SCHLOSS DAGSTUHL
Leibniz-Zentrum für Informatik

Dagstuhl News

January - December 2010

Volume 13
2012

Leibniz Gemeinschaft

Dagstuhl Publications
Dagstuhl News January-December 2010
http://drops.dagstuhl.de/opus/volltexte/2012/3461

ISSN 1438-7581

Schloss Dagstuhl, the Leibniz Center for Informatics is operated by a non-profit organization. Its objective is to promote world-class research in computer science and to host research seminars which enable new ideas to be showcased, problems to be discussed and the course to be set for future development in this field.

Period:	January - December 2010
Frequency:	One per year
Online version:	http://drops.dagstuhl.de/portals/dagstuhl_news/
Associates:	Gesellschaft für Informatik e.V. (GI), Bonn Technical University of Darmstadt University of Frankfurt Technical University of Kaiserslautern Karlsruhe Institute of Technology (KIT) University of Stuttgart University of Trier Saarland University Max Planck Society e.V. (MPG) French National Institute for Research in Informatics and Automatic Control (INRIA) Dutch National Research Institute for Mathematics and Informatics (CWI)
Membership:	The Center is a member of the Leibniz Association.
Funding:	The Center receives federal and state funding
Information:	Schloss Dagstuhl Office Saarland University Campus 66123 Saarbrücken, Germany Phone: +49-681-302-4396 E-mail: service@dagstuhl.de http://www.dagstuhl.de/

Welcome

Here are the Dagstuhl News for 2010, the 13th edition of the "Dagstuhl News", a publication for the members of the Foundation "Informatikzentrum Schloss Dagstuhl", the *Dagstuhl Foundation* for short.

The main part of this volume consists of collected resumees from the Dagstuhl Seminar Reports 2010 and Manifestos of Perspectives Workshops. We hope that you will find this information valuable for your own work or informative as to what colleagues in other research areas of Computer Science are doing. The full reports for 2010 are on the Web under URL: `http://drops.dagstuhl.de/opus/institut.php?fakultaet=01&year=10/`
You may be irritated that you receive the Dagstuhl News 2010 in 2012. Well, not all organizers supply their result digests within the requested time.

The extension building with 7 more rooms has been integrated into our operation It allows us to run two Seminars in parallel. It also offers more capacity for meetings of projects, of research programmes etc.

Thanks
I would like to thank you for supporting Dagstuhl through your membership in the *Dagstuhl Foundation*. Thanks go to Fritz Müller for editing the resumees collected in this volume.

Reinhard Wilhelm (Scientific Director)

Saarbrücken, February 2012

Contents

Chapter 1

Data Structures, Algorithms, Complexity

1.1 Circuits, Logic, and Games

Seminar No. **10061** Date **07.02.–12.02.2010**
Organizers: Benjamin Rossman, Thomas Schwentick, Denis Thérien, Heribert Vollmer

Description of the Seminar's Topic

Starting with the seminal paper by Furst, Saxe and Sipser, the last two decades of the previous century saw an immense interest in the computational model of Boolean circuits. Emerging powerful lower bound techniques promised progress towards solutions of major open problems in computational complexity theory. Within a very short time, further progress was made in papers by Fagin et al., Gurevich and Lewis, Håstad, Razborov, Smolensky, and Yao, to mention only a few. The just mentioned result by Furst et al. was obtained independently by Ajtai making use of model-theoretic arguments, and many further lower bounds in complexity have been obtained afterwards making use of inexpressibility results in logic, very often making use of model-theoretic games. After a decade of active research in this direction things slowed down considerably.

In the same way as during the first seminar on *Circuits, Logic, and Games* (Nov. 2006, 06451), the organizers aimed to bring together researchers from the areas of finite model theory and computational complexity theory, since they felt that perhaps not all developments in circuit theory and in logic had been explored fully in the context of lower bounds. In fact, the interaction between the areas has flourished a lot in the past 2-3 years, as can be exemplified by the following lines of research.

Results of Barrington, Straubing and Thérien show that most circuit classes, if they can be separated at all, can be separated by regular languages – which means that *algebraic properties* of such languages could be used in lower bound proofs. Recent results prove almost linear upper bounds on the size of circuits for regular languages in many important constant-depth circuit classes, implying that an $\Omega(n^{1+\varepsilon})$ lower bound suffices to separate such classes from NC^1. Interesting connections to *communication complexity* have been

obtained in the past two years, showing, e. g., that languages with bounded multiparty communication complexity can be recognized by programs over commutative monoids and thus have very small depth circuit complexity.

While inexpressibility results in finite model theory have been used since the 1980s to obtain circuit lower bounds, recent results make use of circuit lower bounds to separate different logics: The result that number of-variable hierarchy in first-order logic over finite ordered structures is strict was obtained by showing that a certain clique-problem cannot be solved by constant-depth circuits of a certain size.

Further connections between logic and circuits concern *uniformity conditions for Boolean circuits*: It was proved that in a quite general context, when a circuit based language class is characterized using first- order descriptive complexity, the circuit uniformity conditions spring up in the logic in the form of restrictions on the set of numerical predicates allowed. So-called "extensional uniformity conditions" have been studied: Intersecting a non-uniform constant-depth circuit class with a uniform class $\mathcal{L}$ (e.g., a formal language class) in some contexts results in a circuit class that can be characterized by first-order logic with $\mathcal{L}$-numerical predicates. (Intuitively, $\mathcal{L}$-numerical predicates are those predicates definable in first-order logic with one "oracle call" to a language from $\mathcal{L}$, i.e., more precisely, with one appearance of a generalized quantifier for such a language.)

While this in principal points out new ways to separate uniformity conditions via logical means, another line of results goes in the opposite direction: It is shown that for a specific arithmetical problem (division), circuits can be constructed that are uniform under much stricter requirements than was anticipated before. Again, the proofs make heavy use of finite model theory.

A further area of investigation is the structural complexity of dynamic algorithmic problems. There are, so far, no techniques available to prove that a problem does not have AC^0 update complexity. Recent (and forthcoming) work therefore started an investigation of the fine structure of the class of problems with AC^0 updates, yielding lower bound results and uncovering (yet another) surprising characterization of the regular languages as those that can be maintained with quantifier-free formulas.

These results demonstrate the impressive growth of interest and activity in the intersection of finite model theory and Boolean circuit complexity, and as will have become apparent from our above description of recent research, many of these developments rely strongly on game-based methods.

Organization of the Seminar and Activities

The workshop brought together 46 researchers from different areas of logic and complexity theory with complementary expertise. The participants consisted of both senior and junior researchers, including a number of postdocs and a few advanced graduate students.

Participants were invited to present their work and to communicate state-of-the-art advances. Twenty-seven talks of various lengths took place over the five days of the workshop. Around half of these talks were scheduled prior to workshop, including most of the longer morning talks and tutorials. The remaining slots were filled as the workshop commenced.

The morning schedule consisted of survey talks of between 60 and 90 minutes. The presenters and topics were:

- Meena Mahajan, Valiant's classes
- Jean-Eric Pin, A topological approach to recognition
- Jouko Väänänen, Dependence logic
- Kousha Etessami, The computational complexity of Nash and other equilibria and fixed points of algebraic functions

There were an additional eight one-hour talks, as well as fifteen shorter talks of 30 or 45 minutes. These talks covered a wide range of topics in circuits, logic, games and the nexus of these areas. The different approaches discussed above in the seminar description were all very well represented by the different talks given during the five days of the seminar.

Concluding Remarks and Future Plans

The organizers regard the workshop as a great success. The weeklong format was well-suited to a workshop of so broad a scope. Bringing together researchers from different areas of theoretical computer science and logic fostered valuable interactions and led to fruitful collaborations. Feedback from the participants was very positive as well.

Finally, the organizers wish to express their gratitude toward the Scientific Directorate of the Center for its support of this workshop, and hope to continue this series of workshops on *Circuits, Logic, and Games* into the future.

1.2 Data Structures

Seminar No. **10091** Date **28.02.–005.03.2010**
Organizers: Lars Arge, Erik Demaine, Raimund Seidel

The purpose of this workshop was to discuss recent developments in various aspects of data structure research, and also to familiarize the community with some of the problems that arise in the context of modern commodity parallel hardware architectures, such as multicore and GPU architectures. Thus while several attendees reported on progress on (twists on) old fundamental problems in data structures — e.g. Gerth Brodal, Rolf Fagerberg, John Iacono and Siddharrha Sen on search tree and dictionary structures, Bob Tarjan on heaps, Kasper D. Larsen and Peyman Afshani on range search data structures, and Peter Sanders and Michiel Smid on proximity data structures — there were also very inspiring presentations on new models of computation by Erik Demaine and on data structures on the GPU by John Owens. The latter presentation was one of the highlights of the seminar, and provided the attendees a good overview over possibilities and challenges in connection with design of data structures for GPU hardware. The seminar was attended by 45 international researchers, resulting in a congenial and productive atmosphere, which resulted in countless discussions and collaborations. The Dagstuhl atmosphere provided just the right environment for all of this.

1.3 Equilibrium Computation

Seminar No. **10171** Date **25.04.–30.04.2010**
Organizers: Edith Elkind, Nimrod Megiddo, Peter Bro Miltersen, Bernhard von Stengel, Vijay Vazirani

The focus of this seminar was the algorithmic problem of computing equilibria in games and market models, viewed from both the theoretical and practical perspective. The equilibrium computation problem is one of the central topics in the rapidly expanding field of algorithmic game theory.

The seminar was a follow-up to Seminar 07471, on the same topic, but with three new organizers, and with a focus on some of the aspects of this problem that received relatively little attention in Seminar 07471. One of the major themes of this seminar was *dynamics*, i.e., exploring the agents' behavior (at both individual and collective level) that leads to the discovery of equilibria, and, more generally, adaptive changes in the collective behavior. Discussed were the classic game-theoretic approaches to this topic (organizer: von Stengel) as well as more recent computational and simulation-based techniques, as studied by the multi-agent community (organizer: Elkind). Another key emphasis was on algorithms and complexity results for *market equilibria* and their applications to Nash Bargaining Games (organizer: Vazirani). We also compared these approaches with computational and geometric aspects of the central Linear Complementarity Problem (LCP) in mathematical programming (organizer: Megiddo). Finally, since the last seminar there was significant progress understanding the complexity of important algorithms, such as strategy iteration, for solving two-player zero-sum games of infinite duration. Progress in this area has strong connections to mathematical programming and was discussed and extended (organizer: Miltersen).

1.4 Theory of Evolutionary Algorithms

Seminar No. **10361** Date **05.09.–10.09.2010**
Organizers: Anne Auger, Jonathan L. Shapiro, L. Darrell Whitley, Carsten Witt

Motivation and Goals

Evolutionary algorithms (EAs) are stochastic optimization methods that are based on principles derived from natural evolution. Mutation, recombination, and selection are iterated with the goal of driving a population of candidate solutions toward better and better regions of the search space. From a more general perspective, EAs are just one instance of bio-inspired search heuristics. Since the underlying ideas of bio-inspired search are easy to grasp and easy to apply, EAs and different bio-inspired search heuristics are widely used in many practical disciplines, mainly in computer science and engineering.

It is a central goal of theoretical investigations of search heuristics to assist practitioners with the tasks of selecting and designing good strategy variants and operators. Due to

the rapid pace at which new strategy variants and operators are being proposed, theoretical foundations of EAs and other bio-inspired search heuristics still lag behind practice. However, EA theory has gained much momentum over the last few years and has made numerous valuable contributions to the field of evolutionary computation.

The theory of EAs today consists of a wide range of different approaches. Run-time analysis, schema theory, analyses of the dynamics of EAs, and systematic empirical analysis all consider different aspects of EA behavior. Moreover, they employ different methods and tools for attaining their goals, such as Markov chains, infinite population models, or ideas based on statistical mechanics or population dynamics. In the most recent seminar, more recent types of bio-inspired search heuristics were discussed. Results regarding the runtime have been generalized from EAs to ACO and PSO. Although the latter heuristics follow a different design principle than EAs, the theoretical analyses reveal surprising similarities in terms of the underlying stochastic process. New analytic approaches were also surveyed.

Theoretical studies of EAs in continuous domain have recently evoked interest of people working in the field of classical numerical optimization. Although stochastic and deterministic optimization algorithms address optimization of different types of problems—mainly convex and smooth for deterministic algorithms and noisy, multimodal, irregular for stochastic algorithms—the focuses of both fields became closer and closer: on the one hand many hybridizations of stochastic search and gradient-based algorithms have been proposed, on the other hand, derivative-free optimization is now a major part of the research in the classical optimization community.

The goals of the 2010 Dagstuhl seminar were twofold. The first goal was to discuss the potential and limitations of a unified theory of all types of bio-inspired search heuristics with focus on how to analyze the runtime of estimation-of-distribution algorithms, which themselves can be considered as a generalized model of ACO, PSO and EAs. The second goal of the seminar was to bridge the gap with the classical optimization community.

Participants and Results

The seminar brought together 45 participants from 12 countries, and from across the spectrum of EA theory. Talks have been arranged into seven sessions grouped around common themes such as runtime analysis, multi-objective optimization, landscape analysis, black-box complexity, continuous optimization. The length of the different slots were between 30-45 minutes.

In addition, shorter talks called *spotlight talks* were introduced. These gave PhD students the opportunity to present in 5 minutes their main research interest together with questions they are tackling. This allowed PhD students who did not feel ready to give full talks the opportunity to introduce themselves and feel part of the seminar. This addition was very well received.

Another innovation in the 2010 seminar was *working parties*, whose goal was to bring small groups of researchers together to discuss new important results and map out future research directions. Four working parties took place:

Working parties:

- Natural Gradient
- Fitness Landscapes
- Future of Runtime Analysis
- Genetic Programming on General Purpose Graphics Processing Units

Many of the presentations and discussions were dedicated to identifying the limitations of the various approaches, shedding light on their complementarity and arriving at a wider consent with regard to advantages and disadvantages of different techniques. In particular limitations and the complementarity of different approaches for discrete runtime analysis were discussed. This field appears to be reaching a certain maturity. While a great deal of progress has been made on mutation-based search that moves from one point to another, convergence results on more general population based algorithms have been more difficult to obtain. New generalized forms of drift analysis have been developed which now makes it easier to apply this theory to analyze the convergence behavior of randomized search algorithms on a wider range of problems. The method works by aggregating state information in complex population-based evolutionary algorithms. Talks were of high quality and showed remarkable progress in different areas. New results on fitness landscapes analysis were presented which show that certain metrics of search landscapes that usually require exponential time to compute (such as exact statistical moments, and measures such as auto-correlation) can be computed in polynomial time for many NP-Hard problems such as MAX-SAT, QAP and Graph-Coloring. The approach is based on a combination of Walsh Analysis and Elementary Landscapes.

Several new themes that have been largely absent from previous seminars have emerged as prominent research themes from the current seminar. The first theme is the theory for multi-objective optimization algorithms: first results on the linear convergence of a specific MO algorithm, overview of recent results on hypervolume based search algorithms have been presented. The second theme concerns black-box complexity, which provides bounds on the runtime of very general classes of randomized search heuristics. The last theme is the theory of Natural Evolutionary Algorithms: a recent major result presenting the state-of-the-art continuous optimizer CMA-ES as a natural gradient algorithm where a gradient descent is performed on the set of parameter of the search distribution has been discussed during Dagstuhl. This result is making a link between CMA and classical optimization algorithms but more generally the idea offers a very principled design technique for search algorithms that sample from a parametrized distribution.

Conclusion

Fruitful and stimulating discussions among varying groups of participants occurred throughout the week of the Dagstuhl seminar on “Theory of Evolutionary Algorithms”. Besides the presentations, the unconventional format of the working parties was successful in provoking discussions and stimulating the exchange of new ideas. The spotlight talks provided a great opportunity for new PhD students to introduce themselves to the community. The Dagstuhl seminars are firmly established in the community as biannual event, and we hope

to be able to build on this success and continue to promote discussions between researchers in different areas of EA theory at further workshops in the future.

1.5 Exact Complexity of NP-hard Problems

Seminar No. **10441** Date **31.10.–05.11.2010**
Organizers: Thore Husfeldt, Dieter Kratsch, Ramamohan Paturi, Gregory Sorkin

A decade before NP-completeness became the lens through which Computer Science views computationally hard problems, beautiful algorithms were discovered that are much better than exhaustive search, for example Bellman's 1962 dynamic programming treatment of the Traveling Salesman problem and Ryser's 1963 inclusion– exclusion formula for the permanent.

Today we know that all NP-hard problems are unlikely to admit polynomialtime algorithms, yet NP-hard problems must be solved, in some way, for everything from manufacturing and supply-chain optimization to railroad timetabling. Approaches include approximation algorithms, heuristics, average-case analysis, and exact exponential-time algorithms: all are essential. While all NP-complete problems are equivalent from the polynomial-time perspective, their exponential-time properties vary widely. Which problems are easiest or hardest? What are the promising algorithmic techniques? What are the connections with parametrized complexity? How fast an algorithm can we find? What about complexity lower bounds?

Work addressing such questions, both from the algorithmic and complexity theoretic sides, has become known as exact complexity. Despite significant progress, the area is still fairly new and many fundamental problems remain open. Where the approximation algorithms field, for example, has unifying algorithmic techniques such as LP rounding and semidefinite programming, and hardness techniques from probabilistically checkable proofs and the Unique Games conjectures, much exact algorithms work is still specific to a particular NP-complete problem: powerful unified techniques are just emerging.

Exciting new results and directions have been established by scientists on several continents, with important contributions coming from young researchers such as Williams and Traxler. The purpose of this seminar is to accelerate developments in this late-blooming field. Below, we outline several new results and promising directions.

The Tutte polynomial of an n-vertex, m-edged graph can trivially be evaluated in time O*(2m), but no vertex-parameterized algorithm is obvious. The Potts (q-coloring) partition function can trivially be evaluated in time O*(qn), but it is not obvious if one can remove the dependence on q. The Fortuin–Kasteleyn model from statistical physics generalizes both, and a breakthrough result of Björklund, Husfeldt, Kaski, and Koivisto [FOCS 2006, STOC 2007, FOCS 2008] shows how to evaluate it using the inclusion–exclusion method in time O*(2n). It is an intriguing question as to how far these techniques could be extended.

Recently, the color-coding technique of Alon, Yuster, and Zwick [JACM 1995] has been extended by introducing algebraic structures that yield faster fixed parameter tractable

algorithms. Koutis [ICALP 2008] uses "vector coding" for a randomized O*(23k/2) algorithm for the k-Path problem, and Williams [IPL 2009] improves this to O*(2k). Such algorithms from group algebra are a promising direction for further exploration.

Branch-and-reduce is one of the most frequently used methods for solving NP-hard problems, but current analyses of such algorithms may be overly pessimistic. Fomin, Grandoni and Kratsch [ICALP 2005, SODA 2006] used a measure and conquer framework to establish simple and fast algorithms to solve the Minimum Dominating Set and the Maximum Independent Set problem. This and related methods, Eppstein's quasiconvex analysis [SODA 2004] and Scott and Sorkin's linear programming method [Random 2005], have become indispensable, but a need remains for further improvements.

Faster algorithms, notably for Maximum Independent Set, have resulted from computer-produced graphical reductions and case analysis. Can these automated techniques be put on a more general theoretical level, and improved? Can similar automation be applied to logic-based branching rules such as the "clause learning" of Kulikov and Kutzkov [CSR 2007]? What about lower bounds on such local methods?

The meeting was attended by 46 researchers, the maximum possible modulo some last-minute cancellations. The organizers are grateful to all who came, and regret that — due to a gratifyingly high acceptance rate — others who would have contributed could not be invited. The participants came from around the globe, predominantly from Europe as usual for this field, but this time also with a good showing from the US.

1.6 Computational Counting

Seminar No. **10481** Date **28.11.–03.12.2010**
Organizers: Peter Bürgisser, Leslie Ann Goldberg, Mark Jerrum

Introduction

Computational complexity is typically concerned with decision problems, but this is a historical accident, arising from the origins of theoretical computer science within logic. Computing applications, on the other hand, typically involve the computation of numerical quantities. These applications broadly fall into two types: optimisation problems and counting problems. We are interested in the latter, broadly interpreted: computing sums, weighted sums, and integrals including, for example, the expectation of a random variable or the probability of an event. The seminar covered all aspects of computational counting, including applications, algorithmic techniques and complexity. Computational counting offers a coherent set of problems and techniques which is different in flavour from other algorithmic branches of computer science and is less well-studied than its optimisation counterpart.

Specific topics covered by the meeting include: techniques for exact counting, especially moderately exponential algorithms for intractable problems, techniques for approximate

counting including Markov Chain Monte Carlo (MCMC), holographic algorithms and reductions, computational complexity of counting, algebraic complexity of counting, applications to statistical physics, and applications to constraint satisfaction.

The questions addressed include: What algorithmic techniques are effective for exact counting and approximate counting? Do these techniques remain effective in the presence of weights (including negative and complex weights)? What inherent limitations arise from computational complexity? Are there inherent limitations for specific techniques such as MCMC? Our nominated application areas prompted many of those questions and hopefully will benefit from the answers.

Although each of these topics is important in its own right, the real goal of this seminar was to bring them together to allow cross-fertilisation. Here is an example. A key issue for MCMC is the rate at which a Markov chain converges to equilibrium, which determines the length of simulation needed to get a good estimate. An important insight has been that this mixing rate is connected to the phenomenon of phase transitions in statistical physics. But it also seems likely that phase transitions are connected with computational intractability more generally, i.e., resistance to all efficient approximation algorithms, not just those based on MCMC. A further example is provided by the way algebra pervades several of our topics - holographic algorithms, complexity of counting, and constraint satisfaction - and yet the connections between these are only now being explored. For example, algebraic methods permit semi-automatic generation of reductions between counting problems, and open up the speculative possibility of resolving the P versus NP question positively through "accidental algorithms".

We are interested in the complexity of counting in different models of computation. Counting in models of arithmetic circuits is intimately connected with the permanent versus determinant problem. The latter has recently triggered the study of several specific counting problems such as the computation of Littlewood-Richardson coefficients. Another direction of research that is relevant to the meeting is the classification of counting problems in computational algebraic geometry (counting irreducible factors, connected components, etc).

Two key applications areas, statistical physics and constraint satisfaction, have a central role. The problem of computing and approximating weighted sums already arises frequently in statistical physics, where such sums are referred to as partition functions. Constraint Satisfaction is a wide class of problems which arose in the context of AI - many computer science problems can be cast in this framework. Weights are not traditionally considered in CSP, but with this addition, many applications can be viewed in terms of counting CSPs.

Participation and Programme

The seminar brought together 36 researchers from Canada, China, Europe, India, Israel, and the United States with interests and expertise in different aspects of computational counting. Among them there was a good mix of senior participants, postdoctoral researchers and PhD students. Altogether, there were 36 talks over the week including three

overview presentations and 7 ultra short five minute introductions by those participants that did not wish to give a full talk.

The first overview was presented by Jin-Yi Cai (Madison, Wisconsin) on holographic algorithms and the complexity of counting problems. He explained Valiant's notion of holographic reductions by matchgates with its twofold applications. On the one hand for exhibiting new efficient counting algorithms in a surprising way. On the other hand, holographic reductions provide new tools for showing the hardness of problems. The last part of Jin-Yi's survey focused on the various dichotomy results that have been recently obtained by him and his coauthors.

The second overview was devoted to the complexity of arithmetic circuits and delivered by Pascal Koiran (ENS Lyon and Toronto). It was a nice mix of old facts and new insights. Pascal gave a description of the hierarchy of algebraic complexity classes arising by studying arithmetic circuit models under various restrictions (formula, skew, weakly skew), explained the different characterizations and parallelization results. He then led over to newer results on depth four arithmetic circuits and presented his fascinating real tau-conjecture, which would imply that VP is different from VNP.

The third survey was presented by Martin Dyer (Leeds) on the complexity of the counting version of constraint satisfaction problems. The main focus of the talk was on Bulatov's dichotomy theorem, which says that each #CSP problem is either solvable in polynomial time or it is #P-complete, with no intermediate cases. After explaining the history of this result, Martin went on to explain the ideas of his simpler proof (found with David Richerby) that also yields more information, among other things the decidability of the dichotomy.

Other topics covered by the talks included phase transitions, graph polynomials, subexponential time algorithms, approximate (weighted) counting, and exact counting algorithms. The topic of Markov Chain Monte Carlo simulations was less represented than originally conceived, which is probably due to the fact that there was recently little progress in this subject due to its maturity.

One of the main aims of the seminar was to bring together researchers from different, but related fields, covering all aspects of computational counting with the goal of fostering the exchange of knowledge and to stimulate new research. This goal was fully achieved according to our opinion and the participants' feedback. It was an intense week with relatively tight schedule. Still, there were stimulating discussions in the afternoon breaks and in the evenings, some of them even leading to improvements of results that had been presented in the talks. New contacts and maybe even friendships were made.

Due to the snow and heavy cold we abandoned the traditional Wednesday hike. Instead we organized a trip to the UNESCO world heritage "Völklinger Hütte" on which most participants joined and had a good and relaxing time.

The organizers and participants thank the staff and the management of Schloss Dagstuhl for their assistance and support in the arrangement of a very successful and productive meeting.

Chapter 2

Automata Theory

2.1 Quantitative Models: Expressiveness and Analysis

Seminar No. **10031** Date **17.01.–22.01.2010**
Organizers: Christel Baier, Manfred Droste, Paul Gastin, Kim Guldstrand Larsen

Quantitative models and quantitative analysis in Computer Science are receiving increased attention in order to meet the challenges from application areas such as Performance Analysis, Operational Research and Embedded Systems. What is aimed at is a revision of the foundation of Computer Science where boolean models and analyses are replaced by quantitative models and analyses in order that more detailed and practically useful answers can be provided.

