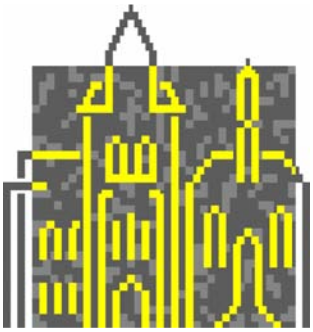

SCHLOSS DAGSTUHL

INTERNATIONAL
CONFERENCE
AND RESEARCH CENTER
FOR COMPUTER SCIENCE

Dagstuhl News

January - December 2005

Volume 8
2007

ISSN 1438-7581

Period: January - December 2005

Frequency: One per year

The International Conference and Research Center for Computer Science 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.

Online version: http://www.dagstuhl.de/files/Reports/News/

Associates (2005): Gesellschaft für Informatik e.V. (GI), Bonn
Technical University of Darmstadt
Johann Wolfgang Goethe University of Frankfurt
Technical University of Kaiserslautern
University of Karlsruhe (TH)
University of Stuttgart
University of Trier
Saarland University
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.

Funded by (2005): The state governments of Saarland and Rhineland Palatinate

Information: Schloss Dagstuhl Office
Saarland University
P.O. Box 15 11 50
66041 Saarbrücken, Germany
Phone: +49-681-302-4396
Fax: +49-681-302-4397
E-mail: service@dagstuhl.de
http://www.dagstuhl.de/

Welcome

Not so long ago, you received the Dagstuhl News for 2004. Here are the ones for 2005, the eigth edition of the "Dagstuhl News", a publication for the members of the Foundation "Informatikzentrum Schloss Dagstuhl", the *Dagstuhl Foundation* for short. Not much has happened in between the productions of both volumes.

The main part of this volume consists of collected resumees from the Dagstuhl Seminar Reports 2005. 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 2005 are on the Web under URL: `http://www.dagstuhl.de/Seminars/05/`

We have switched to publishing online proceedings of our Dagstuhl Seminars on our Dagstuhl Research Online Publication Server (DROPS).
`http://www.dagstuhl.de/publikationen/publikationsserver-drops/`
Authors keep the copyrights to their contributions in order not to harm their rights to submit them to conferences or journals. We hope that the reputation of our Dagstuhl Seminars will make their proceedings a valuable source of information. It encourages us that also external workshops have asked to be hosted on DROPS.

The State and the Activities of the *Dagstuhl Foundation*

The foundation currently has 45 personal members and 7 institutional members.
In 2005, we have supported a number of guests with travel grants and a reduction of the Seminar fees. The supported guests did not have any budget for traveling expenses. Most supported guests were young researchers aged 20-35 years. An increasing number has been older than that. The current funding situation seems to be rather bad.

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, May 2007

Contents

Chapter 1

Data Structures, Algorithms, Complexity

1.1 Algorithms for Optimization with Incomplete Information

Seminar No. **05031** Date **16.01.–21.01.2005**
Organizers: S. Albers, R.H. Möhring, G.Ch. Pflug, R. Schultz

The purpose of this Seminar was to bring together top specialists working in algorithms for optimisation when the decision maker has only partial information. While problem descriptions in the different approaches to optimisation with incomplete information are quite similar, solution concepts and methods of solution may be quite different. Traditionally, the stochastic programming community has focussed on problems, where all uncertainty is due to the fact that concrete realizations are unknown, but the probability distributions from which they stem are fully known. The quality of the solution is typically measured in average case sense. In contrast, the online optimisation community assumes no particular probability model. Therefore the focus is traditionally on worst-case analysis. Recently, new developments made the gap between the two communities smaller. Robust optimisation replaces the assumption of a known probability distribution by an assumption about the range of possible values. Stochastic scheduling incorporates ideas of the competitiveness of algorithms with stochastic models for the demands. The typical assumption in stochastic programming that decisions do not influence the underlying probability distribution can usually no longer be maintained in stochastic scheduling.

To facilitate familiarizing of the different communities with each other's ways of thinking, basic concepts, and basic research questions the seminar was started by four one-hour overview talks. These were delivered by Jiri Sgall (Online optimisation), Andrzej Ruszczynski (Stochastic Programming), Garud Iyengar (Robust Optimisation), and Marc Uetz (Stochastic Scheduling).

The regular program consisted of 38 thirty minutes talks, which could be classified into the following subgroups: Robust and minimax optimisation (Sim, Dupacova); Two- and

multistage stochastic optimisation (Hochreiter, Dye, Sen, Stougie, Tomasgard); Assessing quality of solution (Morton, Rambau); Approximation (Higle, Swamy, van der Vlerk); Algorithmic approaches using game theory and nonlinear programming (Lorenz, Steinbach, Kleywegt, Bastin, Norkin); Applications in Communications and Robotics (Erlebach, Fekete, Epstein, Richter); Dynamic stochastic optimisation (Weiss, Philpott, Nino-Mora); Average case competitive analysis (Fujiwara, Vredefeld); Competitiveness Analysis (van Stee, Schäfer, Ebenlendr, Zhang, Skutella); Risk issues (Dentcheva, Eichhorn); Stochastic online scheduling (Megow, Schulz, Krumke); Probabilistic criteria (Henrion, Hoogeveen).

To assess the results of this seminar, on Thursday afternoon an open discussion was held about different views and perceptions on optimisation with incomplete information. The results of this discussion can be summarized as follows:

What the communities have in common is:

- The desire for optimality.
- The desire for more efficient algorithms, i.e. better/faster results.
- The fact that solutions, which require clairvoyance are not implementable.
- The necessity of comparing the non-clairvoyant solution to the ideal clairvoyant solution by either taking differences (value of perfect information) or ratios (competitive ratio).
- The distinction between individual solutions and solution rules (policies).
- The necessity of approximation.
- The interest in complexity issues.

What distinguishes the communities is:

- The way uncertainty is modelled (from sets of possible values via probability distributions to families of probability distributions).
- The frequency of decision making (once in a while versus online).
- The objective (to look for worst cases, average cases, include risks, chance constraints etc.).
- The class of problems (general as multistage LP, QP, MIP or specialised as scheduling, packing, sequencing).
- The way information is revealed (fixed observation times versus uncertainty about when and if ever information will be available).
- The view on risk.

Some participants brought up their individual views on the topic. It was felt that the advantage of probabilistic modelling lies in the sound concept of probability, developed over centuries, and the clear way of how to obtain and process information (samples). On the other hand, the assumption that a probability model is governing the data process is not always fulfilled, or information is so poor that range sets is all we have. Also, in long term models it is unrealistic to assume that probability distributions do not change over time. Adaptive algorithms in the broad sense are a way to circumvent this difficulty.

This inspired a discussion about bridges and possible collaboration between the communities. As already existing bridges were cited: Minimax and robust approaches, stochastic competitiveness analysis, certain stochastic dynamic models, complexity studies in stochastic optimisation. The need for more real world data and problems was expressed as well as the unanimous wish to study special problem classes which were presented at this seminar in more detail.

1.2 Foundations of Semistructured Data

Seminar No. **05061** Date **06.02.–11.02.2005**
Organizers: F. Neven, T. Schwentick, D. Suciu

As in the first seminar on this topic, the aim of the seminar was to bring together people from the areas related to semi-structured data. However, besides the presentation of recent work, this time the main goal was to identify the main lines of a common framework for future foundational work on semi-structured data.

The workshop was of a very interdisciplinary nature with invitees from databases, structured documents, programming languages, information retrieval and formal language theory. Several of the lectures were presented by PhD students. We had four invited speakers and a panel on research evaluation. Due to strong connections between topics treated at this workshop, many of the participants initiated new cooperations and research projects.

1.3 Design and Analysis of Randomized and Approximation Algorithms

Seminar No. **05201** Date **15.05.–20.05.2005**
Organizers: M. Dyer, M. Jerrum, M. Karpinski

Most computational tasks today that arise in realistic scenarios are intractable, at least if one insists on exact solutions delivered within a strict deadline. Two important means for surmounting that intractability barier are randomized and approximate computations. It is an interesting artifact that these two notions of computation go hand-in-hand.

The Seminar was concerned with the newest developments in the design and analysis of randomized and approximation algorithms. The main focus of the workshop was on three specific topics: approximation algorithms for optimization problems, approximation algorithms for measurement problems, and decentralized networks as well as various interactions between them. Here, some new broadly applicable techniques have emerged recently for designing efficient approximation algorithms for various optimization and measurement problems. This workshop has addressed the above topics and also some new fundamental insights into the design techniques.

The 35 lectures delivered at this workshop covered a wide body of research in the above areas. The meeting was held in a very informal and stimulating atmosphere. Thanks to everyone who made it a very interesting and enjoyable event.

1.4 Sublinear Algorithms

Seminar No. **05291** Date **17.07.–22.07.2005**
Organizers: A. Czumaj, S.M. Muthukrishnan, R. Rubinfeld, C. Sohler

The purpose of the Dagstuhl seminar "Sublinear Algorithms" was to bring together researchers working on the development of algorithms for very large data sets. Over the last few years data sets have become increasingly massive and the need to design special algorithms and data structures that deal with such amounts of data has emerged. For example, the set of all credit card transactions in the world for a month would have been considered a massive data set some time ago. That is comparable to the number of packet transactions a single router processes in **one** hour on an interface and we are now facing problems of analyzing the traffic at a large network of such routers, each with many interfaces! Internet traffic logs, clickstreams, web data are all examples of modern data sets that show unprecedented scale. Managing and analyzing such data sets forces us to revisit the traditional notions of efficient algorithms. The long-held golden standard of "linear algorithms" – algorithms that take time proportional to the input and store no more space than it takes to archive the input – is no longer as efficient as one needs or can afford. Thus, there is now a need for **sublinear** algorithms, that is algorithms that use resources (time and space) significantly less than the input size.

The main areas addressed in the workshop were **property testing, sublinear time approximation algorithms, and data streaming algorithms.** These areas are not only connected by the fact that they require algorithms with sublinear resources but also that they heavily rely on randomization and random sampling. Therefore, we hoped that this workshop helped to exchange ideas between these different areas.

During the seminar one could obtain a good overview of the current state of sublinear algorithms. In many interesting talks new algorithms and models as well as solutions to well-known open problems were presented.

Concluding remarks
The seminar was attended by 52 researchers from eight countries (19 USA, 13 Israel, 10

Germany, 4 Canada, 2 France, 2 United Kingdom, 1 Switzerland, 1 Hungary). From our own experience and the feedback from the participants we believe that the workshop was very successful. Interesting talks, fruitful discussions between researchers working on different areas of sublinear algorithms, and the wonderful working and living environment of Schloss Dagstuhl contributed to the success of the workshop.

1.5 Exact Algorithms and Fixed-Parameter Tractability

Seminar No. **05301** Date **24.07.–29.07.2005**
Organizers: R. Downey, M. Grohe, G. Woeginger

It seems that by now almost everybody in our community has accepted that $P \neq NP$ holds true, although we do not have the slightest idea how to prove it. Regardless, $P \neq NP$ means that there are no polynomial time algorithms for NP-hard problems, and that **super-polynomial** time algorithms are the best we can hope for when dealing with exact algorithms for NP-hard problems.

Recently, there have been some fascinating activities on exact algorithms. For example, there has been a long sequence of papers on exact algorithms for 3-satisfiability. The current best algorithm for this problem is due to Iwama and Tamaki (2003) and needs roughly 1.324^n time for instances with n variables. Fomin and Kratsch (September 2003) have just developed a 1.993^n exact algorithm for determining the tree-width of an n-vertex graph - this is clearly not the last result on this problem. Fedin and Kulikov (2002) gave an exact $2^{m/4}$ algorithm for the Max-Cut problem on graphs with m edges. There also has been a number of ever improving algorithms for the Vertex Cover problem; the current best is from Chen, Kanj, and Jia (2001) and it finds a Vertex Cover of size k in a graph of size n in time $1.285^k + kn$.

However, many very interesting questions are still wide open and have received little if any attention. We mention just two examples here:

- Held and Karp (1962) designed a dynamic programming algorithm for the n-city traveling salesman problem with a running time of roughly 2^n . This running time has not been improved in more than 40 years.
- Nesetril and Poljak (1985) observed that a simple algorithm based on matrix multiplication finds a clique of size k in an n-vertex graph in time $n^{0.79k}$. No improvement has been made since then.

More qualitatively, we ask if there are algorithms solving the satisfiability problem or the traveling salesman problem in time $2^{o(n)}$ or the clique problem in time $n^{o(k)}$. Very little is known about such lower bound questions.

The idea of fixed-parameter tractability is to approach hard algorithmic problems by isolating problem parameters that can be expected to be small in certain applications and

then develop algorithms that are polynomial except for an arbitrary dependence on the parameter. More precisely, a problem is fixed-parameter tractable if its running time is $f(k)p(n)$, where f is an arbitrary function and p a polynomial. Since the choice of suitable parameters allows for a great flexibility, fixed-parameter algorithms have found their way into practical applications such diverse as computational biology, database systems, computational linguistics, and automated verification. The algorithmic methods developed in this area are not far from those used in the exact algorithms mentioned above. As a matter of fact, the fast algorithms for Vertex Cover have been developed in the context of fixed-parameter tractability. But beyond these algorithmic results, parameterized complexity also offers a well developed theory of intractability, and this theory may provide us with the right tools to systematically approach a theory of lower bounds for exact algorithms. For example, parameterized complexity allows us to establish exponential lower bounds on hard problems, modulo complexity assumptions. Indeed, recent work has substantiated the fact that this approach has deep links with the existence of feasible PTAS's, and limited nondeterminism; so we see the exciting emerging interplay between a number of hitherto unlinked groups of researcher.

To summarize, the area of Exact Algorithms is still in a rudimentary stage, but it is full of fascinating and difficult open problems. Fixed-parameter tractability is a branch of algorithms and complexity theory that seems very well suited to approach some of these questions. Connections between the two areas have recently evolved, ranging from very practical questions in algorithms design to the fundamental complexity theoretic problems.

The Seminar

The seminar brought together leading researchers from exact algorithms and parameterized complexity theory. If not before, it became very clear during the seminar that the areas are quickly growing together. Topics of the seminar ranged from very practical aspects of algorithm engineering for applications in computational biology to theoretical and mathematical topics such structural parameterized complexity and matroid theory. The technical program consisted of 7 invited one hour talks and 25 half hour talks. There will be a special issue of the journal **Theory of Computing Systems** , edited by Rod Downey, devoted to work arizing from this seminar.

The main topics of the seminar can be grouped as follows:

- Better Exact and Parameterized Algorithms
- Applications and Algorithm Engineering
- Decompositions of Graphs and Matroids and their Algorithmic Applications
- New Developments in Parameterized Complexity

1.6 Algebraic and Numerical Algorithms and Computer-assisted Proofs

Seminar No. **05391** Date **25.09.–30.09.2005**
Organizers: B. Buchberger, S. Oishi, M. Plum

Recently, a number of problems have been solved by so-called computer-assisted proofs, among them the celebrated Kepler conjecture, the proof of existence of chaos, the verification of the existence of the Lorenz attractor, and more. All those problems have two things in common: first, the computer is used to assist the proof by solving certain subproblems, and second these subproblems are of numerical nature.

There are also many famous results of nontrivial proofs which can be performed automatically by a computer program. For example, Risch's algorithm for integration in finite terms, solution of polynomial systems by Gröbner bases, quantifier elimination and more. Any of those solves a nontrivial mathematical problem which could be quite hard to solve by pencil and paper.

Those algorithms are frequently executed in exact arithmetic over the field of rationals or an algebraic extension field using well known methods from computer algebra. The problems mentioned in the first paragraph are continuous in nature. They can be solved by so-called verification or self-validating methods.

Basically, self-validating methods verify the validity of assertions of certain theorems which are formulated in such a way that validation is possible by means of numerical computations. The main point is that this validation is absolutely rigorous including all possible procedural or rounding errors. On the one hand, the use of finite precision arithmetic implies very fast calculations, but on the other hand it limits the scope of applicability.

In contrast, most algorithms in computer algebra are 'never failing', that is they are proved to provide a solution for any input, and the maximum computing time for this is estimated a priori. To speed up practical implementations, also hybrid methods combining computer algebra with numerical verification methods are used.

