas physicist usually ask for properties of those models such as the presence of phase transitions, there has been an ongoing effort within recent years in computer science to investigate their computational aspects. Especially, two computational tasks are of main interest: (approximate) computation of the partition function and (approximate) sampling from the Gibbs distribution. As it turns out, these computational aspects are not only closely connected to each other, but they also relate to fundamental open questions in complexity theory and statistical physics.

Whereas relationships between physical and computational phase transitions are known precisely for a variety of discrete models, hardly anything is known about computational properties of their continuous counterparts. In our work, we carefully discretize continuous spin systems and adopt different algorithmic techniques from the discrete domain, resulting in rigorous computational results for such continuous models.

Nesting Laser-Cut Objects for Fast Assembly

Muhammad Abdullah (muhammad.abdullah@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Patrick Baudisch

My research focuses on developing a software tool called Roadkill that converts 3D models to 2D cutting plans for laser cutting—such that the resulting layouts allow for fast assembly.

Roadkill achieves this by putting all relevant information into the cutting plan: Thumbnails indicate which area of the model a set of parts belongs to. Parts with exposed finger joints are easy to access, thereby suggesting to start assembly here. Openings in the sheet act as jigs, affording assembly within the sheet. Users continue assembly by inserting what has already been assembled into parts that are immediately adjacent or are pointed to by arrows.

Roadkill maximizes the number of joints rendered in immediate adjacency by breaking down models into "subassemblies." Within a subassembly, Roadkill holds the parts together using break-away tabs. Users complete subassemblies according to their labels 1, 2, 3..., following 1 -> 1 links to insert subassemblies into other subassemblies, until all parts come together.

In our user study, Roadkill allowed participants to assemble layouts 2.4 times faster than layouts generated by a traditional pair-wise labeling of plates.

Personal Small-batch Production

Shohei Katakura (shohei.katakura@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Patrick Baudisch

The goal of my research is to enable non-industrial designers or initial hardware startups who don't have mechanical engineering knowledge to produce 500 - 1000 of their products.

Today, we can design using CAD software and prototype our hardware using 3D printers, laser cutters, etc. 3D printers in particular have few restrictions on manufacturing, allowing people without manufacturing knowledge to quickly fabricate objects they have designed. While there are few problems if we only make a one-off prototype, once we try to mass-produce it, non-expert users encounter the following issues during the 3D modeling and design phase.

- Managing material consumption
- Addressing manufacturing processes
- Catering for machine-specific characteristics
- Designing for easy assembly

In the industrial domain, a professional engineer refines the prototype to be mass producible, the process is called design for manufacturing and assembly (DFMA). This process is critical for mass-production from a cost/manufacturing perspective and is considered early in the design process. However, it is difficult for non-experts to carry out this refinement as it requires a lot of domain knowledge.

I address this by developing a design system with a software agent. In this design system, the user has authority over the function and shape of the product while the agent has authority over the cost of the materials and the manufacturing process. Users can collaborate with the software agent to create products and custom-made tools, facilitating small-batch production.

Self-prediction of epileptic seizures by affective computing using brain activity sensor

Sidratul Moontaha (sidratul.moontaha@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Bert Arnrich

Epilepsy is the most common neurological disorder characterized by unprovoked and unexpected electrical bursts in the brain, which results in seizures. Though most epilepsy treatment lies on anti-epileptic drugs (AEDs), 30% of these patients possess treatment-resistant epilepsy. Some of these patients are capable of reporting self-prediction of their seizures by observing their affectivity. Some patients report no signs of feeling premonitory symptoms, prodromes, or aura. Several studies reported stress as the most frequent seizure trigger. Moreover, affective states such as anxiety, anger, depressive mood, and other negative emotions also correlate with seizures. However, it is crucial to objectify the patients' subjective feeling to provide bio-feedback to prevent seizures long before it occurs. Our approach is to implement such a sensor-based technology that will record physiological parameters and ecological momentary assessment (EMA). Currently, we are focusing on brain activity sensor, i.e., Electroencephalography (EEG) sensor to classify stress and emotion. Though video-EEG recording is widely used for seizure detection in hospital settings, our approach uses wearable EEG device which will associate a pre-emptive therapy to reduce seizure frequency or to eliminate seizure occurrence. To support the idea of self-prediction of seizures by affective computing, we have published a concept study in 2020 [1] which provides an overall picture of the related works in this domain. Additionally, many researches already found the correlation between EEG and emotions under cognitive load and also found that the prodromal symptoms can predict the pre-ictal states of the brain by analyzing video-EEG. Additionally, studies found a significant statistical correlation between EEG data and the patient-reported questionnaires on seizure provocative factors i.e., sleep deprivation, emotional stress, negative emotions, and so on. Therefore, initially we did a survey on the available wearable EEG devices in the market and chose one low-cost wearable EEG device which provides better signal quality and without significant connection issues in the resting state. After significant testing of the device and an established pre-processing pipeline, our following approach is to conduct laboratory experiments with a healthy cohort where participants will wear two devices:EEG sensors and photoplethysmogram (PPG) sensors. The purpose of including the PPG sensor is to include multimodality in affective computing. Eventually, this system will provide objective information on the patients' affectivity by doing statistical analysis, signal processing, and machine learning on the available data from the sensors and EMA.