The purpose of this seminar was to bring together researchers from different communities with their central interest in quantitative models and analysis. The goal was to address three fundamental topics which are closely related: quantitative analysis of real-time and hybrid systems; probabilistic analysis and stochastic automata; weighted automata. These three areas of research have mainly evolved independently so far and the relationship between them has emerged only recently. The seminar brought together leading researchers of the three areas, with the goal of future highly productive cross-fertilizations.

Weighted automata and weighted context-free grammars were first introduced in seminal papers by M.-P. Schützenberger (1961) and N. Chomsky and M.-P. Schützenberger (1963), and a large amount of general results has been developed since then. Weighted automata are classical nondeterministic automata in which the transitions carry weights. These weights may model, e.g., the amount of resources or time needed for executing a transition, or the probability of its successful execution. Very recently, quantitative model-checking, automata with discounting and priced timed automata have been considered motivated by real applications. Also, probabilistic automata are related to weighted automata.

The model of timed automata introduced by Alur and Dill in 1989 has by now established itself as a universal formalism for describing real-time systems. The notion of priced (or weighted) timed automata was introduced independently by Alur et al and Larsen et al in 2001, with the surprising result that cost optimal reachability is decidable. Since

these initial results, efficient tools have been developed and a number of more challenging questions have been considered including multi-priced timed automata, optimal infinite scheduling (both with respect to mean pay-off and discounting), priced timed games and model checking for priced timed automata.

Stochastic models have a long tradition in Mathematics, starting with the work by Markov in the early 20th century and by Bellman around 1950 who introduced the basic principles of Markov chains and Markov decision processes, respectively. Later, distributed randomized systems were modeled by finite-state automata with discrete transition probabilities, and automata- and graph-based algorithms were developed to determine the probabilities of given linear-time properties. However, the concept of "dense time" in most stochastic models is different from that of timed automata. Continuous-time stochastic models support reasoning about time-dependent distributions for the delay of transitions. The seminar discussed methods for the quantitative analysis where the concepts of timed and probabilistic automata have been combined.

In the seminar, 45 researchers from 13 countries discussed their recent research results and developments for quantitative models and their analysis. Four survey lectures and 28 talks were organized in eight sessions with centralized themes. From the beginning, all lectures and talks raised questions of members from the other fields, and lively discussions followed. In particular, the surveys presented the fields of weighted automata, priced timed automata and games, stochastic model checking, and reconciliations between weighted and probabilistic logics. The lectures and talks dealt with, e.g., quantitative modeling formalisms and their semantics, expressiveness of models including quantitative measures for infinite behavior (like discounting, mean payoff, long-run averages), and composition and components of different models, to name only a few topics.

There are a number of open problems concerning the interplay between these fields. For instance, there are many interesting open questions about the combination of real-time and probabilism. Also, weighted automata and probabilistic automata bear similarities but also differences, and one should investigate how known techniques can be transferred in either direction. The interplay between priced timed automata and weighted automata also demands further investigation. Due to these open challenges, several researchers decided to meet again later in the year, e.g. during the workshop in Leipzig on "Weighted Automata: Theory and Applications (WATA 2010)".

During the seminar, there was very much interaction between the participants. It was expressed that a future research collaboration between the different present groups should be highly fruitful and would therefore be very desirable. A Dagstuhl seminar would provide an ideal and unique opportunity for this. The successful collaboration in the present seminar was felt to be due in particular to the superb facilities and excellent organization provided by the Dagstuhl center and its team.

2.2 Advances and Applications of Automata on Words and Trees

Seminar No. **10501** Date **12.12.–17.12.2010**
Organizers: Christian Glaßer, Jean-Éric Pin, Nicole Schweikardt, Victor Selivanov, Wolfgang Thomas

The aim of the seminar was to discuss and systematize the recent fast progress in automata theory and to identify important directions for future research. For this, the seminar brought together more than 40 researchers from automata theory and related fields of applications. We had 19 talks of 30 minutes and 5 one-hour lectures leaving ample room for discussions. In the following we describe the topics in more detail.

Monday

Thomas Wilke gave the opening lecture on functional programs for regular expression matching. He demonstrated how it is possible to develop a non-trivial algorithm online during a talk (by typing Haskell). Bill Wadge talked about degree operations and presented the game/automata characterizations of the Wadge degrees, of the degree multiplication operation of Steele and van Wesep, and reported on the search for degree exponentiation. Friedrich Otto gave a presentation on non-forgetting deterministic restarting automata that are monotone. He presented a hierarchy of language classes that are characterized by various types of non-forgetting, deterministic, and monotone restarting automata, which ranges from the deterministic context-free languages to the so-called left-to-right regular languages.

The space of one-sided infinite words plays a crucial role in several parts of Theoretical Computer Science. For several purposes, topologies other than the one of the Cantor space are useful, e.g., for studying fragments of first-order logic over infinite words or for a topological characterization of random infinite words. For this end, Ludwig Staiger talked about topologies that refine the Cantor topology.

Jacques Sakarovitch showed that the enumerating series brings a structuring vision to abstract numeration systems. The talk showed that it is decidable whether an N-rational series corresponds to a rational abstract numeration system. Christian Reitwießner talked about Boolean grammars, which are an extension of context-free grammars. In contrast to context-free grammars, they can generate quite complicated non-regular languages over a single-letter alphabet. The talk discussed the parsability of Boolean grammars and showed that they can be efficiently parsed. In his talk, Jacques Duparc explained that the Wadge hierarchy of omega-regular tree languages is huge. He defined a weighted monadic second order logic for unranked trees and the concept of weighted unranked tree automata, and investigated the expressive power of these two concepts.

Tuesday

The second lecture of the seminar was presented by Thomas Colcombet and it surveyed important results concerning distance automata and their extensions: Krob's undecidability result, Hashiguchi's result of decidability of the limitedness problem, and the decidability of the star-height problem due to Hashiguchi.

Volker Diekert and Manfred Kufleitner gave a series of talks on small fragments of first-order logic over finite and infinite words. They showed how to combine algebraic and topological properties in order to generalize decidability results to infinite words. Markus Lohrey talked about isomorphism problems on automatic structures and showed several hardness results for the isomorphism problem for transitive relations.

There are developments in automata theory that borrow important ideas from game theory. For this end, Veronique Bruyere gave a talk on equilibria in quantitative reachability games. The talk considered turn-based quantitative multiplayer non zero-sum games played on finite graphs with reachability objectives, and it proved the existence of finite-memory Nash equilibria in multiplayer games. Zoltan Esik talked about axiomatizing regular tree languages and presented complete axiomatizations of regular languages of ranked trees. Christian Choffrut gave a talk on extensions of the theorem by Eilenberg, Elgot, and Shepherdson.

Wednesday

The Wednesday's lecture was given by Howard Straubing who talked about algebras and logics for unranked forests.

One of the fundamental topics in mathematics is the search for relations between local and global regularities. Juhani Karhumaeki analyzed this phenomena in connection with infinite words. Local regularity here means that the word possesses everywhere some local (finitely describable) regularity condition, such as some type of local periodicity, while the global regularity means that the word is periodic (or ultimately periodic).

Victor Selivanov talked about the fine hierarchy of omega-regular k-partitions. The talk developed the theory of ω-regular k-partitions that extends the theory around the Wagner hierarchy of regular ω-languages. In particular, it characterized the structure of Wadge degrees of ω-regular k-partitions, proved the decidability of any level of the corresponding hierarchy, established coincidence of the reducibilities by continuous functions and by functions computed by finite automata on the ω-regular k-partitions, and showed the undecidability of the first-order theory of the structure of Wadge degrees of regular k-partitions for each $k \geq 3$.

Bahareh Afshari talked about new results on the decidability of the mu-calculus alternation hierarchy. The starting point was the open question of whether a given formula is equivalent to a formula with a lower alternation depth. In the talk we learned about partial results for the class of Δ_2 formulae.

Thursday

The seminar started with a lecture by Thomas Schwentick about automata for data words. The talk gave a basic introduction into automata models that have been proposed for data strings and data trees, that is, strings and trees enhanced by data values. The emphasis was on expressiveness and complexity.

Paul Gastin talked about weighted automata with pebbles and weighted FO logic with transitive closures. Here a new classes of weighted automata on words was introduced. Equipped with pebbles and a two-way mechanism, they go beyond the class of recognizable formal power series, but capture a weighted version of first-order logic with bounded transitive closure. The talk also discussed new logical characterizations of the recognizable series.

Manfred Droste gave a talk on weighted logics for unranked tree automata, where he defined a weighted monadic second order logic for unranked trees and the concept of weighted unranked tree automata. It was shown that the weighted tree automata and a syntactically restricted weighted MSO-logic have the same expressive power in case the semiring is commutative or in case we deal only with ranked trees, but, surprisingly, not in general. This demonstrated a crucial difference between the theories of ranked trees and unranked trees in the weighted case. In the last talk of this session, Julian Bradfield explained us what spiders and finite automata have in common.

Damian Niwinski explained the separation problem in the index hierarchy and showed that a certain pair of disjoint co-Buchi recognizable sets is complete for all disjoint pairs of co-analytic sets. The proof involved a construction of a "dichotomic" automaton.

Pierre McKenzie made an excursion to computational complexity and explained us DAG evaluation and the red-blue problem. Alexander Rabinovich talked about decidable expansions of labeled linear orderings, where he proved that if certain monadic second-order theories are decidable, then they have non-trivial expansions that are still decidable.

Friday

The friday's lecture was given by Luc Segoufin, who talked about models of tree walking automata and transitive logic on trees. Ludmila Yartseva gave a talk on definability in the structures with subword order, where she developed a theory of first-order definability in the subword partial order in parallel with similar theories for the h-quasiorder of finite k-labeled forests and for the infix order.

Alexander Okhotin gave a survey on language equations and told us the exciting story of computational completeness. His talk described the research path from encountering the first undecidable properties of language equations in 1998 to establishing the computational completeness of their ultimately simple case: systems over a one-letter alphabet using concatenation only in 2008.

Christof Loeding talked about the nondeterministic parity index problem, which consists in finding for a given regular language of infinite trees the minimal range of priorities needed by a nondeterministic parity automaton accepting the language. The talk presented a

recent approach to tackle this problem. Here the main idea is to translate the problem into a limitedness problem for distance parity automata.

Conclusions

The talks in this seminar ranged over a broad assortment of subjects with the underlying theme of automata on words and trees. It was a very fruitful seminar and has hopefully initiated new directions in research. We look forward to similar meetings in the future!

Chapter 3

Verification, Logic, Semantics

3.1 Decision Procedures in Software, Hardware and Bioware

Seminar No. **10161** Date **18.04.–23.04.2010**
Organizers: Nikolaj Bjørner, Robert Nieuwenhuis, Helmut Veith, Andrei Voronkov

Seminar Goal and Structure

The main goal of the seminar Decision Procedures in Soft, Hard and Bio-ware was to bring together renowned as well as young aspiring researchers from two groups. The first group formed by researchers who develop both theory and efficient implementations of decision procedures. The second group comprising of researchers from application areas such as program analysis and testing, crypto-analysis, hardware verification, industrial planning and scheduling, and bio-informatics, who have worked with, and contributed to, high quality decision procedures. The purpose of the seminar was to heighten awareness between tool and theory developers for decision procedures with the array of applications found in software, hardware and biological systems analysis.

The seminar fell in the week of April 19-23, 2010. In spite of the travel disruptions associated with the Icelandic volcano eruption, 27 researchers from 8 countries (Germany, Austria, Italy, France, USA (who were lucky to arrive late), United Kingdom, Switzerland, and India) arrived and discussed their recent work and future trends. The absence of several attendees from North America, Japan, China and even more distant parts of Europe meant that the seminar was unable to cover planned tutorials on Bio-analysis and constraint solving. Instead it focused heavily on decision procedures in the context of software analysis and to some extent hardware. On the other hand, it allowed for a highly interactive environment and some attendees got a chance to present a talk on more than one topic. The following summary will have difficulties conveying the very nice spirit of the resulting seminar, but will here summarize the main areas covered during the presentations.

Main Areas Covered during the Seminar

Predicate abstraction and interpolants. When performing symbolic model-checking of software systems, a technique known as predicate abstraction has been instrumental in summarizing large programs as finite abstractions. Theorem provers are critical for computing the abstraction mapping by either producing a predicate cover or using interpolants.

This area received very strong attention, perhaps due to a coincidence of the composition of those who where able to attend. There were five presentations related to interpolation:

- Instantiation-Based Interpolation for Quantified Formulae, Jürgen Christ
- Symbol Elimination and Interpolation in Vampire, Krystof Hoder
- Interpolation for Uninterpreted Functions and Linear Arithmetic, Jochen Hoenicke
- Interpolation and Symbol Elimination, Laura Kovacs
- Craig Interpolation for Quantifier-Free Presburger Arithmetic, Philipp Rümmer

Additional abstraction techniques and decision procedures for software analysis were covered as well:

- Computing Abstractions with SMT solvers, Alessandro Cimatti
- The Synergy of Precise and Fast Abstractions for Program Verification, Natasha Sharygina
- Forward Analysis of Depth-Bounded Processes, Thomas Wies
- Software Model Checking via Large-Block Encoding, Alberto Griggio

Hardware verification. Hardware verification has for quite some time now been using propositional logic (SAT) solvers. Even the modern hardware description languages and methods of their use are, to a large extent oriented toward being verifiable by SAT solvers. However, it seems that SAT-based technology will reach its limits in a few years, so there is an extensive search of higher-level languages and approaches shifting verification from bit-level to word-level and higher. The use of SMT-based decision procedures is emerging in this area.

We asked Professor Armin Biere to prepare a one hour tutorial on the subject of decision procedures in hardware. The result was an enlightening tutorial that covered main trends and challenges in hardware verification. Among the covered areas were word-level decision procedures and symbolic simulation techniques. Both are receiving particular heightened recent attention.

- Decision Procedures in Hardware Design, Armin Biere

Verifying compilers and Synthesis. A verifying compiler uses automated reasoning to check correctness assertions of the program that it compiles. Of significant importance are scaling proofs for verification conditions containing thousands of assumptions, and integrating solvers for several domains.

Of recent interest is also using decision procedures for synthesis. Thanks to a contribution of Ruzica Piskac, the seminar touched briefly on this exciting topic of renewed attention.

- Decision Procedures for Security-by-Contract on Mobile devices, Fabio Massacci
- Verifying Functional Properties with Quantified SMT, Michal Moskal
- Complete Functional Synthesis, Ruzica Piskac

Parametric Systems. Within the context of system verification, an important area of parametric system was discussed in:

- The Model Checker MCMT for Array Based Systems, Silvio Ghilardi
- Hierarchical reasoning for the verification of parametric systems, Viorica Sofronie-Stokkermans

Test-case generation. A technique of recent interest in test-case generation combines runtime analysis with static analysis. It is based on converting a sequence of instructions, collected at runtime, into a formula, and using a theorem prover for generating inputs that can be used for exercising new execution paths. Theorem provers must be able to handle very large conjunctions of constraints and produce models, for generating new test inputs; in the case the constraints are satisfiable. Besides the challenge of handling very large sets of conjunctions, these tools commonly require the theorem provers to handle bit-precise reasoning, local or global optimization, heap abstractions, and incrementality.

- HAMPI: A Solver for String Constraints, Vijay Ganesh

Decision Procedure Foundations. The many application areas were complemented by in-depth discussions on tools and foundations of decision procedures. Among the foundational topics, the seminar participants contributed with subjects ranging from Quantified Boolean Formulas, the Bernays Shönfinkel class, Linear Rational Programming, Presburger Arithmetic, (hierarchical) local theory reasoning and combinations of Boolean Algebras with Presburger Arithmetic, C_2 and WS1S.

- Decidable fragments of first-order logic, and combinations, Pascal Fontaine
- Hierarchical Reasoning: Improving Efficiency and Ensuring Locality, Swen Jacob
- Quantifier elimination by lazy model enumeration, David Monniaux

- Variable Dependencies of Quantified CSPs, Marko Samer
- Conflict Resolution, Nestan Tsiskaridze
- Decision Procedures for Data Structures, Thomas Wies

Decision Procedure Implementations The foundational material was to some extent complemented with a few implementation oriented contributions:

- The OpenSMT Solver, Roberto Bruttomesso
- Solvers for String Theories, Vijay Ganesh
- Abstract Groebner Bases and Some Applications in Z3 and RAHD, Grant Olney Passmore

Unfortunately, as the Icelandic volcano eruption prevented several participants from overseas, and even more distant parts of Europe, we were unable to cover some of additional relevant and timely topics that were planned. These were noteworthy:

Bio-informatics. Covering new applications for decision procedures in Biology and Medicine.

We had asked Bud Mishra to anchor a section on Bio-informatics. Unfortunately, all flights from North America were canceled at the time of the seminar and he was unable to attend. We are hopeful that he and several other colleagues in the bio-informatics field will be able to participate and contribute in a future seminar.

Scheduling and Planning. Covering advances in SMT procedures that have enabled combining specialized solvers for difference and octagon constraints with efficient combinatorial search taking advantage of search techniques, such as lemma learning and non-chronological back-jumping.

We had asked Robert Nieuwenhuis (a co-organizer) to anchor a tutorial on scheduling and planning applications. This is particularly relevant to the work undertaken in the context of his research group and it is an important vibrant field where recent advances in decision procedures is influencing constraint solvers. Unfortunately, Robert Nieuwenhuis' flight was canceled as well, and he and his group were unable to arrange transportation to the seminar.

Discussion Session

We arranged a discussion session around the topic of software IP and software licensing. This is increasingly relevant as decision procedure implementations, even as they originate in academia, are finding industrial customers. A number of topics were discussed. We just mention two topics touched in the discussion. Armin Biere explained his licensing model for PicoSat and related tools. The tools are released freely under GPL (Gnu Public

License), which is condusive for academic research use, but is an impediment for industrial users (from the Hardware sector). He then sells a separate license for industrial use. Nikolaj Bjørner explained how the Micrsoft Research licensing model promotes academic research use of research prototypes, but that different licensing models will be needed for commercial uses. A related issue is the vastly different support models that research prototypes enjoy relative to products that are sustained for several years.

3.2 Semantics of Information

Seminar No. **10232** Date **06.06.–11.06.2010**
Organizers: Keye R. Martin, Michael W. Mislove

There have been recent advances in the *applications* of computer science to several areas of science, including for example,

- new models of classical and quantum physics, computing and information that have emerged from work relying on category theory and domain theory,
- the growth from the turn of this century of the application of process algebra and related techniques from concurrency theory – especially those using stochastic models to biology. This comes under the heading "systems biology", although the term also includes the area of computational science, which uses computers more or less as black box computational devices to generate simulations of biological phenomena, and
- increasing evidence that the work on game semantics and its application to computation has important features in common with economic game theory.

These advances have simultaneously generated interest in areas of mathematics that underlie the areas of computer science that have been leading these new applications. The Dagstuhl Seminar 10232, *Semantics of Information* was devoted to talks by researchers in a wide range of disciplines: mathematics, computer science, systems biology, physics, and economic game theory, all of which explored the relationship of computer science and its theory to their area. Because the participants came from such diverse backgrounds, the seminar included only four talks each day, each of which was an hour in length, allowing large amounts of time for researchers to interact with one another in an informal setting. In addition, we held a problem session on the closing day at which researchers discussed problems they were working on, or that they wanted the participants to consider. The overall aim was to generate collaborations among the participants, hopefully forming bridges between their disparate areas of expertise. The list of talks given below shows the diversity of interests represented at the meeting.

The seminar was the latest instance of a series of meetings with the same theme (cf. `http://www.math.tulane.edu/~mwm/clifford`, `http://www.math.tulane.edu/~mwm/WIP2008` and `http://www.math.tulane.edu/~mwm/WIP2009`), all of which have focused on generating similar interactions between researchers actively investigating the applicability and utility of information.

3.3 Game Semantics and Program Verification

Seminar No. **10252** Date **20.06.–25.06.2010**
Organizers: Paul-André Melliès, Andrzej S. Murawski, Andrea Schalk, Igor Walukiewicz

The seminar took place from 20th until 25th June 2010. Its primary aim was to bring together researchers working on modelling programs/proofs using games and the verification community. It was clear to us that both communities could, at this point in time, begin to profit from the methods and insights gained by the other community, and be able to help with some of the other side's unsolved problems. So far the two groups have had very little interaction with each other, although there are some researchers who are active in both areas.

We organized the schedules on a day-to-day basis, in order to be as reactive as possible to the requests and questions coming from the discussions. We were also careful to leave a lot of time for interaction, while offering most participants the opportunity to give a talk. Twenty-two talks were delivered during the meeting.

Scientific content

The field of program verification aims to identify and implement techniques for automatic certification of program correctness or desirable program behaviour. A central task in any software verification project is the choice of a modelling approach and a decidable formalism in which the model will be represented for the purpose of verification.

Game semantics uses the metaphor of game playing to interpret computation, which it views as an exchange of moves between two players. This allows for a very concrete account of interaction consisting of sequences of moves, one that can be readily represented with common formalisms used in verification. As it turns out, this opens up the way to numerous applications. On the more abstract level, game semantics – as a modelling technique – offers a sophisticated abstraction mechanism, which enables one to describe what is observable in a program behaviour rather than what internal symbolic steps the programs make.

Our seminar began with two tutorial talks (Ghica, Murawski), which aimed to introduce the basic principles behind game semantics, outline its place in the landscape of programming language semantics and convey its flavour. Ghica's talk was a survey of domains in which game semantics has been applied to date, including static analysis, equivalence checking and hardware synthesis. He also described verification tools whose engine is founded on game-semantic techniques, and outlined a variety of techniques employed to guard against state explosion in game models. Murawski discussed the anatomy of game models and presented a simple model that could be described with regular expressions. The model has formed the core of several tools implemented so far and is a good point of entry into the area. The concept of strategy composition was then introduced in detail and its crucial role in constructing game models was elaborated. Finally, the principle

of full abstraction was mentioned along with the advantages that it offers in verification tasks.

On the first day we also had extended talks on topics in which game semantics has already proved to be an effective tool or seems to be emerging as a potentially useful technique to ensure further progress. Kobayashi discussed model-checking functional programs via higher-order recursive schemes. The first decidability procedure for the latter was based on games and broke a long-standing stalemate in the field several years ago. Since then, alternative approaches were proposed (notably using suitably crafted type systems) and it remains to be seen what role game models will play in the future in that area. Kobayashi outlined the directions in which research on the analysis of functional programs is proceeding and suggested a few problems which game semantics might help to address. Niwiński later gave a complementary talk about an automata-theoretic approach to analyzing higher-order recursive schemes. Hofmann, in turn, described a problem (side-effect freeness) that he and his collaborators have been attacking using techniques based on logical relations. Quite interestingly, it turned out that the solution they have arrived at can be interpreted in a natural way in the spirit of game semantics, which calls for further investigation. Other unifying talks also included Seidl's perspective on the interplay of game theory and abstract interpretation, and Dal Lago's work on a compositional approach to sublinear complexity, with interesting connections to functional programming.

In the course of the week, approximately half of the talks were devoted to current topics in verification, while the other half concerned developments in game semantics. A variety of topics were covered: Habermehl lectured on regular model checking, Leroux talked about Presburger invariants and Petri-net reachability, Sutre covered the latest results in analyzing pushdown concurrent systems, Lozes presented the latest work in verifying heap-manipulating programs and Tzevelekos talked about reachability in the functional setting.

Concurrency theory featured prominently in several talks. Melliès explained how to clarify the connections between game semantics and verification by applying ideas from concurrency theory: after recalling the tree-automata techniques applied by Ong in order to establish the decidability of mu-calculus formulas on higher-order recursion schemes, he explained how these techniques are inherently connected to the positionality property of innocent strategies in asynchronous games. On the verification side, Müller-Olm gave a survey of his work on analyzing threads and procedures, while Zhang talked about verifying probabilistic concurrent systems. Foundational aspects of model-checking partial-order models in concurrent extensions of the mu-calculus were also discussed by Gutierrez.

A number of talks were also concerned with extending the range of game semantics to new settings. Levy started with a talk describing how to derive strategies from programs via transition systems. Then, Laurent gave an overview of the literature on logic and game semantics, with a special emphasis on the notion of innocent strategies and its relationship to intuitionistic and classical logics. Zeilberger developed this direction, and explained how to think of game semantics in a purely syntactic way, using extensions of traditional sequent calculus. Then, Clairambault discussed calculating least and greatest fixed points in game models, while Laird showed his latest results on modelling polymorphism. Goyet described an extension of the usual syntax of the lambda-calculus in order to obtain a full

definability result for general (not necessarily innocent) strategies on arena games.

The atmosphere during the seminar was very good, clearly all the participants were open to new ideas from 'the other side'. In particular the introductory talks attracted a number of questions asking for clarification on various issues. This showed us as the organizers that people were keen to understand the material that was presented to them, and that our selection of topics was suitable for our purposes. The periods we left unscheduled as well as the meals were then available for further discussion. In particular the young researchers present expressed their delight with the opportunity to talk to established participants in a relaxed atmosphere. Because this was a residential workshop, people did not have to worry about returning to their accommodation, or making arrangements for meals, which greatly facilitated smaller groups having additional discussions, of which we saw quite a few.

The Dagstuhl staff were extremely helpful throughout the meeting and, because most of the organizational tasks were carried out by them, the participants could concentrate on scientific matters. As the organizers we were very grateful for all the support! A number of people also commented positively on this aspect in their feedback forms.