Self-validating and computer algebra methods aim on the reliable solution of certain mathematical problems with the aid of computers. This seems a very natural task. Other areas such as computer geometry and graphics, real number theory, automated theorem proving and more have similar aims. Therefore we think it is very fruitful to create a link of information between experts in those different fields. The common basis or goal is the computer-assisted solution of mathematical problems with certainty.

The choice of organizers also reflects different fields of interest and expertise. The stimulating atmosphere in Dagstuhl was definitely a very fruitful environment for this enterprise.

1.7 Perspektiv-Workshop zur theoretischen Informatik

Seminar No. **05461** Date **14.11.–16.11.2005**
Organizers: W. Thomas, I. Wegener

Ziel des Workshops ist eine Standortbestimmung der theoretischen Informatik in Deutschland mit der Herausarbeitung erfolgversprechender Richtungen (und eventuellen konkreten Schritten). Hierzu dient zunächst die Darstellung der aktuellen Forschungsthemen und ihrer tragenden Ideen sowie die Benennung von Herausforderungen und Entwicklungsmöglichkeiten

- rein wissenschaftlich,
- hinsichtlich Zusammenarbeit mit anderen Theoretikern/Anwendern,
- in nationalen oder internationalen Verbünden,
- jenseits der Fachdiskussion in allgemeinverständlichen Publikationen.

Weitere angesprochene Themen sind die Erwartungen der Informatik und ihrer Nachbarfächer an die Theoretiker, mögliche Weichenstellungen in der Lehre (etwa im Zuge der Einführung von BSc/MSc) sowie die Strukturen in der Wissenschaftsorganisation (insbesondere Theorie-Fachgruppen der Gesellschaft für Informatik).

Chapter 2

Verification, Logic

2.1 Mathematics, Algorithms, Proofs

Seminar No. **05021** Date **09.01.–14.01.2005**
Organizers: T. Coquand, H. Lombardi, M.-F. Roy

This seminar is the continuation of the seminar "Verification and constructive algebra" held in Dagstuhl from 6 to 10 january 2003.

The goal of the seminar is to bring together people from the communities of formal proofs, constructive mathematics and computer algebra (in a wide meaning).

One objective of the seminar is to bridge the gap between conceptual (abstract) and computational (constructive) mathematics, by providing a computational understanding of abstract mathematics.

It is becoming clear that many parts of abstract mathematics can be made constructive and even computational and that abstract mathematics techniques contain an underlying constructive content.

We are not only interested in algorithms however, but also in formal proofs of the correctness of these algorithms.

Computer algebra provides a variety of interesting basic algorithms, from exact linear algebra to various aspects of elimination and real root counting, which are the foundations for much more sophisticated results like nullstellensatz, quantifier elimination etc... It is remarkable that in constructive and computer algebra, progress in sophisticated algorithms often implies progress on basics.

Moreover the scope of computer algebra is now widened by the consideration of seminumerical algorithms. When such algorithms are correctly controlled, they actually deal with real and complex numbers in the constructive meaning of these objects. So computer algebra is lead to fill many objectives of computational analysis.

Providing formal proofs of correctness to the computer algebra community is very useful, specially for algorithms which are basic and used everywhere.

On the other hand, a collection of mathematically non trivial examples is very useful for the formal proof community, which needs also powerful automatic methods from computer algebra.

We observe that the Dagstuhl seminar 03021, which seems to have been the first meeting devoted to the topic, was a success and has been very satisfactory for the participants. They decided to create a group under the acronym "Mathematics, Algorithms, Proofs". We have organised a similar meeting in Luminy, january 2004, and besides the next Dagstuhl meeting, we are trying to organize a MAP summer school in September 2005 in Santander (Spain).

Since the whole field is rather big, we think that it may be indeed a good idea to choose a focal topic. A good such topic for the next 2005 topic appears to be.

General Presentation

This seminar was the third MAP meeting, a continuation of the seminar "Verification and constructive algebra" held in Dagstuhl from 6 to 10 January 2003. The goal of these meetings is to bring together people from the communities of formal proofs, constructive mathematics and computer algebra (in a wide meaning). The special emphasis of the present meeting was on the constructive mathematics and efficient proofs in computer algebra. We were honored to have as invited speakers Helmut Schwichtenberg, proof theorist who is now working on extraction of programs from proofs, Harold Edwards, specialist of the work of Kronecker, and Fred Richman, the specialist of constructive algebra. A sub-theme was on formalization of mathematics, especially on the flyspec project, lead by Thomas C. Hales, and we were fortunate that most people working on this, in particular, besides Hales himself, Jeremy Avigad and Robert M. Solovay were present at the seminar.

Once again, we would like to thank the team of Schloss Dagstuhl. The exceptional working condition we enjoyed there played an important part in the success of this meeting.

2.2 Nonmonotonic Reasoning, Answer Set Programming and Constraints

Seminar No. **05171** Date **24.04.–29.04.2005**
Organizers: G. Brewka, I. Niemelä, T. Schaub, M. Truszczynski

The seminar took place from April 24 to 29, 2005. It was organized by Gerhard Brewka (Univ. Leipzig, DE), Ilkka Niemelä (Helsinki Univ. of Technology, FI), Torsten Schaub (Univ. Potsdam, DE), and Miroslaw Truszczynski (Univ. of Kentucky, US). The seminar was attended by 58 participants from Europe, North America, Asia and Australia.

The technical program consisted of

- Three invited talks:
  Tomi Janhunen: Translating NLPs into Propositional Theories

Thomas Eiter: Extending Answer Set Programming for the Semantic Web
David Mitchell: Progress and Problems in SAT Solving

- 38 contributed talks given by the participants
- Panel discussion on the future of the answer-set programming. The panel was moderated by Ilkka Niemelä. Marc Denecker, Yannis Dimopoulos, Michael Gelfond and Nicola Leone were panelists.
- Special session on the benchmarking system asparagus led by Christian Anger and Miroslaw Truszczynski

The technical program of the seminar demonstrated that since our first meeting in Dagstuhl in September 2002, substantial scientific progress has been achieved in several areas:

- **Theory of answer-set programming**. Much progress has been obtained in understanding encodings of programs as propositional theories through research on the concepts of completion and loop formula. Logic programming with nested expressions and logic of here-and-there and its relatives solidified their position as fundamental formalisms for the development of the theory of answer-set programming. Other notable developments include ID-logic, which expands classical logic with **inductive definitions** represented as a logic program, new results on program equivalence and proof systems for programs with cardinality constraints.
- **Software for answer-set programming**. **Dlv** emerged as an "almost" production-grade package of answer set programming tools supporting program grounding, answer-set computation and integration with database environments. Several new solvers were introduced: **cmodels** enhanced to handle disjunctive programs, **pbmodels**, which uses pseudo-boolean solvers to compute stable models of programs with weight atoms, and **nomore++** – a system implementing new branching and propagation techniques. Asparagus, an environment for systematic and objective testing of answer-set solvers has grown and matured significantly since it was first proposed at the Dagstuhl Seminar 02381 in September 2002. Researchers also have been investigating and developing tools for distributed processing of answer-set programs (Platypus project). Finally, the workshop presented research on program development tools supporting static program analysis and debugging.
- **Applications**. The seminar demonstrated that answer-set programming becomes a viable software tool in several application domains including: semantic web, data integration, systems of boolean equations, planning, security engineering, social modeling, and qualitative decision theory.
- **Strong connections to propositional satisfiability**. It has been clear for quite some time that our field can benefit from closer collaboration with researchers in the SAT community. This seminar had several talks that emphasized that connection, most notably the invited talk by David Mitchell, which will undoubtedly have major impact on the development and implementation of new answer-set programming solvers.

General conclusions from the seminar are very positive. The seminar was dominated by young researchers and students, about 25 of whom delivered presentations. Our community is branching out to related communities of propositional satisfiability and constraint satisfaction, both theory and software development are actively pursued, and there is a strong push towards practical applications.

2.3 Types for Tools: Applications of Type Theoretic Techniques

Seminar No. **05251** Date **19.06.–24.06.2005**
Organizers: F. Henglein, M. Odersky, F. Tip, J. Vitek

Type systems have proven to be the most cost-effective technique for ensuring software systems safety to have gained widespread acceptance. Over the last thirty years, a variety of type systems have been deployed and adopted in programming languages such as Haskell, Java and C#. Types and type systems were originally intended to characterize program properties amenable to mechanical checking for the purpose of preventing certain kinds of run-time errors. In recent years, type-theoretic techniques have been successfully applied to address software engineering challenges and used in software engineering tools that automate tasks related to the maintenance, improvement, translation, restructuring, and upgrading of existing software. These applications are becoming increasingly subtle and application domain specific. Examples of such novel applications of type-theoretic principles include:

- Type inference and type-directed transformation have been used in the translation of large COBOL application for Y2K compliance.
- Linear types have been used to enforce protocols present in the interface between the Windows 2000 kernel and its device drivers.
- Parametric type polymorphism has been retrofitted onto main stream programming languages in a provably sound manner.
- Type constraints have been used as the basis for automating code refactorings related to type generalization.
- Type systems are being used to define the well-formedness of XML documents, with safety guarantees for the input and output of such documents.
- Ownership types which capture the dynamic topology of object graphs have been proposed as a mechanism for enforcing locality in software upgrades of object-oriented databases.
- Behavioral types as automatically checked stateful interface specifications for distributed components such as web services.

- Type for region-based memory management, support for both automatic inference of regions and, for profiling-driven engineering by programmers.

These examples all share the important characteristic that type theoretical ideas have been applied to solve software engineering problems and that the application has been embodied in a practical tool. Although there is a growing number of such applications, the types community, the software engineering community, and the tools development community are still largely distinct, with their own set of conferences and meetings.

This workshop brings together the technical leaders of these communities. The participants, between 30 and 40 researchers, will be selected to cover the fields of type theory, static program analysis, and applied software engineering. The workshhop will consist of a combination of longer tutorial-style talks (1.5 hours) and a collection of short 20 minutes research presentations. Significant time is planned for informal and group discussion of selected topics.

2.4 Deduction and Applications

Seminar No. **05431** Date **23.10.–28.10.2005**
Organizers: F. Baader, P. Baumgartner, R. Nieuwenhuis, A. Voronkov

Formal logic provides a mathematical foundation for many areas of computer science, including problem specification, program development, transformation and verification, hardware design and verification, relational databases, knowledge engineering, theorem proving, computer algebra, logic programming, and artificial intelligence.

Using computers for solving problems in these areas, therefore, requires the design and implementation of algorithms based on logical deduction. It remains one of the great challenges in informatics to make computers perform non-trivial logical reasoning, be it fully automatic, or in interaction with humans. Some progress, however, has been made in the past ten years:

- Automated theorem provers and finite model building programs solved various open mathematical problems of combinatorial nature.
- Model checking, a form of theorem proving over finite models, has become a very successful push-button method for verifying nontrivial safety properties of hardware and software.
- Automated deduction, in particular for so-called description logics, is widely assessed as a core enabling technology for the **Semantic Web**.
- Methods of interactive theorem proving have helped in formally verifying semantic (type) safety aspects of programming languages such as Java. The “Schwerpunktprogramm Deduktion” funded by the Deutsche Forschungsgemeinschaft together with previous Dagstuhl seminars on “Deduction” have been instrumental in obtaining these successes.

The conviction that mathematical logic is a unifying principle in computer science and that methods from different theoretical areas as well as application domains should be brought together as a means to fight fragmentation has lead to successful new conferences like FLoC and IJCAR, and to IFCoLoG, the recently established International Federation for Computational Logic.

This interdisciplinary view of logic in computer science motivated the Dagstuhl seminar. Specifically, we considered several application areas: Software verification, Hardware verification, Cryptographic protocols, Programming languages, Formal methods, Semantic Web, Large knowledge bases.

Chapter 3

Geometry, Image Processing, Graphics

3.1 Computational Geometry

Seminar No. **05111** Date **13.03.–18.03.2005**
Organizers: H. Alt, F. Aurenhammer, D. Halperin

Computational geometry has developed as a subarea of algorithmics, concerned with algorithms and data structures for geometric problems. These problems are motivated by application areas, such as robotics, computer graphics, CAGD, pattern and shape matching and recognition, computer vision, image processing, integrated circuit design, structural bioinformatics, and more. Since the mid 1980s, computational geometry has arisen as an independent field with its own international conferences and journals.

In the early years mostly theoretical foundations of geometric algorithms were laid. Meanwhile, in addition, the area has become application oriented. These latter developments include two related activities of researchers in computational geometry:

1. direct involvement in application domains, and
2. robust implementation of geometric algorithms.

In fact, several software libraries for geometric computation have been developed. In particular, in the CGAL library and in parts of the LEDA library, geometric algorithms have been implemented. Remarkably, this software emerged from the originally theoretically oriented computational geometry community itself, so that many researchers are concerned now with theoretical foundations as well as implementation. Implementation issues like robustness of computation and software design have become an integral part of the research presented at computational geometry conferences and workshops.

The seminar, therefore, should be concerned with fundamentals as well as practical issues of computational geometry.

Dagstuhl seminars on computational geometry have been organized since 1990 in a two year rhythm.

3.2 Graph Drawing

Seminar No. **05191** Date **08.05.-13.05.2005**
Organizers: M. Jünger, S. Kobourov, P. Mutzel

Graph drawing deals with the problem of communicating the structure of relational data through diagrams, or drawings. The field builds on early research in flowchart design, CASE tools, visual programming interfaces, VLSI design, database systems, and software engineering. Graphs with vertices and edges are typically used to model relational data. The vertices represent the objects (or data points) and the edges represent the relationships between the objects. The main problem in relational visualization is to display the data in a meaningful fashion.

The ability to represent relational information in a geometric form is a powerful tool which allows us to perform analysis through visual exploration. With the aid of graph visualization we can find important patterns, trends, and correlations. Graph drawing tools are needed in a growing number of scientific disciplines, including bioinformatics, physics, and sociology. Within the computing disciplines, graph drawing techniques are essential in areas such as networking, internet traffic control, and bioinformatics. For example, Internet Service Providers and web caching providers must be able to quickly identify patterns and trends in internet traffic. Visualization of network topology is used to identify and analyze characteristic patterns leading to better functionality.

Topics of the Seminar

One of the main current challenges in graph drawing research is to deal effectively with very large graphs, and graphs that evolve through time.

Recent technological advances have brought about increased data volumes and increased data complexity. However, visualization of large and complex graphs is difficult, given the constraints imposed by the current technology (limited number of pixels on a screen) and the complexity of the graphs to be displayed (millions of nodes and edges). In many fields, such as telecommunication, databases, and software engineering, the models contain millions of objects and relationships. Models without natural geometric placement often require hours to compute even an initial layout. Reasons are the complex nature of the algorithms and the fact that existing algorithms do not scale well.

In order to deal with huge graphs, we need to develop alternative models and algorithms. Possible techniques include the identification and collapsing of subgraphs, focus and context, and interactive browsing techniques. In addition, the development of new clustering techniques including the identification and representation of clusters (groups of vertices belonging together) plays an important role.

In many applications the models are dynamic, evolving over time, e.g., telephone graphs in which telephones are the objects and calls are the relationships. Even the fastest graphics systems using state of the art drawing and rendering algorithms fail to provide interactive visualization for such complex models.

Since graph drawing is mainly application driven, we focused our research during the seminar on the visualization of large and dynamic graphs in the following application domains: bioinformatics (visualizing biochemical networks such as protein interaction networks, regulatory and signaling pathways), software engineering (e.g., UML class diagrams, memory graphs), internet and telecommunications visualization, and social network analysis.

Good attendance and excellent presentations contributed to the general success of the seminar. The aims and achievements of the seminar are summarized below.

Aims

In the application, we formulated the following aims for the Dagstuhl seminar:

1. To address the long-standing open problems related to the visualization and interaction with large and dynamic networks;
2. To bring together theoreticians and practitioners from the targeted graph drawing application areas: bioinformatics, software engineering, internet and telecommunication visualization, and social network analysis;

The main research topics represent significant long-standing open problems in the area of graph drawing for which no satisfactory solutions are yet known. In order to make progress and obtain new results, it is necessary that theoreticians work together with practitioners and applied researchers as well as with researchers from closely related areas, such as information visualization and software visualization. Moreover, close cooperation with users in the considered application domains is essential for the successful development of effective tools for the visualization and interaction with large and dynamic networks.