[1] Self-prediction of seizures in drug resistance epilepsy using digital phenotyping: a concept study

Shortest Path Enumeration

Stefan Neubert (stefan.neubert@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Tobias Friedrich

When analyzing a problem's complexity, one is usually interested in algorithms that minimize total computation time for large inputs, and in matching lower bounds that rule out faster algorithms. However, most lower bounds are based on conjectures such as the All-Pairs-Shortest-Paths Hypothesis, which claims that there is no ε such that APSP can be solved in $O(n^{3-\varepsilon})$ time.

In current day systems, the input size n tends to be that large, that such an algorithm with cubic theoretical runtime leads to enormous wall-clock waiting time. In addition, systems rarely consist of a single data processing step, but are made of multiple algorithms run in series on potentially many machines. If one of those algorithms takes a lot of time to produce its output, the rest of the pipeline is stalled, and computing power is unused.

To solve this issue, we try to solve problems by enumerating small parts of the solution quickly. Even though this cannot improve the total runtime needed to solve a problem, following steps in the data processing pipeline can work on partial outputs long before the whole solution is provided and by that cut down on the overall runtime of the pipeline.

In our first work in this area, we analyzed the fundamental problem of computing the shortest distances between vertices in weighted and unweighted graphs. We proposed basic terminology and developed suitable techniques for analyzing algorithms and problems in the enumeration setting and now build on this in three follow-up projects:

Firstly, we apply the developed techniques to other problems. Specifically, we are currently interested in scheduling problems: Given a set of jobs to be executed, a schedule enumeration algorithm is expected to assign jobs to machines after little preprocessing time and with small delay, such that machines can start processing jobs even before the complete schedule is computed. The schedule itself has to optimize an objective function. We developed an algorithm that solves the $F2||C_{\max}$ scheduling problem in this setting with linear preprocessing time (which is optimal) and logarithmic delay, thereby matching the total time of the reference algorithm by Johnson.

Secondly, we develop a formal framework for our enumeration setting that provides precise terminology to define and compare problems and is able to capture requirements such as filtered or sorted enumeration and the relation of these problems to their total time counterparts.

Thirdly, within this framework we are working on comparing and linking problems with reductions that transfer lower- and upper complexity bounds in the same way as fine grained reductions do in the total time setting.

Intrinsic Images for Enhanced Neural Style Transfer

Sumit Shekhar (sumit.shekhar@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Jürgen Döllner

Artistic style transfer is a well established problem in computer graphics. It aims to transfer the style of a reference photo/artwork onto another picture. To this end, in the past few years, Neural Style Transfer(NST) has emerged as an efficient tool for transforming an image into a particular style. The basis of it being the separation of deep neural-network image features into content and style. Various transformations have been proposed to the naive NST using additional depth information, by controlling perceptual aspects, controlling stroke parameters, through semantic segmentation etc.. Mostly the NST techniques are explored in the context of non-photorealistic artistic rendering of images. However, such ideas can also be extended in the photorealistic domain. For a photorealistic style transfer the mere separation of deep image features into content and style is not sufficient since we also need to ensure the photorealism of the stylized output. For this purpose we add a photorealism regularizer in the loss function and also augment the style loss with semantic segmentation. The regularizer preserves the structure of the input image by making sure that the image transformations are locally affine in nature. However, neither the non-photorealistic nor the photorealistic approach exploit the physical process of image formation for the purpose of style transfer.

To determine all the physical properties associated with image formation a full inverse rendering in three dimensional space is required. One way to relax this highly ill-posed problem and estimate only few of scene properties in image space is through intrinsic decomposition, which generally refers to the problem of estimating scene characteristics, such as albedo and shading when one view or multiple views of a scene are provided. For this work we also consider scene-properties of specularity, depth, and normals – as intrinsic layers. In case of a rendering pipeline we can obtain the above intrinsic layers as intermediate results. For a real-world photograph we make use of existing techniques to extract these layers given an input image. As part of this work we aim to use the intrinsic layers of source image and naive stylized output to address the challenges associated with NST in a post-processing step. Moreover, we also enable further enhancements – such as relighting, specularity enhancement, using the intrinsic layers. We further propose to extend these techniques to enhance classical stylization approaches such as toon, oil-paint, and water-color.

Efficient Block-based Programming

Tom Beckmann (tom.beckmann@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Robert Hirschfeld

Block-based programming environments have been shown to support programming novices in writing programs. Their directness and guarantees for well-formedness remove common hurdles faced by novices. However, block-based editors are often perceived as too cumbersome or limited to use by professional programmers as they are typically designed for beginners.

In our research, we are investigating the use of block-based programming environments in the context of professional programming. We believe that there are benefits inherent to block-based interfaces for programming tools since blocks can provide a direct mapping between powerful domain abstractions and interactive program elements beyond the level of textual representations. We also investigate block-based polyglot programming to simplify language composition by hiding syntactic idiosyncrasies and focusing on semantic properties of program elements instead.