It is perhaps too early to say how much of an impact our seminar will ultimately have. Because for many participants this was the first sustained encounter with the other community, it will take some time for ideas to be digested and adopted. The main achievement of the meeting is the creation of a platform on which new collaborations can be built in the years to come, leading to even more synergy between game semantics and verification.

3.4 Verification over discrete-continuous boundaries

Seminar No. **10271** Date **04.07.–09.07.2010**
Organizers: Bernd Becker, Luca Cardelli, Holger Hermanns, Sofiene Tahar

The seminar aimed at bringing together researchers working on the analysis of systems, where the analysis uses abstractions or embeddings from discrete to continuous or from continuous to discrete domains. Such analysis across discrete-continuous boundaries appears in a large spectrum of practical and industrially relevant applications. They often play a pivotal role to arrive at useful analysis results. On the other hand, they necessarily incur some error, and make the question how to give proper correctness guarantees for the system behavior a notoriously difficult one.

Seminar Context. Formal models of computation have for long been considered independent of the concrete world, viewing hardware and software as discrete models of computation. However, there is nowadays a striking need to incorporate continuous physical reality, caused by very different trends and challenges, including *embedded and cyber-physical systems*, *deep sub-micron effects*, *biology-inspired computation*, or *analog and mixed-signal circuits design*.

On the other hand there are many application areas of scientific computing, that have traditionally treated their matter as of a mostly continuous nature, but are starting to see the need to consider discrete structures, e.g. in some parts of *cell biology* and *chemical kinetics*, in *numerical mathematics*, and in *distributed control*.

In these areas it also occurs more and more, that a shift from (or to) a discrete interpretation to (or from) a continuous interpretation is a major step in model analysis. Often analyzing a continuous system by computer aided tools requires to switch to an appropriately truncated discrete approximation. Conversely, there are cases where the opposite strategy has proven succesful: A prominent example of this is integer linear programming, where e.g. the cutting plane method proceeds via an iteration over LP problems working on a continuous domain. Other examples e.g. emerge in the area of *mean field analysis* applied to distributed computing, where the interaction of large quantities of discrete components is summarized by an averaging continuous value.

Seminar Objectives. In the future, but even nowadays, it is becoming rather common in modelling and analysis to switch between a discrete and a continuous view on a system. The consequences of such an abstraction step are often overlooked however, especially if several of these switches occur during the modelling. For instance, a fluid mixture of chemical substrates, consisting of a discrete number of molecules, is represented by a differential equation with real valued parameters, which are analyzed by simulating the system in a floating point representation and in discrete time steps. Each of the switches induces an error in the analysis, and the effect on the accuracy of the analysis results might be extreme.

This seminar aimed at bringing together, for the first time, researchers from independent areas working on the boundary of discrete and continuous modelling and verification, with the intention to cross-fertilize their individual research topics.

We were striving for a broad coverage of instances where one or several of these boundary crossings occur, paired with technical discussions about possibilities to quantify induced errors. This created impulses for a cross-fertilizing research agenda that relates scientific and industrial contexts.

3.5 Modelling, Controlling and Reasoning About State

Seminar No. **10351** Date **29.08.–03.09.2010**
Organizers: Amal Ahmed, Nick Benton, Lars Birkedal, Martin Hofmann

Introduction

The combination of dynamically allocated, mutable data structures and higher-order features is present in almost all programming languages, from C to ML, Java and C♯. The search for good models and reasoning principles for, and language features that tame the complexity of, this combination goes back many decades. Recent years have seen a number

of significant advances in our semantic understanding of state and encapsulation, including the development of separation logic, substructural type systems, models using parametric and step-indexed logical relations, and new bisimulation-based reasoning methods.

At the same time, concern about reliability, correctness and security of software has led to increased interest in tool and language support for specification and verification of realistic languages (for example JML and Spec♯), certified and certifying compilation, proof-carrying code, safe systems programming languages (such as Cyclone and CCured), and practical type systems capturing and controlling subtle aspects of state, such as ownership, effects, information flow and protocol conformance. Formalizing the meaning and the soundness of these new languages, analyses and type systems is a major motivation for the development of the theory described above.

This is an exciting and important research area. Mathematically sound reasoning principles for state, combined with recent advances in program analysis and machine-assisted proof, have the potential to lead to improved programming languages, compilers, verification technology and even to new approaches to software deployment and operating system design.

This seminar built on the success of Dagstuhl Seminar 08061 'Types, Logics and Semantics for State', held in February 2008, though with slightly less emphasis on bringing together researchers from very different communities, and slightly more on in-depth technical discussion and collaboration on key technical issues such as the correct modelling of independence, recursive store, step-indexing and purity.

Among the research challenges addressed at the workshop were:

- What are the semantic foundations of existing logics and type systems for ownership, confinement, effects and permissions, and how may such foundations be used not only to understand and improve these systems, but also to relate them formally to one another?
- How can we reason about controlled use of state at multiple levels of abstraction, for example in relating high-level, language-enforced restrictions to low-level guarantees on the behaviour of compiled code?
- What is the right mix of approaches to the control of state and other effects, both in low-level languages and in modern high-level languages with higher-order features? How to balance language design and type systems, automated verification tools and machine assisted proof?
- What is the relationship between the recently appeared step-indexing method for establishing soundness of type systems and fully denotational approaches? In particular, how can denotational methods for mixed-variance equations for predicates and relations be transferred to step-indexed models of types, and how can the respective soundness properties, which are in general not logically equivalent, be compared?
- How can we quantify and use the additional information gained by modular analyses that do not require knowledge of the whole program? Is observational equivalence really the ultimate equivalence?

- How should we deal with the mixed-variance equations for predicates and relations that appear in the denotational modelling of storable procedures ("higher-order store"). In recent developments we have seen such equations that were solvable in an ad-hoc way but escaped the established solution theory.
- How can the algebraic approach developed for global state via Lawvere theories be extended to local state ?

Participation and Programme

The seminar brought together 44 researchers from Europe, Japan, Singapore and the United States with interests and expertise in all aspects of modelling and reasoning about programs with mutable state. There were 32 talks over the course of the week, including invited overview talks on particular topics, shorter contributed talks on recent work, open problems, and issues that arose during the week's discussions. The overview talks were on the use of ultrametric spaces in semantics, the state of the art in logical relations, verifying liveness properties of higher-order programs, and automated verification based on separation logic.

It was clear that modelling and reasoning about state is still a vibrant and fast-advancing research area. Even since the previous seminar on these topics, very significant progress has been made in a number of areas, including semantic foundations, reasoning techniques and the pragmatics of mechanizing both of these in proof assistants and automated verification tools. Many talks concerned, or involved, formalizations using the Coq proof assistant. Parametricity and logical relations continue to be central ideas, and some very impressive recent results were presented. One of these was the construction of very powerful Kripke logical relations models for ML-like languages, dealing with nearly all the 'difficult' features: higher-order functions, polymorphism, general references, recursive types and control operators. Another was the use of logical relations to prove a compositional full functional correctness result for a compiler for such a language. Ideas from separation logic were used in many talks, with a new emphasis on how to make such reasoning more abstract and compositional. The technical device of 'step-indexing' was used in many talks, and is now starting to be better understood and investigated more closely in its own right. A particularly exciting recent development is the use (or rediscovery) of the uses of ultrametric spaces in semantics, in a sense generalizing both step-indexing and the use of sequences of projections in domain theory. The use of algebraic techniques for understanding local state was also an theme of increasing importance.

One of the most interesting developments was renewed attention to game semantics. Whilst game semantics has been an active and extremely successful area of research since the early 1990s, it has not quite broken through into the mainstream of 'applied' semantics. However, several recent developments have brought it to the attention of researchers who have hitherto used other techniques. One is the development of quite novel software model-checking techniques, built from the ground up on game semantic ideas. Another is the combination of nominal structure (which featured in several talks) with games; this has allowed fully abstract models to be built for both ML-like languages and the π-calculus. During the seminar, however, it also became clear that researchers using logical relations

have started to incorporate elements of game semantics into their models. There is an exciting possibility of something of a rapprochement between the different styles of semantics in the near future. Bisimulation-based reasoning techniques were also much in evidence, and although the long-hoped-for unification with logical relations has still not quite happened, we did learn how bisimulation can be used to prove parametricity properties and see a comparison between bisimulations and other methods for reasoning about the nu-calculus. Recent progress in the verification of concurrent algorithms and in the static analysis of resource usage were both the topic of several talks and discussions.

This was an intense and productive week. With a relatively large number of particpants, most of whom wanted to speak, the schedule was relatively tight; though we did this time manage to incorporate the traditional hiking excursion on Wednesday afternoon. Discussions continued late into the night throughout the week, and were still unbelievably technical even at 2am! The proceedings contain five papers on techniques, all of which were inspired or influenced by discussions at the seminar.

The organizers and participants thank the staff and management of Schloss Dagstuhl for their assistance and support in the arrangement of a very successful meeting.

3.6 Runtime Verification, Diagnosis, Planning and Control for Autonomous Systems

Seminar No. **10451** Date **7.11.–12.11.2010**
Organizers: Klaus Havelund, Martin Leucker, Martin Sachenbacher, Oleg Sokolsky, Brian C. Williams

Over the last decade and a half, a phase transition has occurred in the level of processing power that is incorporated in embedded systems. Simultaneously, a phase transition occurred in the scale of problems that can now be solved by automated reasoning methods. This is leading to a revolution in a range of disciplines, including model-based planning and scheduling, verification, diagnosis, and hybrid systems control: each discipline is applying automated reasoning to increasingly complex real-world problems, and is incorporating real-time versions of their respective methods on board embedded systems. These are employed in order to elevate the level at which the embedded system is commanded, to verify correctness of system behavior at runtime, to improve the reconfigurability of the system, and to automatically recover from failure. The future trend is to connect these computationally intensive systems into vast, networked embedded systems, such as nation-wide earth observing systems, coastal cabled observatories, or smart power grids.

The objective of this seminar is to catalyze a new field of model-based autonomous, embedded and robotic systems, with the salient characteristic that these devices incorporate a significant level of the above-mentioned online reasoning, based on a system model. A common vision is emerging of systems that combine varied forms of real-time reasoning on models within comprehensive run-time architectures, and that are programmed using new forms of high-level programming languages. However, while many of the appropriate

languages and modeling formalisms exist, as well as real-time reasoning algorithms for planning and monitoring, these elements are currently spread amongst several disciplines. This seminar will therefore bring together researchers from four complementary disciplines to work towards languages and architectures for model-based autonomy:

- Model-based Diagnosis and Execution
- Runtime Verification
- Continuous Planning and Dispatching
- Control of Hybrid Discrete/Continuous Systems

Through a mix of technical presentations, tutorials, panels, and breakout discussions, we seek to map out the necessary architectures, languages, formal models, and underlying reasoning methods for predictable robust and autonomous embedded systems. Discussions at the seminar will thus help to identify research needs of autonomous systems in terms of capabilities for monitoring, verification, diagnosis, planning and control in the context of compelling applications. At the same time, participants will be able to discuss technical approaches that have emerged in various related research areas, and assess their applicability to this emerging field.

3.7 Perspectives Workshop: Formal Methods – Just a Euro-Science?

Seminar No. **10482** Date **30.11.–03.12.2010**
Organizers: Andrzej Tarlecki, Moshe Y. Vardi, Reinhard Wilhelm

Formal methods are employed during system-development process to improve the quality of the system, to increase the efficiency of the development process, or to derive guarantees about qualities of the system. The term "*formal methods*" has traditionally been used for a number of different approaches, including *modelling* and *specification languages, as well as methods and tools to derive properties of systems*. Because of the vagueness of the term "formal methods", it may perhaps, be desirable to replace it by "*modelling, analysis, and verification*."

A good recent overview of industrial projects concentrating on the early phases of specification and design has been given in a recent survey article: Jim Woodcock, Peter Gorm Larsen, Juan Bicarregui, John S. Fitzgerald: Formal methods: Practice and experience. ACM Computing Surveys 41(4), (2009).

The Dagstuhl Perspectives Workshop, held in December 2010, concentrated mostly on methods for system analysis and verification. These are employed in the design phase as well as in later phases of system development. Model checking, equivalence checking, and abstract interpretation, both developed in academia, are the most impressive success stories.

After a very long gestation period, formal methods for the derivation of program properties have finally gained some measure of industrial acceptance. There are, however, remarkable differences in the degree of this acceptance. There is a clear correlation between the criticality of systems and the costs of failure, on one hand, and the degree to which formal methods are employed in their development, on the other hand. Hardware manufacturers and producers of safety-critical embedded systems in the transportation industry are examples of areas where applications of analysis and verification methods are perhaps most visible. A semiconductor design gone wrong is just too costly for any cost argument against the use of formal design and verification tools be acceptable. Threats of liability costs are strong arguments for the use of formal methods in the development of safety-critical embedded systems. Different application areas often entail different approaches to the use of formal methods. Safety-critical systems call for the use of sound methods to dogmatically ensure correctness. General-purpose software with strong time-to-market pressures encourage more pragmatic attitude, with emphasis on bug-chasing methods and tools.

Industrial domains with certification requirements have introduced tools based on formal methods into their development processes. However, most current certification regulations are still process-based; they regulate the development process and do not state the required properties of the result. Critics describe this as “Clean pipes, dirty water.” The trend to use formal methods will become stronger when certification standards move from process-based assurance to product-based assurance. These new standards will specify the guarantees to be given about system properties. Several current standards for transportation systems highly recommend abstract interpretation and model checking for systems at the highest criticality level. “Highly recommend” actually means “required”. The loophole is the “state-of-practice” argument. The developer can be exempted from using a highly recommended method by arguing that it is not yet the ”state of practice”.

Several participants of the workshop have expressed the important role of champions of a formal method. A champion, enthusiastic about the potential of and competent in the use of verification method, is often needed to introduce the method and associated tool to the development process. Often, once the champion leaves, the degree of adoption declines dramatically.

The expectations towards analysis and verification methods have always been very high, often due to unrealistic promises. These unrealistic promises have mostly been the result of the ignorance of the differentiation of roles. Three distinct *roles* are connected to a formal method: the *researcher* develops the theoretical foundations of the method; the *tool developer* implements the method; and the *users* apply the tool in an industrial setting. The different analysis and verification methods have very different requirements imposed on their users, which has implications for their acceptance in industry. Researchers and tool developers often develop their methods and tools for their own use. Subsequently, they use these tools with a high degree of expertise. The experience of such expert users is quite different from that of industrial users, who do not have such degree of expertise. Thus, reports by expert users often quite rosy and create unrealistic expectations. The expectations towards analysis and verification methods are astonishing in the light of the known undecidability or intractability of the problems they are expected to solve; the methods and tools are expected to be at the same time fully automatic, effective

and efficient, and easy to use. Disappointment is unavoidable. Nevertheless, the border between what can currently be done and what is still out of reach is permanently moving, with significant progress accomplished over the last 30 years.

One challenge for further advances is higher *degree of automation*: the different methods require different degrees of user interaction and of user qualification. Currently, with few exceptions, such as Microsoft Research's Boogie platform, there is little integration among different tools. Nevertheless, advances can be expected in the coming years from *tool integration*, starting with information exchange between tools and common exchange formats. Specifically, there is a high potential for improvement from a synergetic integration of model-based design tools with analysis and verification tools.

Scalability of the methods and tools is still considered a problem. The *exploitation of large-scale parallelism* may increase the size of verifiable systems. A clear *identification of application areas* for the various methods rather than the search for universal methods, doomed to fail, will avoid user disappointment.

The embedded-system industry has already realised that badly structured systems written in obscure programming style cannot be effectively maintained. Similarly, it cannot be expected that verification methods would cope with such systems. Systems should be *designed for verifiability*.

While formal methods have often been dismissed by many as "Euroscience" a rather abstract research with little chance for industrial adoption, decades of research, both basic research and tool development have started to bear fruits, attracting an increasing level of industrial interest. This interest is often accompanied by unrealistic expectations, but, at the same time, provides an opportunity and challenge to researchers working in this area, as more basic research and good tools engineering are needed to solve the challenges outlined above.

Chapter 4

Programming Languages, Compiler

4.1 Relationships, Objects, Roles, and Queries in Modern Programming Languages

Seminar No. **10152** Date **11.04.–16.04.2010**
Organizers: Guido Boella, Erik Meijer, David J. Pearce, Friedrich Steimann, Frank Tip

During the 4 days of the seminar, 21 talks, 4 tutorials and 6 demos were given by the participants. In addition, a beauty contest was run on the last day, where participants were invited to solve a benchmark problem using their system.

On **Monday**, Dr. Stephen Fink gave a tutorial examining state-of-the art practices in programming non-relational stores. This is particularly relevant to web-based systems, and many well-known systems adopt the discussed approaches, including: Facebook, Amazon, YouTube, Craig's List, Yahoo and Twitter. In particular, the main challenge faced by such systems is the potential for significant demand generated from the Internet. Following on from this, were a number talks focusing on the implementation of relationships and roles in programming languages. In the evening, some of these systems were demonstrated, giving the audience a chance to see how far along in development they were.

Tuesday began with a tutorial by Prof. James Noble which surveyed several techniques for implementing roles and relationships in modern programming languages. James focused on the connection between programming languages and models of the real world and, in particular, the current mismatch between them. The morning talk session looked at different techniques for reducing the mismatch between SQL-style querying and modern programming languages. In the afternoon session, further talks were given looking at advances in programming language design which support roles, relationships and querying. In the evening, demos were given on related systems.

Wednesday was a half-day, which saw participants taking an excursion to Trier. A highlight of this was dinner and wine-tasting in the famous Mosel wine region. Prof. Friedrich Steimann gave the morning tutorial on the meaning of roles in programming

languages. Friedrich drew inspiration from several old texts looking at concepts and uses of spoken and written languages.

Serge Abiteboul, a participant from the (concurrent) Dagstuhl seminar 10151 on Enabling Holistic Approaches to Business Process Lifecycle Management, gave a tutorial on **Thursday**. This examined the rise and fall of object-oriented databases — a particularly relevant topic for the participants. Serge brought an interesting and alternative view on this subject, and provided some valuable cross-fertilization of ideas. Following on from this, Stephan Herrmann gave a short tutorial on the Portable Common Tool Environment (PCTE), which is an EMCA standard. Talks on Thursday covered a range of relevant topics, including verification of relationship-based systems, and analysis of object behaviour protocols. Demos were given in the evening of related systems.

Friday saw many participants demonstrate their solutions to the hotly-contest *Beauty Contest*. This consisted of a real-world benchmark problem, put together by Friedrich Steimann, extracted from a (hypothetical) theater application intended to support play rehearsals. Participants were given just ten minutes to present their solution and argue why it was elegant. Highlights included several systems which imported the works of Shakespeare as test data. The solution by Christian Wende using a derivative of the EMFText system was declared the winner.

Chapter 5

Geometry, Image Processing, Graphics

5.1 Information Visualization

Seminar No. **10241** Date **13.06.–18.06.2010**
Organizers: Andreas Kerren, Catherine Plaisant, John T. Stasko

Introduction

Information Visualization (InfoVis) is a research area that focuses on the use of visualization techniques to help people understand and analyze data. While related fields such as Scientific Visualization involve the presentation of data that has some physical or geometric correspondence, Information Visualization centers on abstract information without such correspondences, i.e., it is not possibile to map this information into the physical world in most cases. Examples of such abstract data are symbolic, tabular, networked, hierarchical, or textual information sources. The ever increasing amount of data generated or made available every day confirms the urgent need for suitable InfoVis tools. As prerequisite for building a successful visualization, InfoVis combines several aspects of different research areas, such as Computer Graphics, Graph Drawing, Data Mining, Information Design, Cognitive Psychology, and Human-Computer Interaction (HCI), among many others.

One main goal of this second Dagstuhl Seminar on Information Visualization was to bring together theoreticians and practitioners from the addressed research areas with a special focus on the intersection of InfoVis and Human-Computer Interaction. Many researchers are active in both of these fields, thus the seminar was especially attractive to those people. To support discussions that are related to the visualization of real world data, we also invited researchers from selected application areas, such as Bioinformatics and the GeoSciences.

Seminar Topics

The following themes were discussed during the seminar:

- **Collaboration within Information Visualization:** Collaboration is becoming increasingly important in InfoVis and it can occur in collocated or distributed locations and be synchronous or asynchronous. The development of novel interaction techniques, suitable visual representations, social components, and special display technologies are only a sample of the important issues that were discussed.
- **The Importance of Interaction:** The representational aspects of information visualization often receive the most focus, but the interactive capabilities of an InfoVis system are just as important. The interactive dialog between the human user and the visualization system allows the user to gain new perspectives on the data and ask questions not initially present. Accordingly, what makes for an effective, interactive system? What do powerful interaction capabilities add to a specific visualization? Which interaction techniques best accomplish different analytical goals?
- **The Influence of Display Technologies on InfoVis:** Large displays with high-resolution are one possibility to present increasingly large data sets. On the other hand, small-scale displays, especially in the context of mobile phones or PDAs, are becoming more available and important. The size and the type of a display have a large influence to the user interaction and visual representation within InfoVis.
- **InfoVis for the Masses:** In addition to the typical single-analyst, deep-dive analytical nature of InfoVis, a growing focus of research is examining how to allow large numbers of people to produce, view, and discuss information visualizations as well. This topic emerged during the first Dagstuhl Seminar on InfoVis but insufficient time was available to thoroughly discuss it.
- **Multimodal User Interaction:** Multimodality is referred to as being the combination of several modalities, such as the visual, sensory, or auditory modality. The extension of InfoVis with other modalities is not very well explored, and many open questions exist. For example: when we should use multiple modalities to present a particular data set and which ones? Or, how should they be combined?
- **Prior Knowledge of Users:** Visualization tools should be sensitive to the prior knowledge of their users with respect to both application-specific knowledge and visualization knowledge. Choosing optimal levels of visual abstraction and finding effective visual metaphors are two challenges in this area, where the term "optimal" clearly depends on the user. A visualization system should adjust to the user's needs.
- **InfoVis Aesthetics:** An aesthetically appealing visualization can be more successful than a more efficient but unappealing visualization that presents the same data. A closer look into such phenomena could improve our only vague understanding of this intersection of InfoVis and the Visual Arts, and thus, it could improve the success of information visualization techniques in practice.

The seminar allowed attendees to critically reflect on current research efforts, the state of field, and key research challenges today. Participants also were encouraged to demonstrate their system prototypes and environments relevant to the seminar topics. As a result, further topics emerged and were the focus of deeper discussions:

- **Visualization of Text and Documents:** Textual data is widespread and of importance to many people. While the visualization of text and documents is often treated as a general problem, the problem and suitable solutions can differ strongly depending on the target audience of the visualization and the task support desired.
- **Comparison in Information Visualization:** The visual comparison of data is important for many analysis tasks. A deeper discussion of the underlying models, goals, and challenges helped to facilitate a better understanding of this issue.
- **Data Wrangling:** The transformation of data into alternate forms to enable analysis is an important step in the broad analytical process. Related issues include data quality, how to handle missing data, data cleaning, and normalization, among many others.
- **Analysis Process:** The true analysis process is often significantly different than many assumptions widespread in the information visualization literature. Who are the key stakeholders in this process? How do analysts usually work? What is important to them?

Participation and Program

48 people from 11 countries participated in this seminar. Most attendees were from the US and from Germany, but others came from Canada, Australia, Israel, and other European countries.

The program aimed to generate lively discussions. Presenters were asked not to give talks solely focusing on their own specific research. Instead, the group began the seminar by collecting important themes and then selecting eight specific topics to be discussed in later breakout sessions. The resulting eight group discussions were preceded by provocation talks, and accompanied by scientific talks, a demo session and report sessions given to the entire audience.

5.2 Computational Video

Seminar No. **10411** Date **10.10.–15.10.2010**
Organizers: Daniel Cremers, Marcus A. Magnor, Lihi Zelnik-Manor

Dagstuhl seminar 10411 "Computational Video" took place October 10-15, 2010. 43 researchers from North America, Asia, and Europe discussed the state-of-the-art, contemporary challenges and future research in imaging, processing, analyzing, modeling, and rendering of real-world, dynamic scenes. The seminar was organized into 11 sessions of presentations, discussions, and special-topic meetings. The seminar brought together junior and senior researchers from computer vision, computer graphics, and image communication, both from academia and industry to address the challenges in computational video. Participants included international experts from Kyoto University, Stanford University,

University of British Columbia, University of New Mexico, University of Toronto, MIT, Hebrew University of Jerusalem, Technion - Haifa, ETH Zürich, Heriot-Watt University - Edinburgh, University of Surrey, and University College London as well as professionals from Adobe Systems, BBC Research & Development, Disney Research and Microsoft Research.

Motivating the seminar is the swift development in video imaging technology, processing algorithms, and applications. Similar to the digital revolution in photography of fifteen years ago, today digital methods are revolutionizing the way television and movies are being made. With the advent of professional digital movie cameras, digital projector technology for movie theaters, and 3D movies, the movie and television production pipeline is turning all-digital, opening up numerous new opportunities for the way dynamic scenes are acquired, video footage can be edited, and visual media may be experienced.

For five days, seminar participants discussed the impact of as well as the opportunities arising from digital video acquisition, processing, representation, and display. Over the course of the seminar, the participants addressed contemporary challenges in digital TV and movie production, pointed at new opportunities in an all-digital production pipeline, discussed novel ways to acquire, represent and experience dynamic content, accrued a wish-list for future video equipment, proposed new ways to interact with visual content, and debated possible future mass-market applications for computational video.