Therefore, we invited researchers from the areas of pure graph theory, graph algorithms, and information visualization, as well as from the targeted application areas: computational biology, software engineering, internet and telecommunication visualization, and social network analysis. Our objective was to provide the opportunity to get these groups together in order to work on the emergent problems.

Achievements

Over forty participants from both academia and industry attended the seminar. Over one third of the attendees were graduate and postdoctoral students. There were representatives from more than ten countries, including Germany, Austria, United Kingdom, Ireland, Italy, Slovenia, Turkey, Canada, Australia, and USA. The achievements of the seminar can be summarized as follows:

1. We were lucky to enjoy a number of stimulating presentations on the core topics of the seminar. In particular, the presentations covered new approaches to the layout of large graphs, new ideas about evolving and dynamic graphs, as well as new visualization paradigms.

2. We enjoyed two survey lectures on the graph drawing aspects of bioinformatics and another two lectures on the visualization of social networks. In addition, we learned about novel applications of graph drawing techniques to problems in nano-technology and rank aggregation.

3. The graph drawing e-print archive, gdea was unveiled during the seminar. It provides a powerful depository and search interface for graph-drawing related publications. Staff members for gdea were also appointed during the seminar.

4. Solutions to some open problems were found during the seminar and close interactions led to new collaborations.

Beyond the survey lectures, highlights of the seminar included a lecture on the psychology of visual perception through empirical studies, and lectures on applications of graph drawing in non-traditional areas such as nano-technology. An open problem session led to successful problem solving. The report of the open problem session is included with the seminar materials. New cooperations, e.g., between Irish and German researchers, led to a joint article that is currently in preparation.

In summary, it is our impression that the participants enjoyed the great scientific atmosphere offered by Schloss Dagstuhl, and profited from the scientific program. Several attendees commented on the week of the seminar being one of the most enjoyable research experience they have had. We are grateful for having had the opportunity to organize this seminar.

3.3 Geometric Modeling

Seminar No. **05221** Date **29.05.–03.06.2005**
Organizers: G. Brunnett, G. Farin, R. Goldman, S. Hahmann

Geometric Modeling is the branch of Computer Science concerned with the efficient acquisition, representation, manipulation, reconstruction and analysis of 3-dimensional geometry on a computer. Models and shapes in 3-dimensions can be represented as splines or subdivision surfaces, as well as by polygonal meshes or point clouds. Applications of geometric modeling cover a wide collection of areas from classical computer aided design, reverse engineering and simulation, to computer graphics, scientific visualization, medical imaging, multimedia and entertainment.

The 6th Dagstuhl seminar on geometric modelling was attended by 59 participants. The participants came from 4 continents and 19 countries (!), and included 4 industrial scientists as well as the leading academic experts in the field. Several young invited researchers were funded by the HLSC program of the European community. A very special event during the conference was the award ceremony for the John Gregory Memorial award. This time Prof. Rida Farouki, Prof. Ron Goldman and Prof. Richard Riesenfeld have been awarded with this price for their fundamental contributions to the field of geometric

modelling. After the conference, as with all previous Dagstuhl Seminars on Geometric Modelling, conference proceedings will be published in collaboration with Springer.

There were a total 53 technical presentations at the conference related to the following diverse topics:

- curve and surface modelling
- surface reconstruction
- surface interpolation and fitting
- multiresolution representations, subdivision surfaces
- algebraic methods for curves and surfaces
- 3D meshes
- computational topology
- geometric models for Biomedical application

Despite the very large number of presentations during the conference and the high attendance at these talks, there was time for scientific discussions and research.

3.4 Scientific Visualization: Challenges for the Future

Seminar No. **05231** Date **05.06.–10.06.2005**
Organizers: T. Ertl, E. Gröller, K.I. Joy, G.M. Nielson

Scientific visualization (SV) is concerned with the use of computer-generated images to aid the understanding, analysis and manipulation of data. Since its beginning in the early 90's, the techniques of SV have aided scientists, engineers, medical practioneers, and others in the study of a wide variety of data sets including, for example, high performance computing simulations, measured data from scanners (CAT, MR, confocal microscopy), internet traffic, and financial records. Somewhat as a result of these past successes, matters are changing for research in SV. The data sets are becoming massive in size, complex and multi-dimensional in nature and the goals and objectives of the visualization much less precisely defined, but yet the results are needed with higher urgency and importance. The multiresolution and hierarchical methods of today do not scale to these new data sets. The segmentation and knowledge extraction methods of today need to be completely revamped in order to be useful. Because of the changes that are taking place in SV, it is important that a group of senior researchers meet with select junior researchers to map out the future research agenda for this critical area.

Specific Themes of the Seminar:

Ubiquitous Visualization. As ubiquitous computing is getting increased attention, also visual display of everywhere available data is necessary. Challenges include: heterogeneous output devices, novel interaction metaphors, network bandwidth (availability, reliability), graceful degradation of algorithms with respect to largely varying resources, invivo visualization (real time, no pre-processing, robust).

Categorical Visualization. Information and knowledge is extremely difficult to extract from multi-valued, multi-dimensional, multi-modal and multi-layered categorical data. These data sets abound today and the pay-offs for understanding them are substantial. Mathematical techniques based upon functional relationships break down requiring completely new paradigms to visualize these types of data sets.

Intelligent/Automatic Visualization. Ever-increasing data sizes require semi-automatic methods that concentrate on the typically very small portion of the relevant information in the data. Techniques include model- and knowledge-based segmentation, classification in abstract feature spaces, computation of saliency information from derived data characteristics, automatic detection of important isosurfaces, automatic creation of expressive transfer functions, automatic landmark selection and automatic path and navigation guidance.

Point-based/Mesh-free Visualization. A typical strategy to visualize unorganized multidimensional data sets is to transform the data into standard geometric primitives of triangles and triangular mesh surfaces prior to rendering. This intermediate step is time consuming, but necessary to map the data set to standard (hardware and software) graphics primitives. With the recent advances in point-based rendering, new efficient and creative approaches for visualizing scattered and unorganized data sets are potentially possible.

Chapter 4

Artificial Intelligence, Computer Linguistic

4.1 Probabilistic, Logical and Relational Learning – Towards a Synthesis

Seminar No. **05051** Date **30.01.–04.02.2005**
Organizers: L. De Raedt, T. Dietterich, L. Getoor, S.H. Muggleton

One of the central open questions of artificial intelligence is concerned with combining expressive knowledge representation formalisms such as relational and first-order logic with principled probabilistic and statistical approaches to inference and learning. This combination is needed in order to face the challenge of real-world learning and data mining problems in which the data are complex and heterogeneous and we are interested in finding useful predictive and/or descriptive patterns.

In this context, the terms probabilistic and statistical refer to the use of probabilistic representations and reasoning mechanisms grounded in probability theory, such as Bayesian networks, hidden Markov models and probabilistic grammars and the use of statistical learning and inference techniques. Such representations have been successfully used across a wide range of applications and have resulted in a number of robust models for reasoning about uncertainty. The primary advantage of using probabilistic representations is that well-understood and principled statistical inference and learning algorithms exist.

The term learning refers to deriving the different aspects of the probabilistic model on the basis of data. Typically, one distinguishes various learning algorithms on the basis of the given data (fully or partially observable variables) or on the aspect being learned (the parameters of the probabilistic representation or the structure of the model). Statistical and Bayesian approaches provide a unified framework for learning a model, whether through model selection or explicitly modeling a distribution over the models.

The terms logical and relational refer to first order logical and relational representations such as those studied within the field of computational logic and database theory. The primary advantage of using such expressive representations is that it allows one to elegantly

and naturally represent complex situations involving a variety of objects as well as relations among the objects, which is not possible using the simpler propositional or feature vector based representations. So, probabilistic, logical and relational learning aims at combining its three underlying constituents: statistical learning and probabilistic reasoning within logical or relational representations.

Seminar Goals
The goal of this seminar was to bring together the researchers interested in the area of statistical, logical and relational learning. This allowed the participants to explore the foundations, challenges and research opportunities raised by this important open problem in artificial intelligence.

Conclusion
This workshop brought together a signficant number of researchers from all over the world that are working on all aspects of probabilistic, logical and relational learning. It was also the first workshop on this topic where there was sufficent time for indepth discussions, debates and working groups. It was exciting to see the progression through the week. It was clear that some common ground had been identified, yet this was just the start. There was a general feeling that the workshop was a success, and a lot of enthusiasm for a follow on workshop.

4.2 Machine Learning for the Semantic Web

Seminar No. **05071** Date **13.02.–18.02.2005**
Organizers: F. Ciravenga, A. Doan, C. Knoblock, N. Kushmerick, S. Staab

The Semantic Web has attracted great attention since the vision was first articulated several years ago. In a nutshell, the Semantic Web will augment conventional Web content with explicit machine-processable semantic metadata, enabling a variety of automated content manipulation and aggregation.

As demonstrated by the first two International Semantic Web Conferences, the initial "futuristic vision" has matured into a carefully crafted set of substantive technical proposals, such as the Resource Description Framework (RDF) and the Web Ontology Language (OWL). However, it is widely recognized the Semantic Web will never "take off" until a critical mass of semantic metadata has been deployed. Many SW researchers have therefore built various tools to help developers attach semantic metadata to their content.

More ambitiously, machine learning and other artificial intelligence techniques are being developed that generate the requisite semantic metadata in a semi-automated or even entirely automated fashion. For example, machine learning algorithms for information extraction allow large legacy text repositories to be rapidly enriched with semantic metadata, and machine learning approaches to ontology learning and matching are being developed for the Semantic Web context.

The goal of this seminar is to assemble the leading researchers who work at the intersection of machine learning and the Semantic Web, in order to review progress and identify the

most significant opportunities and challenges over the next several years. We will also invite leading figures from the "conventional" (hand-crafted metadata) Semantic Web community, to ensure both that our technology is fully appreciated by the Semantic Web community, and that the machine learning community focuses on important and realistic problems.

The seminar will focus specifically on the following five topics:

1. Automated document annotation;
2. Ontology learning and maintenance;
3. Ontology mapping and merging;
4. Service discovery; and
5. Content cleaning and normalization.

4.3 Annotating, Extracting and Reasoning about Time and Events

Seminar No. **05151** Date **10.04.–15.04.2005**
Organizers: G. Katz, J. Pustekovsky, F. Schilder

Newspaper articles and other natural-language texts describe actions, events, and states of affairs. A crucial first step toward the automatic extraction of information from these texts – for use in such applications as automatic question answering or summarization – is the capacity to identify what events are being described and to make explicit when these events occurred and which temporal relations hold among them. There has recently been a renewed interest in making use of this kind of temporal and event-based information, with a wide variety of proposals and applications having been presented at recent conferences and workshops. The central goal of the seminar was to consolidate the insights that have been made in recent years and to identify and address issues concerning annotation, temporal reasoning and event identification that remain unresolved.

Much of the temporal information conveyed in a natural language text is left implicit. Significant recent work has focused on developing schema for making this information explicit, typically via annotation. An important result of contemporary research has been the adoption of a **de facto** standard for time and event annotation: TimeML. This XML-based markup language is specifically designed for annotating texts with tags that make explicit the temporal and event-based information conveyed by the text and has been adopted by a number of researchers in this domain. Much of our seminar was concerned with issues specific to this annotation scheme.

There are three basic types of tags used by the TimeML language: TIMEX tags are used to annotate temporal expressions and provide them with a normalized value (e.g.

(TIMEX tid="t1" val="2005-04-21") **April 21st, 2005** (TIMEX)); EVENT tags are used to annotate event expressions, providing "hooks" to relate them to other events and times introduced in the text (e.g. (EVENT eid="ei") **opened** (/EVENT)); So-called TLINK tags indicate the temporal relations that hold between times and events (e.g. **the stock market opened on April 21st, 2005 at 10:00pm** (TLINK event="e1" relatedTime="t1" relation=INCLUDED-BY)). Other tags are used to capture more subtle semantic relations. SLINK tags, for example, are used to indicate various kinds of subordination relations, such as the negation in **The stock market did not open on April 21st, 2005 at 10:00pm** or the only potential event in **Investors hoped that the stock market would open on April 21st, 2005 at 10:00pm.** A small corpus of TimeML annotated documents (**TimeBank**) has been generated, and can be browsed at timeml.org.

The main focus of the seminar was on TimeML-based temporal annotation and reasoning. We were concerned with three main points: determining how effectively one can use the TimeML language for consistent annotation, determining how useful such annotation is for further processing, and determining what modifications should be applied to the standard to improve its usefulness in applications such as question-answering and information retrieval.

Highlights of the Seminar

One of the highlights of the seminar was an annotation exercise which was carried out by all participants in groups. This served both as a touchstone for discussing issues that came up in the course of the seminar and as a source of examples of difficulties to be addressed. As the "target text" we choose a newspaper article from the **Seattle Times** describing the wedding of Prince Charles and Camilla Parker Bowles, an event that had just occurred.

The entire seminar was split up into groups of four or five researchers and each group carried out the annotation in two parts. In the first part, we attempted to identify, making use of the TimeML guidelines, the events and times which were described by the article and to identify the relations that hold among them. We found there to be very clear agreement about what events there were. Issues of event identity (is the **waving** the same as the **greeting** ?) were the foremost problems. Also the temporal relations were fairly well agreed upon. Here again there was very little in the way of disagreement, with the major problems being those surrounding the differentiation among simultaneity, overlap and immediate precedence. What was striking, however, was that there were far more events described (and for which TimeML guidelines require annotation) than participants judged would be likely to be useful for any application.

In a second part of the annotation exercise the same groups attempted to do metrical annotation, of the type described by Hobbes. Here we tried to specify how long each of the events was and how long the intervals between events were. In contrast, here there was wide variation in some cases (how long does the state of the couple being **newly married** hold?), but in other cases fairly close agreement. The highlight of this exercise came when we compared our consensus interpretation of the text to the BBC video of the

event described. The very low correlation between our estimated durations for events (the waving, the walking to the car) and their actual durations as shown on the video raised questions, less for the value of annotation, but for the veracity of newspaper texts.

4.4 Synthesis and Planning

Seminar No. **05241** Date **12.06.–17.06.2005**
Organizers: H. Kautz, W. Thomas, M.Y. Vardi

This meeting has brought together researchers working in two complementary fields: automatic synthesis of (control) programs, and methods for devising planning algorithms in artifical intelligence (AI). Thus, the seminar combines a strong thread of current research in automata theory with an area of possible but so far unexplored applications.

The idea of organizing such a seminar arose during IJCAI 2003, where Vardi gave an invited talk on the automata-theoretic approach to design verification. In discussions between Kautz and Vardi after the talk it became clear that methods of synthesizing strategies for reactive systems is an issue of common interest to automata theory and artificial intelligence.

Automatic Synthesis

The first results on automatic synthesis of control programs go back to the 1960's when Büchi, McNaughton, Rabin, and others showed how to realize specifications for non-terminating reactive computations by finite automata. These results extend the standard equivalence results connecting automata with logic; they are concerned with a specification of an open system (reacting to moves of its environment) and realizations by automata with output, providing the moves of the program component of a system.

Today these results have been recast in the terminology of infinite two-player games. Such a game is played on a directed graph which is the "arena" of the game. Each vertex is associated to one of the two players. A play starts in a given vertex and proceeds along the graph edges; in each step that player to whom the current vertex belongs moves via an edge to a new vertex. The winner is determined by a "winning condition" on the resulting finite or infinite path. Two fundamental algorithmic problems arise in this context: Given a graph and a winning condition, from which start vertices does the first player have a winning strategy, and - if yes - how can one construct a program which realizes auch a strategy?

Building on the classical work, much progress was achieved during the last decade. While in the early papers it was shown that in principle the automatic synthesis of winning strategies is possible (which opens a perspective for automatic controller synthesis), the focus has now shifted to refined and extended questions:

- problems of complexity and efficiency in the construction of strategies (usually in the form of finite automata)

- applications in model-checking (where the games are used in their connection with logic, reflecting the duality of existential and universal logical connectives)
- applications to the synthesis of reactive programs
- expansion of the techniques to further types of games, e.g. with infinite state spaces, and
- games involving continuous parameters, for example stochastic and timed games

A recent GI-Dagstuhl seminar volume (LNCS 2500, edited by Erich Grädel, Wolfgang Thomas, and Thomas Wilke) with survey contributions by young researchers gives an overview of the state of the art.