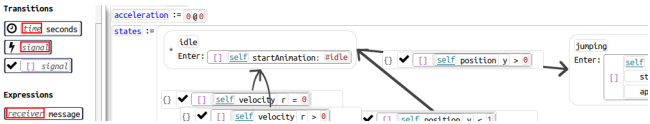


Figure 1: A statemachine embedded in block-based Smalltalk code.

Currently, we are exploring keyboard-directed workflows with blocks that are compatible with direct manipulation interfaces typical to block-based editors. So far, we identified two factors contributing to efficient keyboard-based workflows in block editors: First, a set of key combinations directly expresses most restructuring operations needed when programming. Second, a system we call grammar-assisted input lets programmers enter program elements via the keyboard in the same linear manner when programming in text. The grammar-assisted input system interprets each keystroke based on the context of the active block and transforms the program according to the programmers' input.

Our prototype of a block-based programming environment supports Smalltalk, as well as subsets of JavaScript and Scheme. Further, it implements several visual and domain-specific languages that tightly integrate with their general-purpose host languages, as shown in figure 1.

In a preliminary user study, we found that programmers were able to start using our editor after only little training and needed fewer keystrokes to perform some programming tasks compared to working with contemporary text editors. However, for competing with text editors in task completion time, we expect programmers to need more practice.

Towards Medical Decision Support Systems

Weronika Wrazen and Felix Grzelka
({weronika.wrazen,felix.grzelka}@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Andreas Polze

Nowadays, we can observe a trend that people are being increasingly affected by various diseases and need frequent or even constant medical care. However, that growth is not proportional to the number of available clinical personnel. To solve that issue Medical Decision Support Systems (MDSSs) are introduced as a support not only for medical staff but also for patients. MDSSs are software tools that assist doctors, nursing staff, and physicians in making clinical decisions. They are designed to help to analyze patients' health data and draw conclusions about their health state. They can also decrease the time needed to analyze one patient, which results in higher efficiency of clinical staff. MDSSs can be based on statistical models, like Neural Networks, which are trained on patient data. Alternatively, they can be devised using rules defined by medical experts.

Our research interest lies in combining these two approaches and implementing them in different medical institutions, such as hospitals or nursing homes. Currently, we are working on two projects:

- Telemed5000[1] – conducted in collaboration with Charité, Berlin. The main goal is to develop an MDSS to predict the daily risk of a critical condition for patients with cardiovascular diseases (CVDs).

- Wisemat[2] – conducted in collaboration with Getemedx6. The main aim is to device a mat with pressure sensors, which allows nursing staff to check past and current bed patients' positions, and define the risk of pressure ulcers' occurrence.

Felix will focus his research on the robustness of machine learning systems. So far, building reproducible and reliable ML systems has been hard because big datasets and experiments need to be managed alongside changing code. His research will focus on applying best practices of software engineering such as version management, reproducible builds, and testing to machine learning systems. These practices will ensure high quality standards, so that machine learning based MDSSs can be certified as medical products.

Weronika's spotlight is to devise machine learning models, which will be able to predict and estimate medical events. For medical databases, commonly the number of observations is limited. Usually, not only the number of events per class but also the number of events per patient are non uniformly distributed. This makes it difficult to design a reliable model with satisfactory accuracy. She will work on developing universal methods and approaches which overcome these challenges.

[1] https://telemedizin.charite.de/forschung/telemed5000/
[2] https://www.getemed.de/en/getemed/research/wisemat

Graph Separators and its Applications

Ziena Elijazyfer (ziena.elijazyfer@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Tobias Friedrich

A network graph is a mathematical model consisting of vertices and pairs of vertices, called edges, that can describe structures in nature and technology. They have a lot of practical applications and the research on graph theory is extensive. Graph separators in form of vertices or edges play an important role in this theory and until now they are not completely understood. As the name suggests separators are vertices or edges which separates parts of the considered graph, or equivalently, its removal breaks the graph into several parts (cf. Figure 1). In a lot of real-world

Figure 1: Vertex and Edge Separators marked in red, respectively.

problems like security, supervisory control, and epidemics separators can be used depicted as significant points in a network. Moreover, separators allow solving many graph optimization problems efficiently by using them to design divide-and-conquer strategies or realize parallel processing algorithms. This generality and their wide applicability has made the study of separators a rich and active research field. Generally, scientific progress on separators is highly valuable, since they have the ability to decode structures in graphs. It is interesting for me to understand its relation to other problems and how structures resulting from separators give the possibility to deal with complexity-wise hard problems. For instance, our project group deals with connected subgraph partition problems which demand certain sizes on the searched subgraphs, and through bounded vertex separators (vertices whose removal leads to small connected components) we achieve first constant approximation results for some of these kind problems. Connected subgraphs are one of the most natural structures to encode aspects of a practical task, modeled as a graph problem. On the one side, such subgraphs represent structures we seek to discover, such as territories for postal delivery and similar districting problems. From another perspective, the structures of interest could be operations that scatter a graph into small connected components; a structure e.g. used to model vulnerability in network security. Partitioning a graph into connected components of a given size are also used as a model for task allocation to robots.

Author Index