Examples for viable research areas in computational video identified during the seminar included motion capture of faces, non-rigid surfaces and entire performances, reconstruction and modeling of non-rigid objects, acquisition of scene illumination, time-of-ight cameras, motion field and segmentation estimation for video editing, as well as free-viewpoint navigation and video-based rendering. With regard to technological challenges, seminar participants agreed that the "rolling shutter" effect of CMOS-based video imagers currently poses a serious problem for existing computer vision algorithms. It is expected, however, that this problem will be overcome by future video imaging technology, at least for professional equipment. Another item on the seminar participants' wish list for future camera hardware concerned high frame-rate acquisition to enable, e.g., more robust motion field estimation or time-multiplexed acquisition. Finally, it was expected that plenoptic cameras will hit the commercial market within the next few years, allowing for advanced post-processing features such as variable depth-of-field, stereopsis, or motion parallax.

5.3 Schematization in Cartography, Visualization, and Computational Geometry

Seminar No. **10461** Date **14.11.–19.11.2010**
Organizers: Jason Dykes, Matthias Müller-Hannemann, Alexander Wolff

Motivation. In this seminar, we were interested in computing the layout of complex networks under *angular* restrictions. We refer to this problem as *angular schematization* and subsume under it also the combined effort of network construction and layout. It

is striking that edge directions are being restricted in networks of very different nature and that these networks are constructed in very different communities: graph drawing, information visualization, geographic information science, computational geometry, very-large-scale-integrated circuit (VLSI) layout, and underground mining. In some of these communities (such as graph drawing or VLSI layout), rectilinear connections have a long history, but recently octilinear connections have moved into the spotlight, bringing with them completely new problems and challenges. In other fields of application such as underground mining, it is not the number of slopes that is restricted, but there is an upper bound on the maximum slope.

We believed that it was high time for these communities to meet and to exchange problems and ideas. For example, the drawing of subway maps has been discussed independently in graph drawing, information visualization, and GIS. Manhattan networks are constructed by operations researchers and computational geometers. Octilinear connections are used when drawing subway maps, but also in the new X-architecture in VLSI layout.

We wanted to bring together researchers from different communities that all have to do with (angular) schematization but have little overlap otherwise. These communities have different backgrounds, literature, and histories. Therefore, we needed a forum where we could spend time learning from each other, transferring knowledge and developing common or at least translatable language. This was a difficult process and took time, but Dagstuhl provided an excellent environment for this kind of activity.

Seminar schedule. In order to ignite this process, we started the seminar with a morning of very short (3-minute 2-slide) introductory talks of all participants. These were followed, in the afternoon and the next morning, by a number of survey talks about the role of schematization in different communities: Max Roberts, a cognitive psychologist, talked about "Underground Maps: Design Challenges and Challenging Designs", William Mackaness investigated "Context, Meaning and Schematisation", Jack van Wijk gave a short overview over "Information Visualization and Visual Analytics", Martin Nöllenburg introduced us to "Schematization in Graph Drawing", Heiko Schilling gave us an idea about the potential use of schematization in car navigation, Matthias Müller-Hannemann talked about schematization in the areas VLSI design and underground mining, and Stefan Felsner gave a crash course on "Schnyder Woods and Applications".

Next, we had an open-problem session where some participants introduced their favorite schematization-related problems. We made a list of these problems:

1. Lombardi Metro Maps
2. Dynamic Layout – Remove / Add Edges
3. Geographic Hypergraphs
4. Hexagonal / Octagonal Cartograms
5. Skeletons for Bundled Graph Layout
6. Steiner Trees in Optimal Network Design

7. Which Way is Up?
8. Layout of Alternative Routes
9. Combining Boundary Labeling and Graph Drawing with High Vertex Resolution

Then we asked each participant to select a problem he or she wanted to work on. Various working groups formed and walked off into different corners of the castle. The group structure changed somewhat over the remaining days of the seminar, but many participants simply stuck to their group.

In order to keep discussions lively, we set up an 'open mic' session each morning where participants could present their own software or point to interesting projects, not necessarily related to schematization. Presenters had to sign up for 10-minute slots beforehand; the slots filled quickly.

On Friday morning, we had a progress-report session where all groups reported back to the plenum. The results varied from group to group. Some groups had already very concrete ideas about publications, whereas others had just arrived at the point where they thought they were beginning to ask the right questions. We asked each group to nominate a group member responsible for stimulating the discussion after the end of the seminar.

Conclusion. One of the key workshop objectives was to "develop a common language" across communities. This is difficult to achieve, and even the identification of "open problems" and the use of a problem-led approach drew attention to the different approaches and expectations of groups from the participating disciplines.

The beauty of Dagstuhl is that it provided a means of enabling colleagues from different disciplines and with different backgrounds and expectations to communicate and share perspectives: whether in an "open problem" description, an "open mic" demonstration of work in progress, or during a game of table tennis!

Lots of knowledge about how the different groups operate was shared... along with many suggestions of work relating to some of the approaches and open problems. The Dagstuhl library proved to be an excellent source of information to support this activity.

In summary, we did manage to get people from different communities to contribute to the groups in which open problems were discussed... certainly in some cases. Those with backgrounds in disciplines that do not have a tradition of working in this way tended to move between groups – meaning that groups were able to both focus on the problem in hand and benefit from a range of perspectives, whilst these individuals were able to sample the scope of problems being addressed and apply their knowledge in diverse contexts.

There was some "retreat" in a few cases, where participants with shared background focused on known and specific problems. But even here some sharing of knowledge, approaches and understanding took place, particularly through the intermittent participation of "floating" group members, the stimulating "open mic" sessions, informal discussion during the very useful breaks and the plenary reporting. And some of this intra-disciplinary effort may be beneficial to the wider community – involving, for example, members of a

single discipline identifying the need to describe and communicate established knowledge within their domain more widely.

Undeniably, those who participated in the meeting acquired valuable knowledge, new perspectives and insights into the ways in which domains relating to their own discipline operate. New and lasting contacts were forged as ideas were shared and understanding of the similarities and differences between subject areas and their associated approaches were established.

5.4 Scalable Visual Analytics

Seminar No. **10471** Date **21.11.–26.11.2010**
Organizers: Daniel A. Keim, Stefan Wrobel

The goal of the seminar was to bring together the participants of the DFG strategic research initiative SPP 1335 "Scalable Visual Analytics: Interactive Visual Analysis Systems for Complex Information Spaces" and international members of the visual analytics community. Topics addressed include the analysis of bio-molecular data, spatio-temporal data, document analysis, and streaming data.

The different scenarios had in common that approaches which work either on a purely analytical or on a purely visual level do not help due to the dynamics and complexity of the underlying processes or due to intelligent opponents. During the seminar, the participants did not only present each other the status of their own work but also discussed more general issues that are common to all projects. This includes the evaluation of visual analytics approaches, building up suitable infrastructures to improve dissemination and re-usability of the results, or aspects of cognitive science that are critical for effective user-centric analysis processes. Furthermore, special interest groups were formed to encourage and foster joint projects and collaborations.

The Scalable Visual Analytics seminar was a fertile meeting in which researchers from diverse backgrounds met. It included industry and academia, senior and junior researchers, multi-national representation, and people coming from several disciplines. The diversity resulted in interesting and useful discussions, which will help to shape the future of the versatile research area of Visual Analytics.

The seminar included multiple presentations and discussions which helped to exchange domain knowledge and steer future research activities. Besides, several working groups during the seminar not only identified future research directions in the field of scalable visual analytics but also initiated new joint projects. In total, plans for three position papers, two overview papers to outreach to other communities, and three EU FET Open Projects were drafted. Furthermore, three workshops as satellites of conferences that cover specific application areas were planned to further disseminate the work and provide a platform for ongoing discussions and activities.

This seminar clearly illustrated the diversity, relevance, and fertility of the topics within the field of Visual Analytics. The intensity of the participants' involvement is a clear

indication that the interactions fostered by the seminar will generate significant follow-up research, and eventually lead to practical use of Visual Analytics as well. All participants of the seminar were encouraged to reference Dagstuhl in future activities that were conceived during the seminar; in particular, the papers if accepted will contain acknowledgments of their inception during the Dagstuhl seminar.

5.5 Representation, Analysis and Visualization of Moving Objects

Seminar No. **10491** Date **05.12.–10.12.2010**
Organizers: Jörg-Rüdiger Sack, Bettina Speckmann, Emiel Van Loon, Robert Weibel

This seminar is a successor to the Representation, Analysis and Visualization of Moving Objects seminar in 2008 (seminar 08451). The major goal has been to bring together the diverse and fast growing, research community that is involved in developing better computational techniques for spatio-temporal object representation, data mining, and visualization of massive amounts of moving object data. The participants included experts from fields such as computational geometry, data mining, visual analytics, GIS science, transportation science, urban planning and movement ecology. Most of the participants came from academic institutions, some from government agencies and industry. The seminar has led to a fruitful exchange of ideas between different disciplines, to the creation of new interdisciplinary collaborations, concrete plans for a data challenge in an upcoming conference, and to recommendations for future research directions

People, wildlife, material, food, data and even ideas move in increasing volumes at increasing speeds over increasing distances, hence mobility and movement are key processes in our present world. Understanding of mobility patterns is essential to substantiate decision making in public and private sectors, in application domains such as fleet management, transportation modeling, urban planning, tourism, wildlife ecology, spatial epidemiology, location-based services, flight safety, and marine safety. It is needed, for instance, for the prediction and monitoring of individual and group behaviors in response to and mitigation of security threats over short and long time scales. Traffic management can greatly benefit from the analysis of movement data, for example through better movement simulation (leading to better road network designs) but also by incorporating advanced detection sensors in vehicles. As a final example, mobility patterns of endangered species are prerequisites to devising protective measures in nature conservation and successfully managing interactions between tourism and conservation.

Moving object data typically include trajectories of concrete spatial objects (e.g. humans, vehicles, animals, and goods), as well as trajectories of abstract concepts (e.g. spreading diseases, gaze points in eye movement tracking). Technologies for object tracking have recently become affordable and reliable and hence movement records are nowadays generated in huge volumes on a routine basis, using diverse technologies such as radio telemetry, GPS, analysis of video sequences, Doppler radar, or infrared eye tracking. Despite this plethora of readily available tracking data, methods for extracting useful information

are still immature, due to fragmentation of research and lack of comprehensiveness from monodisciplinary approaches. Overcoming these limitations calls for improved networking of the type that can be facilitated by Dagstuhl seminars.

The main obstacle in movement research is that it is still a young field, facing problems of a predominantly interdisciplinary nature. The fact that the field is still young is illustrated by the fast development of new analysis algorithms and the ever increasing data availability (in terms of diversity as well as quantity). At the preceding seminar 08451, many interesting research results were presented, demonstrating the progress in this field, and an agenda for future research was compiled. While the participants were highly satisfied with that seminar, it was also felt that future meetings should involve an increased representation of domain specialists from relevant application domains. This need has been addressed by inviting representatives from diverse applications domains, including animal ecology, transportation, urbanism, tourism, and mobile information systems.

Interdisciplinary collaboration requires a long-term investment from the different disciplinary experts to learn about the problems at hand and the terminology used by their counterparts from other disciplines. Bringing together computing scientists with domain experts at this seminar helped to develop concrete case studies, identify suitable example data sets, and sketch out guidelines for benchmarking. The availability of example data, as well as concrete case studies help to speed up the process of bridge-building between different disciplines. This seminar has served as a catalyst in this respect, and has stimulated research on interdisciplinary topics. Nevertheless, continuity over a sufficiently long period is still important to achieve real progress. The establishment of sustainable and long-term projects (and project funding) for this type of research remains a challenge and deserves utmost attention. This seminar thus also paid specific attention to informing the participants about potential funding opportunities, and to stimulating joint grant proposals. A long-term perspective is also provided by the recently established network of the COST action IC0903 MOVE (`http://www.move-cost.info/`), which was joined by many of the participants, owing to the open participation framework of COST actions.

Chapter 6

Artificial Intelligence, Computer Linguistic

6.1 Cognitive Robotics

Seminar No. **10081** Date **21.02.–26.02.2010**
Organizers: Gerhard Lakemeyer, Hector J. Levesque, Fiora Pirri

Research in robotics has traditionally emphasized low-level sensing and control tasks including sensory processing, path planning, and manipulator design and control. Research in Cognitive Robotics, on the other hand, emphasizes those cognitive functions that enable robots and software agents to reason, act and perceive in changing, incompletely known, and unpredictable environments. Such robots must, for example, be able to reason about goals, to choose actions and to focus on patterns, objects and events according to the task execution and the cognitive states of other agents, by taking into account time, resources, and the consequences of their decisions. In short, cognitive robotics is concerned with integrating reasoning, perception, and action within a uniform theoretical and implementation framework.

The term cognitive robotics and the vision that knowledge representation and reasoning plays a fundamental role in the design of cognitive robots was first laid out by the late Ray Reiter in his lecture on receiving the Research Excellence Award by the International Joint Conference on Artificial Intelligence (IJCAI) in 1993. Since 1998, biannual Cognitive Robotics workshop with Dagstuhl being the seventh in this series.

While the earlier workshops were largely a forum for presenting state-of-the-art research results, the purpose of the Dagstuhl event was to broaden the view and bring together people from various disciplines to shed new light on the issues in cognitive robotics. In this respect we were very fortunate to have participants from areas such as robotics, machine learning, cognitive vision, computational neuroscience, and knowledge representation and reasoning.

Given the diversity of the group, we spent the first day with tutorial-style presentation, starting out with an overview of Cognitive Robotics in the sense of Ray Reiter's vision by

Gerhard Lakemeyer. This was followed by presentations on Computational Neuroscience by Laurent Itti, Planning and Execution Monitoring by Brian Williams, Probabilistic Reasoning by Eyal Amir, Cognitive Vision by Jim Little, and Human-Robot Interaction by Geert-Jan Kruijff. The rest of the Workshop consisted of research presentations, a panel, and three breakout discussion groups on the following topics: the nature of perception, symbolic and numerical uncertainty, and the role of automated reasoning.

6.2 Spatial Representation and Reasoning in Language: Ontologies and Logics of Space

Seminar No. **10131** Date **28.03.–01.04.2010**
Organizers: John A. Bateman, James Pustejovsky, Anthony Cohn

The goal of this seminar was to bring together researchers from diverse disciplines to address the spatial semantics of natural language, the interface between spatial semantics and geospatial representations, and the role of ontologies in reasoning about spatial concepts in language and thought. There are five themes that were addressed in this seminar:

- Designing and reasoning with spatial ontologies;
- Representing and processing spatial information in language;
- Identifying appropriate spatial logics for linguistic expressiveness;
- Mapping and normalizing spatial representations for geospatial tasks and domains;
- Integrating temporal and spatial ontologies and logics for reasoning about motion and change.

To this end, we invited researchers from the following areas: spatial and temporal logics, qualitative reasoning, ontologies and knowledge representation, natural language processing, geographic information systems, and computational semantics. As a result of the discussion from this seminar, we expect the following milestones and agreements to emerge:

- Coordination on ontologies for space and time;
- Initial consensus on spatial representations derived from language;
- Strategies for mapping linguistically derived spatial information to GIS baseline representations.

6.3 Learning paradigms in dynamic environments

Seminar No. **10302** Date **25.07.–30.07.2010**
Organizers: Barbara Hammer, Pascal Hitzler, Wolfgang Maass, Marc Toussaint

The seminar centered around problems which arise in the context of machine learning in dynamic environments. Particular emphasis was put on a couple of specific questions in this context: how to represent and abstract knowledge appropriately to shape the problem of learning in a partially unknown and complex environment and how to combine statistical inference and abstract symbolic representations; how to infer from few data and how to deal with non i.i.d. data, model revision and life-long learning; how to come up with efficient strategies to control realistic environments for which exploration is costly, the dimensionality is high and data are sparse; how to deal with very large settings; and how to apply these models in challenging application areas such as robotics, computer vision, or the web.

Machine learning techniques such as neural networks and statistical counterparts constitute particularly successful tools with numerous applications ranging from industrial tasks up to the investigation and simulation of biological systems. Unlike their biological paragon, however, they are often restricted to very narrow settings such as dedicated classification and regression tasks and simple data structures such as real vectors, neglecting the rich repertoire of biological networks in complex dynamic behavior, hierarchical organization, and implicit evolution of structures. In particular, classical learning models such as neural networks, decision trees, or vector quantizers are often restricted to purely feedforward settings and simple vectorial data instead of dynamic environments which include rich structure. Further, the learning scenario which is usually considered in technological settings is by far too restricted not only concerning recurrence, but concerning the integration and emergence of structure, cooperation, and autonomous behavior. Although promising approaches can be found in the literature, we are lacking general learning paradigms for complex dynamical systems which go beyond the standard comparably narrow setting as usually formalized and successfully tackled in classical learning theory. The goal of the proposed seminar was to further the understanding of complex dynamic learning scenarios and to bridge the gap from powerful biological systems towards learning in technical applications by bringing together researchers in the fields of autonomous learning, structure learning, and biological or psychological investigations of dynamical behavior, structure formation and structure perception.

Within the seminar, a number of dedicated techniques such as echo state networks, evolutionary learning, deep learner, or techniques from statistical relational learning have been discussed as well as diverse learning paradigms which help to bridge the gap towards more realistic settings in autonomous learning. Further, a few dedicated problem areas which are closely connected to the rather fundamental issue of autonomous learning and which focus the problem complex towards relevant aspects which can be fruitfully discussed and addressed within the limited time span of a Dagstuhl seminar have been considered:

- How can structures be dealt with in autonomous environments and how does structure emerge to shape complex environments accordingly? This question concerns

topics related to neural-symbolic integration, statistical relational learning, the automatic formation of hierarchies, and similar.

- How can we deal with statistically sparse settings and high dimensionality, how can we deal with non i.i.d. settings? This question concerns topics such as inference from very few samples, online algorithms for large data sets, dimensionality reduction in complex environments, etc.?

- How can we develop control strategies in complex environments, what are possible strategies capable of dealing with only limited reinforcement signals, ill-posed domains, or partially underspecified settings?

Important application areas of these challenges concern diverse topics such as robotics, computer vision, or the semantic web, as discussed in the seminar.

26 experts from 6 different countries joined the seminar, including a mixture of established scientists and promising young researchers working in the field. According to the topic, the main subjects of the researchers covered heterogeneous fields including computational neuroscience, pattern recognition / neuroinformatics, reinformcement learning, and logic / relational learning. In spite of the diverse backgrounds, the participants shared a common strong interest in the application of machine learning system to complex environments. This setup allowed us to discuss salient issues in a way that integrated perspectives from several scientific disciplines, thereby providing valuable new insights and research contacts for the participants. Correspondingly, a wide range of topics was covered during discussions and brainstorming in the seminar.

Overall, the presentations and discussions revealed that the topic of learning in dynamic environments constitutes a highly diverse and evolving field which opens interesting perspectives to machine learning.

6.4 Learning, Planning and Sharing Robot Knowledge for Human-Robot Interaction

Seminar No. **10401** Date **03.10.–08.10.2010**
Organizers: Rachid Alami, Rüdiger Dillmann, Thomas C. Henderson, Alexandra Kirsch

Autonomous robots acting in unstructured environments need sophisticated cognitive capabilities including perception, manipulation and navigation, planning and reasoning as well as knowledge acquisition and processing skills. Only when all these capabilities are combined can the robot show intelligent behavior and act appropriately in environments that are primarily designed for humans. Given the growing availability and interconnectivity of modern agents and robots, better mechanisms to define, learn and share knowledge must be developed. Advances in robot platforms have led to their integration into society for a variety of functions, and there is a pressing need to understand how they can acquire and exploit knowledge required for their specic activities. In order to be

useful, robotic agents must be able to recognize a wide range of objects, relations, and situations in their environment, and to understand the semantics of these. In addition, it is imperative that procedural and process knowledge also be known so that interactions are safe and meaningful. Cooperation and collaboration are also essential as well as the ability to perceive human and animal emotions and intentions to the largest degree possible.

This seminar brought together 25 scientists with experience in planning, learning, knowledge processing and human-robot interaction to discuss how the high-level control concepts are to be used and adapted to human-robot interaction. Because high-level control of robots interacting with humans is of special practical importance, we invite experts with experience in assistive technology and those examining social acceptance of such systems.

The seminar consisted of 8 sessions, which were focussed on discussions of the topics knowledge sharing, learning and interaction. In addition, there were 3 short presentations on topics that arose in the discussions and 7 full presentations.

6.5 QSTRLib: A Benchmark Problem Repository for Qualitative Spatial and Temporal Reasoning

Seminar No. **10412** Date **10.10.–13.10.2010**
Organizers: Anthony G. Cohn, Jochen Renz, Geoff Suthcliffe, Stefan Wölfl

The Dagstuhl seminar "QSTRLib" was planned as a successor event of the previously mentioned AAAI Spring Symposium. The intention for organizing this seminar was to discuss requirements on a problem library in a small group of researchers with expertise in benchmarking, in formal approaches to qualitative reasoning, or in specific application areas of QSTR.

More specifically, the main objectives of the seminar were to identify (a) a set of typical spatial and temporal problems and query tasks used in real world applications, (b) significant benchmark domains and problem instances, (c) measures to compare different QSTR formalisms in terms of expressiveness and efficiency, and (d) parameters to evaluate the performance of reasoning systems. In the following we provide a brief report on the course of the meeting and summarize the results of the discussions in the working groups and plenary sessions.

The plan for the seminar was to first collect spatial problems from various application scenario, then to develop classifications of these problems according to different viewpoints, and finally to analyze for selected problem instances how these could be represented in a problem library for QSTR. The seminar started on Monday, Oct 11, with a welcome session. In a series of short talks all seminar participants introduced themselves, highlighted their research interests, and explained their expectations on the seminar. After this S. Wölfl presented more background information on the benchmarking initiative and sketched a first draft of the formal language for the problem library, which will be required to represent problem domains and problem instances. A major point in the following discussion was the expressive power of the proposed language, the underlying formal logic,

and the abstraction level at which the repository language allows for expressing spatial representation and reasoning tasks. A broad agreement could be achieved in that the proposed language may serve as a starting point, but that further developments of the language should be backed up on the experience of users working with the repository.

The aim of the two afternoon sessions was to collect a wide range of spatial representation and reasoning problems from different application scenarios. To feed in material, most seminar participants contributed flash presentations each focussing on a different use case, reasoning problem, tool, or data set. Afterwards, tools and demos were discussed in a more informal setting with the presenters. In the following breakout session small working groups were formed to collect spatial reasoning problems in different application scenarios including spatial reasoning in GIS, human spatial reasoning and reasoning with natural language, and spatial reasoning with sensor data.

Results of this seminar will be integrated into the proposed benchmarking library step-by-step. It is planned to start with a specification of a rather limited formal language that covers the symbolic formalisms and prominent reasoning tasks in QSTR research. Currently, a first draft of the language proposal is compiled and is planned to be available to the public by the end of April 2011. In the following months also the development of tools (in particular, parsers) supporting the language and the integration into existing reasoners such as SparQ and GQR will play an important role. A crucial point for the further development of the library will be to bridge the gap between potential expectations towards a problem library (i.e., the kind of problems users of the problem library would like to see included) and the level of reasoning support that users can expect from currently available reasoning engines. Nevertheless, the discussions in the seminar confirmed that interesting problem instances for the problem repository may be contributed from the following research areas:

- *Geographic Information Systems.* Qualitative reasoning methods are promising methods to check the integrity of information to be added to geographic knowledge bases and to rewrite possible queries against such knowledge bases.
- *Ontological reasoning.* Many knowledge bases (e.g., medical knowledge bases) could be enhanced by spatial relations between objects and/or temporal relations between events. Hybrid methods integrating ontological and spatial reasoning may allow for answering queries efficiently against such knowledge bases.
- *High-level agent control.* Qualitative representation formalisms provide natural representations of (spatial or temporal) situations arising in high-level agent control systems. In such systems the application of specialized qualitative spatial and temporal reasoning methods (used as external reasoning methods) can show considerable performance gains.

As a final remark, it should be mentioned that this 3-day seminar with 19 participants benefited greatly from the specific facilities at Schloss Dagstuhl, namely easy access to the relevant literature as well as the flexibility to work in small groups on specific topics and discuss the results in plenary sessions.

Chapter 7

Software Technology

7.1 Service-Oriented Architecture and (Multi-)Agent Systems Technology

Seminar No. **10021** Date **10.01.–15.01.2010**
Organizers: Monique Calisti, Frank Dignum, Ryszard Kowalczyk, Frank Leymann, Rainer Unland

Today's world not only has become more complex but also much more dynamic. From the business point of view this requires an IT that is not only affordable but also highly flexible and adaptable. In order to achieve this, the idea of software services composition (also called choreography or orchestration) came up. Here, enterprise applications are assembled from comparatively simple pre-existing building blocks that are loosely integrated in an efficient and meaningful way - especially tailored to the specific needs of the underlying business process. This leads to highly distributed and fluid software systems that may even cut across the boundaries of enterprises.

Software services are typically designed, built and deployed independently from each other, however, are meant to follow common standards to enable dynamic interoperability and loose coupling. If we take a look at what SOA is expecting to deliver we find, among others, features like flexibility, adaptability, autonomy, cooperation, or interoperability. However, if we take a look at current implementations of SOA it becomes clear that especially the mentioned features are far away from being realized yet. Implementations are usually static, provide comparatively little fault-tolerance and lack dynamism, versatility, and adaptivity. Moreover, it takes substantial human efforts to build such systems. This is where agent technology comes into the picture.

An agent is an autonomous and encapsulated software system that is situated in a particular environment. Among the properties that are associated with agents and multi-agent systems and which are also relevant for our seminar are autonomicity, adaptivity, reactivity, pro-activeness, mobility, goal-orientedness, cooperation, and interactivity, as can be implied from the above service-oriented architecture and multi-agent systems technology share a number of features on an abstract level.