In this situation, where a solid body of constructions and nontrivial results is available and the further development is somewhat open, it is essential to expose, try, and adjust the methods in application areas. Planning in AI is one of them, another (not excluded for the seminar) is the connection with researchers on discrete event systems.

AI Planning

Planning is a sub-field of artificial intelligence which is concerned with the generation of a rational course of action given a declarative specification of the environment, the goals, and the possible actions. The field can be further subdivided by the kinds of problems considered:

Classical planning considers single-agent deterministic domains where the initial and goal states are specified by sets of logical formulas. Classical planning corresponds to reachability analysis in large state spaces. Research thus focus on algorithms that can perform such an analysis without actually enumerating the state space. Classical planning is closely connected to the area of verification called model checking, and in recent years there have been fruitful exchanges of techniques and algorithms between the two fields.

Universal planning involves synthesizing a reactive control program that can direct an agent toward a goal state from any possible situation. It is thus universal in that no fixed initial state is assumed. Furthermore, universal planning problems often include non-deterministic actions, which can be used to model action failure and/or changes in the world induced by nature. Universal planning can be viewed as control program synthesis where every computation terminates in a goal (or failure) state.

Decision-theoretic planning adds two features to universal planning: first, non-determinism (or nature) is modeled by a probability distribution over the result of each action; and second, a positive or negative reward is associated with each state. The goal of the agent is to maximize the sum of rewards that the agent receives over its lifetime (or in the case of an infinite lifetime the discounted or average reward) rather than to reach a particular goal state. If the agent is able to observe all variables in the domain the problem becomes that of solving a Markov Decision Process (MDP); if part of the state is hidden from the agent, the problem is that of solving a partially-observed MDP (POMDP). As with

classical planning much research on decision theoretic planning focuses on techniques for handling large state spaces in a factored form, thus avoiding enumeration of all states.

Game-theoretic planning is a recent and fertile area of activity in AI research. While the previous approaches model the actions of other agents or other natural events simply as sources of uncertainty, a game-theoretic planner explicitly reasons about the choices other agents make in order to maximize their own utility. In terms of the synthesis of control programs as games against nature, this line of work allows us to consider cases where nature is actively hostile or actively helpful to purposes of the system.

Finally, any of the preceeding areas can be generalized to consider the case where the full specification of the problem in terms of a world model, an action model, and a reward function (or goal specification) is not known to the system in advance. The planner must act while learning about the environment on the basis of the feedback it receives from (possibly infrequent) rewards. This research, called reinforcement learning, has deep roots in both control theory and models of animal behavior.

Links

This description of the field of AI planning should make clear that it is closely linked to the problem of synthesizing reactive control programs: in fact, one can argue that the two fields have the same subject matter, and are distinct only because of historic conditions. As we have noted, the strongest connection in terms of scientific dialog between different communities has occurred between classical planning in AI and model checking in formal methods. For example, researchers in AI have found uses for BDD (Boolean decision diagram) algorithms from model checking, and techniques for reducing planning to satisfiability testing that were originally developed in the AI planning world are now used for hardware verification.

The workshop helped to increase the awareness of the researchers working in one field of the problems and methods in the other one, and thus to increase the interaction and collaboration of the two research communities, and the transfer of methodologies from one field to another.

4.5 Belief Change in Rational Agents: Perspectives from Artificial Intelligence, Philosophy, and Economics

Seminar No. **05321** Date **07.08.–12.08.2005**
Organizers: J. Delgrande, J. Lang, H. Rott, J.-M. Tallon

Introduction

The area of *belief change* studies how a rational agent may maintain its beliefs when obtaining or perceiving new information about the environment. This new information

could include properties of the actual world, occurrences of events, and, in the case of multiple agents, actions performed by other agents, as well as the beliefs and preferences of other agents. Not surprisingly, this area has been of interest to researchers in different communities.

The initial research in belief change came from the *philosophical* community, wherein belief change was studied generally from a normative point of view (that is, providing axiomatic foundations about how rational agents *should* behave with respect to the information flux). Subsequently, *computer scientists*, especially in the *artificial intelligence* (AI) and the *database* (DB) communities, have been building on these results. Belief change, as studied by computer scientists, not only pays attention to behavioural properties characterising evolving databases or knowledge bases, but must also address computational issues such as how to represent beliefs states in a concise way and how to efficiently compute the revision of a belief state. More recently, the *economics and game theory* community, in particular the emerging field of *cognitive economics*, has become active in belief change research, adopting a normative point of view, like philosophers, but paying more attention to the "cognitive plausibility" or "fitness" of the belief change operators.

The goal of the seminar is to bring together researchers from these areas. This would allow the identification and addressing of problems of common interest in this highly challenging and relevant area, as well as an exploration of ways in which one area may contribute to another.

Goals and Content of the Seminar

The area of belief change can be regarded as originating in the philosophical logic community. This work provided abstract, formal, and precise specifications of desirable properties for belief change operators, as well as the identification of distinct types of change. However, this research says nothing about specific implementable operators nor computational issues – issues of fundamental importance to computer scientists. Researchers in artificial intelligence and computer science have followed up on these latter issues, as well as developed other specific operators (addressing e.g. sensor fusion and belief base merging) and examined their complexity characteristics. In artificial intelligence, the relatively recent emergence of the field of *cognitive robotics*, which is concerned with endowing artificial agents with cognitive functions that involve reasoning, for example, about goals, actions, the states of other agents, collaboration and negotiation, etc., has given impetus to the development of computational operators for belief change and the identification of issues arising from concrete, evolving sets of knowledge. As well, more recently, economists have been using work in belief revision, and applying it to notions of mistaken and changing beliefs among interacting and negotiating agents. Such work is also of obvious interest to researchers in artificial intelligence.

To date, there has been limited interaction among these communities. Clearly however there are deep problems of common interest, and results in one area will contribute to another. We have already mentioned that research in economics has made use of the work from the philosophical community, and that such results will be of use to researchers in AI. As well, contributions may also flow back from economics to research in the foundations of

belief revision: For example, recently it has been suggested that that economic principles (dealing with choice, preferences, and utility) may provide a more appropriate foundation for belief change. Computational issues raised and addressed by researchers in computer science and AI will be of use to economists addressing related problems; as well such work can contribute to the other areas by further elucidating the abstract area of belief change, as well as providing implementations and identifying philosophically-interesting "pragmatic" or "practical" problems.

Thus we see researchers in three broad areas (philosophy and logic, artificial intelligence and computer science, and economics and game theory) addressing highly related (in some cases, the same) problems, in which work in one area in all likelihood will benefit research in another. Hence for the Dagstuhl seminar, we feel that there would be valuable interactions and contributions that would be anticipated by bringing people together in these areas.

Conclusion

We found the workshop successful, especially on the following two achievements: first, the seminar made participants aware of a commonality of interests across different disciplines; second, it suggested new directions for research that will probably be taken up by researchers in the next couple of years.

Where is the field going? We can mention at least two emerging issues:

- the field is broadening both with respect to theoretical underpinnings, and so begining in incorporate notions from game theory and social choice theory. As well, it is broadening wrt application areas, moving beyond traditional areas in AI and database systems, to including areas in description logics, the semantic web, and in economics.
- As well, there is an emerging focus on epistemic notions having to do with communicating, negotiating, competing, and collaborating agents in belief change. Dynamic epistemic logic seems to have an important role to play here.

Moreover, it looks like belief merging and iterated belief revision are still hot topics and will remain so for the next few years.

4.6 Principles and Practices of Semantic Web Reasoning

Seminar No. **05371** Date **11.09.–16.09.2005**
Organizers: F. Bry, F. Fages, M. Marchiori, H.-J. Ohlbach

The seminar "Principles and Practice of Semantic Web Reasoning" took place from Septem-ber 11-16, 2005. It was organised by F. Bry (Univ. München, DE), F. Fages

(INRIA Rocquencourt, FR), M. Marchiori (MIT - Cambridge, US) and H. J. Ohlbach (Univ. München, DE). The seminar was a forum for discussing various forms of reasoning that are or can be used on the Semantic Web. Moreover, it addressed both reasoning methods for the Semantic Web and Semantic Web applications relying upon various forms of reasoning.

The seminar was attended by 50 researchers that work in the area of "Reasoning on the Web". 29 of these participants are members of the European Network of Excellence REWERSE ("Reasoning on the Web with Rules and Semantics") that is funded by the EC and Switzerland. REWERSE is a leading project in the area of rule-based reasoning and applications for the Semantic Web. REWERSE started on March 1, 2004 and is scheduled for four years (cf. http://rewerse.net).

The technical program of the Dagstuhl seminar consisted of talks, tutorials, panel discussions, demo sessions and general discussion sessions.

Emerging research topics from the seminar

While the presented talks, demos and tutorials showed significant progress in various specific research areas of "Reasoning on the Web" in particular the seminar panels and discussions revealed a number of challenges for the current research on rules and reasoning for the Semantic Web. Some of the addressed challenges were:

- Rules need well-defined semantics
- Rules necessary for making ontologies inter-operable
- Semantic Web requires general high-level rule languages
- Relevance of Rule Interchange Languages
- Concrete foundational research issues
- Viewpoints on Rule Languages
- Truth on the Web is "context-dependent"
- Semantic Web community has to be aware of previous research
- Scalability of reasoning in a completely decentralized setting
- Trends towards mainstreaming of rules
- Central role of use cases
- Start with simple but relevant use cases

General Conclusions from the Seminar

General conclusions from the seminar are very positive. The seminar contained very good presentations also by young researchers, had very lively and fundamental discussions profiting from the contributions of senior researchers, fostered the co-operation between different communities working in the area of "Reasoning on the Web with Rules and Semantics" and showed the relevance of further meetings and workshops of this type. In particular, showcasing the relevance for practical applications was frequently commented as a central task for the coming months.

4.7 Spatial Cognition: Specialization and Integration

Seminar No. **05491** Date **04.12.–09.12.2005**
Organizers: A.G. Cohn, C. Freksa, B. Nebel

Representation and processing of spatial information about our environment is an essential requirement of everyday cognition. We can find our way in the environment and learn the layout of a building or a city. We can infer the location of objects from the location of other objects. We can reach out and manipulate objects in the environment. We can envisage spatial arrangements to find creative solutions for complex technical problems, for construction planning of buildings, and for the creation of pieces of art. Today there is a great body of evidence on how humans (and animals) reason about space, how they navigate through familiar and unknown environments without getting lost, how they act in spatial environments, how they interact in space, and how they communicate spatial information. One of the major challenges for current research is how these abilities can be accomplished by technical systems.

A variety of disciplines are involved in the spatial cognition enterprise: besides computer scientists / AI researchers who develop and analyze calculi for qualitative spatial (and temporal) reasoning there are cognitive psychologists and biologists who study human spatial navigation behavior and other spatial task performance and their neural correlates; cognitive geographers who study the use of spatial knowledge in large-scale spatial environments and appropriate representations of geographic knowledge; philosophers of the mind who study conceptions of spatial entities and their formal description; cognitive roboticians who employ spatial representations for autonomous robot navigation and develop systems that autonomously acquire knowledge about their spatial environments; computational linguists who study human spatial concepts through the analysis of natural language and formalize this knowledge to support human-robot communication; architects who design spatial environments for human use and must configure these spaces according to functional requirements and according to human conceptions of space; informaticians try to make use of all these insights to develop appropriate representation and reasoning tools and to build assistance systems that support and complement human capabilities.

The intention of this Dagstuhl Seminar is to bring together researchers working on different aspects of spatial cognition and from the perspectives of various disciplines to discuss

the state of the art in spatial cognition. A focus of the discussions will be the trade-off between specialized representations and general approaches and the integration of different approaches into a common representational framework.

Specialized representations are needed for efficient spatial reasoning; for example, incomplete knowledge about spatial situations must be represented in such a way that not all possible extensions of an under-specified situation need to be computed. Specialized representations are also found in the communication between human and artificial cognitive agents whose cognitive and perceptual capabilities differ critically. Communication serves here as an integration process between two different spatial representations. To be successful we must establish ways to transform between different ontologies and abstractions in such a way that we can switch the perspective on a given segment of spatial reality.

A goal of the workshop is to clarify the relation and integration of different specialized representations of a segment of a spatial environment under different perspectives. For example, depending on the class of tasks to be solved, we may need representations of (3-dimensional) objects, (2-dimensional) regions, (1-dimensional) routes, or (0-dimensional) landmarks. For certain spatial tasks, e.g. navigation, there may be substantial advantages to abstract from most of the possible perspectives in order to maintain a single consolidated view of the environment, for example in a route graph representation.

The discussion about issues of integrating spatial representations is expected to result in a better understanding of the relationships between spatial environments, cognitive agents and their actions in spatial environments, interactions among cognitive agents and between agents and their environments, and the representations and processes involved. The interdisciplinary character of the seminar opens up the possibility of discussing the various contributions offered by different research efforts and for evaluating to what extent they are alternative approaches towards the same goal or necessary complementary efforts to explain and understand spatial cognition in terms of a computational process model.

The contributions from different lines of research in spatial cognition will be critically evaluated and discussed. To what extent are optimization criteria from informatics applicable to cognitive performance? Do we have to take additional dimensions into account? How useful are empirical studies of populations of cognitive agents for understanding their computational mechanisms? How much variation and variability can we expect in spatial cognitive functions and more generally: in cognitive abilities? What is the role of formal systems for spatial cognition and spatial cognition research? What is the relation between visual and linguistic forms of spatial representations? The roles of multiple conceptual and spatial reference systems, of low-level and high-level structures, of multiple spatial ontologies, and of the problem context may turn out to be of particular relevance in this discussion.

A valuable result of the Dagstuhl Seminar would be an assessment of the relative importance of various virtues of spatial cognition systems – like completeness, uniformity, consistency, precision, crispness, tractability, formality, and others – and of the various methods employed for spatial cognition research – formal approaches, empirical studies, computational models, robot implementations, etc. Should we invest in efforts to integrate these different approaches in a more systematic fashion?

Chapter 5

Programming Languages, Compiler

5.1 Verifying Optimizing Compilers

Seminar No. **05311** Date **31.07.–05.08.2005**
Organizers: J. Knoop, G. Necula, W. Zimmermann, L. Zuck

There is a growing awareness, both in industry and academia, of the crucial role of formally proving the correctness of systems. Most verification methods focus on the verification of a specification with respect to a set of requirements, or of high-level code with respect to a specification. However, if one is to prove that a high-level specification is correctly implemented in low-level code, one must verify the compiler which performs the translations. Verifying the correctness of modern optimizing compilers is challenging because of the complexity and reconfigurability of the target architectures, as well as the sophisticated analysis and optimization algorithms used in the compilers.

The introduction of new families of microprocessor architectures, such as the EPIC family exemplified by the Intel IA-64 architecture, places an even heavier responsibility on optimizing compilers. Static compile-time dependence analysis and instruction scheduling are required to exploit instruction-level parallelism in order to compete with other architectures, such as the super-scalar class of machines where the hardware determines dependences and reorders instructions at run-time. As a result, a new family of sophisticated optimizations have been developed and incorporated into compilers targeted at EPIC architectures.

The increasing popularity of embedded systems brings with it a demand for fully automatic certifiers for a wide range of compilers, ensuring an extremely high level of confidence in the compiler in areas, such as safety-critical systems and compilation into silicon, where correctness is of paramount concern. Consequently, we have witnessed, in the past few years, a surge in application of formal method to the development and verification of optimizing compilers.

The goal of the seminar is to bring together researchers in all areas concerning the development of correct and verifiable state-of-the-art compilers, and provide a forum in which they can discuss, study, and develop cooperation, and advance the knowledge in all issues

concerning development of verifiable optimizing compilers. By encouraging discussions and cooperations across different, yet related fields, the seminar strives for bridging the gap between the communities, and for stimulating synergies and cross-fertilizations among them. The seminar is a consequence of the recent COCV workshops that are held under ETAPS, the idea of which was conceived in a Dagstuhl seminar (00381) in 2000.

Typical topics of interest include: Translation Validation, Self-Certifying compilers, Verification of Optimizations, Design of verifiable optimizations, Verification of Compilation, Use of logic in designing optimizations, and Impact of verification on compiler construction.