However, if we take a look at the research areas of each community, their main focus seems to be different. Right now, the SOA community concentrates quite a lot on software services engineering. Service engineering methodologies with appropriate new service engineering methods, techniques and tools are developed. Active topics in the multi-agent systems community are collaboration, self-organization, adaptability, pro-activeness, and interoperability. However, in order to get the best out of it we need to combine these research efforts. If we consider IT to be the skeleton of an enterprise then SOA can be seen as the bones and agent technology as the cartilage that makes sure that the joints are properly connected, even if the mere bones may look completely incompatible when we just look at them.

Thus, agent technology may not help to realize the core functionality of a service but it may help substantially to wrap its functionality in a way that it becomes highly adaptable, more intelligent, more cooperative, and self-organizing. Replaceability, compatibility, and process conformance checks are tasks that can be performed well by agents on a higher level than just checking the interfaces of service descriptions. However, despite of all these potentials the research communities do not overlap much. As a consequence, possible synergies are not exploited and solutions are developed in each community that may not harmonize; i.e., the communities do not benefit much from each others work. One of the main goals of this seminar is to bridge the gap between these communities and to stipulate a fruitful and long-lasting collaboration.

7.2 Practical Software Testing: Tool Automation and Human Factors

Seminar No. **10111** Date **14.03.–19.03.2010**
Organizers: Mark Harman, Henry Muccini, Wolfram Schulte, Tao Xie

The main goal of the seminar "Practical Software Testing: Tool Automation and Human Factors" was to bring together academics working on algorithms, methods, and techniques for practical software testing, with practitioners, interested in developing more soundly-based and well-understood testing processes and practices. The seminar's purpose was to make researchers aware of industry's problems, and practitioners aware of research approaches. The seminar focused in particular on testing automation and human factors:

Tool automation. Automation of testing is a crucial concern in industry. It is only with automation that testing becomes practical and scalable to the size of a typical system with which the industry has to deal. Test automation or tool support spans the spectrum from test planning, generation, minimization, execution, oracle checking, to management. Test automation can exploit not only knowledge from the code under test but also from available models or specifications.

Human factors. Human factors play important roles in software testing. Given the code under test, tools can try to automate the generation of test inputs as much as possible, but test oracles still need to come from testers, who specify them in the form of specifications,

properties, or test assertions, or directly inspect the actual test outputs for correctness. In addition, tools are not always perfect to deal with software complexity; testers need to cooperate with tools to effectively carry out testing tasks, by giving guidance to tools and interpreting results produced by tools. Thus testers need to be well trained.

In the week of March 14-19, 2010, 40 researchers from 11 countries (Canada, France, Germany, Italy, Luxembourg, the Netherlands, Sweden, Switzerland, South Africa, United Kingdom, United States) discussed their recent work, and recent and future trends in software testing. The seminar consisted of five main types of presentations or activities: *topic-oriented presentations, research-oriented presentations, short self-introduction presentations, tool demos, and working group meetings and presentations.*

In summary, the seminar accomplished all the expected goals, generating a great deal of forward momentum. The discussion and working groups allowed participants to form better understanding of open challenges and future directions in software testing. During the seminar, academic researchers and industrial researchers fully exchanged ideas for attempting to bridge the gap between research and practice. A number of participants exploited the substantial interactions at the seminar to foster future collaborations. After the seminar, the seminar organizers and participants compiled a bibliography by collecting a list of papers discussed or mentioned during the seminar. Several of the groups indicated that they intended to continue the discussion process after the seminar. We hope that the ideas and collaborations initiated at this Dagstuhl seminar in March 2010 will find fruition in papers, funded research projects, and technical innovations in the years to come.

7.3 Perspectives Workshop: New Frontiers for Empirical Software Engineering

Seminar No. **10122** Date **21.03.–24.03.2010**
Organizers: Victor R. Basili, Nachiappan Nagappan, H. Dieter Rombach, Andreas Zeller

Software engineering still is an ever changing and to some degree immature discipline. To achieve higher software quality and productivity, we must introduce well-understood and tested practices and technologies in practical software development. But what is a "well-understood" technique, and in which contexts can it be applied?

Over the last decade, it has become clear that software development practices and technologies must be investigated by *empirical means* in order to be understood, evaluated, and deployed; and empirical research in software engineering has made considerable advances to support decisions related to software development. Yet, as software evolves, so must research in empirical software engineering.

The first challenge for empirical software engineering, as for software engineering in general, is that modern software becomes much more a *service* in an interconnected world, with distributed development and operation, short release cycles, fast correction deployment, and failure models that are very different from traditional, monolithic systems. All these features challenge traditional findings on what makes software development successful, and

due to the complex nature of development, require new research methods. Such frontiers were first identified in the Dagstuhl Seminar 06262, "Empirical Software Engineering" (June 2006); it is now time to see how to transcend them.

The second challenge for empirical software engineering is more of an opportunity, but brings problems in itself. As more and more development is fully computer-supported (especially in distributed and open-source systems), there is a wealth of data available that can be exploited automatically. While this opens the door for further observational and ethnographic studies, it also presents several challenges due to the wealth of data involved. At the Dagstuhl Seminar 07491 "Mining Programs and Processes" (December 2007), it became clear that the essential techniques of mining facts from archives as well as relating them at a low level are now well-understood; but the challenge resides in appropriate deep empirical assessment.

In this Dagstuhl perspectives workshop, we will address these challenges by bringing together senior researchers from empirical software engineering, program analysis, software mining, and novel software architectures with the added view of new areas in emerging software systems.

Our main aim is to locate synergies between these communities, to identify future perspectives for the field, and to forge strategies for the future. As a result, we envision a manifesto that will guide and inspire research in software engineering for the next decade.

7.4 Program Composition and Optimization: Autotuning, Scheduling, Metaprogramming and Beyond

Seminar No. **10191** Date **09.05.–12.05.2010**
Organizers: Christoph W. Kessler, Welf Löwe, David Padua, Markus Püschel

Seminar Motivation

Components are a well-proven means of handling software complexity. Reusable components and software composition support the construction of large and reliable software systems from pre-defined and tested partial solutions. When maximizing reusability, we end up with components that are very general and do not fit one particular scenario perfectly. Therefore, adaptation, especially optimization, is established as a technique to deal with such mismatches.

A main motivation for us to initiate this Dagstuhl seminar was our observation that composition and optimization are discussed almost in isolation in two different scientific communities. Composition is understood as a problem in software engineering for general-purpose computing; techniques include meta-/generative programming, configuration and (self-)adaptation, and architecture and component systems. Optimization is typically understood as a problem of high-performance and stream computing, and compiler construction; techniques include auto-tuning, scheduling, parallelization, just-in-time compilation, run-time optimizations, and performance prediction.

The goal of this Dagstuhl seminar was thus to bring together outstanding researchers from these two communities. We focused on a common topic of interest: the *optimized generation and composition of programs from components*, which includes the automatic adaptation to advanced execution platforms. Hence, the topic addresses solutions for both sequential and parallel components and target systems. This issue requires the integration of techniques from both subdomains.

A longer introduction to the main issues in composition and optimization is given in a separate paper.

Seminar Organization and Results

As a baseline, the seminar reviewed key results achieved so far in the two communities. In the discussion sessions, we combined expert insight from both groups to identify common issues and point out challenges for advancing the state of the art in composition and optimization.

The seminar started with a round of short self-presentations by all 26 participants. Based on the 17 presentation proposals submitted by the participants in advance, the 2.5 days of the seminar were organized into three main topic areas: *Software composition* (monday), *Autotuning* (tuesday) and *Multicore issues* (wednesday). Each topic area was opened by a survey presentation of 30 minutes, followed by sessions with 30-minutes technical presentations, and concluded with a panel discussion each.

After some initial effort of understanding specific terminology and concepts used and problems discussed by the respective other group, there were lively and productive discussions across the two communities. In particular, we collected, during discussions, a longer list of research challenges and open problems in optimized composition, which we intend to condense into a common paper in the next months.

At the end of the seminar, we received positive feedback from several of the participants indicating that the seminar took up relevant topics, provided sufficient room for discussion in spite of the tight schedule of a short seminar, and inspired concrete plans for cooperations both within and beyond community boundaries.

In summary, we conclude that the seminar was constructive and largely met its goals. Dagstuhl provided a productive and interactive atmosphere that stimulated a cross-fertilization between the two groups. We consider the seminar a first step in a beginning process of getting the two communities closer to each other.

7.5 Algorithm Engineering

Seminar No. **10261** Date **27.06.–02.07.2010**
Organizers: Giuseppe Italiano, David Johnson, Petra Mutzel, Peter Sanders

Algorithm engineering (AE) consists of the design, theoretical analysis, implementation, and experimental evaluation of algorithms, with the aim of bridging the gap between theory

and practice in the area of algorithms. In the last decade, this approach to algorithmic research has gained increasing attention. The aim of this seminar was to bring together researchers with different backgrounds, e.g., from combinatorial optimization, algorithmic theory, and algorithm engineering, in order to strengthen and foster collaborations in the area of algorithm engineering and to identify key research directions for the future.

The seminar was attended by 29 participants from both academia and industry. Much was accomplished, fostered by the productive atmosphere of the Dagstuhl Center. Here we describe some of the more important achievements.

The program consisted of a wide variety of presentations and discussion sessions. The presentations included several survey lectures, in addition to more specialized talks. David Bader and Roman Dementiev presented surveys of the challenges of multi-core and many-core architectures for algorithm engineers, and convinced us that the new computer architectures will strongly influence future algorithmic research. Rüdiger Schultz surveyed the area of stochastic programming, including bi-level problems (e.g., Stackelberg games) and risk aversion. A key issue is that, for at least some of the input data, only the probabilistic distribution is known in advance. This area was initially studied by mathematicians and experts in mathematical programming, but has been recently discovered by computer scientists. Catherine McGeoch and David Johnson provided surveys on experimental procedures. Giuseppe Italiano gave an overview of resilient algorithms and data structures, and Philippas Tsigas discussed lock-free data structures, both areas of increasing algorithmic interest. Peter Sanders described the aim of the German priority program SPP 1307 *Algorithm Engineering* which began running in 2007.

Beyond the survey lectures, highlights of the seminar included lectures on routing in networks (e.g., transit networks for public transportation, and highway networks), on specific stochastic optimization and game-theoretic problems (e.g., 2-stage stochastic Steiner tree, stochastic ad allocation, pricing lotteries, and risk-averse models), on mixed integer linear programming approaches (e.g., MIP domination), and on clustering algorithms.

Arguably the most-appreciated features of the Seminar were the four lively open discussion sessions, which led to several concrete proposals for the future of the field which, as a result of the workshop, are now being actively pursued.

David Johnson moderated a discussion on topics for future DIMACS implementation challenges. (DIMACS is the Center for Discrete Mathematics and Theoretical Computer Science, and has a long-running series of such challenges.) We agreed that the next two challenges should be on *clustering and graph partitioning* and on the *Steiner tree problem* and its variants. The clustering challenge has now been approved by the DIMACS leadership, and will be co-led by Workshop participants David Bader, Dorothea Wagner, and Peter Sanders.

Catherine McGeoch moderated a discussion on methodological questions in algorithm engineering. These include runtime measurements, tuning for algorithm comparison, and guidelines (standards) for reporting. One major point of agreement was that page limits for experimental papers in electronic proceedings should be increased to provide room for more complete reporting of results.

Catherine also led a discussion on benchmark instance sets and code repositories, following an excellent talk by Dorothea Wagner on some of the desirable properties for benchmark

sets, such as variety, relevance to practical performance, and parameterization (to allow for controlled experiments). It was suggested that testbeds should be encouraged to grow and evolve, but the question of how best to encourage this process remains open. As to code repositories, again a major problem is getting people to contribute – many are unwilling to post unpolished code (the "embarrassment factor") but also unable to spend the time needed to do the polishing.

The last discussion focused on the area Algorithm Engineering itself, and came up with several concrete actions we could take to improve its impact and outreach. In order to stimulate research in the area and to broaden its visibility, we decided to work to establish an annual prize for the best (or most influential) paper in the field. Eligibility would be restricted to a fixed but extended period (e.g., the last five years) so as to allow for some historical perspective. Several organizations (e.g., EATCS, SIGACT, ACM/SIAM) have since been approached as potential sponsors of the award, and we expect it to be established soon. In addition, we decided that it would be useful to have a website specifically for the Algorithm Engineering community, and, shortly after the workshop, Catherine McGeoch set up a home page with the domain name *algorithm-ee.org*, which will be useful for increasing the profile of the field as a discipline, expanding communication among researchers, and providing outreach to related areas. We also hope to form an official group which will support to reach the above aims (for instance, as an ACM SIG).

It is our impression that the participants enjoyed the great scientific atmosphere offered by Schloß Dagstuhl, and profited from the scientific program and the fruitful discussions. We are grateful for having had the opportunity to organize this seminar. Special thanks are due to Carsten Gutwenger for his assistance in the organization and the running of the seminar.

7.6 Resilience Assessment and Evaluation

Seminar No. **10292** Date **18.07.–23.07.2010**
Organizers: Alberto Avritzer, Katinka Wolter, Aad van Moorsel

Resilience of computing systems includes their dependability as well as their fault-tolerance and security. Resilience defines the ability of a computing system to perform proper service in the presence of all kinds of disturbances and to recover from any service degradation. These properties are immensely important in a world where many aspects of daily life depend on the correct, reliable and secure operation of computing systems. Considered systems include, but are not limited to, infrastructures, computer networks, including adhoc and mesh networks, web-based systems, or service-oriented architectures, embedded systems, manufacturing systems, control systems and more.

Much work has been done already on modelling, measuring and evaluating performance and dependability of systems but a lot remains to be done. The combination of performance and dependability, performability, has been widely studied. But aspect such as benchmarking performability, or the combination of security and performance are still on-

going work. Similarly, systems to be studied change over time, as do their characteristics. This creates the need to continue and expand work in this area.

The workshop has addressed methods and tools to describe, measure, evaluate, benchmark, guarantee and improve resilience as well as case-studies and experimental work concerning resilience of computing systems. We have aimed at collecting existing work into a monograph with profound investigation of the subject as well as discuss open problems and research directions for the future.

7.7 Model-Based Testing in Practice

Seminar No. **10421** Date **17.10.–22.10.2010**
Organizers: Wolfgang Grieskamp, Robert M. Hierons, Alexander Pretschner

Model-Based Testing

Software testing is one of the most cost-intensive tasks in the modern software production process. Model-based testing is a light-weight formal method which enables the automatic derivation of tests from software models and their environment. Model-based testing (MBT) has matured as a rich research area in the last decade, with a significant body of research and applications. The academic community is well established with many conferences, workshops, and research projects. Tools for model-based testing have been developed both as research prototypes and as commercial or semi-commercial applications brought to users by midsize and enterprise-level companies, and applied in large scale projects.

In the family of model-driven approaches, model-based testing can be seen as a success story in particular with respect to the degree of mechanical processing and automation that has been achieved, and the adoption in practice. The successful deployment of model-based testing in industrial settings can be seen in the telecommunication domain, chip cards, specific Windows components, and embedded systems in general. An interesting issue is under which circumstances we can expect these successes to carry over to other domains and families of systems as well (e.g., distributed systems; testing 'the cloud').

The Seminar

This Dagstuhl Seminar brought together top researchers, young scientists, and practitioners to discuss the state of the art, compare it with practical experiences, and derive future directions for model-based testing research and industrialisation. Model-based testing has proved promising even in industrial terms and currently seems at the verge of large-scale deployment. While previous Dagstuhl seminars around model-based testing (04371 in 2004 and 98361 in 1998) were dedicated to bringing research results into practice, in this seminar we aimed to take advantage of the relative successes of model-based testing and have the discussion guided by available industrial experience. For this reason, this seminar

was designed with a particularly high industry participation rate. We also started with a day of talks from industrialist with the focus being on the use of MBT in industry and the challenges faced.

Challenges from 2004

In Seminar 04371, in 2004, 10 main MBT challenges were identified and we discussed the progress that has been made on these challenges.

[The list of the 10 old challenges was deleted by the Dagstuhl News editor.]

We can see that there has been progress on meeting several of the challenges, but some also remain. Given that these challenges were identified only five years ago, it is natural that not all challenges have been overcome. However, we are encouraged by the fact that MBT has clearly become industrially applicable in this short period of time; this is a sign of very lively and productive communities of researchers and test engineers. Naturally, new challenges have arisen.

Challenges Identified

During the Seminar we discussed the key challenges for the community. Within this discussion we identified three groups of challenges and specific challenges within these groups. We now discuss the challenges identified by group.

Challenges for conformance and test selection

- **Concurrency and Distribution.**

  Concurrency and distribution in the SUT affect the ability of the users or testers to observe behaviours. In particular, if there are distributed testers then we obtain a set of local observations rather than a global observation. There is a need to develop new conformance relations that reflect this reduced observational power. Naturally, this also affects the oracle problem: if we have different conformance relations then we need new ways of checking observations against a model.

  Traditional deterministic models are unsuitable for distributed and concurrent systems. There is therefore the question of what type of models are suitable. In addition, there is potential to reflect some of the issues met, such as communications being asynchronous, in either the conformance relation or in the model (by e.g. modelling the queues). It is unclear which approach leads to least complexity and greatest potential to reuse current methods and tools. Naturally, there are also consequences for test selection and test execution is radically changed if we have separate independent testers at the different system interfaces.

- **Model analysis and test selection.**

  There are clear benefits to using model analysis techniques to detect anomalies as early as possible. For specifications and design models, this can lead to the early

identification of potential problems. The analysis of test models can also identify requirements errors but can also avoid testing from an incorrect model. However, such analysis techniques can have an additional benefit: they can be used to drive test selection. There has already been some work along these lines but there is scope to extend this. In particular, most work in this area has used model checkers and there has been relatively little work using other analysis techniques. There is also the question of whether models can be modified to simplify the process of test generation and possibly also to lead to simpler test cases.

- **Non-determinism in models and SUTs.**

  Many models and systems are non-deterministic. Non-determinism leads to additional challenges in testing. One practical problem is that testing often has to be on-the-fly: methods that devise preset test cases suffer from a combinatorial explosion. However, in order to apply on-the-fly methods one requires a test tool/infrastructure that is sufficiently quick and this can be difficult for systems that rapidly interact with their environment.

  There are alternative sources of non-determinism, such as concurrency and abstraction. It seems likely that these different sources of non-determinism will require different test methods.

- **Coverage and fault detection.**

  Many test selection methods are based on a notion of coverage. However, there is only limited evidence regarding the relationship between coverage and fault detection; the community would benefit from the development of an elegant theory that explains this relationship. In addition, there are often many alternative models at different levels of abstraction. It is unclear how the level of abstraction influences the effectiveness of a coverage criterion and how we can evaluate models with regards to potential fault detection.

- **Fault models.**

  Fault models can be used to represent possible or likely faults. When using a model M for testing, a fault model can be represented as a set of mutations of M: versions of M with faults/differences introduced. This allows one to reason about test effectiveness. However, there is a need to better understand what types of fault models are suitable and how these relate to faults in real systems.

MBT and Cloud Computing

The Cloud is special because is brings together the following aspects: it is large scale; there is significant amount of concurrency; it is highly distributed; and it is dynamically changing. These properties introduce significant challenges for any engineering method and here we describe some of the challenges for using MBT for cloud computing.

- **Controllability, observability and test oracle**

It is known that controllability and observability are affected by distribution: if the system interacts with its environment at physically distributed interfaces then a tester at one interface only observes observes events at its interface and no tester has a global view. As a result of testers not having a global view, a tester might not know when to supply an input. These issues make testing more difficult. Sometimes we can overcome these problems by allowing the testers to communicate during testing but this can make testing more expensive since there is a need for an external communications network. It can also make testing take longer since testers have to wait to receive messages. There is a need to understand the trade-off between the cost of including communications between testers and the benefits it brings. There is also the question of whether there should be a global test oracle of whether each tester should have its own local oracle. Finally, there is the potential to have a Cloud infrastructure that is designed for testability.

- **Testing against simulator vs. real cloud**

  In testing software designed to run on a Cloud infrastructure we might simulate the infrastructure. This can make testing much simpler but has some clear disadvantages: the simulation might be incorrect and even if it is correct it will be an abstraction of the Cloud infrastructure and this abstraction could lead to misleading results. If we do not use a simulator then we may need to deploy the test cases into the Cloud. This makes testing more complex and relies on properties of the Cloud. There are also likely to be additional costs where the provider of the Cloud charges for its use. There are challenges relating to the development of appropriate simulations and also the understanding of the trade-off between cost and effectiveness when using a simulation.

- **Non-functional testing becomes more important**

  Cloud systems will typically have important non-functional requirements contained in Service Level Agreements. There is the challenge of producing MBT methods for testing such requirements, a problem that is complicated by the nature of Cloud systems. There is also a potential opportunity for MBT since MBT methods might be used to certify cloud service quality?

- **How to get repeatability?**

  Cloud systems can be dynamic and non-deterministic and so we can lose repeatability in testing. For example, we might run a test, observe a failure, but not be able to reproduce this failure. This can make it more difficult to locate and fix a fault. There is the challenge of devising methods to help make testing repeatable.

- **Applying existing methods of concurrency testing**

  Cloud systems are likely to be highly concurrent. Several testing methods have been produced for testing concurrent systems. However, these are typically for testing multi-threaded systems. There is a need to devise such methods that are specialised for Cloud systems: they need to scale to large systems and also to work with systems that are highly distributed. The properties of Cloud systems are likely to introduce significant theoretical and practical challenges.

- **MBT vs. Runtime Monitoring**

  In runtime monitoring we observe the behaviour of the system and check that the observations are consistent with certain requirements. There is the potential to use the same model in MBT and runtime monitoring. Runtime monitoring could be a good complement for testing. It is potentially less costly and easier to conduct when compared to MBT since there is not the need to generate and apply test cases. However, runtime monitoring is a form of passive testing and so the tester does not decide which parts of the system are to be tested: this depends on how the system is used. There is a need to understand how MBT and runtime monitoring can complement and where each is suitable.

- **Strong collaboration between research and industry**

  There is a need for more significant real collaboration between academia and industry. In order to make experimental results more relevant, academia needs real environments and subjects. These have to come from industry. All should gain from such collaboration: if companies can be persuaded to provide such environments and subjects then the results of experiments will be more relevant and will be particularly relevant to these companies.

Adoption, Model-Driven Development, Functional and Extra-Functional Requirements

- **Ensure Model Quality.**

  The quality of the model used is important, but how can we measure this? Are there criteria that we can use to evaluate the quality of models? It appears that there are no standard criteria and it seems likely that criteria will vary with problem and/or domain. There is also likely to be a trade-off between factors, such as clarity/abstraction and containing all of the important aspects.

  There are also issues related to the modelling language since different modelling languages can lead to different tests being derived. The goals of tests generated from models depend on the features of the modelling language. It is unclear how we can decide which modelling languages are most suitable and also how we measure the suitability of a language.

- **How to Teach / Learn Modelling?**

  It is often said that many developers and testers find it difficult to produce appropriate models. Assuming this is the case, there is the challenge of finding effective ways to teach modelling. It might be helpful if researchers and practitioners published 'good' models and possibly also 'bad' models. Naturally, what is meant by a 'good' or 'bad' model depends on the intended use of the model and so there may be a need to produce several sets of examples. Naturally, tool support for model authoring and analysis is also important.

- **Is There a Methodology for Modeling?**

The development of good models would be facilitated by there being suitable methodologies. The methodology may well depend on the type of language used and, for example, whether it is compositional. It may also depend on the intended use of the model since the model can completely change for different concerns.

- **Do we need Multi-Faceted Models?**

  Different kinds of models are needed for non-functional concerns and it seems that this is particularly the case for distributed systems. As it turned out, this same question was, in a slightly different context, discussed in the context of the cloud; a second listing of the insights is omitted here for brevity's sake.

- **Adoption of Model-Driven Engineering**

  Much of industry still does not use model-driven engineering (MDE) despite the significant amount of research in MDE. In part this may be because the MDE community has not been effective in promoting some of the advantages of using formal models. There should be particular emphasis on this in MBT since the use of a formal model can allow test generation and execution to be automated but adoption is still limited. We are starting to see empirical results that show significant savings resulting from the use of MBT in industry and this may help promote the use of MBT and, in turn, MDE. However, the MDE community would benefit from additional case studies, possibly in an area such as the nuclear power plans in which there are several different concerns, such as safety and security.

  There are several alternative types of modelling languages, For example, the UML has the advantages of being standardised and having significant tool support. However, some prefer domain specific languages (DSLs). In addition, different languages are appropriate for different stakeholders. For example, one might use UML for designers and a language such as C++ for testers.

- **Relate requirements and models.**

  Requirements are typically text documents that cover several kinds of aspects such as security, safety, and timing. There is the question of how models can be derived directly from requirements and there are some methods, such as sequence enumeration, for doing this. If we can relate parts of the model to requirements then we gain traceability. This should help the tester to communicate the purpose of a test case (what requirements it tests) and also for failures to be traced back to requirements. Such traceability is crucial if MBT is to be used more widely and there is therefore a need for the development of methods for deriving a model in a manner that facilitates traceability.

Conclusions

There has been significant progress in MBT since the previous Dagstuhl seminar in 2004. In particular, many more tools are available and there is significant industrial uptake. However, many challenges remain. Some of these challenges were identified in 2004, an

example being understanding the relationship between coverage and test quality. In addition, new challenges have arisen. It appears that the move towards highly distributed systems, such as Cloud systems, introduces many interesting scientific and engineering challenges for the community. However, it also provides a great opportunity: these systems are extremely difficult to test in a systematic manner and if effective MBT approaches can be developed for such systems then this should further promote the industrial use of MBT.

7.8 Software Engineering for Self-Adaptive Systems

Seminar No. **10431** Date **24.10.–29.10.2010**
Organizers: Holger Giese, Hausi A. Müller, Mary Shaw, Rogério de Lemos

Introduction

The simultaneous explosion of information and integration of technology together with the continuous evolution from software intensive systems to systems of systems to ultra-large-scale (ULS) systems requires new and innovative approaches for building, running and managing software systems [ULS2006]. A consequence of this continuous evolution is that software systems are expected to become more versatile, flexible, resilient, dependable, robust, continuously available, energy-efficient, recoverable, customisable, self-healing, configurable, or self-optimising by adapting to changing requirements and contexts/environments [Roadmap2008]. One of the most promising approaches to achieving such properties is to equip software systems with self-managing capabilities using self-adaptation mechanisms.

Research on the theory and practice of self-adaptation is highly interdisciplinary, and it draws ideas and solutions from many diverse fields, such as control engineering and dynamical systems, automation and instrumentation, machine learning and planning, fault-tolerance and reactive systems, and many others. The applications of self-adaptation also span a wide range: autonomic computing [Huebscher2008], dependable computing, autonomic communications and networks [Dobson2006], mobile ad hoc networks, sensor networks, ubiquitous computing, computing systems management [Hellerstein2004], biologically-inspired computing, user-interface customisation, embedded computing, service-oriented architectures, web-service composition, embedded systems, mechatronics, mobile and autonomous robots, multi-agent systems, to financial systems.

The seminar focused on software engineering aspects of building self-adaptive and self-managing systems. The topic of self-adaptive systems has been studied independently within the different research areas of software engineering, including requirements engineering, modelling, architecture and middleware, event-based, component-based and knowledge-based systems, testing, verification and validation, as well as software maintenance and evolution [Dagstuhl2008]. Recently several workshops have emerged to bring these independent efforts together by concentrating on the software engineering aspects of self-adaptive systems: WOSS (Workshop on Self-Healing Systems), WADS (Workshop

on Architecting Dependable Systems), DEAS (Design and Evolution of Autonomic Application Software), SEAMS (Software Engineering for Self-Adaptive and Self-Managing Systems), and Dagstuhl Seminar 08031 on Software Engineering for Self-Adaptive Systems [Dagstuhl2008].

The flexible nature of software provides an ideal platform for self-adaptation. However, the proper realisation of the self-adaptation functionality remains a formidable intellectual challenge. In the long run, we need to establish the foundations that enable the systematic development of future generations of self-adaptive systems. Therefore the current achievements have to be integrated into a more comprehensive overall research effort from which generic approaches should be devised. Building self-adaptive software systems cost-effectively and in a systematic and predictable manner is also a major engineering challenge.

The goal of this seminar was to bring together the leading software engineering experts and other distinguished experts from related fields on self-adaptive systems to discuss the fundamental principles, models, methods, techniques, mechanisms, state-of-the-art, and challenges for engineering self-adaptive software systems.

Topics of discussion

The aim of the Seminar was not so much to be comprehensive concerning the topics associated with software engineering for self-adaptive software systems, but to be focused on key and challenging topics.

- **Design spaces**: The development of complex software involves models at different levels of granularity. Also all approaches to self-adaptation which plan the effect of changes ahead of their application require models as a means for prediction. Similarly, approaches to estimate the quality of a solution at development-time also needs models to evaluate the quality of the adaptation. Following different trends these models may be oriented towards control theory featuring related analysis capabilities [Hellerstein2004] or support a more architecture centric view [Kramer1996], [Kramer2007]. Other relevant aspects are their partial or global character as well as the capability to adjust the models to observations of the system and environment to also address un-anticipated dynamic changes. Therefore, the question which models are required for the different development steps of self-adaptive systems is of paramount importance. Design spaces represent the set of design alternatives as decisions to be made, along with the alternatives for each decision and criteria that guide the choice. The decisions are often organised hierarchically.

- **Verification and validation**: The vision of self-adaptive systems is promising but also risky. Giving systems the freedom to self-adapt at run-time obviously weakens our capabilities to assure their proper operation. Therefore, an important question is how much run-time assurance is needed in addition and above design-time assurance. Moreover, it is crucial to identify properties which are only assessable at run-time (e.g., stability) and thus irrelevant in the contact of traditional static systems. What

are the validation requirements associated with each type of adaptation? How can these obligations become a visible part of the development process?

- **Processes**: Activities such as selecting a single execution strategy which are traditionally handled at development-time may be deferred to run-time in a self-adaptive system to gain more flexibility and the capability to adapt to the current situation at run-time. It seems important to identify which activities are often deferred to run-time as well as for which activities this seem still not possible respectively useful.

- **Decentralisation**: Self-adaptive systems can be viewed from two perspectives, either individual systems or cooperative systems. Individual self-adaptive systems evaluate their own global behaviour and change it when the evaluation indicates that they are not accomplishing what they were intended to do, or when better functionality or performance is possible realising self-adaptation. In a centralised fashion. such systems typically operate with an explicit internal representation of themselves and their global goals. On the other hand, self-adaptive systems can be composed of a large number of components that interact according to local and typically simple rules in a decentralised manner. The global behaviour of the system emerges from these local interactions, and it is difficult to deduce properties of the global system by studying only the local properties of its parts. Such systems do not necessarily use internal representations of global properties or goals. These two perspectives of self-adaptive system represent two extremes going from full centralized to a fully decentralized. In practice, there is clearly a continuum of designs and implementations where both system types are intertwined to achieve their behavioural goal.

Outcomes

The two concrete outcomes from this Seminar will be a new roadmap paper and a new book. The roadmap paper, which will follow the same format of the previous paper will be structured according to the new topics to be identified during the seminar, namely: design spaces, processes, verification and validation, and decentralisation. For each topic, the objective is to summarise the current state-of-the-art, discuss its limitations, and identify future challenges for the field.

The book will contain state-of-the-art contributions from participants of the seminar and some invited contributions. In addition to these contributions, the roadmap paper will be the introductory chapter of the book, which should be followed by four chapters containing extended versions of the topics discussed in the roadmap paper. The book will be published by Springer as Lecture Notes in Computer Science volume on their State-of-the-Art series.

[Dagstuhl 2008] Schloss Dagstuhl Seminar 08031. Software Engineering for Self-Adaptive Systems, Wadern, Germany (2008) `http://www.dagstuhl.de/08031/`

[Dobson 2006] Dobson, S., Denazis, S., Fernandez, A., Gaiti, D., Gelenbe, E., Massacci, F., Nixon, P., Saffre, F., Schmidt, N., Zambonelli, F.: A Survey of Autonomic Communi-

cations. ACM Transactions on Autonomous and Adaptive Systems (TAAS) 1(2):223-259 (2006)

[Hellerstein 2004] Hellerstein, J.L., Diao, Y., Parekh, S., Tilbury, D.M.: Feedback Control of Computing Systems. John Wiley & Sons (2004)

[Huebscher 2008] Huebscher, M.C., McCann, J.A.: A Survey of Autonomic Computing-Degrees, Models, and Applications. ACM Computing Surveys, 40 (3):7:1-28 (2008)

[Kramer 1996] Kramer, J., Magee, J.: Dynamic Structure in Software Architectures. ACM SIGSOFT Software Engineering Notes 21(6):3-14 (1996)

[Kramer 2007] Kramer, J., Magee, J.: Self-managed Systems: An Architectural Challenge. In: Future of Software Engineering (FoSE 2007), pp. 259-268, IEEE Computer Society, Washington, DC, USA (2007)

[Roadmap 2008] Cheng, B.H.C., et al: A Research Roadmap: Software Engineering for Self-Adaptive Systems. Schloss Dagstuhl Seminar 08031 Report on Software Engineering for Self-Adaptive Systems, Wadern, Germany, 12 pages (2008)

[ULS 2006]* Northrop, L., Feiler, P., Gabriel, R., Goodenough, J., Linger, R., Longstaff, T., Kazman, R., Klein, M., Schmidt, D., Sullivan, K., Wallnau, K.: Ultra-Large-Scale Systems: The Software Challenge of the Future. Technical Report, Software Engineering Institute, Carnegie Mellon University, 134 pages (2006) `http://www.sei.cmu.edu/uls`

Chapter 8

Distributed Computation, Networks, Architecture

8.1 Perspectives Workshop: Digital Social Networks

Seminar No. **10041** Date **24.01.–29.01.2010**
Organizers: François Bry, Clemens Cap, Ingo Dahm, Julia Maintz, Sebastian Schaffert

Summary

The Perspective Workshop on "Digital Social Networks" held in between January 25th and 29th, 2010 at the research centre Schloss Dagstuhl focused on technological, socio-economical, and political aspects of digital social networks, and, more generally, digital social media. Digital social media give rise to users and communities to collaboratively generate and exchange content and to interact. They enable social computation, i.e. computations that involve both software and groups of people. They are operated by specific software systems called social software and use information and communication technologies such as the Internet and Web technologies. Digital social media ease and strengthen social interactions by overcoming physical limitations in communication (like distance and synchronicity) and alleviating human limitations like the number of people with whom one can maintain relationships. Digital social media build, and/or rely upon, social networks that might be the primary purpose of the media. The workshop participants identified the following issues as today's major socio-economical, political, and technological challenges of digital social media:

Social Media Literacy

Digital social media re-launch social communication and expression. Therefore, they challenge to re-consider self-expression, self-perception, social identity, and social participation. Awareness of both, the possibilities and dangers of digital social media, is still insufficient and needs to be strengthened. In order to avoid a digital social media divide, social media literacy must be promoted.

Digital Social Media Governance

The reshaping of our society, culture, and economy through digital social media implies a need for novel policies and therefore represents major political challenges. Debated current issues include privacy, the ownership of user-generated content, and update and exploitation rights on such contents. There is the challenge to find an appropriate societal balance between regulation and free media usage and to deliver translations into technical solutions.

Technological Challenges

Digital social media must move to new levels of social scalability. As one-to-many distribution models move on to many-to-many architectures, they must avoid an information overloading of their users. Research and further development in the areas of content selection and user attention management are essential. Moreover, the improvement of social media usability, in particular the development of appropriate user interfaces, is required. The problem is most pressing in the mobile domain and for elderly and impaired audiences. Digital social media, especially today's social networks, predominantly grow and persist as 'segregated data silos'. As interoperability is a pre-requisite to user control of user-generated data, the development of interoperable systems and definition of novel standards are required. Today's media build more and more upon pervasive digital-physical systems able of so-called ambient intelligence and of adaption to geographical and social contexts. Conceiving and bringing ambient intelligence and adaption to their full potentials represent considerable technological challenges for the further development of digital social media. How news and ideas widespread in a community and how such a wide spreading can be initiated or enhanced by algorithms is, possibly, one of the most challenging technological issues.

The workshop participants identified an urgent need for action in education, fundamental and applied research:

- Digital social media should become both, subject of teaching and tools, at primary and secondary schools.
- Interdisciplinary study courses and doctoral schools in social and human computation should be started at universities.
- A "Priority Programme" of the "German Research Foundation" in social and human computation should be initiated.
- European and national funding programmes should open up to applied research in digital social media, i.e. allow for a joint academic-industrial research accessible especially to small and medium sized enterprises.

The workshop participants are convinced that a fast deployment of actions in education, fundamental and applied research like the aforementioned is necessary for giving Germany and Europe a good position to be at the cutting-edge of digital social media research and education.

8.2 Semantic Challenges in Sensor Networks

Seminar No. **10042** Date **24.01.–29.01.2010**
Organizers: Karl Aberer, Avigdor Gal, Manfred Hauswirth, Kai-Uwe Sattler, Amit P. Sheth

There has been significant progress in the number and capabilities of mobile devices, wireless sensors, and sensor networks. These developments, combined with the improved ability to bridge between the physical and cyber world in a more seamless way, have fostered the broad availability of sensor data capturing the state of the physical world. Promising and already successful examples are applications in environmental monitoring, agriculture, surveillance and intrusion detection, public security, and supply chain management. Furthermore, ideas towards a Web of sensors have been proposed, which is to be understood as a (large scale) network of spatially distributed sensors. In particular, terms like "Internet of Things", "Collaborating Objects" and "Ambient Intelligence" emphasize the trend towards a tighter connection between the cyber space and the physical world.

The existence of a huge number of sensor sources producing data continuously results in tremendous data volumes which are often valid or useful only for certain period of time and are never inspected by humans. In order to make sensor data useful despite the lack of human supervision in the loop, semantic annotation and analysis becomes a key component in setting up sensor data-based applications: Only if sensors and sensor data are annotated and enriched by information describing their meanings, source, and validity scope, they can be automatically discovered, processed and combined with other data in an open world. The kind of useful semantics ranges from technical metadata describing the sensors and the measurements (time, location, sensor type, validity, measurement error etc. as partially captured by standardization proposals like SensorML) to emergent semantics derived by aggregating, combining, analyzing, and enriching the raw data, e.g., in the form of analytical models, annotations, correlations etc. on the other spectrum, the data collected by human-in-the-loop sensing is small but of significant verity and complexity (e.g., language nuances, and capturing sentiments and emotions), which offer additional challenges to annotation, integration and analysis of such data.

Modeling, representing, discovering and deriving as well as using semantics for sensor data raise several challenges which are related to different aspects of developing, deploying, and using sensor network based applications. Thus, the goal of this seminar was to bring together researchers from relevant areas, such as:

- sensor node providers and sensor networking,
- data fusion and data stream processing,
- sensor middleware,
- geospatial and uncertain data management,
- semantic integration and Semantic Web, and

- social computing and collective intelligence.

Semantics plays an important role in all of these areas, either by producing and enriching data with explicit semantics or by exploiting semantics for data processing. Therefore, sharing and exchanging knowledge and experiences among disciplines could result in significant synergy effects.

The seminar focused on the following major issues:

- methodologies and languages for modeling and representation, issues of sensing-perception-semantics,
- standards, ontologies, and middleware for semantic sensor networks, ¡item¿ semantic annotation of high throughput machine sensor data as well as social/human-in-the-loop sensing data,
- emergent semantics in sensor networks and sensor data processing,
- exploiting sensor data semantics for geospatial and uncertainty data management, and
- specific use cases and applications of semantic sensor networks.
- review of related community efforts that are directly relevant to the seminar topic, especially W3C's XG on Semantic Sensor Networking,

The objectives set out by the organizers were to analyze the state of the art in the different areas with respect to semantics, discuss problems, specific methodologies and applications of semantic-aware sensor networks and emergent semantics as well as to identify future trends and research directions.

The seminar was well attended: 27 researchers from Europe, Asia, and North America actively contributed to the seminar. During the week, two special discussion sections were organized where the seminar was split in smaller groups. The topics of these discussion groups were "Data Representation & Semantics", "Query Models", "Architectures for Semantic Sensor Networks", and "Application Requirements".

At the end of the seminar a joint session with the parallel Dagstuhl seminar "Digital Social Networks" was held to explore research topics at the intersection of both research domains, for example, the use of social network infrastructures to discover and publish sensor data, the problem of privacy at the intersection of sensor networks, mobile phones and social networks, and the use of social networking in social sensing. There was enthusiasm to organize a follow-on seminar bringing together the two areas and a proposal for a Dagstuhl Seminar is planned.

8.3 Quantitative and Qualitative Analysis of Network Protocols

Seminar No. **10051** Date **31.01.–05.02.2010**
Organizers: Bengt Jonsson, Jörg Kreiker, Marta Kwiatkowska

The Dagstuhl seminar *Quantitative and Qualitative Analysis of Network Protocols* was held in the week from January 31, 2010 to February 5, 2010. It was the first Dagstuhl seminar to bring together researchers from the network and systems community as well as from the verification community. 10 female and 33 male researchers from institutes in 12 different countries including two industrial participants contributed to the success of this event.

The working hypothesis of this seminar was based on a perceived imbalance. On the one hand network researchers may tend to - in the absence of a suitable modeling discipline - build and then measure rather than to model, verify, and then build. On the other hand, network models analyzed by the formal verification community may be overly simplified. Below we explain how we went about this imbalance during the seminar.

During the Monday morning session each participant spent a few minutes on her background, her modelling and verification challenges, and her expectations regarding the seminar. In the afternoon four tutorials provided state-of-the-art information representing each of the four communities, where the verification community is viewed as three different sub-communities.

On the remaining days we enjoyed 29 talks well distributed over the different communities. Talks addressed plenty of methodologies like simulation, graph theory, graph rewriting, process algebras, static analysis, model checking, quantitative model checking, theorem proving and control theory. All of them were successfully applied to specific problems. One of the expected outcomes of the seminar was a benchmark collection. Indeed, a number of the presented case studies may serve this purpose: Chess WSN clock synchronization (Vaandrager), multipath routing (N.Walker), Fraglets, in particular the alternating bit protocol written in Fraglets (Tschudin, Vaandrager), DYMO (Jonsson), AODV (Nanz), ZigBee key establishment (Yuksel), API Authorization (Lee), or Gossip (Bakshi).

Other than the talks we had two long and insightful discussion sessions on Tuesday afternoon and on Friday morning. One key observation from the discussions was that simulations are hard to trust and often do not work. Many people agreed that models cannot be trusted either. On the other hand, it was pointed out that many probabilistic systems tend to behave deterministically in the limit. This allows for continous approximations of discrete behaviour, like the mean field method, the chemical master equation, or statistical model checking. All of these approaches are promising with respect to huge state spaces, which prevent scalability for many discrete, finite abstractions-based methods. A better use of such methods might be the exploration and discovery of corner cases, which can be hard to detect using statistical methods.

We conclude by a few personal remarks. First, it should not go unmentioned, that the seminar suffered from the unexptected absence of the one organizer representing the net-

work community, Timothy Griffin from Cambridge. It came as a blow to us, just days before the seminar. Second, the seminar took place during a week with heavy snowfall. So much snow, in fact, that virtually all outdoor activities including the traditional Wednesday excursion had to be cancelled. However, ending on a positive note, we were glad to observe a number of inter-community collaborations getting sparked. We are optimistic that these will be perpetuated.

8.4 Program Development for Extreme-Scale Computing

Seminar No. **10181** Date **02.05.–07.05.2010**
Organizers: Jesus Labarta, Barton P. Miller, Bernd Mohr, Martin Schulz

The number of processor cores available in high-performance computing systems is steadily increasing. A major factor is the current trend to use multi-core and many-core processor chip architectures. In the November 2009 list of the TOP500 Supercomputer Sites, 98.4% of the systems listed have more than 2048 processor cores and the average is about 9300. While these machines promise ever more compute power and memory capacity to tackle today's complex simulation problems, they force application developers to greatly enhance the scalability of their codes to be able to exploit it. This often requires new algorithms, methods or parallelization schemes as many well-known and accepted techniques stop working at such large scales. It starts with simple things like opening a file per process to save checkpoint information, or collecting simulation results of the whole program via a gather operation on a single process, or previously unimportant order $O(n^2)$-type operations that now quickly dominate the execution. Unfortunately many of these performance problems only show up when executing with very high numbers of processes and cannot be easily diagnosed or predicted from measurements at lower scales. Detecting and diagnosing these performance and scalability bottlenecks requires sophisticated performance instrumentation, measurement and analysis tools. Simple tools typically scale very well but the information they provide proves to be less and less useful at these high scales. Clearly, understanding performance and correctness problems of applications requires running, analyzing, and drawing insight into these issues at the largest scale.

Consequently, a strategy for software development tools for extreme-scale systems must address a number of dimensions. First, the strategy must include elements that directly address extremely large task and thread counts. Such a strategy is likely to use mechanisms that reduce the number of tasks or threads that must be monitored. Second, less clear but equally daunting, is the fact that several planned systems will be composed of heterogeneous computing devices. Performance and correctness tools for these systems are very immature. Third, the strategy requires a scalable and modular infrastructure that allows rapid creation of new tools that respond to the unique needs that may arise as extreme-scale systems evolve. Further, a successful tools strategy must enable productive use of systems that are by definition unique. Thus, it must provide the full range of traditional software development tools, from debuggers and other code correctness tools

such as memory analyzers, performance analysis tools as well as build environments for complex codes that rely on a diverse and rapidly changing set of support libraries.

Many parallel tools research groups have already started to work on scaling their methods, techniques, and tools to extreme processor counts. In this Dagstuhl seminar, we wanted participants from Universities, government laboratories and industry to report on their successes or failures in scaling their tools, review existing working and promising new methods and techniques, and discuss strategies for solving unsolved issues and problems.

This meeting was the forth in a series of seminars related to the topic "Performance Analysis of Parallel and Distributed Programs", with previous meetings being the Dagstuhl Seminar 07341 on "Code Instrumentation and Modeling for Parallel Performance Analysis" in August 2007, Seminar 02341 on "Performance Analysis and Distributed Computing" held in August 2002, and Seminar 05501 on "Automatic Performance Analysis" in December 2005.

The seminar brought together a total of 46 researchers and developers working in the area of performance from universities, national research laboratories and, especially important, from three major computer vendors. The goals were to increase the exchange of ideas, knowledge transfer, foster a multidisciplinary approach to attacking this very important research problem with direct impact on the way in which we design and utilize parallel systems to achieve high application performance.

8.5 Flexible Network Design

Seminar No. **10211** Date **24.05.–28.05.2010**
Organizers: Anupam Gupta, Stefano Leonardi, Berthold Vöcking, Roger Wattenhofer

Network design with its many variants is one of the most active research areas in theoretical computer science involving researchers from Algorithms and Complexity, Combinatorial Optimization, Distributed Computing and Algorithmic Game Theory. This area is fertile ground for new problems and the development of new techniques, since modern communication systems require a flexible and permanent adaptation of the network structure to highly dynamic access pattern especially in the context of mobile and adhoc networks. This brings many new aspects into network design such as flexible construction, dynamic scheduling of recources, adaptive bandwidth or spectrum assignment, and routing with respect to varying demands.

The goal of this seminar was to bring together experts from the several different communities above, and to give them the possibility to discuss recent advances, understand current trends, identify understudied areas, and formulate new directions for further investigation in this area. The seminar focussed on emergent problems in the area of flexible network design like online and oblivious network design, network design with game theoretic approaches, positioning in wireless sensor networks, and scheduling interfering signals in wireless networks.

The seminar provided an opportunity for information sharing and collaborations, and to identify new problems and areas for future collaboration. Indeed, to facilitate the

exchange of ideas, some of the participants were requested to give overview talks or surveys of subjects of cross-cutting interest; this was complemented by shorter talks by other participants on specific research results.

8.6 Dynamically Reconfigurable Architectures

Seminar No. **10281** Date **11.07.–16.07.2010**
Organizers: Peter M. Athanas, Jürgen Becker, Jürgen Teich, Ingrid Verbauwhede

Dynamic and partial reconfiguration of hardware architectures such as FPGAs and coarse grain processing arrays bring an additional level of flexibility in the design of electronic systems by exploiting the possibility of configuring functions on-demand during run-time. When compared to emerging software-programmable Multi-Processor System-on-a-Chip (MPSoC) solutions, they benefit a lot from lower cost, more dedication and fit to a certain problem class as well as power and area efficiency. This has led to many new ways of approaching existing research topics in the area of hardware design and optimization techniques. For example, the possibility of performing adaptation during run-time raises questions in the areas of dynamic control, real-time response, on-line power management and design complexity, since the reconfigurability increases the design space towards infinity.

This Dagstuhl Seminar on Reconfigurable Architectures has aimed at raising a few of the above topics. In the methodological area, on-line placement, pre-routing/on-line routing and dynamic compaction algorithms and techniques were presented. In the architectural realm, novel interconnection schemes as well as hypermorphic architectures for the future of reconfigurable computing systems in general were introduced and discussed. A major question here was whether programmable multi-core Systems-on-a-Chip (MPSoC) will win margins for typical application domains such as rapid prototyping and emulation and how to scale reconfigurable hardware further in the threat of diminishing returns due to more and more unreliable components of future nano-electronic devices. Also questioned were tool maturity and recent developments in the usage of reconfigurable computing for increase of fault-tolerance. Finally, for the first time, a special focus day was spent on the area of embedded security and the role of reconfigurable hardware in this emerging and important application area.

8.7 Perspectives Workshop: Service Value Networks

Seminar No. **10301** Date **25.07.–30.07.2010**
Organizers: William E. Hefley, Steffen Lamparter, Christos Nikolaou, Stefan Tai

Summary

Services are receiving increasing attention in both economics and computing. This trend is due to two observations:

1. From an economics viewpoint, services today are contributing the majority of jobs, GDP, and productivity growth in Europe and in other countries worldwide. This includes all activities by service sector firms, services associated with physical goods production, as well as services of the public sector.

2. From an ICT viewpoint, the evolution of the Web enables provisioning software-as-a-service (and, services-as-software). Modern software systems are designed as service-oriented architectures consisting of loosely coupled software components and data resources that are accessible as Web services.