5.2 Beyond Program Slicing

Seminar No. **05451** Date **06.11.–11.11.2005**
Organizers: D. Binkley, M. Harman, J. Krinke

The aim of the "beyond program slicing" seminar was to explore emergent applications of program slicing and ways in which slicing techniques and ideas could be combined with those from other areas of program analysis and manipulation.

To achieve this goal, the seminar gathered together 36 people, including experts in the theory and practice of program slicing and those working on closely related areas, such as model checking, measurement, analysis, debugging, program comprehension, testing, reengineering and semantics.

The seminar was structured to provide a mix of pre-prepared talks and talks on work developed by the participants during the seminar. To achieve this, time was set aside for group working in groups of three. Groups were chosen to facilitate cross pollination of ideas from different fields. There was also time provided for preparation and networking and for tutorials and demonstrations of practical systems. The discussions and collaborative work continued into the small hours every morning, yet all the participants remained energetic and enthusiastic throughout the event.

Several new topics and ideas emerged at the workshop, both through formal presentations by the formally constituted groups of three and through unplanned serendipitous collaboration between the participants. The organisers are confident that several of the abstracts the reader will find under the DROPS proceedings of the workshop will become extended papers, forming the seeds of on-going collaboration and work.

Conclusion
The "beyond program slicing" Dagstuhl seminar was a resounding success with many technical outcomes which will continue to be developed by the inter-locking collaborative working groups formed during the seminar. The strong spirit of co-operation and collaboration which permeated the seminar is also expected to lead to a number of valuable, on-going, infrastructural, efforts to help in the support, facilitation and maturation of this growing community of researchers within source code analysis and manipulation and its application to software engineering.

Chapter 6

Software Technology

6.1 Transformation Techniques in Software Engineering

Seminar No. **05161** Date **17.04.–22.04.2005**
Organizers: J. Cordy, R. Lämmel, A. Winter

Abstract:
TrafoDagstuhl brought together representatives of the research communities in re-engineering, XML processing, model-driven architecture and other areas of software engineering that involve grammar- or schema-driven transformations. These various existing fields and application contexts involve widely varying transformation techniques — the tradeoffs of which are worth analysing. This seminar initiated a process of understanding each other's transformation techniques — their use cases, corresponding methods, tool support, best practises, and open problems. This process makes it possible to exchange knowledge and experience between these various communities. This effort should also help in transposing transformation concepts from established application fields to new fields.

Transformations everywhere

The idea for this seminar began with the observation of a discrepancy:

> While *software transformation* is a crosscutting theme in software engineering, the various fields in which it is used are only passingly aware of each other.

It would therefore make sense to bring together leading representatives from the different fields so that they can share problems and solutions related to their use of transformations and begin a dialogue on understanding transformation itself as a whole. Without claiming completeness, the following (somewhat overlapping) communities can be identified:

1. Program calculation

2. Language implementation
3. Model-driven development
4. Grammar(ware) engineering
5. Modelling and meta-modelling
6. Generative software development
7. Code restructuring and refactoring
8. Database reverse and re-engineering
9. Co-evolving designs and implementations
10. Data integration incl. semi-structured data
11. Design recovery and architectural recovery
12. Intentional and aspect-oriented programming

Most of these communities know of more than one kind of transformations. Also, transformation techniques are not always tied to a specific community. So it makes sense to abstract a little from the communities, and to identify some of the dimensions of variation for transformation techniques.

- The kind of grammars or schemas involved.
- The degree of automation of transformations.
- The degree of interactive transformations.
- The degree of formalisation of transformations.
- The degree of programming language support.
- The computational framework for transformations.
- The nature of transformation properties.
- The kinds of artifact: programs, data, and schemas.
- and so on.

During a week of intensive discussion, 47 participants from 12 countries attended the seminar, contributed presentations and participated in and/or organised discussion groups and panels. (International statistics for participants: Germany (14), Canada (12), U.S.A. (5), Belgium (4), the Netherlands (3), France (2), United Kingdom (2), Hungary (1), Ireland (1), Italy (1), Japan (1) and Switzerland (1).)

Handling diversity

How does one make sure that participants successfully share problems and solutions? It is clearly challenging to bring together a group with so much diversity — there is a chance that attendees simply lack enough shared common concepts to understand each other, that speakers can become confused and defensive when the established axioms in their fields are not observed by the mixed audience, and that different interpretations of technical vocabulary can lead to severe misunderstandings.

This risk was anticipated, and it was addressed in two ways.

1. The focus was moved away from "religious" issues.
   Instead, speakers were encouraged to focus as follows:

   - The presentation of technical content.
   - The discussion of accepted, best practises.
   - The presentation of foundations such as principled formalisms.
   - The selection of material that is *clearly* relevant across the community.

2. Particular technical concerns of interest were identified up-front:

   - Language-parametric transformation.
   - Reuse of transformation components.
   - Interdisciplinary transformation scenarios.
   - Coupled transformations of separate artifacts.
   - Language support for software transformation.
   - Properties of (semi-) automatic transformations.
   - Design patterns for transformational programming.
   - Validation of transformations by testing and others.
   - Best practises for the underlying grammars and schemas.

As a means to communicate the interdisciplinary character of the seminar well before the actual meeting, all participants were asked to prepare a very short introduction complete with an indication of their expectations from the seminar. By mining this information from the participants, it was possible to identify a number of themes with broad support:

- Modularity and composition of transformations.
- Correctness of transformations (in several respects).
- Evaluation and comparison of transformation setups.
- Adoption of transformation techniques ("ready for prime time").
- Transformations crossing boundaries (e.g., coupled transformations).
- Language independence and genericity in transformation techniques.

Scientific program: introductory presentations

The seminar week started with 6 introductory seminar presentations. These invited presentations were meant to introduce basic concepts and vocabulary so that some common ground was established. The remainder of the week consisted of regular presentations by participants, a small set of focused discussion groups built around questions that emerged early, and a panel session based on a suggestion that emerged at mid-week.

The actual selection of introductory areas was necessarily a compromise: it was not possible to have a dedicated overview talk for every one of the 12 identified communities. Titles and abstracts of the presentations follow in the order given.

Jean-Marie Favre (LSR – IMAG)
Etymology of Model Driven Engineering and Model Transformation

This presentation just attempts to put emphasis on the importance of defining words and concepts before the start. What is a model? What is a transformation? What is a model transformation? To be honest, I don't know. While everybody could agree with her- or himself, agreeing with others is usually more challenging, especially when going into the details. We should admit that we don't have precise and unique answers to these questions and that further work is required. This presentation just aims at raising the terminological problem considering not only the MDA approach but also other technical spaces.

Gabriele Taentzer (TU Berlin)
Specifying and Analysing Model Transformations based on Graph Transformation

Nowadays the usage of model transformations in software engineering has become widespread. Considering current trends in software development such as model driven development (MDD), there is an emerging need to develop model manipulations such as model evolution and optimisation, semantics definition, etc. If a model transformation is described in a precise way, it can be analysed later on. Models, especially visual models, can be described best by graphs, due to their multi-dimensional extension. Graphs can be manipulated by graph transformation in a rule-based manner. Thus, we specify model transformation by graph transformation. This approach offers visual and formal techniques in such a way that model transformations can be subjects to analysis. Various results on graph transformation can be used to prove important properties of model transformations such as its functional behaviour, a basic property for computations. Moreover, certain kinds of syntactical and semantical consistency properties can be shown on this formal basis.

Don Batory (University of Texas at Austin)
Understanding Aspect Composition Using Program Transformations

Program transformations can play a pivotal role in future models of software design, software architectures, and generative programming. This talk will examine AspectJ from this perspective.

AspectJ is among the more popular "new modularisation" technologies that are being explored in Software Engineering. AspectJ has clear advantages over traditional modular-

isation technologies (e.g., classes, packages), but has equally clear drawbacks as well. We argue that the drawbacks stem from the way in which aspects are composed, and show that step-wise development (and aspect reuse) is difficult with AspectJ. We propose an alternative model of composition, again based on program transformations, that supports step-wise development, retains the power of AspectJ, aligns AspectJ with existing results in Component-Based Software Engineering (CBSE), and simplifies program reasoning using aspects.

Eelco Visser (University of Utrecht)
Strategies for Rule-Based Program Transformation

In this talk I will sketch the development of rule-based approaches to program transformation using the program transformation language Stratego as a vehicle. The talk will cover techniques from pure rewriting, through strategies for the control of rewrite rules, to techniques for context-sensitive transformations. The techniques will be illustrated by means of examples of program transformation that prompted their development.

Andy Schürr (TU Darmstadt)
Model-Based Software Development - A Survey

Model-Based Software Development (MBSD) and OMG's recently invented three letter word MDA (Model-Driven Architecture) are the new hypes which tend to replace former terms like "object-oriented programming", "component-based software development", and the like. In this talk we will discuss the relationships between MDA and MBSD in general as well as their relationships to model and code transformation techniques. We will see that MBSD requires the combination of quite a number of today's existing transformation techniques and tools. Furthermore, we will present a long list of problems which have to be solved in order to make MBSD and transformation-based software development a "real" success story in practise. Finally, we will make some propaganda for so-called triple grammars, a special kind of declarative, bidirectional model transformation approach.

Alberto Pettorossi (CNR - Rome)
Formal Aspects of Program Transformation: Rules and Strategies

We present some program transformation techniques for: (i) deriving efficient programs from less efficient ones, (ii) synthesising programs from specifications, and (iii) proving program properties. In our examples we consider various programs taken from the area of discrete mathematics, optimisation, combinatorial mathematics, searching and sorting problems, and protocols for concurrent programming. The techniques we present are based on the syntax of the language (in particular, we consider functional and logic programs) and on theorem proving ideas such as those of proof rules and proof strategies.

6.2 Multi-Version Program Analysis

Seminar No. **05261** Date **26.06.–01.07.2005**
Organizers: T. Ball, S. Diehl, D. Notkin, A. Zeller

Change is an inevitable part of successful software systems. Software changes induce costs, as they force people to repeat earlier assessments. On the other hand, knowing about software changes can also bring benefits, as changes are artifacts that can be analyzed.

In the last years, researchers have begun to analyze software together with its change history. There is a huge amount of historical information that can be extracted, abstracted, and leveraged:

- Knowing about earlier versions and their properties can lead to **incremental assessments**.
- Analyzing the history of a product can tell how changes in software are **related** to other changes and features.
- Relating properties to changes can help **focusing** on changes that cause specific properties.

In this Dagstuhl seminar, researchers that analyze software and its history have met and discussed for a full week, exchanging their ideas, and combining and integrating the techniques to build a greater whole. Clearly, understanding history can play a major role when it comes to understand software systems.

Scientific Highlights

The main concern of the seminar was the synergy of the individual approaches. Themes that emerged during the seminar included:

- The use of **bug databases** to judge whether changes were beneficial or not;
- The use of **advanced visualization techniques** that integrate program analysis and history; and
- The use of **version histories** to conduct empirical research, as in the study of clone genealogies.

The latter point – leveraging version histories to conduct empirical research – was maybe the strongest highlight of the seminar. As a direct result of the seminar, a mining challenge was introduced at the Workshop of Mining Software Repositories.

All in all, Software engineering is full of anecdotical evidence, often relying on insufficient or proprietary data. Publicly available change and bug histories may change this, providing reproducible benchmarks for empirical research, and allowing anyone to assess hypotheses about what works in software engineering and what does not.

Perspectives
The analysis of programs across multiple versions has a bright future. The workshop on mining software repositories has never been more active; researchers begin to recognize the great potential of software history to understand development processes. With this Dagstuhl Seminar, we are very happy to have contributed to this momentum. We also look forward to see this research being applied in practice, as it is already the case with IBM and Microsoft.

6.3 Perspectives Workshop "Challenges for Software Engineering Research"

Seminar No. **05402** Date **06.10.–08.10.2005**
Organizers: M. Broy, M. Jarke, M. Nagl, H.D. Rombach

Software ist der fundamentale *Werkstoff des Informationszeitalters*. Innovative Produkte und Dienstleistungen sind ohne Software nicht mehr denkbar. Die Wettbewerbsfähigkeit der deutschen Wirtschaft hängt entscheidend von der Fähigkeit ab, Software-intensive Produkte und Dienstleistungen mit höchster Qualität zu erstellen. Software Engineering über dem Weltniveau ist die Voraussetzung dafür, dass Deutschland seine führende Stellung im Ingenieurbereich, etwa im Maschinenbau, halten und ausbauen und entsprechende Positionen in neuen Sparten, etwa im modernen Gesundheitswesen (e-Health), aufbauen kann.

Software wird in der Zukunft *integrierter* – in vielen Fällen sogar dominierender – Teil großer komplexer Systeme sein. Nicht nur in der Automobil- und Luftfahrtindustrie wird dieser Trend bereits heute deutlich sichtbar. Die erforderliche Integration von Mechanik, Elektronik und Software und die Vermeidung unerwünschter Wechselwirkungen kann nur durch die frühzeitige Integration der Modellierungskompetenzen des Software Engineering in den Entwicklungsprozess beherrscht werden.

Diese neue Positionierung von Software Engineering als systemische Disziplin erfordert eine neue Ausrichtung und Stärkung der Bereiche Forschung, Lehre und Technologietransfer. Es ist die auf sorgfältige Analysen abgestützte Überzeugung der Unterzeichner, dass verstärkte Anstrengungen in allen drei Bereichen notwendig sind, um die Herausforderungen des Informationszeitalters und der Globalisierung anzunehmen. Dieses Manifest stellt Forderungen inhaltlicher und organisatorischer Art für Forschung, Lehre und Technologietransfer auf, deren Erfüllung Voraussetzung für die Realisierung künftiger weltweiter wirtschaftlicher Potentiale ist.

Der Standort Deutschland wird sich in diesem Markt nur behaupten, wenn Software Engineering sowie Softwareingenieure auf höchstem Qualitäts- und Produktivitätsniveau verfügbar sind.

Die Ergebnisse des Perspektiven-Workshops, bei dem Wissenschaftler aus der universitären Forschung in Dagstuhl zusammenkamen, sind in einem Manifest zusammengefasst: "Strategische Bedeutung des Software Engineering in Deutschland":
http://drops.dagstuhl.de/opus/volltexte/2006/585/

Chapter 7

Applications, Multi-Domain Work

7.1 Perspectives Workshop: Multimedia Research – where do we need to go tomorrow

Seminar No. **05091** Date **01.03.–04.03.2005**
Organizers: S. Boll, T.-S. Chua, N. Dimitrova, R. Jain

The idea for this seminar arose during the ACM Multimedia Conference 2003. In a Strategic Retreat meeting of the SIG Multimedia of the ACM just before the 2003 conference, the question of whether the research community is sufficiently contributing to "real" multimedia research has recently been very vividly discussed at this year's ACM Multimedia conference 2003. From this meeting and its outcome it was clear that there was high demand in discussing what the future of multimedia research is and where it should go. So with this Dagstuhl Perspectives Workshop we wanted to invite leading researchers in the field to an open but yet focused forum to formulate and consequently establish future research directions in multimedia in the context of a Dagstuhl perspectives workshop. As the seminar has a broad and visionary goal, we aimed at bringing together not only researchers from different fields but also from different institutions and those with potentially controversial research opinions – to really achieve a leap forward. Venturous position papers were invited for submission before the seminar that would state their view and controversial issues that they want to discuss.

The course of the seminar. During the seminar the positions have been presented and discussed as a starting point for forming discussion groups that would go into details with specific topics and areas in the field. The seminar was structured as such that there was still enough time to discuss research opinions and selected topics both in working groups as well as in plenary. The short presentations and discussions were intended to identify and finally select the "brave new topics" that are used as the starting points for the working groups' discussions for the next two days. The goal is to stimulate discussions throughout the entire workshop. After the presentation of the participant's position statements on the first day, in a plenary debate we cluster the topics identified but also of controversy research positions and formed working groups. Each of the groups was going into discussion on the following questions:

- What are the real problems?
- What are the small what are the global problems?
- What do day to day users really need?
- What are the fundamental issues of MM here?
- What are the burning open issues in the field?
- Why are they not solved yet?
- Where are the challenges?
- Where do we need to go the next 10 years?
- What will be the results and applications in 2010 and 2015?