The notion of "service" used in the corresponding research communities is different; however, they are not independent but have a strong symbiotic impact on each other. ICT services create and enable new ways of business process management and value co-creation for service providers and service consumers. They enable companies to spread their planning, design, manufacturing, distribution, and delivery functions. The modularization of corporate functions takes place in a wide range of industries (electronics, car manufacturing, aerospace, retail, etc.). Competitive markets evolve best of breed functions, which in turn encourage deconstruction of formerly vertically organized companies and industries into **service systems**, also referred as **value networks**, to capitalize on this advantage. One of the key aspects distinguishing service systems from the traditional product-centric view is the importance of value co-creation, i.e. customers act as co-producers in the service provisioning process or as co-innovators in the evolution of services. The concept of **service value networks** (SVNs) captures this idea by modeling the business structures and inter-relationships and dependencies between service providers, consumers, and intermediaries (or enablers). They facilitate representation of flexible, dynamic supply and demand chains together with their social, technological, and economic context. Thus, service value networks are a promising conceptual framework to better understand and manage the operational, strategic, and technological challenges of business design and business alliance formation.

The goal of the Dagstuhl Perspectives Seminar on Service Value Networks was to bring together researchers and experts from the various relevant fields and discuss the different existing approaches to modeling service value networks, identify shortcomings and open challenges, and to suggest a research agenda that leads to a better understanding of the functioning of complex service systems. This research agenda [1] identifies SVN research challenges. To address these challenges, the Manifesto provides a synthesis of outstanding research questions, with an emphasis on modeling and analysis of SVNs and the role of coordination in SVNs, and presents an outlook to future work to address these outstanding research questions and move the SVN field ahead.

References

[1] Hefley, B. (ed.). (2010). Dagstuhl Perspectives Workshop Manifesto: Service Value Networks Reflections and Future Directions. (Schloss Dagstuhl Perspectives Workshop 10301). Schloss Dagstuhl - Leibniz Center for Informatics, Wadern, Germany.

The Web of Services

Hans Akkermans (Vrije Universiteit - Amsterdam, NL)

At the cross-section of Web Science and Service Science, we investigate how the Web may develop into an open, common and universal platform of services, such that - ideally - end customers can submit a need/request and then get appropriate service bundle offerings in response.

These offerings may come from existing multi-supplier SVNs or may be formed ad hoc on-the-fly.

What are needed elements to achieve such a Web [Web 4.0] acting as a big service cloud or rather marketplace, as seen from a value perspective?

(a) service semantics: services are to have self-descriptions as socio-economic entities (cf. [1]), either in some form of catalogue or fully decentralized as meta-info in service components themselves.

(b) pragmatics: there must be facilities for interactive, but structured and knowledge-based, customer dialogue [based on service semantics], such that high-level needs are progressively refined, adapted and ultimately matched to actually available service bundles.

(c) network "socio-digital" formation logics: services on the Web need to be modular and self-configurable for automated service bundling triggered by customer interaction.

We have developed a number of ontological theories, computational methods and tools for such an eco-system of self-organizing services. These developments are based on a range of practical and industrial use cases and case studies we work on in different sectors (e.g. W4RA Regreening in Africa, e-health dementia care, smart power grids).

In our experience, Service research poses challenges to interdisciplinary scientific method. Neither the typical CS engineering science approach nor the social science empiricist schools are up to this challenge. Recent MIS attempts at defining a design science, rooted in H. Simon's views on researching artefacts, are also too narrow and found wanting.

References

[1] Hans Akkermans, Ziv Baida, Jaap Gordijn, Nieves Peña, Ander Altuna, Iñaki Laresgoiti: Value Webs: Using Ontologies to Bundle Real-World Services. IEEE Intelligent Systems 19(4): 57-66 (2004)

8.8 User-Centric Networking

Seminar No. **10372** Date **12.09.–15.09.2010**
Organizers: Jon Crowcroft, James Kempf, Paulo Jorge Mendes, Rute Sofia

In the latest years the highly nomadic lifestyles that Internet users experience, and the strong entanglement between society and technology, lead to the appearance of community networks where the end-user has, most of the times, an active role in terms of sharing Internet access. Such networks range from basic functionality, such as the ability to create a wireless (ad-hoc) network on-the-fly with a simple PC (e.g., Internet Connection Sharing functionality from Microsoft), or more elaborate cases of commercial success, e.g. FON. Wireless networks provided by end-users are expected to grow, despite the limitations imposed by traditional operator-centric Internet communication models. In this new scenario the end-user (or a community of end-users) is a micro-operator in the sense that he/she shares his/her subscribed broadband Internet access based on some incentive scheme. Besides Internet access sharing, being a micro-operator also means providing other network functionality, such as local mobility management and store-cache-forward mechanisms, based on the right set of incentives as well as on adequate information concerning the way people interact and move.

User-provided networks disrupt Internet communication models due to its user-centricity. First, any regular end-user device may behave as a relay of information and consequently becomes part of the network, which has an impact on the Internet architecture: the central building block of the Internet, the end-to-end principle, describes a clear splitting between network and end-user systems. Second, user-provided networks grow spontaneously based on the willingness of users to share subscribed Internet access and to relay data. Such willingness is sustained by trust management and incentive mechanisms that mimic social behavior. Third, between communities of users connectivity is intermittent. User-provided networks have to consider mechanisms that support routing in intermittent connected networks, as well as quick and transparent mobility management between micro-operators. Fourth, people mobility has an impact on connectivity provided by user-provided networks, since human carried devices will also be willing to operate as networking devices. This characteristic means that a realistic modulation of human mobility and social patterns is fundamental for the optimization of user-centric networking technologies. Fifth, most of the devices used by humans will use wireless technologies to communicate, which means that taking advantage of the broadcast nature of the wireless media by means of cooperative networking techniques is of core importance to an efficient deployment of user-centric technologies, in terms of throughput gain and energy saving for instance. The purpose of this seminar is to bring together researchers from core disciplines for the future of user-centric networking: Internet architecture, human behaviour and mobility modeling, as well as wireless and routing optimization.

The presentation and brainstorming sessions allowed all participants to discuss the role of the user-centric networking in several different perspectives such as user-centric requirements (e.g. social behaviour; energy efficiency), mobility management, access aspects, cloud architecture, trust management, network fault detection and repair, cooperative networking, identity management and applications.

Based on the results of the brainstorming session, it was decided that it would be interesting to create a User Centric Networking Working Group (UCN-WG) to continue the discussion started in this Dagstuhl seminar. To start wit it was created a mailing list (ucn@uitc.ulusofona.pt) and a wiki page. The latter should be used to collect links about relevance information, from presentations to publications, running code, etc.

8.9 Inter-Vehicular Communication

Seminar No. **10402** Date **03.10.–06.10.2010**
Organizers: Falko Dressler, Frank Kargl, Jörg Ott, Ozan Tonguz, Lars Wischhof

The management and control of network connections among vehicles and between vehicles and an existing network infrastructure is currently one of the most challenging research fields in the networking domain. Using the terms Vehicular Ad-hoc Networks (VANETs), Inter-Vehicle Communication (IVC), Car-2-X (C2X), or Vehicle-2-X (V2X), many applications - as interesting as challenging – have been envisioned and (at least) partially realized. In this context, a very active research field has developed. There is a long list of desirable applications that can be grouped into four categories:

- eSafety applications that try to make driving safer, e.g. road hazard warning;
- traffic efficiency applications aiming at more efficient and thus greener traffic, e.g. detection of traffic jams;
- manufacturer oriented applications, e.g. automatic software updates; and
- comfort applications, e.g. automatic map updates.

While there are some similarities with fields like mobile ad-hoc networks or wireless sensor networks, the specific characteristics of vehicular networks require different communication paradigms, different approaches to security and privacy, or different wireless communication systems. For example, the nodes usually do not have severe power and form factor constraints, and they might be always on. On the other hand, due to high relative speeds, wireless connections may not be stable for a longer time period and the network density is expected to vary from sparse to very dense networks.

Another challenging issue is the efficient use of available infrastructure, such as road side units or even cellular networks. Furthermore, IVC has strong links to other research domains, e.g., geo-informatics as it requires very precise localization and precise maps or highly scalable simulations that are a requirement for analyzing traffic systems with hundreds or thousands of vehicles.

In the past, many specific solutions for IVC have been identified and now, industry and other stakeholders are already calling for standardization. Still, we believe that many important research questions have only been partially answered and the approaches discussed in the standardization bodies are based only on a minimum consensus of simplest solutions. Security and privacy, scalability, use of advanced communication patterns like aggregation, transmit power control, and optimal medium access are just a few of such issues.

The main goal of this seminar was to bring together leading researchers both from academia and industry to discuss and evaluate the state of the art and to highlight where sufficient solutions exist today, where better alternatives need to be found, and also to give directions where to look for such alternatives. Furthermore, it was the goal of this workshop to go on

step beyond and identify where IVC can contribute to the basic foundations of computer science or where previously unconsidered foundations can contribute to IVC.

For example, IVC has triggered active research on reactive and dynamic security systems that do not try to provide security in a cryptographic sense at usually high costs, but create a tunable security-performance trade-off using reputation and consistency-checking mechanisms that are not unlike human and social mechanisms to estimate trust in information. It remains to be seen if such mechanisms can be generalized and be applied to future networks that will be dynamic and self-organizing in nature.

We organized four working groups on some of the most challenging issues in inter-vehicular communication:

- Fundamental Limits (Hannes Hartenstein),
- Communication Principles and Patterns (Ozan Tonguz),
- Security & Privacy (Elmar Schoch), and
- Simulation and Modeling (Martin Treiber and Christoph Sommer).

The workshop gathered a roster of highly qualified senior participants and several talented young researchers from both academia and industry, who convened to discuss issues in the listed working groups. We kept a very loose schedule with four invited speeches by leading experts in their respective domains as a starting point for each working group. Most of the time was spend in working group discussions. The key outcome of our working groups is available as a report in the Dagstuhl Seminar proceedings.

8.10 Impact of Human Mobility on Communications : Measurement, Analysis, Modeling, and Simulation

Seminar No. **10403** Date **06.10.–09.10.2010**
Organizers: Kevin Almeroth, Gunnar Karlsson, Cecilia Mascolo, Jörg Ott

Human mobility can be classified using three levels: strategic, tactical and operational mobility. At the strategic level humans decide their daily movement patterns and activities, such as go to work or to walk in the park. The tactical level considers the implementation of a strategic decision, such as choosing a way of travel. At the operational level, human movement is considered, including speed, physical size of nodes and interaction with others due to queuing or for avoiding collisions.

Capturing mobility in all its facets is crucial to the evaluation of mobile communication systems since it affects the quality and availability of a radio channel. There are two trends that motivate a closer look at mobility. First, the spatial dimensions of cellular wireless communication systems are shrinking with their evolution to enable spatial reuse of spectral bands and higher data rates per node. The second trend is the interest in

self-organizing ad hoc networks, as connected multi-hop networks or sparse intermittently connected delay-tolerant networks. When cell sizes shrink and multi-hop ad hoc communication becomes more prevalent, it is apparent that mobility is not only affecting the radio channel but is also causing churn in mobile networks and intermittency.

The performance of wireless communication systems is likely to be affected in a different way by each level of mobility. Decisions at the strategic and tactical levels determine the regularity and routines in movement which in turn affect how and when nodes interact with each other or with infrastructure. Some routing protocols for delay-tolerant networks try to take advantage of such non-randomness to route messages to a destination node. Mobility at the operational level presumably affects node connectivity and individual contact durations. This determines the amount of data that can be transferred per contact. Predictability and repeatability are thus two important aspects of movement patterns.

The mobility used in performance evaluation is synthetic and has several shortcomings including: 1) Migration of people in and out of a modeled area is missing; 2) There are few models that consider obstacles; 3) Periodicity and predictable patterns are not captured; and 4) User behavior, social relationships, and community structure is absent. Recently, there has been interest in measuring mobility, albeit in small and controlled groups (e.g., participants at a conference, or a sample of students on a campus). The measurement results allow for realism but are difficult to generalize from the often small scale and infrequent measurements.

This seminar highlighted human mobility and its connection to communication processes. We considered a wide variety of mobility types and discussed means to capture mobility on the strategic, tactical and operational levels. In addition to radio communication, we discussed the traffic patterns from mobile nodes and potential correlation between mobility and communication patterns with respect to the three levels. The seminar brought together researchers who take different approaches to mobility and pursue different methodologies: from measurements and analyses in social sciences to mobility modeling and simulations for understanding the impact of mobility on communication protocols and systems.

8.11 Information-Centric Networking

Seminar No. **10492** Date **05.12.–08.12.2010**
Organizers: Bengt Ahlgren, Holger Karl, Dirk Kutscher, Börje Ohlman, Sara Oueslati, Ignacio Solis

Information-Centric Networking (ICN) is one of the significant directions of current networking research. In ICN, the principal paradigm is not end-to-end communication between hosts - as it is in the current Internet architecture. Instead, the increasing amount of content that must be distributed requires alternatives: Architectures that work with information objects as a first-class abstraction; focusing on the properties of such objects and receivers' interests to achieve efficient and reliable distribution of such objects.

Such architectures make in-network storage, multiparty communication through replication, and interaction models such as publish-subscribe generally available for all kinds of applications, without having to resort to dedicated systems such as peer-to-peer overlays and proprietary content-distribution networks.

The ICN approach is currently being explored by a number of research projects, both in Europe (4WARD, SAIL, PSIRP) and in the US (DONA, CCN). The Delay Tolerant Networking (DTN) community has developed a message-oriented architecture that has been used along with ICN addressing and routing concepts. While these approaches differ with respect to their specific architecture, they share some assumptions, objectives and certain structuring architectural properties. In general, the aim is to develop network architectures that are better suited for content distribution, the currently prevailing usage of communication networks, and that better cope with disruptions in the communication service. The basic idea of ICN still leaves room for many variations. The Dagstuhl ICN seminar was intended as a catalyst for these variations and as a forum for discussing the following research topics:

- The relationship of networking architecture innovation vs. so-called over-the-top approaches in the application layer
- The support of an Internet of Things and Services by an ICN architecture
- How to migrate towards an information-centric architecture, and whether and how to use it as a migration enabler for, e.g., an IPv4/IPv6 technology step
- The role of and needs for naming and addressing and name resolution systems, along with the necessary security aspects of a naming scheme; a fundamental dichotomy between flat and hierarchical naming schemes needs to be resolved
- Efficiency and robustness of ICN data dissemination vs. specific content distribution overlay solutions
- The desirability of using specific transport protocols for ICN vs. the use of standard protocols like TCP or disruption tolerant protocols like the DTN Bundle protocol
- The integration and placement of caches inside a network
- Can the introduction of a new ICN architecture enable new types of applications that were too complex to create/operate/deploy/maintain in traditional networks?

Organization of the seminar

The seminar was organized as a 2.5 days seminar that provided room for presentation of approaches, results so far, as well as presentation and discussion of new ideas and selected specific topics.

The seminar was structured in 4 main blocks:

1. Presentation of on-going research activities

2. In-depth presentations and discussion of *naming*, *security*, and *routing and resolution* for ICN (Group Discussion 1)
3. In-depth presentations and discussion of *resource management and transport*, *ICN APIs and ICN hour glass waists*, and *deployments aspects, business models and incentives* for ICN (Group Discussion 2)
4. Discussion of seminar results and next steps

The seminar started, in the first block, with a set of presentations of on-going research activities:

- Teemu Koponen: DONA (Data-Oriented Networking Architecture)
- Jim Thornton: NDN (Named-Data Networking)
- Bengt Ahlgren: NetInf (Network of Information) in the 4WARD project
- George Xylomenos: PURSUIT project

The seminar then addressed important specific ICN topics such as naming, security, routing and resolution. For that, a set of discussion starter presentation set the scene by summarizing important issues and by providing new ideas:

- Christian Dannewitz: Naming and Security in Information-centric Networking
- Kevin Fall: Discussion on Information Centric Networking with a Security Focus
- Jarno Rajahalme: What's in a Data Name?
- Jussi Kangasharju: Naming and Search in Information-Centric Networks

These topics were then discussed in smaller groups (Group Discussion, part A), and the results of these discussions were presented and discuss in a plenary session.

In the second block of specific ICN topics discussion, several discussion starter presentations on resource management, congestion control, and ICN in challenged networks have been given:

- Van Jacobsen: Congestion Control and Transport in ICN
- Sara Oueslati: Ideas on Traffic Management in CCN
- Volker Hilt: Energy Consumption of Content-Centric Networks
- Joerg Ott: Delay-tolerant Networking: Elements of ICN
- Stephen Farrell: ICNing DTN

- Armando Caro: Content Based Networking in DTNs
- Christian Esteve Rothenberg: Compact Forwarding in Content-Oriented Networks
- Henrik Lundqvist: Deployment of Information Centric Networking from a Mobile Operator Perspective: Service Program Mobility
- Antonio Carzaniga: Content-Based Publish/Subscribe Networking and Information-Centric Networking

Aspects of these presentation were then discussed in *dedicated* groups on *resource management and transport*, *ICN APIs and the ICN hour glass waist*, and *deployment aspects, business models, and incentives*.

The seminar was wrapped up by a discussion of common concepts, future research topics and next steps for the ICN community.

Outcome of the seminar

The seminar delivered a comprehensive analysis of the state of the art in information-centric networking, progress on specific technical issues such as scalable addressing and content distribution, a better understanding of the legal requirements and application developer needs. It also touched upon possible next steps in research and helped to form an ICN community. The seminar has led to the organization of a SIGCOMM workshop `http://www.neclab.eu/icn-2011/` on the same topic that is co-organized by seminar organizers and participants.

Chapter 9

Modelling, Simulation, Scheduling

9.1 Scheduling

Seminar No. **10071** Date **14.02.–19.02.2010**
Organizers: Susanne Albers, Sanjoy K. Baruah, Rolf K. Möhring, Kirk Pruhs

Scheduling is a form of decision making that involves allocating scarce resources over time to achieve some objective. The primary objectives of this seminar were to bring together leading researchers working on scheduling problems in three different research communities – operations research, theoretical computer science, and real-time systems – to expose each community to the important problems addressed by the other communities; to enable and encourage cooperation among the researchers; and to facilitate a transfer of solution techniques from each community to the others. This is the second Dagstuhl seminar organized to further these objectives (the first – Dagstuhl Seminar 08071 – was held two years ago, in February 2008).

There were approximately sixty participants at the seminar, roughly evenly split between the three communities. Several of these participants had also attended the previous seminar. There was one common session each morning and one each afternoon of the seminar. During the first morning, there were presentations describing some of the research outcomes of the previous scheduling seminar. These presentations highlighted the success of the previous seminar in fostering collaborations between the communities. These talks also provided succinct snapshots of the remaining open problems in the domains addressed in these projects. The remaining sessions mostly consisted of one tutorial/survey talk presenting a line of research or a solution technique of a particular community in a manner that is accessible to researchers from the other communities, and many open problem talks, in which multiple short (5-10 minute) presentations that invited collaboration with the speaker on one of his/her favorite open problems. Write-ups of these open problems were collected and published.

Several clusters of seminar participants formed around common research interests. Ample time was built into the schedule to enable these clusters to meet multiple times to get to know each other better, to work on problems together, and to develop plans for continuing some of these collaborations after the seminar. We expect that, as happened in the

last seminar, that several successful collaborations will have been formed that result in publications in prestigious conferences and journals.

In essence, this seminar continued the process initiated in Seminar 08071, of getting the real-time systems community on the one hand, and the operations research and theoretical CS communities on the other, better acquainted with each others' formal models, interesting problems, and solution techniques. We consider these objectives to have largely been met.

9.2 Grand Challenges for Discrete Event Logistics Systems

Seminar No. **10102** Date **07.03.–12.03.2010**
Organizers: Peter Lendermann, Leon F. McGinnis, Lars Mönch, Arnd Schirrmann

Discrete Event Logistics Systems (DELS) are networks of resources through which material flows. Each node of the network corresponds to some resource (or set of resources) by which the materials are either converted in some way (refined, shaped, assembled, disassembled, etc.), moved (transported within one facility or between facilities), or simply held for some period of time (as work-in-process or stored in a warehouse). Material handling and transportation are key components of DELS. DELS are "discrete" in part because they move material in discrete quantities, and in part because their behavior can be characterized effectively in terms of events happening at discrete points of time, i.e., the start or end of some conversion, transport, or storage process. A DEL system may take the form of a single warehouse, a portion of a factory, a complete factory, or a global supply network.

DELS have been the subject of a large body of analytic research. A huge variety of specific models exists that generally require application by model and/or solver experts to answer narrowly-defined logistics questions about inventory, sourcing, scheduling, routing, etc. It has proven difficult to integrate these models in any comprehensive way into information systems like Enterprise Resource Planning (ERP) systems, Advanced Planning and Scheduling (APS) systems, Manufacturing Execution Systems (MES) or Supply Chain Management (SCM)systems, because of the lack of conceptual alignment between the models produced by researchers and the information systems deployed in practice with which they should be integrated.

This difficulty is magnified enormously by four factors: (1) the scale and scope of global supply networks, such as those developed to support airplane, automobile or telecommunications systems manufacturing, and service systems, which may involve literally thousands of individual enterprises; (2) the dynamics of these networks, which are constantly changing as firms enter and leave, products change, markets change, etc.; (3) the broad range of information and communication systems deployed; and (4) the very high density of real-time decision making. Today, there is little base of theory or methodology for addressing decision problems of this scope, scale, and complexity.

It seems clear that methods from computer science, industrial engineering, information systems, and operations research must be used together to address critical issues in architecting, configuring, planning, managing, and controlling DELS. In the past, researchers in industrial engineering and operations research have quite actively investigated DELS problems, but there has been less direct engagement from computer science and information systems. However, there is an ongoing trend also in computer science towards more business-related application domains, and as a consequence, the main slogan of the 'Gesellschaft für Informatik (GI)', the German Chapter of the ACM, 2007 annual meeting was 'Computer Science meets Logistics'.

Recognizing this, the authors of this report sought to bring together a group of researchers from Europe, North America, and Asia spanning the spectrum of industrial engineering, operations research, and computer science, to consider the following question: What are the grand challenges (for their combined research communities) in supporting decision making in the DELS environment. These efforts were generously supported by the Leibniz Center for Informatics, resulting in a Dagstuhl seminar held March 7-12, 2010, and attended by 28 participants representing universities, research centers, and companies in Europe, North America, and Asia.

Chapter 10

Cryptography, Security

10.1 Distributed Usage Control

Seminar No. **10141** Date **06.04.–09.04.2010**
Organizers: Sandro Etalle, Alexander Pretschner, Ravi Sandhu, Marianne Winslett

In general, access control defines who may access which data, and under which circumstances. A good access control system is at the base of every process which handles confidential information. As an extension to access control, usage control is about defining and enforcing how data may or may not be handled *after* it has been accessed (e.g., "do not disseminate," "delete after thirty days," "notify me when accessed," "use only for scientific purposes".) Usage control is particularly relevant when it comes to privacy, protection of trade secrets or intellectual property, digital rights management, and auditing/compliance in the context of regulatory frameworks. Usage control is hence both relevant for society and economics.

While there is a pressing need for usage control, existing solutions are partial — e.g., via access control mechanisms — and often specialized. The problem is particularly challenging in distributed environments where servers, which give away data, can neither see nor control what clients do with the data after their reception. In this setting, enforcement can be accomplished in one of two ways: by ensuring that policies are not violated, or by detecting and reporting violations, online or off-line. These two approaches apply in different technological environments, and they apply to different underlying trust and business models.

With about 50 attendants, the Dagstuhl seminar on Distributed Usage Control has had an overwhelming response to the invitations that were sent out. One noteworthy characteristics of the seminar was its multidisciplinary nature. Security is not only technical; it is a multidisciplinary field that has legal, regulatory and societal aspects too. This makes security research particularly challenging. This Dagstuhl seminar had a technical core, but sparked discussions also from neighboring fields, in particular a plethora of issues related to privacy. This gave rise to three days of lively discussion, with a regular interleaving of general agreements and disagreements.

In sum, the seminar enjoyed a somewhat unexpected focus on privacy-related issues and intense discussions on the general subject of security research and its connection or disconnection with real-world problems. To the surprise of some, there continues to be disagreement on whether 100% security is a desirable goal, even though it is unlikely to be reached, or if pragmatic considerations including cost, feasibility, usability, innovation and fun should rather lead to a risk-based approach that aims at imperfect security, and if the community shouldn't strive to understand what the risks are, and what imperfect security really is.

10.2 Insider Threats: Strategies for Prevention, Mitigation, and Response

Seminar No. **10341** Date **22.08.–26.08.2010**
Organizers: Matt Bishop, Lizzie Coles-Kemp, Dieter Gollmann, Jeff Hunker, Christian W. Probst

The Dagstuhl seminar "Insider Threats: Strategies for Prevention, Mitigation and Response" was held on August 22–26, 2010 (Seminar 10341), to advance our understanding of ways of reducing insider threats. The insider threat is cited in many studies as the most serious security problem facing organizations. Insider threats are particularly difficult to deal with because insiders have legitimately empowered knowledge of the organization and its systems, and therefore malicious and benign actions by insiders are hard to distinguish.

The 2010 seminar built on the results of its predecessor from 2008 (Countering Insider Threats, 08302). In this seminar we developed a shared,inter-disciplinary definition of the insider 1 and a good formulation for a taxonomy or framework that characterizes insider threats. The seminar also began to explore how organizational considerations might better be incorporated into addressing insider threats.

The purpose of the 2010 seminar was to make progress towards an integrated framework for selecting among and evaluating the impact of alternative security policies against insider threats. An integrated framework, we recognized, needs to include issues not considered in insider work before, such as the economics of insider threats, and the role of law as both a preventative and punitive instrument. We saw the need for creating and testing alternative integrated frameworks so that practitioners and researchers could make informed choices as to combinations of actions targeted at insider threats, and also the need for methods to evaluate the effectiveness of these actions.

The Dagstuhl seminar on strategies for prevention, mitigation, and response with respect to insider threats explored all these areas through discussions and presentations based on input from different and divert communities.

The goal of the seminar was to develop a taxonomy for identifying insider threats and an integrated approach that allows a qualitative reasoning about the threat and the possibilities of attacks. We expected this result to allow us to develop a deeper understanding of security policies and how to evaluate them.

During the seminar, all these issues were inspected and scrutinized, resulting in a better appreciation of social and organizational factors relevant to insider threats, and addressing important questions in related areas.

We would like to thank all participants of the seminar for making it a fruitful and inspiring event — and especially Dagstuhl's wonderful staff, for their endless efforts, both before and during the seminar, to make the stay in Dagstuhl as successful as it has been.

Chapter 11

Data Bases, Information Retrieval

11.1 Automation in Digital Preservation

Seminar No. **10291** Date **18.07.–23.07.2010**
Organizers: Jean-Pierre Chanod, Milena Dobreva, Andreas Rauber, Seamus Ross

Digital Preservation has evolved into a specialized, interdisciplinary research discipline of its own, seeing significant increases in terms of research capacity, results, but also challenges. However, with this specialization and subsequent formation of a dedicated subgroup of researchers active in this field, limitations of the challenges addressed can be observed. Digital preservation research may seem to react to problems arising, fixing problems that exist now, rather than proactively researching new solutions that may be applicable only after a few years of maturing.

Recognising the benefits of bringing together researchers and practitioners with various professional backgrounds related to digital preservation, a seminar was organized in Schloss Dagstuhl, at the Leibniz Center for Informatics (18-23 July 2010), with the aim of addressing the current digital preservation challenges, with a specific focus on the automation aspects in this field. The main goal of the seminar was to outline some research challenges in digital preservation, providing a number of "research questions" that could be immediately tackled, e.g. in Doctoral Thesis. The seminar intended also to highlight the need for the digital preservation community to reach out to IT research and other research communities outside the immediate digital preservation domain, in order to jointly develop solutions.

Despite its rapidly increasing role, digital preservation has not yet reached a level of maturity similar to the constituent research domains. It is a multidisciplinary area involving researchers and practitioners from several fields ranging from Information Retrieval to Library and Archival Science, Content Management, Modelling, Simulation, Human-Computer Interaction, Scholarly Communication and Natural Language Processing. A consistent theory of preservation is not yet in place, and as a result the automation efforts in digital preservation most often address specific tasks, and are far from being contextualized.

This preliminary exercise has produced a number of proposals and ideas ranging from short to medium term research questions to long term and highly speculative research challenges. These ideas need to be organized in a coherent road map, with priorities and a judgement of relevance. In many cases these ideas have highlighted the need for the digital preservation community to reach to other research communities for cross-fertilization and joint development of solutions.

In conclusion, we can say that while the field of digital preservation is still young, it has matured to a level of considerable complexity and specialization. In order to solve the challenges ahead of us, however, the preservation community needs to ensure it remains open and manages to attract professionals from different backgrounds, including but definitely not limited to, computer science experts, to jointly address the challenges that our information society is facing.

11.2 Robust Query Processing

Seminar No. **10381** Date **19.09.–24.09.2010**
Organizers: Goetz Graefe, Harumi Anne Kuno, Arnd Christian König, Volker Markl, Kai-Uwe Sattler

Motivation and Goals

In the context of data management, robustness is usually associated with recovery from failure, redundancy, disaster preparedness, etc. Robust query processing, on the other hand, is about robustness of performance and scalability. It is more than progress reporting or predictability. A system that predictably fails or obviously performs poorly is somewhat more useful than an unpredictable one, but it is not robust. This is comparable to an automobile that only starts in dry weather: it is predictable but not nearly as useful or robust as a car that starts in any weather.

Robust query processing performance has been a known problem for a long time. It also seems common to most or all database management systems and most or all installations. All experienced database administrators know of sudden disruptions of data center processing due to database queries performing poorly, including queries that had performed flawlessly or at least acceptably for days or weeks.

Some techniques are meant to alleviate problems of poor performance, e.g., automatic index tuning or statistics gathered and refreshed on-demand. However, they sometime exacerbate the problem. For example, insertion of a few new rows into a large table might trigger an automatic update of statistics, which uses a different sample than the prior one, which leads to slightly different histograms, which results in slightly different cardinality or cost estimates, which leads to an entirely different query execution plan, which might actually perform much worse than the prior one due to estimation errors. Such occasional "automatic disasters" are difficult to spot and usually require lengthy and expensive root cause analysis, often at an inconvenient time.

A frequent cause of unpredictable performance is that compile-time query optimization is liable to suffer from inaccuracy in cardinality estimation or in cost calculations. Such errors are common in queries with dozens of tables or views, typically generated by software for business intelligence or for mapping objects to relational databases. Estimation errors do not necessarily lead to poor query execution plans, but they do so often and at unpredictable times.

Other sources for surprising query performance are widely fluctuating workloads, conflicts in concurrency control, changes in physical database design, rigid resource management such as a fixed-size in-memory workspace for sorting, and, of course, automatic tuning of physical database design or of server parameters such as memory allocation for specific purposes such as sorting or index creation.

Numerous approaches and partial solutions have been proposed over the decades, from automatic index tuning, automatic database statistics, self-correcting cardinality estimation in query optimization, dynamic resource management, adaptive workload management, and many more. Many of them are indeed practical and promising, but there is no way of comparing the value of competing techniques (and they all compete at least for implementation engineers!) until a useful metric for query processing robustness has been defined. Thus, defining robustness as well as a metric for it is a crucial step towards making progress.

Such a metric can serve multiple purposes. The most mundane purpose might be regression testing, i.e., to ensure that progress, once achieved in a code base, is not lost in subsequent maintenance or improvement of seemingly unrelated code or functionality. The most public purpose might be to compare competing software packages in terms of their robustness in query processing performance and scalability as a complement to existing benchmarks that measure raw performance and scalability without regard to robustness.

Outcome and next steps

The seminar was well attended: 34 researchers (with 16 researchers from industry) from Europe, India, and North America actively explored metrics and tests in the context of physical database design, query optimization, query execution, and workload management. Participants investigated many approaches to measuring and testing robustness without being able to unify them into a single metric. It became clear, however, that continuous parameters such as sizes of tables and intermediate results are much more tractable than discrete parameters such as presence or absence of specific indexes.

At this time, we are pursuing multiple steps based on this seminar. First, several groups of participants are researching and authoring papers on robust query processing, its causes and appropriate metrics. Second, we have been invited to edit a special issue of a journal. Third, we have been invited to organize a panel on robust query processing in an international database conference. Fourth, we have applied for a follow-on seminar in Dagstuhl that will focus on continuous parameters (such as table size), on turning discrete parameters (such as existence of a specific index) into a continuous one, and on scalability and problems in high parallel query processing including cloud servers.

Working Agenda

Each day had a theme related to seminar goals, and several breakout sessions were organized where the seminar was split in smaller groups. The topics of these sessions were structured along the main areas of robustness in query processing: test suite design, query optimization, query execution, physical database design, and system context.

The goal of the first day's breakout sessions was test suite design, i.e. which aspects of robust query processing should be tested and which metrics are appropriate. The second day's session were devoted to robustness issues in query optimization. Measuring the contribution of query execution and access methods to robust query processing was the goal of day 3. On the fourth day, two topics were addressed in parallel. In the first track of sessions, robustness of physical database design tools such as index selection tools and design advisors was discussed. The second track of sessions addressed a broader system context by discussing the interaction of various system components in the relationship to robustness.

Chapter 12

Bioinformatics

12.1 Structure Discovery in Biology: Motifs, Networks & Phylogenies

Seminar No. **10231** Date **06.06.–11.06.2010**
Organizers: Alberto Apostolico, Andreas Dress, Laxmi Parida

This seminar was intended to focus on combinatorial and algorithmic techniques of structure discovery from biological data that is at the core of understanding a coherent body of data, small or large. In biological systems, similarly to the tenet of modern architecture, "form and function are solidly intertwined". Thus, to gain complete understanding in various contexts, the curation and study of forms turns out to be a mandatory first phase.

Biology is in the era of the Ome's: Genome, Proteome, Transcriptome, Metabolome, Interactome, ORFeome, Recombinome and so on. Thus it is no surprise that biological data is accumulating at a much faster rate than it can be understood. While on one hand, the sheer size of data can be daunting, this provides on the other hand a golden opportunity for testing (bioinformatic) structure-discovery primitives and methods. In spite of the difficulties of structure discovery, there are reasons to believe that evolution based on reproduction, variation, and selection endowed biological systems with some underlying principles of organization (involving redundancy, similarity, and so on) that appear to be present across the board. Correspondingly, it should be possible to identify a number of primitive characteristics of the various embodiments of form and structure (using for instance notions of "maximality", "irreducibility", etc.) and build similarly unified discovery tools around them. At a core level in these efforts, there is the need for novel techniques enabling the automated discovery of structures, whether syntactic, such as patterns or motifs, or semantic, such as phylogenies. It is therefore a worthwhile effort to try and identify these primitives.

The workshop gathered a roster of highly qualified senior participants and several talented young researchers who convened to discuss issues of modeling, formalization, and algorithmic design as they emerge in the discovery of biological structure. The schedule unraveled in form of a busy sequence of sessions, intermixed with small group-work

sessions and guest lectures. The opening consisted of an extended round-robin aimed at letting participants give detailed accounts of their background and interests. Following opening remarks by Andreas Dress and completed by concluding comments by Alberto Apostolico, three broad-spectred lectures were delivered by Benny Chor, Laxmi Parida, and Raffaele Giancarlo in the afternoon of the first day. A special lecture on epigenetics offered on Tuesday morning in the concurrent Workshop on "The Semantic of Information" provided the opportunity to nicely complement the presentations already scheduled. The morning session resumed with three lectures centered on networks, offered in turn by Axel Mosig, Ina Koch, and David Gilbert. The afternoon session was devoted to phylogenies, with featured presentations by Stefan Gruenewald, David Bryant, Peter F. Stadler, and Andeas Spillner. After dinner, a guest lecture by Michael Clausen exposed, not without entertaining notes, the analogies between motifs in biosequences and music and methods for searching for them. The morning session of Wednesday revolved on regulation, with presentations by Rahul Siddharthan, Esko Ukkonen, Matteo Comin, and Ion Mandoiu. Thursday was entirely dedicated to motif discovery in various contexts. Presentations were given in the morning by Jens Stoye, Peter Erdos, Asif Javed, and Nadia Pisanti, followed in the afternoon by Cinzia Pizzi, Fabio Cunial, Matthias Gallé and Benjarath Pupacdi. After dinner, a thorough introduction to toponomics by Andreas Dress was followed by a lecture and software demonstration by Peter Serocka. The participants reconvened in the lecture hall on Friday morning for a general discussion, an assessment of the experience and recommendation for (enthusiastically endorsed) possible encores in the future.

Chapter 13

Applications, Interdisciplinary Work

13.1 Pervasive Public Displays

Seminar No. **10011** Date **03.01.–08.01.2010**
Organizers: Nigel Davies, Antonio Krüger, Marc Langheinrich, Albrecht Schmidt, Martin Strohbach

This Dagstuhl seminar has focused on bringing together researchers from a diverse set of fields of Computer Science to discuss the next generation of pervasive public display environments. The state-of-the-art in software control of display environments is best represented by commercial products that enable advance scheduling of content on a network of displays. Essentially such systems offer a traditional broadcast model based on linear playout of content and offer no support for user recognition or interaction with displays. Other display environments, e.g., in offices or conference centres, might simply run single, isolated applications, such as video conferencing or video players. In addition, current systems typically function as small isolated networks consisting of a limited number of displays under a single management domain (e.g. in a single shopping centre). In this respect, a parallel can be drawn with the state of computing prior to the invention of the Internet - machines were networked together in small clusters to facilitate resource sharing, control and communication but there were no mechanisms for interconnecting these networks.

The focus of the seminar was on future systems designed to be fundamentally different to the current display systems that are being deployed in public, semi-public, and private spaces. Specifically, instead of displays showing linear presentations that have been created and scheduled in advance it is possible to imagine that future displays will be highly interactive and show information personalised to viewers and their context. Such personalization could be achieved by, for example, detecting the proximity of individuals to the display through detection of the presence of a personnel device such as a mobile phone with Bluetooth enabled. Secondly, it is envisaged that these advanced display systems will transition from small, isolated islands to a single global network of interactive displays and associated sensors. Such a platform would enable the removal of the traditional static, tight coupling between display owners and content producers and create a new dynamic

system in which display ownership and content production are decoupled. This would, in turn, create the potential for new business models to emerge within the context of a new global communications medium.

The seminar has considered key areas related to next generation display networks including:

- Protocols and Services to Support Next Generation Display Networks
- User Interaction with pervasive displays
- Tensions between privacy and personalisation
- Business and legislative requirements for pervasive display networks

This Dagstuhl seminar presented an unique and exciting opportunity to bring together representatives from these fields with the explicit aim of advancing the state of the art in public display systems. The main goal was to help kickstart international collaborative research and deployments and to begin to define the shape of future display networks. There will be plenty of opportunities for participants to present and discuss their own work as well as participating in discussions. We hope that the seminar has the potential to be noted as the birth-place of new ideas in an area with enormous potential impact in academia, industry and everyday life.

13.2 Computational Foundations of Social Choice

Seminar No. **10101** Date **07.03.–12.03.2010**
Organizers: Felix Brandt, Vincent Conitzer, Lane A. Hemaspaandra, Jean-Francois Laslier, William S. Zwicker

Classical social choice theory deals with the design and analysis of methods for collective decision making and plays an important role within economic theory as witnessed by the Nobel prizes awarded to social choice theorists such as Kenneth Arrow and Amartya Sen. Examples of collective decision making mechanisms are voting rules and methods for the fair division of one or more goods among several agents. For a few years now, researchers from computer science, in particular AI and theory, have been taking an ever increasing interest in social choice. There are two main reasons for that, leading to two different lines of research.

The first of these is concerned with importing concepts and methods from theoretical computer science and AI for solving questions originally stemming from social choice. The point of departure for this line of research is the fact that most of the work in social choice theory has concentrated on establishing abstract results regarding the existence of procedures meeting certain requirements. However, computational issues have rarely been considered. For instance, while there may not exist a voting protocol that makes it impossible for a voter to manipulate it in one way or another, it may well be the case

that any such manipulation can be rendered computationally infeasible, and may thus be deemed an acceptable risk.

The second line of research within computational social choice goes in the other direction, i.e., concepts and procedures from social choice theory are imported to solve questions that arise in computer science and AI application domains. This is, for instance, the case for managing societies of autonomous software agents, which calls for negotiation and voting procedures. Other examples are the application of techniques from social choice to the development of webpage ranking systems for Internet search engines or recommender systems for electronic commerce.

The seminar brought together 44 researchers who have worked on various aspects of computational social choice, and who have come to this area from very different backgrounds: theoretical computer science, artificial intelligence, economics, operations research, mathematics, and political science. Only half of the participants were computer scientists. Despite—or maybe because of—this heterogeneity, every talk was followed by a long and lively discussion. There was a total of 34 talks plus a rump session consisting of about ten short (5-minute) presentations. Participants originated from 19 different countries with the majority being from France, Germany, and the USA.

A wide variety of topics were discussed during the seminar. Common research themes that emerged included manipulability, approval voting, cake-cutting algorithms, tournaments, and abstention. We are currently preparing a special issue of the journal *Mathematical Social Sciences* (edited by Felix Brandt and Bill Zwicker) consisting of work presented at this seminar.

13.3 Computational Transportation Science

Seminar No. **10121** Date **21.03.–26.03.2010**
Organizers: Glenn Geers, Monika Sester, Stephan Winter, Ouri E. Wolfson

Computational Transportation Science has made its first steps of consolidation. A PhD program on the subject, funded by the National Science Foundation, was established at the University of Illinois at Chicago in 2006. Two international workshops on CTS were held (2008 in conjunction with the 5th Annual International Conference on Mobile and Ubiquitous Systems, and 2009 in conjunction with the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems). A third one will be held with the 18th ACM SIGSPATIAL International Conference in 2010. With the first workshop appeared a preliminary publication exploring a research agenda in this area (Geers 2008). Then a Dagstuhl Seminar on Computational Transportation Science was held in 21-26 March 2010 to characterize the discipline and identify its research agenda. The seminar was attended by 25 invited researchers from USA, Australia, Germany, Belgium, and Switzerland, with nationalities also from China, India, Greece and former Yugoslavia. This report presents the highlights of this Dagstuhl Seminar. Major steps at the seminar have been:

- Collaborative definition of CTS, vision of CTS, and core research agenda for CTS
- Set up of a Wikipedia entry for the definition and vision
- Set up of a webpage as a bulletin board for the growing community
- Plans for the third international workshop on CTS later in 2010
- Engagement with funding bodies promoting CTS as a discipline (outreach)
- Establishing collaboration by developing some larger joint research project proposals
- Publishing the (first) core research agenda via this report

A discipline is only as good as its academic community. If this paper finds your support or meets your interests you are cordially invited to participate and engage. The infrastructure set up so far is a beginning but requires your collaboration, be it the Wikipedia entry, the CTS webpage, or the CTS workshop series. These are all small seeds that — if they grow — can lead to conferences and journals on CTS, not only in the content but also in name.

Finally, the community should shape its own academic programs or introduce core subjects on computational transportation science into the programs on transport engineering, electrical engineering, software engineering, and geographic information engineering. The spread demonstrates the inter-disciplinarity of computational transportation science, illustrates that engineering problems do not present themselves any longer wholly contained in one traditional discipline, and supports the fundamental concern that engineering disciplines have grown to be too narrow (National Academy of Engineering 2004)

13.4 Enabling Holistic Approaches to Business Process Lifecycle Management

Seminar No. **10151** Date **11.04.–16.04.2010**
Organizers: Serge Abiteboul, Andreas Oberweis, Jianwen Su

Process modeling tools and techniques are used in different phases of a business process management project. At the level of modeling, the basic formalisms proposed for modeling business processes, tend to focus very heavily on the side of process, without enabling equally rich modeling of the data and information that is being manipulated by the processes. More broadly, the research is largely independent of the overall lifecycle of business process management, including evolution of business processes over time, and design of business processes from already existing business processes.

The purpose of this seminar was to bring together a cross-disciplinary group of academic and industry researchers from the areas of Information Systems/Business Process Management, Service and Software Engineering, Semantic Technologies and Human Computer Interaction to foster a better understanding of how to manage the lifecycle of business process models, ranging from initial design, evolution, implementation, and monitoring. The

primary emphasis was on paradigms and technologies that enable a more holistic perspective, including new modeling techniques, new conceptualizations and visualizations, applications of recommendation system and other social networking techniques, and more flexible implementations. While there have been a number of Dagstuhl Seminars focused on modeling issues pertaining to distinct phases of a modeling project, such as Process, Service or Software Engineering, there has not been a seminar to gather specialists from across several disciplines to study this lifecycle management challenge.

We believe that each of these technologies is separately important for advancing the state of the art in managing the lifecycle of business processes, and that a dramatic improvement can be obtained if the four areas work in a coordinated way.

13.5 Event Processing

Seminar No. **10201** Date **16.05.–21.05.2010**
Organizers: K. Mani Chandy, Opher Etzion, Rainer von Ammon

Event processing (EP) is an area in the field of information technology that is central to many systems on which our society depends. These systems include energy, healthcare, the environment, transportation, finance, services, and manufacturing.

Event processing consists of methods and tools to filter, transform, and detect patterns in events, in order to react to changing conditions, typically under some time constraints.

We present this Manifesto to introduce the area of event processing, explain its pertinence to other fields, and to provide information to enable relevant business opportunities. We also aim to establish guidelines for how event processing can fit into current standards and to put forth short- and long-term goals for event processing professionals in industry and academia.

Event processing systems perform the following four main functions:

- Obtain data from multiple sources in real or near-real time
- Aggregate and analyze this data to detect patterns that indicate the presence of critical situations requiring a response
- Determine the best response for such situations
- Monitor the execution of that response
- Why is event processing of increased importance now, when even the earliest rule engines and business processes had mechanisms used to detect critical situations and respond accordingly?

Today's world is much more dependent on IT systems than it ever was. All of us are much more interconnected and interdependent than ever. Systems must be able to react to events

anywhere on the globe. An outbreak of Ebola on one continent, for example, demands a response in countries everywhere. Responses must occur ever quicker, sometimes in milliseconds, as the pace of the stock exchange illustrates.

The costs of inappropriate responses can be staggering, as we see in the cases of certain defense applications. In many telecommunication systems, the volume of data that must be analyzed in near-real time is torrential. In addition, the variety and types of data that must be analyzed in event processing systems is enormous. Such data may be in the form of structured text, natural language, images, audio, or video. The data may be delivered to the system, or it may have to be extracted by the system. In many systems, security is an overarching concern.

Coming decades will see many more applications with event processing capabilities, as society demands smarter ways for managing electric power, water, health, retail and distribution, traffic, and safety — smarter meaning responding better and faster to changing conditions. The interconnected nature of the modern world means that researchers, designers, and students can no longer develop event processing for a single domain, such as the smart grid, without incorporating developments related to event processing technologies in other domains, such as smart healthcare. To step up to these challenges, we are in urgent need of event processing theory, design methods, and tools.

13.6 Dynamic Maps

Seminar No. **10371** Date **12.09.–17.09.2010**
Organizers: Claus Brenner, Wolfram Burgard, Marc Pollefeys, Christoph Stiller

In recent years, the advent of car navigation systems has laid the ground for an entirely new industry sector, consisting of map producers, car/ personal/ smart phone navigation manufacturers, and service providers. It has probably gone unnoticed that navigation systems mark a major change in the way we use maps. Partially, they are still just a replacement for traditional maps, providing a means to store and visualize a representation of the environment. In contrast to the traditional use of maps, however, navigation systems perform computations using the map's data structures, such as shortest route, map matching, and route guidance. That is, from an abstract point of view, part of the map is made for machine use only — the user has no direct access to it but rather is only presented the outcome of the computations.

This development will accelerate in the near future, as sensor, processing, and communication capabilities of cars, or in general, robots, increase. Maps will evolve from their traditional meaning towards virtual representations of the environment, containing information specifically tailored for the employed sensors, algorithms, and questions to be solved. For example, a map may contain geo-referenced feature descriptors which are recognized by a given algorithm in the live video stream of a car's camera, allowing to solve the specific problem of highly accurate positioning.

Such maps will not only be very different from today's maps, they also will have to be produced by entirely different mechanisms. Since they contain very many details, it is

infeasible to use (partially) manual processes to produce them, as is done today. Also, the mapping of details usually implies a high change rate, calling for frequent reacquisition, which is not economically feasible using traditional surveying and mapping techniques. What is required is an approach which is both, inexpensive and accurate, detailed and up to date, using a hierarchy of measurement systems which include the map users themselves.

The purpose of the 'dynamic maps' seminar was to bring together researchers from academia and industry, from several fields of computer science, including computer vision, photogrammetry, robotics, computer graphics, geoinformatics, and driver assistance. Central objectives of the seminar were to alert the different communities to the efforts in the respective other communities, to exchange our ideas about such a future, 'dynamic' map, and to foster future research proposals in this area.

Conclusions

The organizers thank all participants for their lively presentations and the fruitful discussions. We think that the seminar was a great success and served its main purpose, namely, making the different communities aware of each other, to a great extent.

It is clear that 'dynamic maps' is, and will continue to be, an active research topic for many years to come. Several attendants of the seminar have applied for research grants, as part of a larger German Research Foundation bundle project, in December 2010. It is worth to note that Wolfram Burgard has been granted an ERC advanced grant in 2010 for the topic of 'LifeNav – Reliable Lifelong Navigation for Mobile Robots', confirming the importance of the topic.

It is also clear that this area is so wide and complex, involving sensors, vision algorithms, in-car systems, acquisition strategies, data modeling, storage, and serving, that it will require a combined effort of all players to move research forward, and finally develop systems that will ultimately change the lives of all of us. That said, the organizers are looking forward to further seminars on the topic of dynamic maps.

13.7 Demarcating User eXperience

Seminar No. **10373** Date **15.09.–18.09.2010**
Organizers: Jettie Hoonhout, Effie Lai-Chong Law, Virpi Roto, Arnold Vermeeren

Thirty user experience (UX) researchers and practitioners spent three days in Dagstuhl in order to bring clarity to the concept of user experience. The participants represented different perspectives to user experience from holistic to modeling approach, from real-time psychophysiological research to investigating user experience after a long period of time, and from standardization and research to consultancy work.

By 'demarcating' user experience, the organizers wanted to make the relation clearer to the neighboring concepts of usability, interaction design, consumer experience, etc. The term created a lively discussion on whether this field needs demarcation: many researchers

do not want their research field to be limited, while some industry people need a sound judgment on what user experience work includes. Despite the different needs, the participants seemed to agree on the need for bringing clarity to the vague concept of user experience. The participants also identified the need for further work on clarifying the different theoretical perspectives behind the different interpretations of user experience, and their impact on user experience work both in industry and academia.

Instead of dismissing the whole term of user experience, we believe UX can be defined to cover a certain set of cases and to make its scope clear. In this seminar, we will demarcate the field of user experience.

The main result of the seminar is a white paper, which aims to clarify some core concepts of user experience. As can be seen from the abstracts in this collection, it has been challenging to come up with a white paper that would serve all needs and do justice to all the different perspectives. This work was on conceptual level, so the paper does not provide direct practical guidance for UX work. Nevertheless, thanks to the wide variety of perspectives to user experience represented, the seminar was an eye-opening experience for the participants.

The User Experience White Paper is based on the discussions between the participants in Dagstuhl and is available at `http://www.allaboutux.org/uxwhitepaper`