The spirit of the Dagstuhl Seminar remains. During the month after the Dagstuhl seminar one could see presentations and publications that referred to the seminar and make it a lasting success. Interestingly before the seminar the participants that did not know Dagstuhl so far were a bit skeptic about the seminar and its potential value, to them it was a pleasure to see the lively and interesting discussions during the seminar. Now, that the discussions started many of the participants would have liked to continue for one or two more days. And even beyond this, from the researchers and work that refers to this seminar we can see the real value of the seminar.

The future of multimedia research. Motivated by the outcome, the participants and the results we plan to continue with this seminar in the sense of a series in a "multimedia research" series in adequate regularity – in case we get accepted again with a new proposal at Dagstuhl. The organizers would take the opportunity to thank Dagstuhl for the wonderful ambience and their pleasant organization and all our participants for coming to Dagstuhl and making this multimedia research discussion such a pleasure and success.

7.2 Power-aware Computing Systems

Seminar No. **05141** Date **03.04.–08.04.2005**
Organizers: L. Benini, U. Kremer, C.W. Probst

Rapidly increasing chip densities and processor speeds have made energy dissipation a leading concern in computer design. The problem raised by energy consumption is especially severe for a whole class of computing devices which has recently become almost ubiquitously available – mobile devices like notebooks, PDAs, or mobile phones. On the one hand, these are only equipped with a very limited power supply, so any computation on such a device should be especially careful about resource usage. Even worse, the battery technology for these devices has not kept pace with advances in processor technology and the growing complexity of software. On the other hand, cooling mechanisms become more

and more important. Recent trends suggest that processor power consumption doubles every four years – and cooling costs rise exponentially with heat increases. The future processors will require energy management solutions more cost effective than the cooling fans used today.

It has been the goal of the seminar to bring together researchers from the main communities working on reducing power consumption, namely hardware, operating systems, virtual-execution environments, compilers, and applications. The main seminar result is a classification of the obstacles, and therefore research directions, with respect to power consumption seen for different classes of devices, ranging from very low power devices, over handheld devices, to servers and work stations. In a next step the seminar identified the impact different levels of dealing with power concerns can have.

The program of the seminar featured presentations of about 35 participating researchers from academia and industry. They were chosen to represent major areas in targeting the energy consumption of a computing system – Applications, Compilers, Virtual-execution Environments, Operating Systems, and Hardware.

In order to identify problem areas and future research areas, discussion groups were formed that resulted in four working groups. In addition, abstracts of the presentations as well as work-in-progress papers are published in the proceedings. Some of the work presented at the seminar will be published in a special issue of the International Journal of Embedded Systems on power, energy, and thermal topics.

7.3 Efficient Text Entry

Seminar No. **05382** Date **21.09.–24.09.2005**
Organizers: K. Harbusch, K.-J. Raiha, K. Tanaka-Ishii

The range of electronic devices which don't provide a full keyboard is increasing (e.g., cellular phones, tablet PCs, palm- and watch-sized devices). This deficiency imposes the challenge of efficient alternatives to typing with a full keyboard.

The same problem has to be tackled by motor impaired persons. It holds in general that typing is slow and cumbersome for these users. Persons with motor and speech impairments unconditionally depend on typed and synthesized utterances produced by a communication aid in order to communicate.

Similarly, typing of Asian languages with huge alphabets (e.g. Chinese or Japanese) requires suitable and efficient methods to access all the letters with a standard keyboard.

In all these application areas, various text entry methods have been suggested to provide a more efficient input of texts with lower motor demands. Usually, they combine specific typing devices with methods that aim at reducing the number of necessary keys or key strokes, respectively. Among these are word prediction, abbreviation expansion, ambiguous typing, text compression and text compansion (i.e., short telegram style or even pure semantic concepts are automatically expanded into complete sentences).

At their core, these methods rely on statistical and to a lesser extent rule-based language models to predict and complete the user input and thus save keystrokes. Unlike in speech recognition, the language models are also used to change and evaluate the way how to enter text.

In this workshop, we invite researchers of the different application areas to share their results and ideas with the other communities in computational linguistics.The thirteen presentations covered the full range of the above mentioned areas. In the detailed discussions with every presentation, generalizations, similarities and differences of the specific viewpoints in the different application areas were pointed out.

Chapter 8

Distributed Computation, Networks, VLSI, Architecture

8.1 Computing and Markets

Seminar No. **05011** Date **03.01.–07.01.2005**
Organizers: D. Lehmann, R. Müller, T. Sandholm

Markets are institutions used by economic agents to trade goods. Progress in information and communication technologies, in particular the development of the Internet, have changed the way information is shared amongst market participants, how market clearing is organized, and last but not least, who is aware of and has access to a particular market. Internet auctions, like eBay, provide a prominent example. Industrial applications like supply chain management and procurement auctions challenge computer scientists by their need for complex negotiation and clearing protocols. In recent years this has stimulated a plethora of computer science research in market design, in particular for combinatorial auctions.

The infrastructure underlying this development, the Internet with its services for communication and content, becomes itself a huge market with millions of agents sharing resources like bandwidth, server CPU cycles, memory, and content. The performance of such systems depends not only on their implementation, but more and more on the behavior of its users. Therefore, analysis and design of information systems need to incorporate a game-theoretic analysis in order to accurately predict system performance. Central to such analysis is the notion of independent agents which act selfishly in order to maximize their own utility. The implementation of the system defines the rules of a game. Game theory provides us with the tools to analyze the behavior of agents in such environments. Agents themselves can again be computerized in the form of adaptive software agents, using learning techniques developed in AI to outperform other agents.

The Dagstuhl seminar Markets and Computing brought together researchers from Economic Theory, Game Theory, Artificial Intelligence, Theoretical Computer Science and Operations Research to present and discuss their approaches towards understanding the

interrelation between advances in computing and the design of markets. The main threads of the seminar were the following.

A couple of sessions dealt with algorithmic issues arising in combinatorial auctions. We had sessions on the topics pricing in auctions and iterative auctions, analyzing the (pricing) information that can be efficiently provided during and after an auction, and the convergence of iterative auctions towards a pricing equilibrium. A session on discrete convexity presented the relation between this notion from discrete mathematics and the existence of particular pricing equilibria in combinatorial auctions.

Related to these topics were sessions on bidding and simulation. Important issues here were the implementation and experimental evaluation of bidding strategies, and experimental evaluation of efficiency and other market objectives under such bidding strategies.

A third thread of sessions dealt with mechanism design for particular market situations. One of these sessions investigated designs for online situations, in which bidders arrive over time, or supply of items changes over time. Other papers in this thread dealt with specific domains, approximate mechanism design, and automated mechanism design.

There was a thread of sessions with an emphasis on game-theoretic issues. A prominent topic there was the impact of selfish agents on system performance with respect to various performance measures. Finally we had sessions on social choice, dealing with the possibility of the design of mechanism with specific desirable properties.

Similar to previous Dagstuhl seminars on related topics, this seminar facilitated a very fruitful interaction between economists and computer scientists, which intensified the understanding of the other disciplines' tool sets, and which is likely to lead to joint research projects across the disciplinary borders. This seminar helped to pave the way to a unified theory of markets that takes into account both the economic and the computational issues – and their deep interaction.

8.2 Foundations of Global Computing

Seminar No. **05081** Date **20.02.–25.02.2005**
Organizers: J. Luiz Fiadeiro, U. Montanari, M. Wirsing

Today software undergoes a fast technological progress where not only the complexity of software systems is considerably increasing but also the computing infrastructure is dramatically changing. The internet serves as platform for globally distributed applications; objects we work with are more and more equipped with processors and embedded software; in the Grid approach computing power is shared among many distributed processors; physical devices are mobile, data can migrate, and applications come and go in a dynamically changing ad-hoc way. Thus computation becomes global in the sense that it is distributed over the net and highly dynamic, with the network often changing dynamically.

These developments lead to enormous engineering challenges for

1. constructing global computing systems with predictable and desirable behaviour and

2. ensuring reliability, safety, security, and availability of global computing systems.

Current engineering techniques such as object-orientation, modelling languages such as UML, and actual programming languages such as Java and CASE tools can only partly cope with these problems. These approaches support only partly the construction of global computing systems and have to be extended with new appropriate concepts for software architecture, coordination, distribution and mobility. Moreover, these pragmatic approaches lack scientific foundations for analysing and validating global computing systems. On the other hand, formal techniques such as process calculi, type systems and logics for specifying, analysing and validating global computing systems are only partially available and mostly not well connected with systematic development methods.

The aim of this seminar is to study foundational theories, calculi and techniques for global computing systems and to bring them together with methods for systematically developing global computing systems. The topics of the seminar include:

1. calculi, models, and semantic theories of concurrent, distributed, mobile global computing systems,
2. types and logics for security, safety, resource-control and availability of global computing systems,
3. systematic development techniques, modelling and programming languages for global computing.

8.3 Scheduling for Parallel Architectures: Theory, Applications, Challenges

Seminar No. **05101** Date **06.03.–11.03.2005**
Organizers: E. Altman, J. Dehnert, C.W. Kessler, J. Knoop

Scheduling, the task of mapping computation units to time slots on computing resources for execution, is important for the effective use of resources in all kinds of parallel systems, ranging from the level of more coarse-grain tasks in multiprocessors, clusters and computational grids, to medium-grain tasks at the loop level, down to instruction-level parallelism (ILP).

Scheduling issues are of crucial importance in very diverse areas ranging from operating systems and realtime systems to network management to static and dynamic program optimization and code generation. Likewise, they evolve on very different levels of granularity, from coarse grain task and job scheduling over loop scheduling to fine-grain instruction scheduling. Though highly interrelated, these fields are tackled by usually independently working communities. However, emerging processor architectures such as chip multiprocessors will demand effective hybrid scheduling strategies that unify previously separate scopes of scheduling.

In practice, scheduling problems often do not appear in isolation but come with a domain-specific context that — explicitly or implicitly — introduces interdependencies with other optimization problems. For instance, when compiling for parallel execution platforms, decisions made in scheduling depend on and influence other aspects of the problem of generating efficient parallel code, such as resource allocation, clustering, or program transformations, such that scheduling can rarely be considered as an isolated problem. Such interdependences, even though perhaps most apparent for instruction-level parallelism, appear at all levels of parallelism and are solved by various techniques, including heuristics, integer programming, dynamic programming, or genetic programming. Integrated approaches are generally more flexible but suffer from an increased problem complexity.

Interestingly, the research communities for task-level, loop-level and instruction-level scheduling appear to be quite separated from each other. Furthermore, there appears to be a gap between the theoretical foundations of scheduling, formulated in terms of abstract machine models, and the algorithms developed in both academia and industry for concrete scheduling problems in compilers and run-time systems for parallel computer architectures. This gap is exacerbated by requirements that practical schedulers deal with the complexities of irregular architectures.

The purpose of this seminar was therefore to gather leading experts from these scheduling communities, to identify common approaches, techniques, frameworks and tools, and to stimulate cross-fertilization between the various scheduling communities. Moreover, we intended to bridge the gap between scheduling theory and methods currently applied in compilers and run-time systems for parallel architectures. A third goal was to encourage a constructive dialog between scheduling algorithm designers and developers of parallel architectures, specifically in the embedded systems domain.

31 researchers accepted the invitation to the seminar and met 6–11 March 2005 at Schloss Dagstuhl, Germany. The seminar participants represented a broad spectrum of research on scheduling, including instruction scheduling, job scheduling, task scheduling, loop scheduling, parallel computer architecture, and scheduling theory.

With the invitation and the opening address, we provided the following guiding questions:

- How can we bridge the gap between scheduling theory and practice?
- Can the practical ILP scheduling problems broaden the theory models?
- In particular, do recent micro-architectural trends such as clustered architectures add fundamentally different factors to the problem?
- Is there current theory that can lead to interesting practical algorithms?
- What are the interference effects between task-level, loop-level and instruction-level scheduling?
- Can existing scheduling approaches be transferred to other problem domains, granularities, or architecture models?

- What are the phase-ordering effects, and the techniques, potential, and limitations of integrating scheduling with other transformations or code generation phases?
- Can we define generic scheduling approaches for flexible optimization goals (execution time, stack/register space, energy consumption)?

A central goal of this seminar was thus to bring together leading experts of the various communities to foster discussions on the usability and usefulness of approaches developed for specific areas and the impact they may have to others. By means of cross-fertilization and synergy the seminar should contribute to both a better understanding of the key issues of scheduling and to further advancing the state-of-the-art in the various fields. The specific atmosphere of Dagstuhl Seminars, which can be characterized by openness, accessibility and cooperation, and which is supported by Schloss Dagstuhl's architecture, its services and facilities, encourages both formal and informal meetings and discussions and therefore provided a perfect environment to achieve these goals.

The seminar started with a short introductory presentation from each participant, with his/her view on the topic and expectations for the seminar. During the seminar week, there were 26 presentations by the participants, discussions in plenum and in smaller working groups, and work-in-progress sessions. The presentation abstracts are given in the remainder of this seminar report.

At the end of the seminar, the general impressions brought up in the final wrap-up discussion were thoroughly positive: The representatives of the various scheduling communities, albeit being different in scope, application domain, and even in the terminology used, really understood each other and could learn and got inspirations from each other's presentations. For instance, a constructive dialog was initiated between compiler scheduling and job/task scheduling theory researchers, providing in the one direction new problem formulations and application areas for the theory community and in the other one the wish from the practitioner's side to the theory community for better communication of the constraints and limitations of theoretical solutions, such as assumptions in the model. There was an insight into the problem of different scheduling communities and of insufficient communication of results e.g. between theory and the various application domains, leading to undesirable effects such as reinvention of the same or similar ideas in different fields and with different terminology. The need for bridging the gap between scheduling theory and practice was recognized, although a fundamental solution to this problem, of course, cannot be provided within the scope of a single seminar week. There were suggestions to create a common web page collecting resources, existing results, ongoing work, open problems etc.; to compile a common list of most important literature references, to agree on a common benchmark suite for scheduling problems that covers a broader range of problems and in particular includes malleable tasks, and even to build a common experimental compiler platform that covers all types of scheduling problems in a single framework and therefore allows e.g. to empirically study trade-off effects between instruction-level, loop-level, and thread-level scheduling, which turn out to be hard to define in a generic way because they heavily depend on the underlying hardware platform. It may be premature to speak of the birth of a new, merged scheduling community, but at least there exist concrete plans to organize a successor workshop with the same broad scope of scheduling issues for parallel architectures, prospectively in 2007.

8.4 Disruption Tolerant Networking

Seminar No. **05142** Date **03.04.–06.04.2005**
Organizers: M. Brunner, L. Eggert, K. Fall, J. Ott, L. Wolf

During the past 40+ years, numerous architectures were developed for network communication, including the ISO OSI reference model and its related protocol specifications and – of course – the Internet architecture. These network architectures all have been designed with some implicit assumptions about specific target applications and deployment scenarios. Among the most important assumptions are specific characteristics of the underlying network (= link layer) technologies, such as relatively short transmission delays, low error probability and the existence of end-to-end paths.

In certain advanced scenarios, these assumptions no longer hold. Examples of such advanced scenarios include networks with frequent connectivity disruptions, extremely long transmission delays or loose connectivity. Consequently, the existing network architectures fail to support communication in these scenarios, resulting in either significant inefficiencies or complete loss of connectivity.

Several localized, ad hoc solutions attempt to improve specific aspects of existing network architectures to better support these advanced scenarios, such as long/fat pipe extensions to individual transport protocols. These fixes can be successful in limited scenarios, but often lack broad applicability, i.e., they often address the symptoms of the issues instead of considering the causes. General architectural considerations are needed to approach the issues from a more fundamental and long-term perspective, rather than adding to a growing collection of short-term "patches".

Disruption Tolerant Networking is a new area of research in the field of networking that deals with extending existing protocols or inventing new ones in a coordinated, architecturally clean fashion, to improve network communication when connectivity is periodic, intermittent and/or prone to disruptions.

Among the challenges of this field of research are large transmission delays. These may result either from physical link properties or extended periods of network partitioning. A second challenge is efficient routing in the presence of frequently disconnected, prescheduled, or opportunistic link availability. A third challenge is high link-error rates that make end-to-end reliability difficult. Finally, heterogeneous underlying network technologies (including non-IP-based internetworks) and application structure and security mechanisms capable of limiting network access prior to data transit are required in environments with very large round-trip-times. In some cases, an end-to-end path may not even exist at any single point in time. From a mobility perspective, DTN relaxes the "always on" paradigm, which would be extremely costly or even impossible to realize in challenged environments.

These challenges can decrease the reliability and performance of communications at essentially all layers of the protocol stack, ranging from packet-based forwarding and routing, to reliability and other features provided at the transport layer, to the application protocols (and applications) themselves. In addition, traditional mobility approaches may have

to be revisited to accommodate users in networking environments prone to connectivity disruptions.

Numerous research activities over the past three years have focused on various facets of communications in challenged environments. Architectural concepts have been devised, prototype implementations were developed and research results are available from analysis, simulations and real-world experiments. The Dagstuhl seminar brought together researchers working in otherwise at least partly disjoint areas and established an intense dialogue across the variety of application domains.

In summary, this Dagstuhl seminar has sharpened the understanding of the very different perspectives from which researchers approach the problem space of disruption-tolerant networking, their assumptions and requirements, and the short- and longterm solutions they envision. This has broadened the view on DTN at large and contributes further issues to the present DTN research topics such as naming, security, service differentiation and efficiency. Assuming the traditional well-connected Internet architecture and its (interactive) applications as one extreme and the DTNRG architecture for purely asynchronous communications as another, the middle ground of mobile and partly (dis)connected operation may be approached from either edge. Future research will need to determine how far the DTNRG architecture can and should reach towards traditional Internet applications while maintaining its architectural integrity.

8.5 Perspectives Workshop: Peer-to-Peer Mobile Ad Hoc Networks – New Research Issues

Seminar No. **05152** Date **09.04.–12.04.2005**
Organizers: M. Gerla, C. Lindemann, A. Rowstron

Scope

A mobile ad hoc network (MANET) is a collection of mobile wireless nodes that can dynamically form a network without any pre-existing infrastructure. Due to the distributed nature of MANET, its networked applications typically employ the peer-to-peer (P2P) paradigm. Multiple P2P networks, corresponding to different applications, may coexist on a large ad hoc network structure (eg, battlefield, urban vehicular grid, etc). MANETs may be standalone, or may grow as an "opportunistic" extension of the wired Internet. In contrast with the wired Internet counterparts, P2P applications for MANETs are still in its infancy. Beyond popular P2P systems like KaZaA, one of the most promising recent developments for the Internet is a set of proposals for P2P look-up services based on Dynamic Hash Tables. In this invited workshop, we would like to focus on the applicability of the wired P2P models to the wireless scenarios and more generally to the challenging research problems emerging in the ad hoc P2P area. Particular topics of this workshop include, but are not limited to:

- Quantitative characterization of mobile P2P application scenarios (e.g., emergency, infotainment, military, ...)

- Design challenges of purely mobile P2P applications (e.g., mobile file sharing, mobile instant messaging, mobile gaming, ...)

- Design challenges for ad hoc networked extensions of wired P2P systems (e.g., Gnutella, KaZaA, dynamic hash tables, multi-player-gaming, ...)

Format

Participation in the workshop is by invitation only. Attendance is limited to 30 people to foster extensive discussion. The workshop will comprise of both plenary sessions and working-in-group sessions. Thus, we are taking the term "workshop" literally and concentrate on discussing existing results and future research challenges rather than just presenting research papers. The goal of the workshop lies in taking stock of the challenges imposed by, and state-of-the-art results, already available for P2P applications for MANET. The workshop will start with a plenary session in which all participants will provide a short introduction to their work. Following the plenary, we will form working groups of 4-8 participants for each of about 4 sub-areas. An area captain will head each working group and will report the results of working-in-group sessions in subsequent plenary sessions. Throughout the workshop, we will focus on several key issues including, but not limited to the following:

- Which results known for wired P2P systems can be adopted for developing and deploying P2P applications for ad hoc networked systems?

- What kinds of new results are needed for developing and deploying P2P applications for ad hoc networked systems?

- What are first success stories and failures of P2P applications?

- What are current burning questions of P2P applications?

Anticipated Results

After the workshop, the organizers will put together a write up summarizing the discussion results of the workshop (i.e., write a manifesto on the topic) and try to publish this paper in a broadly distributed technical journal. The aim of this paper lies in taking stock of the state-of-the-art in peer-to-peer applications for ad hoc networked systems and in developing recommendations for future research directions in Germany, Europe, and the US. Furthermore, the organizers will apply for guest-editing a special issue on the workshop topic at a journal like Mobile Networking and Applications (ACM MONET). Contributions for this specials issue will be solicited from the workshop participants, though, there is also an open call-for-papers. All submissions will go through a rigorous review process to ensure technical quality.

8.6 Mobile Computing and Ambient Intelligence: The Challenge of Multimedia

Seminar No. **05181** Date **01.05.–04.05.2005**
Organizers: N. Davies, T. Kirste, H. Schumann

Mobile computing gains in last years an increasing importance. It penetrates into new areas of everyday life. With increasing computing power that is available in recently developed mobile devices the mobile computing approach is used in new areas (where its use would not be possible few years ago). Due to specific properties of mobile devices the traditional approaches to the traditional schemes for communication between users and the devices are not applicable. This fact requires both development of new approaches for the design of user interfaces and also new approaches how the new possibilities of mobile computing should be used in particular applications, should be explored.

Another important issue that is linked with new applications in the field of mobile computing is derived from the fact that portable information appliances are pervading the everyday life and ambient intelligence is starting to surround us. Personal mobile guides and intelligent meeting rooms are examples of future smart environments that provide us with information and assistance, tailored to our individual needs, anytime, and anywhere. However, this also means that future infrastructures for multimodal interaction with multimedia information will be distributed and heterogeneous. Adapting multimedia applications and services to these delivery environments and enabling ensembles of multimedia appliances to organize themselves spontaneously and ad hoc will be major technical issues that have to be solved in near future. These problems were discussed during the Dagstuhl Seminar 05181 that took place in Schloss Dagstuhl from 01.05.05 to 04.05.05.

8.7 Semantic Grid: The Convergence of Technologies

Seminar No. **05271** Date **03.07.–08.07.2005**
Organizers: C. Goble, C. Kesselman, Y. Sure

The scientific paradigms of the Semantic Web, Web Services, Agents, Peer-to-Peer Networks and Grid Computing are currently receiving a lot of attention in the research community, and are producing solutions to important problems ranging from e-science to e-business. The United States DAML program, the European Commission and other organisations have also been investing heavily in these technologies. This Dagstuhl Seminar brought together world-leading experts from the diverse organizations and research areas. It strengthened the international collaboration with the aim to realize the vision of the Semantic Grid.

The main achievements of the Seminar include:

- creation of a strong and vivid Semantic Grid community, which shares understanding of principle ideas

- foundations for a WSRF-resource ontology
- foundations for a Virtual Organisation (VO) ontology, based on the EarthSciences Grid ontology
- a shared understanding of what a Semantic Grid is, and a need to encapsulate this in a bumper sticker
- an understanding of what a VO is, and using Dagstuhl as an example,
- the role of agents and Grids,
- a coordinated exchange of staff between Inteligrid, K-WfGrid and OntoGrid, which has already begun
- a book (working on it ...), and
- a number of visits stemming from the spin of the seminar (Frank Sieberlist visited Hannover and Manchester from Globus, for example)

The results of the seminar also contributed to presentations at the KnowledgeWeb Summer School on the Semantic Web 2005 on 22nd July, and the OECD Grid Global Science Forum in Sydney, Australia 25th September, both given by Carole Goble. Material developed in the VO discussions have contributed to a case study in the OntoGrid EU STREP and the WSRF/VO ontology has contributed to OntoGrid's Reference Architecture for Semantic Grids. Two EU IP proposals (BIG and BREIN) on the Semantic Grid have also be submitted in Sept 2005 by partners attending the seminar, based in part on discussions therein.

8.8 Algorithmic Aspects of Large and Complex Networks

Seminar No. **05361** Date **04.09.–09.09.2005**
Organizers: F. Meyer auf der Heide, D. Wagner

Information systems like the Internet, the World Wide Web, telecommunication networks, sensor networks, or peer-to-peer networks have reached a size and a dynamics that puts them beyond our ability to deploy them and to keep them functioning correctly through traditional algorithmic techniques. Their complexity arises from their growth and dynamics: They change their size and structure over time, large components may be modified, deleted or replaced. Challenges for algorithms research include development of models for the dynamics of such systems, the design of decentralized, scalable and adaptive mechanisms for regulation and improvement, and the design, the analysis, and the experimental evaluation of algorithms for analyzing and optimizing such systems.

Today, networks play an important role in many areas of our society. Information acquisition and dissemination as well as the further growing mobility is based on the understanding and employment of such huge networks. Interesting relations between networks in areas like traffic and transport, politics and society make it promising to strengthen the joint effort to tackle them. Dealing with large networks is a big topic also in areas like physics, biology, and economy, with very different, mainly non-algorithmic methods and different optimization goals.

The seminar brought together 46 researchers from different European countries, Australia and USA. Most participants were from algorithms research, but several also came from other branches of Computer Science as well as from Physics and Biology. As in previous meetings there were not only scientific talks on results respectively ongoing research, but also fruitful and stimulating discussions.

The seminar also offered an opportunity to meet and intensify collaboration for resrachers involved in the DFG research cluster "Algorithmic Aspects of Large and Complex Networks" (DFG Schwerpunktprogramm 1126) and from the EU Integrated Project DELIS (Dynamically Evolving Large-scale Information Systems).

8.9 Form and Content in Sensor Networks

Seminar No. **05381** Date **18.09.–23.09.2005**
Organizers: L. Guibas, U.D. Hanebeck, T.C. Henderson

Advances in computing hardware and wireless networking technologies have enabled low-cost, low-power miniature sensor devices. By combing these tiny sensor nodes, which comprise sensing, data processing, and communication components, a sensor network for distributed sensing is obtained. For example, thousands of these nodes could be spread across a large geographical area in order to perform weather monitoring. Since every node suffers from limited sensing, computing, networking, and energy resources, collaboration between nodes is required, which stimulated the development of new types of communication and information processing algorithms.

Sensor networks have received increasing attention over the last few years. For example, DARPA's SensIT program envisioned the development of dense fields of simple, inexpensive, micro-sensor-processors to exhibit emergent behavior which should be collectively brilliant, operationally effective, long-lived, and robust to failure. Researchers have begun to explore the rich design space of low-power processors, communication devices, and sensors and the development of algorithms for these distributed wireless sensing systems

Some examples of issues addressed by these various projects include self-configuration, data handling, systems issues, power minimization, and fault tolerance. In general, higher-level exploitation of sensor networks requires the application of distributed algorithms to the data. Corresponding applications include distributed source localization, calibration, and habitat monitoring. Areas of particular interest include: (1) Algorithms for sensor tasking and control, including distributed techniques for the formation of sensor collaboration

groups, tracking large-area phenomena, information storage and in-network aggregation, identity management, and relational tracking and reasoning, (2) the application of optimal estimation techniques to multisensor data analysis, nonlinear filtering, position estimation and tracking, and the application to mobile robotics, cellular mobile phones, and virtual reality, and (3) the creation of an information layer on top of the sensor nodes, including distributed algorithms for leadership protocols, coordinate frame, and gradient calculation, and distributed signal, feature and structural analysis methods.

The aim of this seminar was to summarize the current state of the art in the field of algorithms for sensor networks and how content and structure impact information processing in the network, and to identify open problems in the following areas

- Low-level Processing (e.g. distributed signal processing, compression, estimation, detection, error handling)
- High-Level Processing (e.g. network structure, dynamic self-organization, tracking, monitoring, system behavior evaluation)

The seminar brought together workers from these areas to establish a stronger dialog with the goal of achieving a coherent view of the communication and information processing aspects which are typically treated separately today.

8.10 Anonymous Communication and its Applications

Seminar No. **05411** Date **09.10.–14.10.2005**
Organizers: S. Dolev, R. Ostrovsky, A. Pfitzmann

Throughout history encryption was used to hide the contents of transmitted data. The rapid growth in the use of the Internet only increased the necessity of encryption. However, encryption does not hide all relevant information, for example, it does not hide the identity of the communicating parties. That is, it does not prevent traffic analysis.

Modern cryptographic techniques are extremely good in concealing all the contents of data, by means of encrypting the messages. However, concealing the contents of the message does not hide the fact that some message was sent from or received by a particular site. Thus, if some location (or network node) A is sending and/or receiving a lot of messages to/from B, and if an adversary can monitor this fact, then even if the adversary does not understand what these messages mean, just the fact that there are a lot of messages between A and B reveals that and when they cooperate.

The objective of anonymous communication is to show how to hide, in an efficient manner, the identity of users who transmit (or receive) a data to (or from, respectively) other sites in the network. The workshop is aimed both at discussing various techniques of anonymity and also exploring applications of anonymity.

The question of anonymous communication was studied both by academia and industry for more than two decades.

The seminar will serve as a platform for establishing a community that blends theoreticians and practitioners interested in the important aspect of anonymity in communication and its applications.

We plan to discuss, study and investigate several aspects of the fields including:

- Anonymous communication in the Internet
- Anonymous communication in sensor and ad-hoc networks
- Anonymous communication to hide control and command
- Privacy issues and the social interests in anonymous communication
- Unidirectional encryption schemes

The scientific relevance of the field to the academia and industry is bold: there are leading researchers in cryptography and networking that are interested and active in the area.

The results so far include possibility/impossibility results, efficiency issues, and system architectures that support anonymity.

The Dagstuhl seminar will be the first time in which researchers and practitioners in the area are gathered to interact, no doubt that the meeting will boost the activity in the area.

8.11 Data Always and Everywhere – Management of Mobile, Ubiquitous, Pervasive, and Sensor Data

Seminar No. **05421** Date **16.10.–21.10.2005**
Organizers: G. Alonso, C.S. Jensen, B. Mitschang

Following decades of rapid and sustained advances in computing and communication technologies, we have reached a stage where it is becoming feasible to embed computing and communication functionality in the physical objects that surround us. For example, such functionality may be embedded into dishwashers, refrigerators, coffee machines, heating systems and even clothing and jewelry. Other, perhaps more mundane examples include PDAs, mobile phones, MP3 players, and car navigation systems. Many of these computationally enabled objects are small and mobile. As they are also able to communicate with their surroundings, e.g., via low-cost transceivers that allow them to spontaneously interconnect with other objects or via a cellular network, they will have access to the Internet and will be accessible from the Internet. Put differently, their data will be part of a global information space, and they will be able to exploit this space.

As a consequence, data is spread all over, which offers a host of new challenges to data management technology. In particular, data synchronization and consistency become substantial challenges. This is also true for data placement that comprises replication techniques as well as caching technology. This situation is further complicated by these data

stores most likely being inherently heterogeneous in terms of data models, storage formats, and access technologies.

This development makes it worthwhile to rethink whether the currently available data management solutions are appropriate for these new scenarios, and how data management technologies that support the whole spectrum would look like. It is exactly the focus of this seminar to discuss issues related to the management of mobile, ubiquitous, and pervasive data.

The seminar brought together representatives from different communities (researchers, software vendors, and users) from different areas (mobile application, middleware, sensor systems, distributed systems, database systems) for joint, in-depth discussions of emerging data management challenges, key objectives being to identify research challenges and standardization needs, and to better understand open problems.

Workshop Itinerary

The seminar started with the following list of potential discussion points phrased as questions and given to the participants up front:

- How to synchronize the data and how to achieve consistency?
- How to integrate the data?
- Where to place the data and which technologies to use?
- How to manage, store, and access the data?
- What does a suitable processing model looks like?
- What kind of communication technology is needed?
- What about a platform approach?
- How to build these kinds of applications?
- What are the killer applications?
- How would a comprising technology for moving objects look like?

The five days were filled with discussions, workgroup meetings, and presentations. The organizers decided on purpose to leave substantial room for discussions and workgroup meetings, to enable the participants to cover new problems and topics that emerged as the seminar progressed and that were considered important for the development of the field.

Workshop Résumé

At the end of the seminar, the participants categorized the issues that were considered during the presentations and discussions throughout the seminar week. The following list of general topics resulted. These topics capture well the breadth of the seminar. The topics were considered to be particularly interesting by the participants, primarily due to research challenges they embody or their potential relevance for future practice.

- Location-based and moving-objects-based applications
- Data/content integration, federation, and management
- Context management and context-aware services
- Data in sensor networks, data streams, and sensor fusion
- Modeling and querying of mobile and spatial databases
- Indexes for high update rates, including the indexing of the past, current, and near-future positions of moving objects
- Replication and caching
- P2P database middleware
- Privacy Issues

It is our belief that the seminar has improved the participants' understanding of the seminar's topic area in general and of the abovementioned topics in particular, has built new collaborations among the seminar participants, and will stimulate further collaborations among members of the different communities involved.

8.12 Service Oriented Computing (SOC)

Seminar No. **05462** Date **15.11.–18.11.2005**
Organizers: F. Casati, B. Krämer, M.P. Papazoglou, F. Cubera

Service-Oriented Computing (SOC) is a new computing paradigm that utilizes services as the basic constructs to support the development of rapid, low-cost and easy composition of distributed applications even in heterogeneous environments. The visionary promise of Service-Oriented Computing is a world of cooperating services where application components are assembled with little effort into a network of services that can be loosely coupled to create flexible dynamic business processes and agile applications that may span organisations and computing platforms. SOC is being shaped by, and increasingly will help shape, modern society as a whole, especially in the areas of dynamic and ondemand business and education, health and government services.

The subject of SOC is vast and enormously complex, spanning many concepts and technologies that find their origins in diverse disciplines that are woven together in an intricate manner. In addition, there is a need to merge technology with an understanding of business processes and organizational structures, a combination of recognizing an enterprise's pain points and the potential solutions that can be applied to correct them. The material in research spans an immense and diverse spectrum of literature, in origin and in character. As a result research activities at both worldwide as well as at European level are very fragmented. This necessitates that a broader vision and perspective be established – one that permeates and transforms the fundamental requirements of complex applications that require the use of the SOC paradigm. This will further enhance the value proposition of Service-Oriented Computing and will facilitate the formulation of a Services Research Roadmap leading to more effective and clearly inter-related solutions and better exploitation of research results.

The research Roadmap provides the necessary background for deciding on potential future research programmes in SOC and places on-going research activities and projects in the broader context of a SOC Roadmap. It launches four pivotal, inherently related, research themes to SOC: service foundations, service composition, service management and monitoring and service-oriented engineering. The viewpoints presented in the research Roadmap partly result from intensive discussions experts with various backgrounds led in parallel workgroup and in plenary sessions.

8.13 Automatic Performance Analysis

Seminar No. **05501** Date **12.12.–16.12.2005**
Organizers: H.M. Gerndt, A. Malony, B.P. Miller, W. Nagel

The Workshop on Automatic Performance Analysis (WAPA 2005, Dagstuhl Seminar 05501), held December 13-16, 2005, brought together performance researchers, developers, and practitioners with the goal of better understanding the methods, techniques, and tools that are needed for the automation of performance analysis for high performance computing.

High Performance Computing is a crucial component of current and future advances in science and engineering. These advances depend on simulations that run on massively parallel systems to obtain results within a reasonable response time. HPC is becoming commonplace, but exists in a wide variety of forms, and the forms dictate what kind of computation will be running efficiently. Most current HPC systems are clustered SMP architectures with a private address space for each node. A few large-scale shared memory systems exist with a NUMA characteristic. Architectural aspects, like the memory hierarchy within a node, the latency and bandwidth of the communication network, and the connection of the node to the network are exposed to the programmer. Their individual complexity and mutual dependencies make the process of iterative program tuning, combining performance measurement, analysis, and optimization, an essential technique for high performance.

The workshop approached this problem by addressing techniques for automating the performance analysis process. Automatic performance analysis will: enable analysis of large application runs with a different behavior than scaled-down versions, allow triggering on-the-fly optimizations, and free the application expert from technical details of the analysis process. New advances in performance analysis automation are timely and achievable, and are the first step towards a fully automated tuning process.

WAPA 2005 builds on the successful Dagstuhl Seminar 02341 on Performance Analysis and Distributed Computing held in August 2002. The workshop is also a continuation of the European Working Group on Automatic Performance Analysis which ended in July, 2005. The experiences from those two activities played an important role in shaping the WAPA focus.

The WAPA program allowed a variety of opportunities for interaction among the workshop participants. Many leaders in the HPC performance community were able to attend WAPA and contribute to the technical content and discussion. As outlined in the attached agenda, the workshop was loosely organized around four themes:

- Performance analysis for large-scale parallel systems
- Automated performance analysis and diagnosis
- Automated performance tuning and performance prediction
- Performance tools and technology

Each theme session consisted of a set of technical talks. In addition, we were fortunate to have two perspective talks by John Levesque (Cray Inc.) and Phil Roth (Oak Ridge National Laboratory, on behalf of Jeff Vetter). Complementing the technical presentations, two panels provided an interesting perspective on the state of performance analysis:

- What works? What's missing? What's all this about productivity?
- Wizard Tales and Santa's Wish List

To complete the program, there was the opportunity for performance tool demonstrations. Eight tools were presented covering the techniques for HPC distributed memory analysis , scalable tracing, open performance technology, performance databases, and performance data mining.To complete the program, there was the opportunity for performance tool demonstrations. Eight tools were presented covering the techniques for HPC distributed memory analysis , scalable tracing, open performance technology, performance databases, and performance data mining.

The goals of WAPA were to increase the exchange of ideas among the tool developers, to transfer knowledge on existing and planned automatic techniques, to engage people supporting application development, and to start a dialog between researchers, developers, and users of automatic performance analysis methods and tools. By all accounts, the WAPA meeting was a tremendous success. Many participants commented on the high

quality of the technical talks and the fruitful discussions they had during the week. And of course, everyone enjoyed our excursion and dinner in the historic city of Trier.

It is the energy and meaningful involvement of each participant that is most responsible for the positive WAPA experience. The organizers would like to thank everyone who came to Schloss Dagstuhl and we look forward to when our paths cross again in the future.

Chapter 9

Modelling, Simulation, Scheduling

9.1 Simulation & Scheduling: Companions or Competitors for Improving the Performance of Manufacturing Systems

Seminar No. **05281** Date **10.07.–15.07.2005**
Organizers: J. Fowler, B. Nelson, M. Pinedo, O. Rose

Motivation

Scheduling is a critical need for manufacturing and service systems, and it is not exaggeration to say that effective and timely scheduling can be the difference between success and failure in an era when customers demand rapid response, product customization and low prices. As a result, there has been an explosion of Deterministic Scheduling papers over the decade with a strong group of researchers from Computer Science, Operational Research, and Industrial Engineering departments.

Discrete Event Simulation (DES) has become a widely used technique to predict and ultimately improve the performance of manufacturing and service systems. In particular, DES is often used in capacity planning, to evaluate order release policies, and to evaluate dispatching policies. However, deterministic scheduling approaches are not often evaluated in this way. There is a well established research community focused on DES issues in Manufacturing and Services, mainly from Computer Science and Industrial Engineering departments. Further, advances in simulation design and analysis methodology is making optimization of DES models feasible.

While both of these research communities are working to improve manufacturing and service system performance, the two groups rarely work together. Our goal was to bring these two communities together to see if synergistic results from interactions between them can be identified. In particular we set out to investigate issues common to both communities such as:

- Should these techniques be used in combination? There is a clear indication that using simulation to evaluate scheduling approaches in a dynamic factory environment can be very fruitful. In the project "Scheduling of Wafer Fabrication Facilities", which was funded by the Semiconductor Research Corporation and International Sematech, a scheduling prototype based on the Shifting Bottleneck Heuristic was developed. There, the majority of the performance testing of this scheduling approach was done by simulation because it was important to see how well the scheduler behaves in an almost realistic environment.

- If the techniques are used in combination, what is the right way to combine these techniques to obtain an optimal result with respect to factory performance, i.e., improved cycle times and on-time delivery? Due to the fact that little research is done in this area of combining scheduling and simulation there are a lot of open issues, including how to provide both simulator and scheduler with a consistent model of the system on which they are working. It is unclear whether both approaches need the same system model or if different levels of abstraction are useful.

- Is there more than simply validating scheduling approaches by simulation? For instance, can simulation be used to schedule as well as to evaluate scheduling algorithms? Or can simulation be used to determine appropriate planning horizons for scheduling algorithms in the presence of uncertainty? Or could simulation be used to evaluate schedule robustness to uncertainty? Runtimes for computer simulations are becoming smaller as computers get faster. Thus, it becomes possible to use simulation not only for validating schedules but also for decision making inside a scheduling method. This is quite different from using simulation or sampling to optimize a deterministic scheduling problem (as occurs in Genetic Algorithms, for instance). Rather, simulation of the schedule is a component of the search for a good schedule.

- Why has there been so little interaction between the scheduling and simulation groups? What are impediments to more interactions? We see a lot of potential in this interaction, in particular, when practitioners from industry and simulation software developers join the group. The practitioners bring up-to-date problems, while the software developers provide the conduit for technology transfer from research to practice.

The desired outcomes of the conference were:

- Ideas for collaborative research

- Plans to organize sessions at future open conferences

- Special issue(s) of journal(s) based on the conference

- Decision on whether or not to apply for future Dagstuhl Seminar

The Seminar

A total of 25 researchers from academia and industry attended the seminar, 32% which were young researchers under the age of 35. Of the participants, 14 were affiliated with institutions from EU member states, 8 from the US, 2 from Singapore, and 1 from Israel. Of the EU participants, 12 were from Germany, 1 from the UK, and 1 from Poland. Five of the participants were from industry and the rest from academia.

Given that there were only 25 attendees, it was planned for every participant to give a talk. The seminar was divided into five main sections: 1) keynote overview talks, 2) detailed individual talks, 3) breakout sessions , 4) report out from breakout groups, and 5) wrap-up and discussion of next steps. Each of these will be discussed below.

Keynote Overview Talks

After brief introductory remarks by Oliver Rose and introductions by all participants, John Fowler gave a talk that provided a framework for seminar discussions on the relationship between deterministic scheduling and simulation. Five basic elements were discussed: a) simulation-based schedule generation and refinement, b) emulation of deterministic scheduling via simulation, c) evaluation of deterministic scheduling via simulation, d) deterministic problem instance generation through simulation methods, and e) simulation for support of scheduling.

Next, two overview talks were given to provide basic knowledge of simulation to the deterministic scheduling participants and to provide basic knowledge of deterministic scheduling to the simulation participants. Barry Nelson gave the simulation talk and John Fowler (substituting for Mike Pinedo who could not attend at the last minute) gave the deterministic scheduling presentation. These talks filled Monday morning.

Detailed Individual Talks

Monday afternoon and all day Tuesday were devoted to individual talks by the participants. The talks were almost evenly divided between talks on deterministic scheduling, talks on simulation, and talks that discussed aspects of both. Please see the seminar web page for abstracts and PowerPoint slides.

Breakout Sessions

Late Tuesday, the participants decided that there should be four breakout groups. A list of eight possible breakout themes was developed and then reduced down to four. The breakout groups were formed around the following themes:

- Simulation-Based Scheduling
- Emulation of Scheduling via Simulation

- Evaluation of Scheduling via Simulation
- Infrastructure for the Support of Simulation.

The breakout groups were given the following charge:

- Identify issues relevant to topic
- Determine key issues to address
- Develop a list of future needs
- Generate ideas for future collaboration

The groups spent Wednesday and Thursday mornings meeting among themselves.

Reports from Breakout Groups

On Thursday afternoon, the groups reported on their activities. The PowerPoint presentations are all posted on the seminar web page. All of the groups did a good job of meeting the charge they were given. Each group decided that they would develop a journal article around their topic.

Wrap-up and Next Steps

Friday morning was devoted to discussing next steps from the seminar. The group decided that the following next steps were appropriate:

- Organize sessions for the Simulation-Based Scheduling track of the Winter Simulation Conference
- Organize sessions at the next MISTA conference
- Continue to look for other opportunities for organizing sessions at major simulation, deterministic scheduling, computer science, and operational research conferences
- A Special Issue of Journal of Scheduling devoted to the seminar
- A Special Issue of Simulation: Transactions of the SCS devoted to the seminar
- A proposal for another Dagstuhl Seminar on this topic in 2007

Efforts are currently underway on all of these recommendations.

Chapter 10

Bioinformatics

10.1 Managing and Mining Genome Information: Frontiers in Bioinformatics

Seminar No. **05441** Date **30.10.–04.11.2005**
Organizers: J. Blazewicz, J.Ch. Freytag, M. Vingron

Bioinformatics has evolved at the interface of biology (especially molecular biology), mathematics, and computer science. Its main goal is to develop mathematical models of biological phenomena, especially at a molecular level. The models are then used to construct algorithmic methods for the analysis of biomolecular sequences, structures, and more recently functional data and gene networks. In particular, it has become clear during the last years that only a small fraction of the human genome encodes proteins. This leaves the non-coding DNA responsible for the regulatory functions, i.e. the encoding when and where a gene becomes active. However, while algorithms development has constituted an emphasis of bioinformatics for a long time, data used to be stored in flat files with little importance attributed to data base and knowledge management issues. Under the pressure of incoming data, this is currently changing and increasing efforts are dedicated to knowledge management in combination with data analysis in molecular biology and genome research.

Bioinformatics in Dagstuhl. Since 1992, several Dagstuhl Seminars on Molecular Bioinformatics took place. Some of these dealt with Bioinformatics in general, while others focused on the more specific topic of Metabolic Pathways. These seminars successfully brought together computer scientists and applied mathematicians with biochemists and molecular biologists in order to discuss possibilities of cooperation in the growing field of analysis of biomolecular sequences and structures. The seminars were seen to be extremely fruitful by the participants. Many contacts were established that provided the basis for future cooperation. In particular, a number of application problems could be formulated providing computer scientists with a sound basis for further work. Many German computer scientists who participated in these seminars have since taken on the challenge of working in bioinformatics .

This seminar understands itself in the tradition of the earlier Dagstuhl Bioinformatics meetings. At the same time the scientific focus shall be shifted to the interplay between biological data management and analysis methods for biological data. This particular area has not constituted the focus of a Dagstuhl Bioinformatics Seminar before. It is of particular relevance today due to the rapid developments of novel high-throughput methods for probing gene function ("functional genomics"). These methods have produced a data deluge unprecedented in the life sciences. Traditional as well a newly developed analysis methods need to be applied to these data. In this situation it is of prime importance to interlink data management and analysis methods in the most efficient and flexible manner.

10.2 Computational Proteomics

Seminar No. **05471** Date **20.11.–25.11.2005**
Organizers: C. Huber, O. Kohlbacher, K. Reinert

The Dagstuhl Seminar on Computational Proteomics brought together researchers from computer science and from proteomics to discuss the state of the art and future developments at the interface between experiment and theory. This interdisciplinary exchange covered a wide range of topics, from new experimental methods resulting in more complex data we will have to expect in the future to purely theoretical studies of what level of experimental accuracy is required in order to solve certain problems. A particular focus was also on the application side, where the participants discussed more complex experimental methodologies that are enabled by more sophisticated computational techniques. Quantitative aspects of protein expression analysis as well as posttranslational modifications in the context of disease development and diagnosis were discussed. The seminar sparked a number of new ideas and collaborations and resulted in joint grant applications and publications.

Seminar Conclusion
The seminar on Computational Proteomics was a full success, as has been confirmed by its participants. Bringing together scientist from different communities – from computer science and life sciences – turned out to be fruitful indeed. Traditionally, proteomics and bioinformatics/computer science are mostly disjoint communities with separate meetings and conferences. The chance to get insights into the problems and challenges both of the experimental and computational world, the need to learn and understand the idiosyncratic "languages" and "vocabulary" of the different disciplines was well appreciated by the attendants. Validation of proteomics data generation and evaluation was spotted as one of the most challenging issues in the application of proteomics as a technology for clinical diagnosis and monitoring. Participants from the two communities were exposed to new ideas, concepts, and techniques – both experimentally and computationally – they were not previously aware of. These ideas were then discussed over a glass of wine or two until late at night. The seminar produced a number of personal contacts which was positively remarked by the participants. In addition to the interaction and personal contacts of the

attendants, the quiet atmosphere of the location also allowed ample time for developing new ideas for solving proteomic challenges.